

# Welcome to Math 50!

For all course information see the course webpage.

Math 50 is about:

- Extracting and interpreting relationships between variables from data.

Math 50 is about:

- Extracting and interpreting relationships between variables from data.
- How do we determine if two variables are related? How can we understand the underlying source of this relationship?

What type of background do you need to take the course?

What type of background do you need to take the course?

- “Comfortable” coding. ChatGPT allowed for HW, but will need to understand simple code on a closed book exam

What type of background do you need to take the course?

- “Comfortable” coding. ChatGPT allowed for HW, but will need to understand simple code on a closed book exam
- Some: Take simple derivatives, integrals, knowing basic matrix operations is helpful, but not required (e.g. multiple a vector by a matrix)

What type of background do you need to take the course?

- “Comfortable” coding. ChatGPT allowed for HW, but will need to understand simple code on a closed book exam
- Some: Take simple derivatives, integrals, knowing basic matrix operations is helpful, but not required (e.g. multiple a vector by a matrix)
- Not stats background is needed (e.g. can skip Math 10), but it might be helpful.

# Comparison to other courses

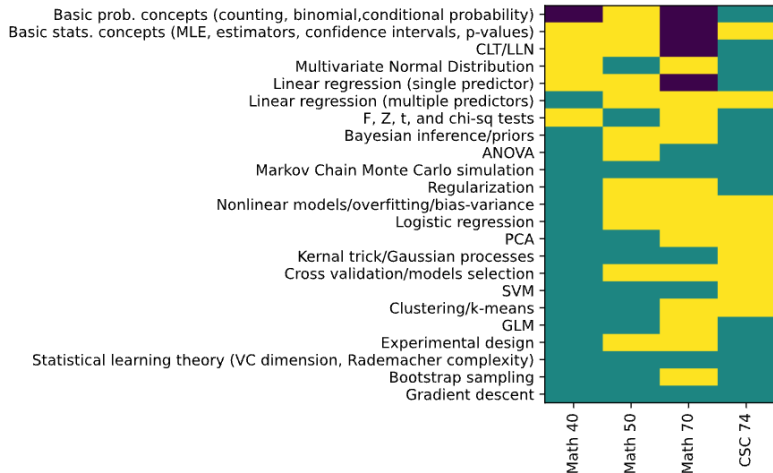


Figure: Topics covered in similar Dartmouth courses – work in progress



# Comparison to other courses

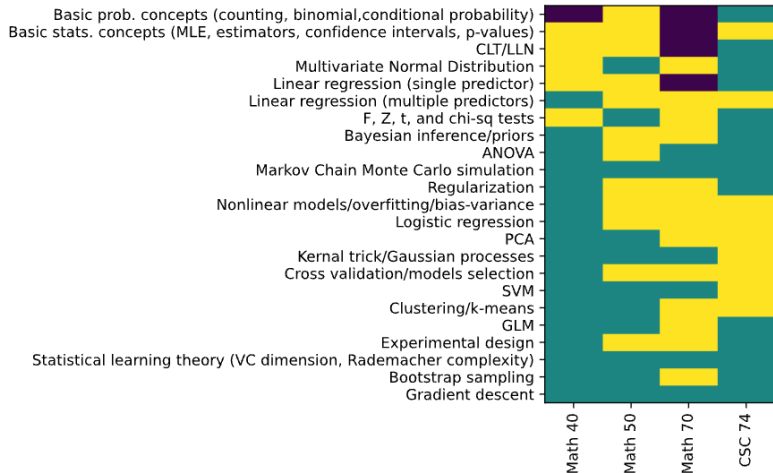


Figure: Topics covered in similar Dartmouth courses – work in progress

Lot's of overlap, but key differences

- **Math 40:** Deeper dive into basics. More rigorous math, less coding.
- **Math 70:** More broadly focused on multivariate stats (including clustering, PCA). Less emphasis on computational/coding, hands on examples. Less connection to ML.
- **Math 74:** Less probabilistic focus, more broadly focused on ML methods (including kernel methods and SVM). More focus on prediction vs. inference.

Lot's of overlap, but key differences

- **Math 40:** Deeper dive into basics. More rigorous math, less coding.
- **Math 70:** More broadly focused on multivariate stats (including clustering, PCA). Less emphasis on computational/coding, hands on examples. Less connection to ML.
- **Math 74:** Less probabilistic focus, more broadly focused on ML methods (including kernel methods and SVM). More focus on prediction vs. inference.

For MDS, but take 40,70 and 50. 40 before 70.

In addition to the specific topics mentioned above, there are some general themes which will color how we explore this world of regression modeling

- **Correlation vs. causation:** When can we tell the difference?
- **Bayesian vs. frequentist:** Different ways of seeing the same thing?
- **Classical statistics vs. machine learning:** Is statistics “just” machine learning?
- **Experimental mindset:** Learn by running computational experiments rather than doing complex mathematical calculations

# Course structure

## Class time:

- **Lecture:** Usually Monday/Wednesday. **No phones or computers please!**
- **Hands on coding:** Usually Friday, sometimes Wednesday, bring laptop, or tablet, but don't need to – see schedule for which day.
- **Exams:** Midterm and Final in class.

# Course structure

## Class time:

- **Lecture:** Usually Monday/Wednesday. **No phones or computers please!**
- **Hands on coding:** Usually Friday, sometimes Wednesday, bring laptop, or tablet, but don't need to – see schedule for which day.
- **Exams:** Midterm and Final in class.

## Outside of class:

- Readings  $\approx$  3-4 hours/week.
- Supplemental resources (e.g. videos)  $\approx$  1 hours/week
- Practice/reflection (HW, practice problems, practice exams)  $\approx$  5-8 hours/week

## How to succeed in this course

- **Do the readings:** Don't just read, ask yourself questions, work through examples, try to challenge claims made in the text.
- **Come to class:** Don't look at your phone during class. Ask questions.
- **Come to office hours:** Helpful if you have concrete questions, but okay if not.
- **Have an open mind:** Just because you've seen something in a different course doesn't mean you won't benefit from thinking deeply about it again.

## What to expect from me

- **Availability:** Office hours and xhours (4 hours total outside of class available to both sections).
- **Responsiveness:** I will respond in 24 hours on weekdays.
- **Grading:** Detailed grading of midterm within 2 weeks (hopefully 1). No detailed grading of problem sets (graders will review for completeness).
- **Material:** I put a lot of thought and effort to design an novel course which caters to the diverse needs to students who take math 50; however, over the years I've realized I don't have the time to prepare original material which is as polished as students like. This why I'm heavily relying on the two textbooks this term. My course notes should be understood as a summary of material covered. Refer to textbooks/readings for technical details.