EBIO 5460
Data Science for Biological Research

Dr Brett Melbourne
Associate Professor, EBIO
brett.melbourne@colorado.edu
Office hours: Any time by appointment
Office: Ramaley N336 and Zoom
Pronouns: he, him, his

Today

- Introductions (20 mins)
- How are we going to do this class (15 mins)?
- Pretest & survey

Introductions

- Name
- Masters or PhD (what year)?
- Advisor
- Department
- What fascinates you (your research)?
- Hopes for the course

Data Science

- You will be confident to use a range of skills and concepts to:
- plan for, acquire, manage, analyze, infer or predict from, and report about datasets of any size in your area of biological research
- Focus: workflows and algorithms to learn from data

Algorithms and models

- Understand the broad classes
- Frequentist, Bayesian, likelihood, information theory, predictionist
- Emphasize multi-level linear models
- We'll start by considering a simple linear relationship between y and x from all 5 perspectives and a range of algorithms

Learning format

- Flipped, mostly. Short video lectures. Sometimes short live lectures.
- Collaborative learning. Work in small groups or share in small groups.
- Slack: collaboratively discuss the preclass work. Such collaborative learning is not only allowed but **encouraged** in this class!

Git & GitHub

- Class Github organization
- Bookmark this:
- https://github.com/EBIO5460Fall2021

Texts

- All on Google Drive or open source
- I'll provide all materials.
- This one is worth buying:
- McElreath, R (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan.*

Grading

- GitHub portfolio
- 50% continuous Github code commits
- 50% individual assignment

Pretest - survey

- https://cuboulder.qualtrics.com/jfe/form/SV_3P0lJYiV8ErUFgy

Homework

- Posted to GitHub
  - "preclass4wed"
- Update R & R studio
- Set up GitHub
- Reading for Wednesday's discussion
  - Send me an email with your thoughts on these before class