Business Experimentation and Causal Methods

Winter 2024
Instructor: Andrey Fradkin
Lecture 1: Introduction

Welcome! The goals for today are:

- Convince you that this class will be **interesting** and **impactful** for your career.
- Discuss the logistics of the class.

Causal claims are everywhere!

Joe Rogan, confined to Spotify, is losing influence

by Ashley Carman, theverge.com August 25, 2021 08:30 AM

Photo by: Illustration Alex Castro / The Verge | Vivian Zink/NBCUniversal via Getty Images

Eight Self-Help Books That Actually Help

ELEANOR CUMMINS
8:00 AM ET

Tesla Stock Drops After Vehicle Delivery Falls Short

Shares were off 14%, making them the biggest decliner on the S&P 500. Tesla said Monday it delivered fewer vehicles than it initially targeted.

- **Heard on the Street: Sales Put Spotlight on Margins**

Why skyscrapers are so short

Words by **Brian Potter**

The height of skyscrapers is limited by physical, economic and regulatory barriers, but we should want to overcome them and build taller. Here's how we can do it.

Read more →

Economics

This class is about evaluating and measuring causal claims.

- A causal claim takes the following form.

  - If a treatment was given to a person (or group), they would have one outcome.

  - If instead, a different treatment would have been given, they would have another outcome.

World A

All platforms

World B

Spotify

Important philosophical point

- We can never see for the same person and time what would happen if they were treated and if they were not treated.

- To do so, we would need a parallel universe.

Example of a causal question in business:

- What would happen to the profits of Starbucks if they raised the price of a latte?

Example of a causal question in business:

- What would happen to the profits of Starbucks if they raised the price of a latte?

- In this class, we want to make causal claims precise!

  - From what old price to what new price?
  - For what locations?
  - How are consumers informed of price change?

Can you come up with a causal question?

About Me

- Ph.D. in economics from Stanford.
- Data Scientist at Airbnb, worked with Meta, Indeed, Rover, Pinterest.
- Many (but not all) class examples from digital platforms:

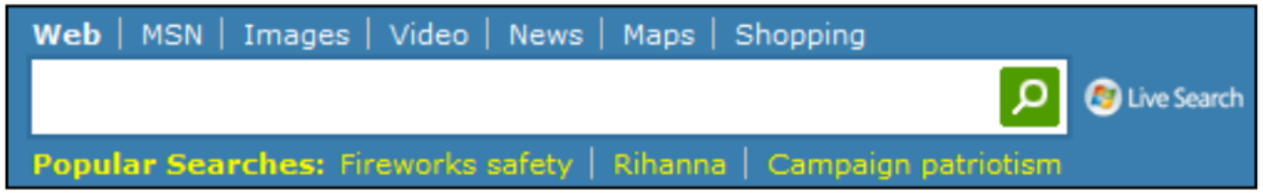
For Fun: Climbing, Strategy Games

Predicting the effects of a change is hard, let's try!

Three experiments from Microsoft Bing, you have to guess the result.

Example 1: MSN Home Page Search Box

A



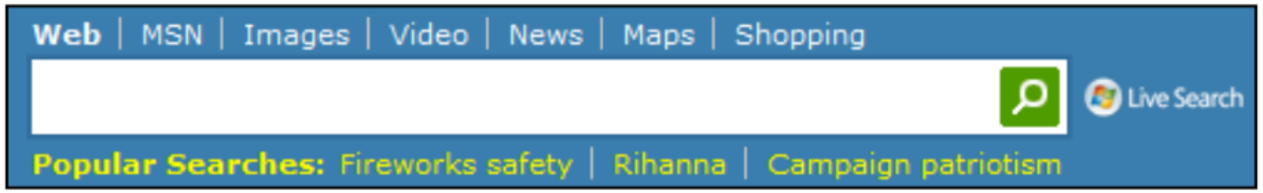
B

Overall Evaluation Criterion (OEC)?

Kohavi, 2017

Example 1: MSN Home Page Search Box

A



B

Overall Evaluation Criterion (OEC):
Clickthrough Rate for Search Box and Popular Searches

Which one is better? A B Same

Kohavi, 2017

Result: No Effect

Example 2: Bing Ads with Site Links

A

B

Overall Evaluation Criterion (OEC):
Revenue (constrained to same amount of page space)

Which one is better? **A** **B** **Same**

Kohavi, 2017

Result: Millions of dollars in profit.

Example 3: Number of Search Results

A: 10 Results

B: 8 Results

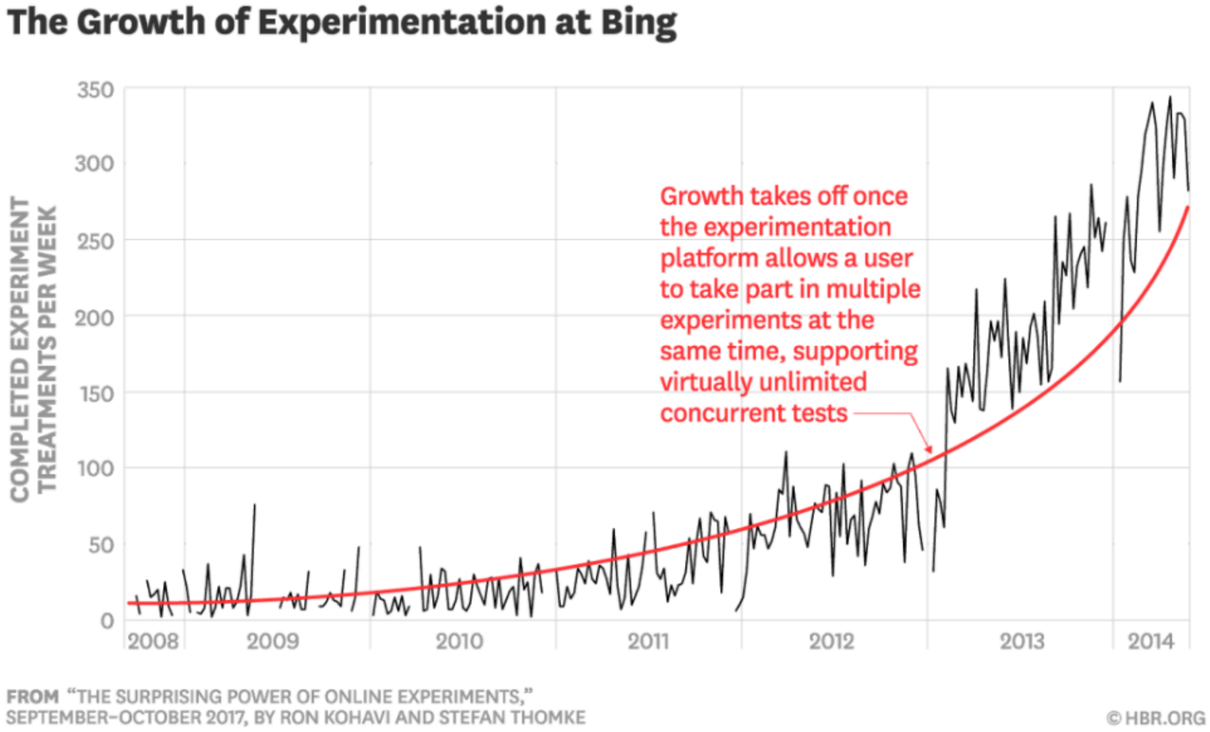
Overall Evaluation Criterion:
Click rate on page.

Which one is better? **A** **B** **Same**

Kohavi, 2017

Result: No Effect!

Business experimentation has grown over the past 15 years



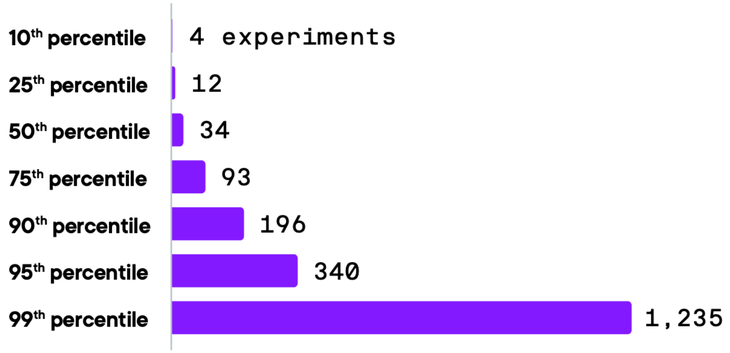
"Our success at Amazon is a function of how many experiments we do per year, per month, per week, per day."

- Jeff Bezos

Data from Optimizely

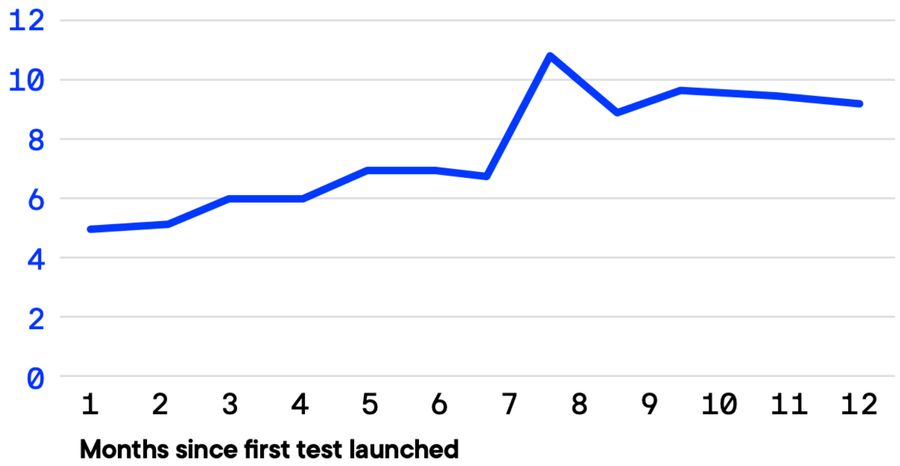
To be in the top 10% of experiment velocity, companies need to run around 200 tests annually

Experiment velocity by company
Experiments created in 2022, over 900 companies



- The median company runs around 3 experiments per month.
- Reaching the top 10% of velocity requires scaling to 16+ tests per month.
- Only 3% of companies are in the elusive 500 tests club.
- The top 1% run over 1,000 tests per year.

Experiments created per month from first test launch on Optimizely
116 companies creating their first experiment 1st December to 1st June 2022



Predicting impact without experiments is hard!

Why Randomize?

Randomization solves selection bias.

- It ensures that as the sample grows large, the only **relevant** difference between treatment and control groups is the treatment!

Selection bias

- Selection bias is any difference in outcomes between the treatment and control group **not caused by the treatment.**
- For example, if all of the treated people were old and all of the not treated people were young, their survival rates may be different because of age.
- Train you to recognize when selection bias may be present.

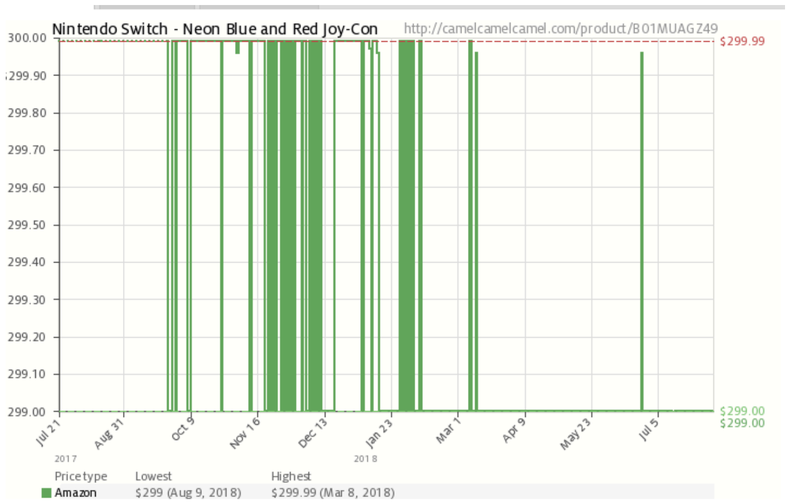
Another reason to experiment: the situation you are interested in has never occurred before!

A product whose price has never changed.

Price of Nintendo Switch (1 year period)

$300

$299



Causal Methods vs Machine Learning?

Machine Learning in 1 Slide

- Training Data: Pictures with muffins or dogs, with **labels**.

- Goal, predict if unlabeled images are muffins or dogs.

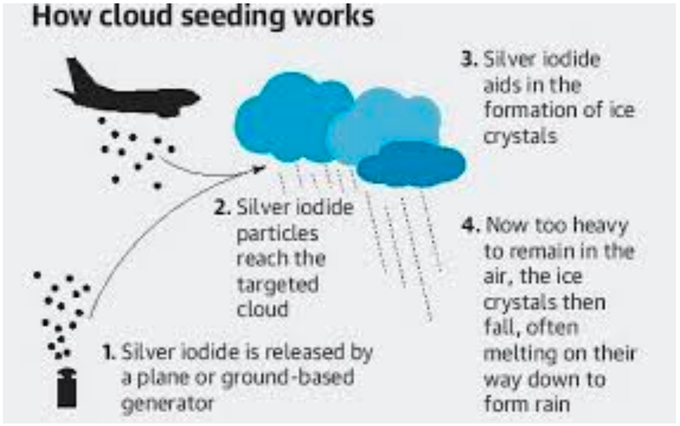
- **Simplified** version of what machine learning does:
  - Finds training data close to test.
  - Takes an average of muffin vs dog for similar training data.
  - That average is the prediction.

“Big Data” = More images.

- Can find closer matches for unlabeled data.
- Generally better prediction accuracy.

Prediction vs Causality

- We can use the machine learning prediction to decide if we need to bring an umbrella!
- Harder to use machine learning to tell us how to create the rain!
- Can use experiments to study if cloud seeding works.



Programming in this class!

Github Codespaces

- Github Codespaces are an easy and self-contained way to code.
- You can choose to use your local machine, but we can provide less support with that.
- We will primarily use Jupyter Notebooks, accessed from Github Codespaces, for assignments and analysis.

Our contract:

- I will do my best to help you learn and use python.
- You must be enthusiastic about learning and using python!
- In the beginning of class, there is a lot of syntax to learn. This may feel overwhelming but you'll get the hang of it.

A note on math.

- This class has some math but it is not a math class.
  - Mathematical notation is very useful in stating what assumptions an analysis makes.
  - This class will **not** focus on the technical details of these statistical procedures
  - More technical and less practical treatment in stats and economics departments, see me if interested.

At the end of the course you will:

- Understand why measuring impact is useful and tricky.
- Understand the basics of experimental design.
- Use Python and Jupyter Notebooks to analyze an experiment, produce figures, and write-up analyses.
- Be able to recognize the causal aspects of business decisions.

Course Logistics

Syllabus is on Notion (link is available on Blackboard / Github).

BA830 Syllabus and Course Resources (Winter 2024)

Business Analytics 830 - Business Experimentation and Causal Methods

Instructor: Professor Andrey Fradkin

Office: Questrom 617, Zoom

Office hours:

YOU MUST BOOK BEFOREHAND AT THIS LINK:
https://calendly.com/afradkin/15min
Typical availability:
TBD

Meeting Times:

Tue, Thursday

12:30 - 3:15 (Afternoon Section), A1

6:30 - 9:15 (Evening Section), B1

In-Person Room:

HAR 324 (Afternoon)

HAR 222 (Evening)

Zoom link for Prof. Fradkin's office hours:
https://bostonu.zoom.us/my/afradkin

TA Office Hours:

Head TA: Liyan Wang (lyanwong@bu.edu)

Friday (10am, room 427 Questrom)

Press 'space' for AI, '/' for commands...

TA: Chandrahas Aroori (charoori@bu.edu)

Monday 2pm (Zoom)

TA: Tamanna Jain (tamannaj@bu.edu)

1pm Friday

https://bostonu.zoom.us/my/tamannajain

Role of in-class time

- Class is for the following activities:
  - Lectures.
  - Group discussions and small lessons to reinforce materials.
  - Working on assignments and group project.
  - Guest Lectures.
  - Doing cases - often the session after the assignment is due.
  - Asking me questions.
  - 1 Exam.

To see what we're doing in class, check the syllabus!

Schedule

Dates	Type	Description	Assignments Due	TODO Before Class	Guest Lecture
January 18, 2024		Introduction to Business Experiments, Course Logistics, Potential Outcomes			
January 23, 2024	Case	Case on Nudges, Intro to Python, Work on Assignment 1		Prepare Nudge Case, Review Python	
January 25, 2024		Discuss Assignment 1, What Makes a Plot Look Good, Lecture On Measuring Uncertainty, Assumptions of causal Inference	Assignment 1 Due	Finish Assignment 1	
January 30, 2024		Lecture on Ads, In-class exercise, work on assignment 2			
February 1, 2024		Experiments in practice lecture, Guest Lecture (Dean Eckles)			Dean Eckles
February 6, 2024	Case	Case, intro to final project, assignment solutions	Assignment 2 Due	Finish Assignment 2	
February 8, 2024		Regression + Analyzing Experiment Lecture, in-class regression exercise. Create regression tables.			
February 13, 2024		Heterogeneity, p-hacking, ethics, Guest Lecture (Madhav Kumar)			Madhav Kumar
February 15, 2024	Case	Case, assignment solutions, give out practice midterm	Assignment 3 Due	Watch John Oliver (https://www.youtube.com/watch?v=0Rnq1NpHdmw), Finish Assignment 3	
February 20, 2024		Midterm Review + Work on final Project			

Two awesome guest lectures lined up:

- Dean Eckles - MIT Professor and world famous data scientist.
  - Work with Facebook, Twitter, Pinterest, GoFundMe

- Madhav Kumar - MIT Postdoc experience with Jet/Walmart and Pinterest.

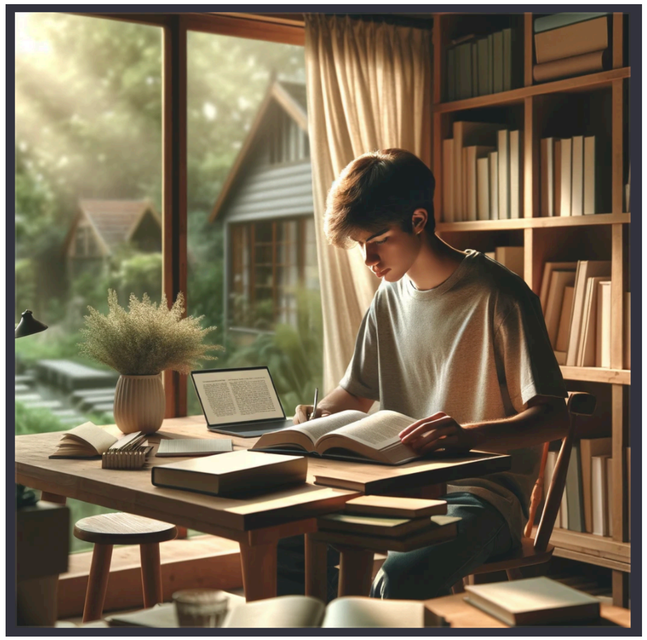
Advancing social studies at MIT Sloan

Associate Professor Dean Eckles studies how our social networks affect our behavior and shape our lives.

Peter Dizikes | MIT News
August 30, 2023

Role of non-class time

- Read course materials:
  - Python notes
  - Causal inference notes
  - Cases
- Finish up parts of assignment not done in class.
- Work on final project.



Course communication:

- Please use slack.
- Post general questions in the channel.
- Ask specific questions about assignments / course organization to the TAs, who can then forward them to me.
- Questions about life, overall course progress, etc… Direct to me.

Office Hours:

- TA office hours in syllabus. Can contact them for additional time.
- I will also have office hours, please book those via the calendly link.

Instructor: Professor Andrey Fradkin

Office: Questrom 617, Zoom

Office hours:

YOU MUST BOOK BEFOREHAND AT THIS LINK:
https://calendly.com/afradkin/15min
Typical availability:
TBD

Zoom link for Prof. Fradkin's office hours:
https://bostonu.zoom.us/my/afradkin

TA Office Hours:

Head TA: Liyan Wang (lyanwong@bu.edu)

Friday (10am, room 427 Questrom)

Press 'space' for AI, '/' for commands...

TA: Chandrahas Aroori (charoori@bu.edu)

Monday 2pm (Zoom)

TA: Tamanna Jain (tamannaj@bu.edu)

1pm Friday

https://bostonu.zoom.us/my/tamannajain

Course Resources

- Python and Github guides are linked to from the syllabus and Github.
- Causal inference notes on Github.
- HBS Course Pack - Please purchase.
- Additional textbooks listed in the syllabus.

Course Materials

📦 Course Pack

I will distribute recorded lectures, notes, slides, and articles as needed throughout the class. We will also be doing several cases, which you must purchase from the Harvard Business School course pack (see link at the top of this document).

📕 Texts

There is no required textbook for this class. I'll provide written notes on the class materials. There are some optional texts you may look at if you'd like:

Optional reading:

- The Effect:

  The Effect: An Introduction to Research Design and Causality | The Effect

  The Effect: An Introduction to Research Design and Causality | The Effect is a textbook that covers the basics and concepts of research design, especially as applied to causal inference from observational data.

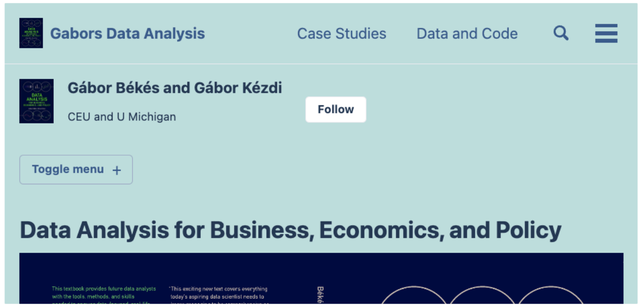
  https://theeffectbook.net/

- *Field Experiments* by Gerber and Green.
  This book focuses on running and analyzing field experiments in political science and economics. It has clear examples and explanations of the potential outcomes framework.

- *Mastering Metrics: The Path From Cause to Effect* by Angrist and Pischke.

  This book has a more regression based take on causality but is still very accessible for those without a math background.

- Data Analysis for Business, Economics, and Policy has many code examples.



Grading

Grade Breakdown

Type	Percent	Group Size
Participation (Including Cases)	12%	Individual
Assignments	28%	Individual
Exam	30%	Individual
Final Project	30%	Groups of 4 - 5

+ New

COUNT 4

Participation

- Very important to practice for interviews + being effective in an organization.
- Biggest failure modes of data scientists:
  - Not understanding project requirements and working on something useless due to bad communication.
  - Failing to convince others their work is useful.

Participation

- Especially important: Class participation during cases. TA will be in the room to track participation.
- Participate during other times too.
- Ask questions and answer questions on **Slack**.
- On-time attendance!
- Work productively during in class exercises.

Assignments

- Mix of conceptual and data analysis questions. Goal is to be short and focused.
- You can ask others for help, but you **must** acknowledge them when you hand in an assignment and you must write your own assignment.
- Due date is on the syllabus and in the assignment.
- Submit on Gradescope.
- Grading based on a check +, check, no credit system.

ChatGPT

- In the real world, you will use chatGPT to code. As a result, I will not ban the use of chatGPT on assignments. That said, chatGPT can be very wrong!

- You will not be able to use chatGPT on the exam, in case discussions, or in class exercises.

> *You may use one or more generative AI tools for the following assignments: homework assignments and final project. In each case, you are required to provide appropriate acknowledgment of the AI tools. This specific approval is given to allow you to become familiar with AI tools, to increase your AI literacy, and to understand how to use AI in appropriate ways within an academic context.*
>
> *You may not use generative AI tools during the case discussions, in class exercises, and exam. You may not use generative AI to answer my questions during class time.*

Midterm

- Goal is to make sure you know the basics of experimentation.
- Important to know this material to do a good job in designing the experiment for your final project.
- Will be done on paper, with 1 page of notes you can bring to class.
- February 22nd.

Final Project (More details later!)

- Groups of 4 - 5 people.
- Run an experiment!
- The experiment should involve you gathering data about the impact of an intervention that you randomize. It should not involve pre-existing data.
- Write about the limitations of the approach.
- Present the project and write it up.

Examples from last year

- Does having your camera on Zoom affect your learning performance?
- Does putting your phone away while completing a task decrease distractions?
- Do people retain knowledge better from reading, listening, or watching (or watching at 1.5x) based on the same content?
- Does having a group of people in a photo of a tourist destination make respondents more or less interested in going there?

Let me know how I’m doing!

- I want to help you by improving my teaching. I can’t know how well I’m doing unless you tell me.
- Contact me with suggestions: fradkin@bu.edu
- Office hours:
  - https://calendly.com/afradkin/15min
  - Please ask class related questions on Slack **first**!
  - If these don’t work for you, just tell me.

For next class:

- Read the HBS Case about Nudges.
- **Write a new tax letter that HMRC could use and be prepared to discuss.**

Assignment 1 is due on Thursday morning at 12:30pm!

Recap:

- Measuring impact is critical for making good decisions.

- Randomization and deliberate intervention allows us to learn impact.

- This class how to run and analyze experiments to learn impact (causal effects).