

# MS&E 125: Intro to Applied Statistics

## Introduction

Professor Udell

Management Science and Engineering  
Stanford

March 3, 2023

# Outline

## MS&E 125: Intro to Applied Statistics

want to take this class?

▶ **ASAP:**

- ▶ enroll (or drop) (or get on wait list)
- ▶ fill out course survey
- ▶ sign up for discussion forum
- ▶ sign up for polleverywhere

▶ **Thursday 9/2/2021:** homework 0

links on course website:

<https://people.orie.cornell.edu/mru8/orie4741/>

## Course staff

- ▶ Prof. Madeleine Udell
- ▶ CA: Mike Van Ness (MS&E PhD)
- ▶ CA: Josh Grossman (MS&E PhD)

\* who am I? \* who are you? \* what is this class about? \* intro  
- prereqs: basic coding, basic stats \* applied - I want to teach  
the most useful things - the most applied: data injection and  
preprocessing, assessing data, selecting data, making sense of  
results - there is nothing more useful than a good theory: how  
do we know? can we be sure? \* statistics - quantitative  
assessment for important questions - prediction - models and  
inference - uncertainty - causality \* walk through topics,  
examples, and applications \* rolling out a new feature to  
ChatGPT \* predicting and mitigating heart failure \* \* grading  
and policies \* questions for students \* what's your major \* is  
this course required for you? \* what's a job you're interested in?  
\* how should we use generative AI in this course? rate each use  
on a scale of appropriate/useful vs cheating: \* ok for  
homework? \* ok for projects? \* ok for quizzes? \* should we do  
oral exams or presentations? \* what fraction of assessment  
weight should be in-class? \* allocate weight to each of the  
following: \* homework \* quizzes \* remote \* in class \* projects \*  
report \* presentation \* exams \* midterm \* final \* participation

\* in-class questions \* discussion forum

## Who am I?

academic

- ▶ B.S. in Mathematics and Physics at Yale
- ▶ Ph.D. in Computational and Mathematical Engineering at Stanford
- ▶ postdoctoral fellow at the Center for the Mathematics of Information at Caltech
- ▶ professor OR at Cornell / in MS&E at Stanford

# Who am I?

## academic

- ▶ B.S. in Mathematics and Physics at Yale
- ▶ Ph.D. in Computational and Mathematical Engineering at Stanford
- ▶ postdoctoral fellow at the Center for the Mathematics of Information at Caltech
- ▶ professor OR at Cornell / in MS&E at Stanford

## applied work

- ▶ finance: Goldman Sachs, BlackRock, Capital One, Schonfeld, Two Sigma, ...
- ▶ tech: Google, Retina.ai, Marketing Attribution
- ▶ cybersecurity: DARPA, Expanse (formerly Qadium)
- ▶ clean energy: Aurora Solar
- ▶ politics: Obama 2012
- ▶ also healthcare, supply chain, ...



## Who are you?

- ▶ Majors: MS&E? Other majors?
- ▶ Future jobs: PMs, finance analysts, consultants, founders, data scientists, ???
- ▶ Year: 1st, sophomore, junior, senior, grad student?

## What is this class about? Intro

prereqs:

- ▶ basic coding
- ▶ basic probability
- ▶ basic linear algebra
- ▶ basic calculus

prereq means: if it's your first time seeing it, it will go very fast!

## What is this class about? Applied

the most useful things

- ▶ data injection and preprocessing
- ▶ assessing data
- ▶ selecting data
- ▶ making sense of results

but also: there is nothing more useful than a good theory

- ▶ how do we know? can we be sure?

# What is this class about? Statistics

- ▶ quantitative assessment for important questions
  - ▶ hypothesis testing, confidence intervals, . . .
- ▶ prediction
- ▶ models and inference
- ▶ uncertainty
- ▶ causality

# Outline

## Course objectives (I)

- ▶ plot
- ▶ predict
- ▶ choose
- ▶ understand

## Course objectives (II)

this course is about

- ▶ learning to ask the right questions
- ▶ learning to understand the answers

at the end of the course, you should know

- ▶ at least one method to solve any problem
- ▶ when **not** to trust your solution

## Course objectives (II)

this course is about

- ▶ learning to ask the right questions
- ▶ learning to understand the answers

at the end of the course, you should know

- ▶ at least one method to solve any problem
- ▶ when **not** to trust your solution

the rest you can learn online. . .



# Outline

## Tech stack

- ▶ In person or Zoom for lectures, section, and office hours
- ▶ Course website for course materials  
(syllabus, schedule, homework, project, etc)
- ▶ poll everywhere for polls [poll Everywhere.com/madeleineudel1824](https://poll Everywhere.com/madeleineudel1824)
- ▶ Ed for Q&A
- ▶ Gradescope for quizzes, submitting homework, grades, solutions
- ▶ Github for code (demos, projects, and hw starter code)

## Course requirements and grading

### **course website:**

(grading, course requirements, lectures, homework, etc)

<https://people.orie.cornell.edu/mru8/orie4741/>

- ▶ (15%) Participation: for every lecture (after this one), use
  - ▶ iClicker for sync lectures
  - ▶ participation form for async lectures
- ▶ (30%) Homework
  - ▶ due every two weeks or so
  - ▶ first one due next Thursday
- ▶ (15%) Quizzes
  - ▶ 30 min quiz every week or so
- ▶ (40%) Project

# Questions

during lecture:

- ▶ ask out loud
- ▶ zoom chat (to everyone, or to a TA)

outside of lecture:

- ▶ ask at office hours
- ▶ ask on discussion forum
- ▶ don't send email

## MS&E 125: Intro to Applied Statistics

want to take this class?

▶ **ASAP:**

- ▶ enroll (or drop) (or get on wait list)
- ▶ fill out course survey
- ▶ sign up for discussion forum
- ▶ sign up for iClicker

▶ **Thursday 9/2/2021:** homework 0

links on course website:

<https://people.orie.cornell.edu/mru8/orie4741/>