# DBA4712: Causal Analytics for Managerial Decisions

NUS Business School Department of Analytics & Operations (DAO)

23 May, 2022

### **Administrative Information**

• Instructor: Long ZHAO

- Office: BIZ1 8-61

Office Hours: By AppointmentEmail: longzhao@nus.edu.sg

• Prerequisite: DAO2702 & DAO3803

• Evaluation:

Individual Assignments: 20%
Two Projects: 20% + 20%
Two Quizzes: 15% + 15%
Class Participation: 10%

• Coding: R or Python

#### Course Outline and Schedule

We often use machine learning to find the association of two events, such as: the frequency to the hospital is strongly negatively correlated with life expectancy; or the number of bedrooms strongly positively correlates with HDB price. Such strong associations might lead to non-sensible decisions like not seeing a doctor when getting sick or building a wall to create a new bedroom. Other than the warning "correlation is not causation," can we have a systematic way to disentangle correlation from causation?

This course will enable you to apply causal analysis to a wide range of problems, such as why the experts are confident that the COVID vaccine is effective, what business policy boosts the economy, and which ad recommendation algorithm generates higher revenue. Based on such solid causal relationships, one could make sensible life and business decisions.

Compared with Econometrics I & II, this course depends much less on linear regression. Instead, we will focus much more on intuition of 'like v.s. like' comparison to derive causal relationship.

| Week | Date | Topic                            | Remark           |
|------|------|----------------------------------|------------------|
| 1    | Jan. | Introduction to Causal Analytics | -                |
| 2    | Jan. | Potential Outcome                |                  |
| 3    | Jan. | Random Controlled Trials         |                  |
| 4    | Feb. | A/B Test                         | Assignment 1 Due |
| 5    | Feb. | Multi-Armed Bandit               | In-class Quiz 1  |
| 6    | Feb. | Project 1                        |                  |

| Week | Date | Topic                                   | Remark           |
|------|------|---|------------------|
| _    | Feb. | Recess Week                             |                  |
| 7    | Mar. | Matching                                | Project 1 Due    |
| 8    | Mar. | Event & Regression Discontinuity Design |                  |
| 9    | Mar. | Instrumental variable                   |                  |
| 10   | Mar. | Difference in differences               |                  |
| 11   | Mar. | Machine Learning in Causality           | Assignment 2 Due |
| 12   | Apr. | Project 2                               | In-class Quiz 2  |
| 13   | Apr. | No Class                                |                  |
| -    | Apr. | Reading Week                            | Project 2 Due    |

#### Resources

Optional Videos & Books:

- Causal Inference Bootcamp by Matt Masten
  - These videos are well explained and non-technical. Highly recommended.
- Mastering Econometrics with Joshua Angrist
  - Wonderful videos with excellent animation.
  - Joshua Angrist Nobel Prize Lecture 2021
  - Thank Joven for the recommendation.
- Causal Inference for the Brave and True
  - Very accessible. Lots of code examples.
  - Thank Min Da for the recommendation.
- Causal Inference for Statistics, Social and Biomedical Sciences.
  - Most theoretical materials come from this book.
- Mostly harmless econometrics: an empiricist's companion
  - Regression discontinuity design and instrumental variable depends heavily on this book.
  - Videos: Mastering Mostly Harmless Econometrics

## Individual Assignments & Quizzes

Individual assignments help you to prepare for the quizzes. All of them are conceptual. Here is one example.

An Asian flush reaction is a condition that a person develops flushes on the face, neck, and sometimes entire body after consuming alcoholic beverages. A gene mutation on chromosome 12 causes this reaction. About 30-50% of east Asians carry this mutation, while quite rare among Europeans and Sub-Saharan Africans. The gene mutation is random and has no other known effect. However, we found that the people with the mutation have a lower risk of heart disease than those without. Which method should you use to obtain a causal estimation of drinking on heart disease? Please provide the rationale for the answer.

## Two Group Projects

The first project is quite challenging despite of having detailed instructions. There are two reasons.

1. The setting is so practical that theory can not tell you what to do. You need to carefully think about the tradeoff and try to strike a balance.

2. Sometimes you need to write certain parts of code from scratch.

You could choose the second project. The topic shall be interesting and could be investigated using the methodologies taught in this course.