

# Applied Causality

David M. Blei  
Columbia University  
Spring, 2019

## Description

We will study applied causality, especially as it relates to Bayesian modeling and probabilistic machine learning. Topics include causal graphical models, potential outcomes, and counterfactual inference. We will study applications, such as in economics, genetics, and healthcare. Each student will embark on a semester-long project around applied causal inference.

## Prerequisites

- You have taken STCS6701 *Foundations of Graphical Models*.
- You can derive and implement approximate posterior inference algorithms for new models.
- You have a causal inference research problem in hand.

## Organization

The course is open to doctoral students; no auditors and no pass/fail.

Our meetings are based around discussion. The discussion will focus on the readings and on your projects, including “best practices” for research. I want us to share with each other the tools and techniques that we use to develop ideas, software, and writing. At a higher level, this is a course about doing research.

In the first part of each session, we will discuss the reading. In the second part, one or two students will discuss an aspect of the applied side of the material. This discussion can include

- a tutorial about a tool you like (e.g., Stan, Jupyter, RStudio)
- background about your project and a technical issue
- an important paper or idea that we are not otherwise covering
- anything else you want to talk about related to the course

## Coursework and grade

The coursework is the following:

- weekly readings

- weekly responses to the readings
- a weekly commit to your final project repository (see below)
- a final project

You are graded on completing the responses, working consistently on your project, the project itself, and general participation in the class community.

## **Final project**

The project has two components: a report and a git repository.

The repository contains what you need for the report (except files that are too large like raw data and output of an analysis). It is organized like this:

- `readme.md`
- `journal.md`
- `doc/`
- `src/`
- `etc/`

This repository will document your progress through the semester.

The file `readme.md` simply contains an abstract of the project. At first, it is an “aspirational abstract,” one that describes the project you want to complete. You’ll refine it through the semester.

The file `journal.md` is a diary of your progress. It contains dated entries with a description of what you are up to, what you found, what you are thinking, and so on. It is mainly a resource for you, but I will glance at it too (at the end of the semester). Please have the habit of updating and committing it at least once per week.

The `doc/` directory contains latex documents that you are writing, a subdirectory for each one.

The `src/` directory contains the code you are writing.

The `etc/` directory contains anything else—materials, notes, photos of whiteboards, and so on—that you want to keep track of.

(Feel free to have other directories too, such as `dat/`, `out/`, and `fig/`. But keep the top level directories to three characters.)

Commit often, at least every week. You are graded on the quality of the project and the path that you took to get there.