A/B Testing, Design and Analysis (A/B Testing)

(outline subject to minor changes)

Course Statement and Motivation

When is making a change to a webpage, an algorithm, or products worthwhile? Should I increase or decrease prices? Is my advertising cost-effective? Rather than simply relying on the intuition of managers, businesses increasingly strive to accurately identify the impact of their actions. This is particularly the case in the digital economy where firms have access to large amounts of data.

However, properly identifying the causal effect of a business change requires knowledge of the appropriate methods. One such method is conducting field experiments, also called A/B tests. Yet, implementing experiments is not always straightforward; instead, it requires a thorough understanding of the underlying business problem and the challenges associated with identifying causal effects.

This class teaches methods for measuring the true impact of business changes in the digital economy. Because of the prevalence of A/B testing in practices, we especially emphasize how A/B tests work and how to apply them. We will have hands-on discussions of how managers can measure impact in business situations and how to evaluate claims of impact made by others. While the focus will lie on A/B testing, we will also discuss other methods that can be used when experiments are not feasible. Methods are illustrated with examples from a variety of digital businesses including Microsoft, Amazon, eBay and Uber. We will cover examples in a variety of topical areas including e-mail marketing, technology adoption, advertising, pricing, among others.

Guest speakers from the New York Times and Spotify will discuss the practical importance and challenges of implementing A/B tests.

Learning Outcomes

At the end of the course, students should be able to:

(1) Explain why measuring impact by using the correct causal methods is important
(2) Recognize whether claims are causal when they encounter them in the real world and to be appropriately skeptical of claims.
(3) Identify the conditions that allow one to learn a causal effect from data
(4) Describe why experimentation solves the causal inference problem and identify when natural experiments can approximate actual experiments

(5) Implement field experiments in a variety of empirical contexts, such as advertising or pricing.
(6) Understand the challenges of why firms do not experiment more often and how to potentially overcome those.

Class activity

There will be 14 semi-weekly sessions of 1hr 20min each. The schedule is MW 9.30 am (HBH 1202).

During the lectures there will be open-ended discussions, data exercises, and the presentations by guest speakers. The discussion sessions within lectures are often accompanied by assigned readings. The readings are considered an integral part of the course and will facilitate participation by everyone. The case studies cover a range of different real-life case examples of analysing data for causal inference. You are expected to prepare all case discussions. Depending on the case, this involves reading the case, downloading the data in order to be ready to analyze the data in class, or start the analysis at home. We will welcome multiple guest speakers, each with in-depth experience in measuring impact in the digital economy.

Pre-requisites

The course assumes that you have a working understanding of basic statistical concepts such as hypothesis testing, statistical significance, t-tests and regression. I build on this knowledge and will not be able to teach it from scratch. If you do not have knowledge or are uncertain how much you recall, I strongly encourage you to review these concepts prior to the start of the class.

The course likewise assumes that you are able to manipulate data and use software to plot data, conduct t-tests and run regressions. Again, if you are unsure how to do so or how much you remember, you should familiarize yourself with the techniques prior to class. I do not require that you use a specific statistical software (you can use Stata, R, Excel, or other tools) but you will be asked to document how you came to your results.

The course will provide a data tutorial for these statistical concepts. I will also go over these concepts but not in too much detail. I do not expect students to be experts before starting the course, but hope that they have working knowledge. The course does not require sophisticated knowledge of statistical concepts.

Assessment structure:

There are four assignments which will be made available after lectures 1, 2, 3, 5. All homework deadlines are published on the course website and set since the first lecture. Late homework is not permissible because they are often case-based and discussed in detail in the class. No homework is received with a delay unless it is due to illness (contact me in such cases ahead of the deadline).

Assignments 1, 2, 3, and 4 are to be submitted individually (i.e. each student submits their work). Assignment 1 counts for 5% of the grade while assignment 2 counts for 10%. Assignment 3 counts for 15% of the grade while assignment 4 counts for 20% of the grade. The assignments have to be submitted on Gradescope.

A **group project** (in teams of 2) will count for 30% of the grade. Consult a separate document for details on this.

In-class quizzes: 10% of the grade will be for short quizzes administered through the course during the lectures. The quizzes will last for a maximum of 10 minutes. The aim of the quiz is to simply keep you updated about the material covered in the lectures. No extra reading or preparation is required if you have been following the material and discussions in the class.

Class participation: will account for 10% of the grade. The quality of our learning environment is contingent on the effective and informed participation of each class member. The goal here is not to dazzle us with your individual brilliance or maximize the number of comments you make, but rather to help make the class learn. Note that, in addition to providing new insights to the discussion, it is possible to make the class smarter by asking the right question or by assimilating comments from other students or sources. Contribution is assessed on an individual basis. This is inherently a subjective assessment but will take the following factors into account: preparation, ability to coherently present arguments and to contribute to a discussion. Attendance in class is necessary but not sufficient to get credit for the contribution grade. A Teaching Assistant will sit in every class and record class participation.

Integrity Policy:

All students are urged to follow CMU's policies for academic integrity and plagiarism. In general, students are encouraged to discuss lecture material among themselves. For individual home assignments, students should not consult among themselves. For the assignments that require group study each team is required to submit their team answers without consulting other teams.

Use of AI in coursework: To ensure all students have an equal opportunity to succeed and to preserve the integrity of the course, students are not permitted to submit the output from artificial intelligence (AI) systems such as ChatGPT or any other automated assistance for any classwork or assessments. This includes using AI to generate answers to assignments or using AI to complete any other course-related tasks. Violations of this policy will be treated as academic misconduct. There are tools to detect whether the submitted text is written by ChatGPT or other similar tools. If you have any questions about this policy, please do not hesitate to ask for clarification.

Course attendance:

Class attendance is mandatory for taking the quizzes and getting a minimum grade for class participation. All the take-home assignments will be based on material and discussions in the lectures.

Contacts:

Instructor: Ananya Sen

More information: https://sites.google.com/view/ananyasen/home

Contact via email: ananyase@andrew.cmu.edu

Teaching Assistants: Naveen Bavasaraj (naveentb@cmu.edu), Prachee Talwar

Office Hours: Fridays 1pm-2pm (Naveen Bavasaraj)

Tuesdays 12pm-1pm (Prachee Talwar)

Tentative Calendar:

Lecture 1 (March 11). INTRODUCTION AND EXAMPLES
Introduction to the course. What are experiments and why should we run them.

Lecture 2 (March 13). CAUSAL INFERENCE: OLS AND THE NEED FOR RANDOMIZATION
Introduction to the econometric tools used for causal inference.

DATA ANALYSIS TUTORIAL (MARCH 16 1pm). ON ZOOM
Using Stata, R and Excel to carry out simple statistical analysis (e.g. t-tests, linear regressions). Recording will be made available.

Lecture 3 (March 18). BOOKING CASE
A case highlighting the use of experimentation within companies.

Lecture 4 (March 20). EMAIL MARKETING A/B TEST: DATA EXERCISE I
Analysis of email messaging A/B test with a focus on heterogeneity of treatment effects.

Lecture 5 (March 25). ROCKET FUEL CASE
Measuring the Effectiveness of Online Advertising.

Lecture 6 (March 27). *GUEST LECTURE* ROHIT SUPEKAR, NEW YORK TIMES
The role of A/B testing at NYT focusing on the complementarity of ML and A/B testing.

Lecture 7 (April 1). UBER DATA ANALYSIS CASE I
Innovation at Uber: The Launch of Express POOL

Lecture 8 (April 3). UBER DATA ANALYSIS CASE II
Innovation at Uber: The Launch of Express POOL

Lecture 9 (April 8). *GUEST LECTURE*: EMMA ZETTERDAHL, SPOTIFY
Experimentation at Spotify and introduction to the role of surrogate indices in A/B testing.

Lecture 10 (APRIL 10). DIFFERENCE IN DIFFERENCES ANALYSIS
What happens when you cannot randomize? The use of natural experiments in observational data.

Lecture 11 (APRIL 15). DATA EXERCISE II
Identifying causal effects in observational data using diff-in-diff.

Lecture 12 (April 17). GROUP PROJECT PRESENTATIONS

Lecture 13 (April 22). GROUP PROJECT PRESENTATIONS

Lecture 14 (April 24). ROUND UP OF PRESENTATIONS AND SUMMARY