Welcome to DS5030!

We are glad you are joining us 🎉

Sarah Groves

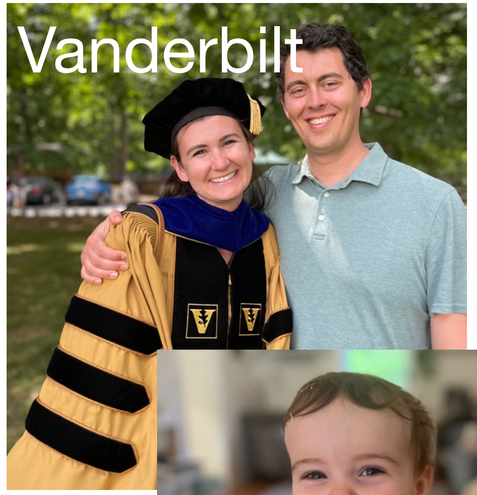

W&M






Coffee


Baking


Vanderbilt


Hannah, 1 yo

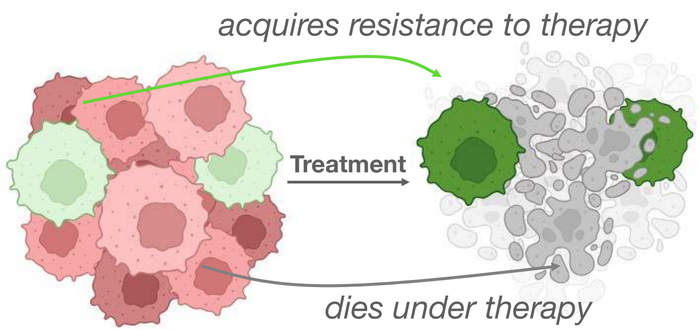

Mary, 3.5yo


Hiking

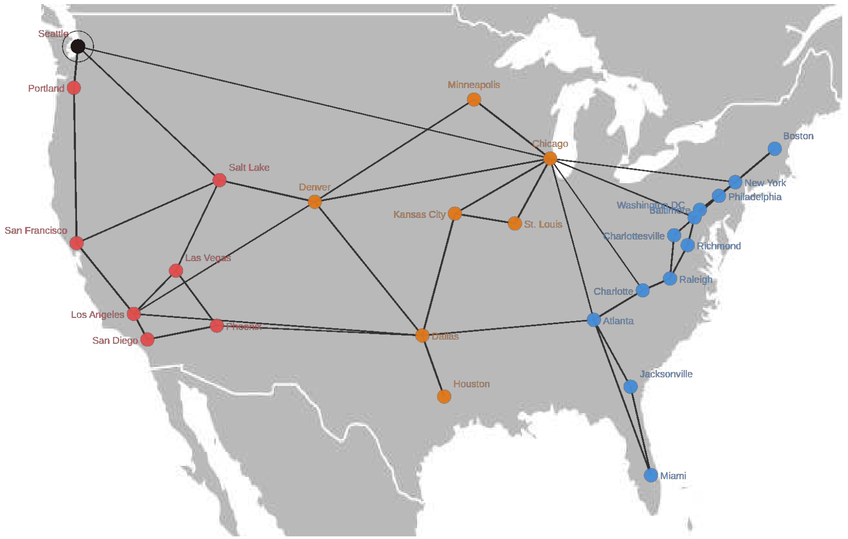

Card & board games


Cooking

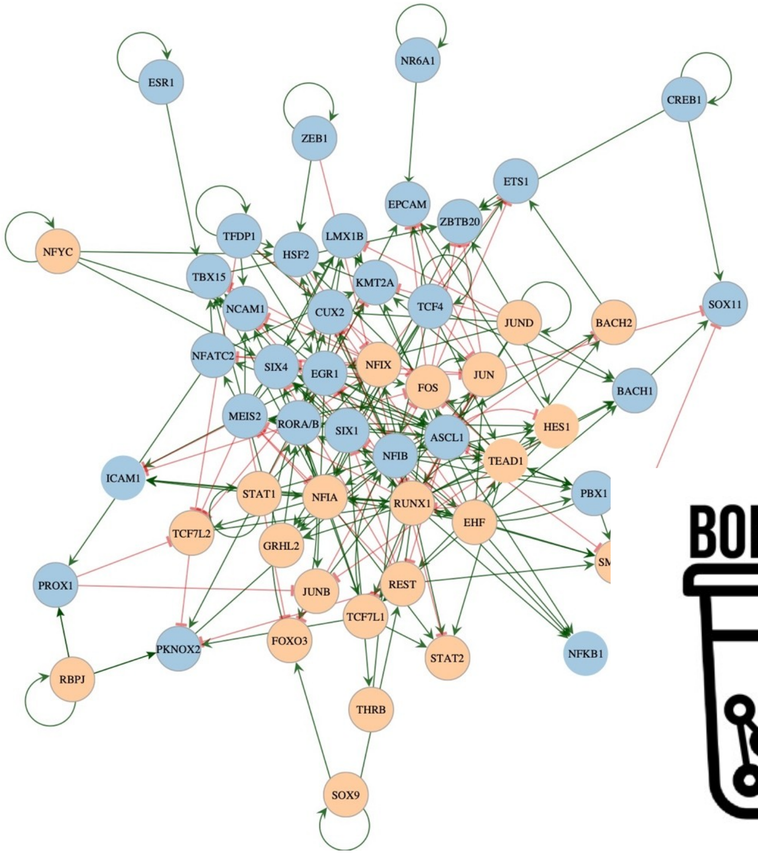
My research: biomedical network science



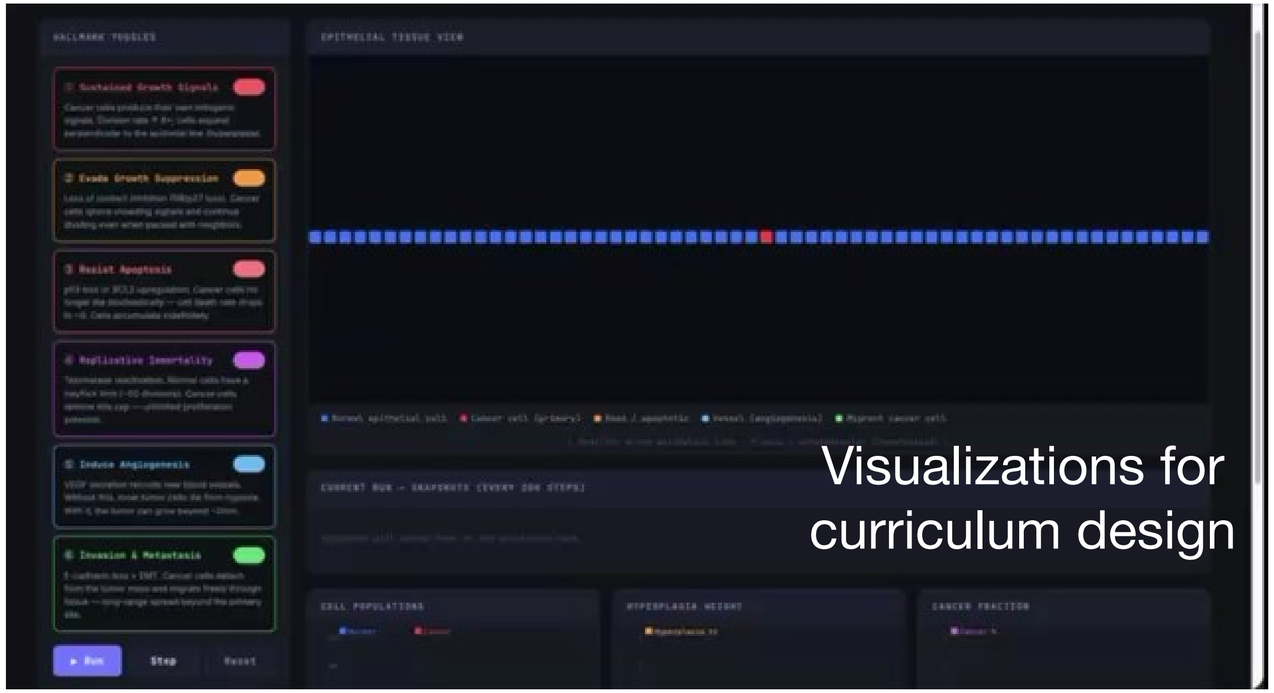
“NETWORKS, NETWORKS EVERYWHERE”



Transit networks and traffic





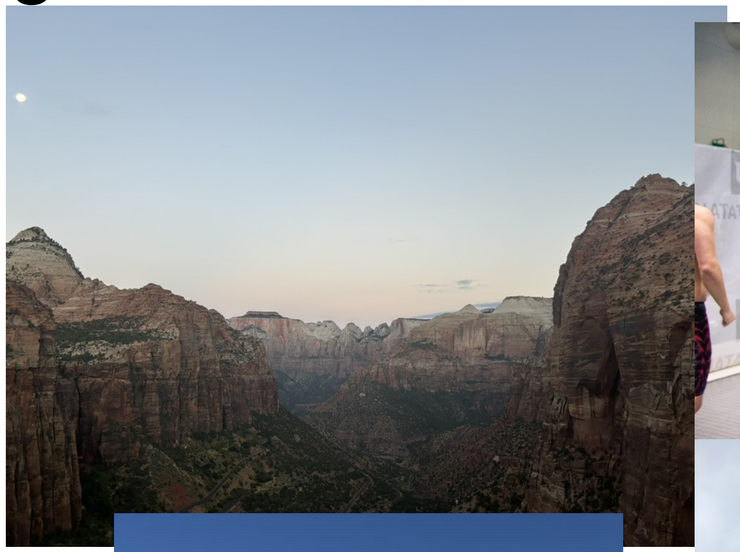


Visualizations for curriculum design

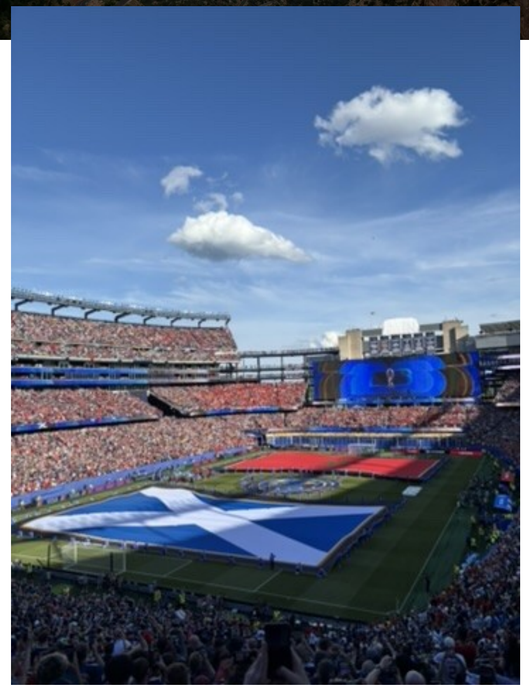
Instructional Assistant: Jeffrey Zhang

About Me!

Hometown: Sudbury, MA
Undergrad: Stats @ UVA
Grad: MSDS @ UVA ’26
ECs: Club Swim, Graduate Data Science Council, Peer Financial Counseling
Interests: Watching my favorite soccer team lose on Saturday mornings (Chelsea FC), World Cup
Hobbies: Visiting National Parks
- Favorite: Zion
- Dream: Patagonia
Things I Did This Summer: Data Science Internship, Hawaii, Noah Kahan @ Fenway









Make a name card!

- Your name
- What field are you in or interested in?

What is this class?

Course Learning Objectives:

- **Complement** and **foreshadow** your machine learning courses
- **AI-proof your career**, as you build **foundational knowledge of the inner mathematical workings** of modeling approaches that you will learn in other classes and use in your career as a data scientist (or data science-adjacent field)
- Develop a **mathematical mindset** and feel comfortable manipulating variables, working with mathematical equations, and reading/understanding statistical plots
- Understand how to work with **random variables** and **uncertainty** to mitigate risk
- Develop your **probabilistic toolbox** of common distributions and densities, axioms, and theorems needed for data science
- Understand and use **maximum likelihood estimation**, **Markov chain** modeling, and **optimization**

Shape of the semester

1. Introduction to describing data: data wrangling, vectors, matrices, and probability
2. Probability, in depth: random variables, densities, and distributions
3. Uncertainty about estimates: sampling distributions, bootstrapping, the CLT, simulations
4. Modeling uncertainty: likelihood, MLE, regression
5. Conditional probability in action: Markov chains
6. Optimization under uncertainty

almost everything this course estimates is the answer to a minimization problem

Graded assignments in pink
Not all graded for correctness

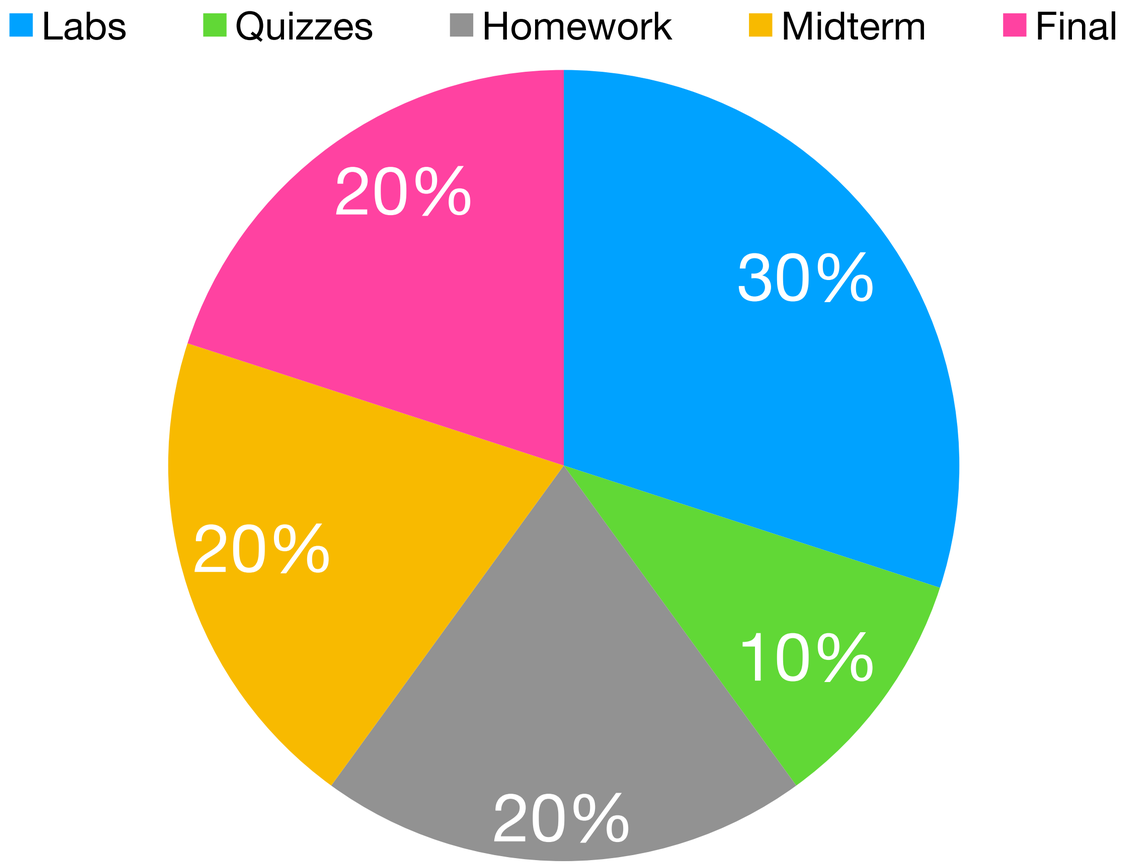
Class Schedule & Assignments

- **BEFORE CLASS EACH WEEK**
  - You will have a video to watch
  - Problem set at home, assigned Tuesday and turned in the **following Wednesday**
- **TUESDAY: QUIZ, LECTURE, MATH [FIRST QUIZ IS NEXT TUESDAY 9/1]**
  - Starts with a short quiz; recall of previous weeks' material, lowest 2 dropped
  - No make-ups for unexcused absences
  - In class lecture, working through problems
  - Don't let AI learn for you!
- **THURSDAY: RECAP, LAB [FIRST LAB NEXT CLASS 8/27]**
  - Short lecture/review and lab (I do, we do, you do): due the **following Wednesday**
- **Midterm/Final**
  - You will have two exams: **October 8** (after Fall break) and on **December 10** (9-12)

Grading

- 20%: Problem Sets
- 10%: Quizzes [drop lowest 2]
- 30%: Labs
- 20%: Midterm
- 20%: Final

If you can’t make class, please email me as soon as possible (preferably before class!)



My expectations of each student:

- Show up
- *Engage*
- Be kind to peers and yourself; let yourself fail
- Trust that the course structure and assessments are intentional for your learning but ask me for my reasoning if it isn’t clear!
- Ask questions and reach out for help when you need it
- Give me feedback about how your learning is going and how we can help
- Be patient with me as I teach this course for the first time

We have a social contract: I will teach to the best of my ability and you will learn to the best of your ability. Neither of us can do our part without the other

Student expectations of me & this class:

- Clear communication
- Office hours/availability
- Making sure you are introduced to the material before an assignment
- Flexibility/adaptability
- Challenge

AI: the good and bad

GOOD

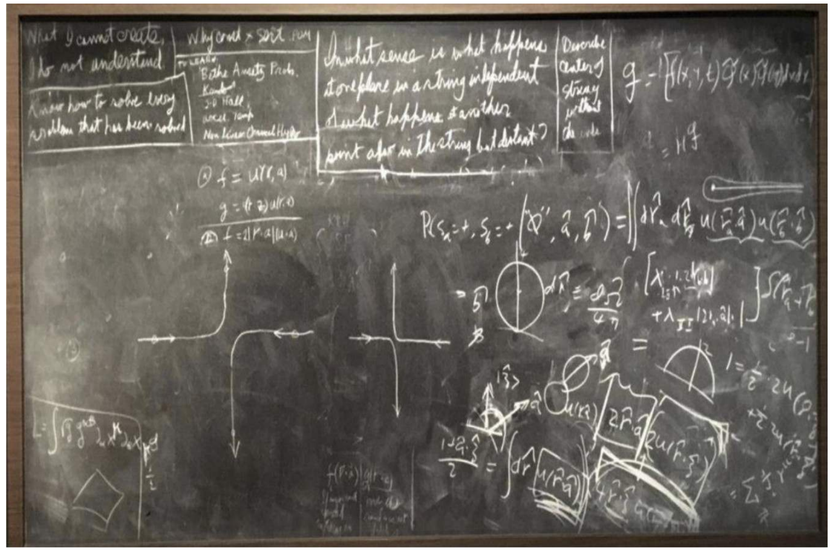
- Researching and editing/tool
- Learning tool/critiquing tool
- Brainstorming
- Personalization/memory

BAD

- Don’t use it to learn for you
- We need to fail to learn
- Takes out critical thinking
- Hallucination

AI Policy: my expectations for you

- See the syllabus
- In general, you can use AI to help your learning, but don’t use AI to replace your learning
  - AI convinces you that you *know* something, without ever encoding it in long term memory
  - Using AI to do things for you handicaps the procedural learning process severely
- If you do use AI, cite it



“What I cannot create, I do not understand”
Feynman’s blackboard

AI: your expectations for me

- Don’t use it grade