

Causal Inference

Lecture 1 Introduction to the course

Dr. Xi Chen

6/15/2024



Welcome!





What is causal inference?

- Inferring the effects of any treatment / policy / intervention / business strategy etc.
- Examples:
 - ...Omitted because there are too many of 'em...





Why causal inference?

- Correlation is not causation.





What is so special about causal inference?

- You may have heard of instrumental variables, natural experiments, difference-in-difference etc.
- What is their commonality?



Correlation is not causation



This is the idea from **Karl Pearson** – the inventor of correlation!

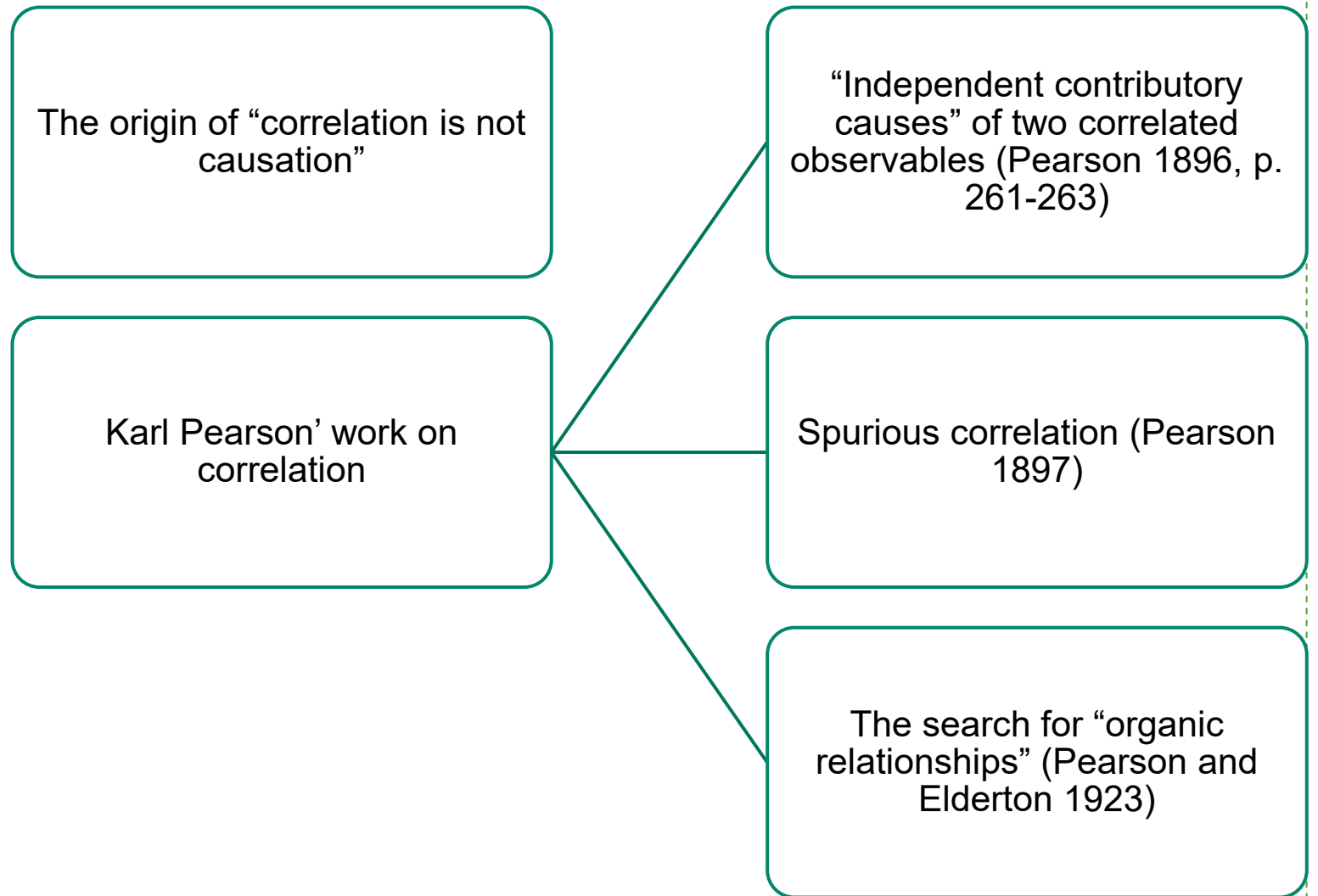


“If the **unit A** be always preceded, accompanied or followed by **B**, and without **A**, **B** does not take place, then we are accustomed to speak of a causal relationship between A and B.” (Pearson and Lee, 1897, p. 459)



“It is the conception of correlation between two occurrences embracing all relationship from absolute independence to complete dependence, which is the wider category by which we have to replace **the old idea of causation.**” (Pearson 1910, p. 157)

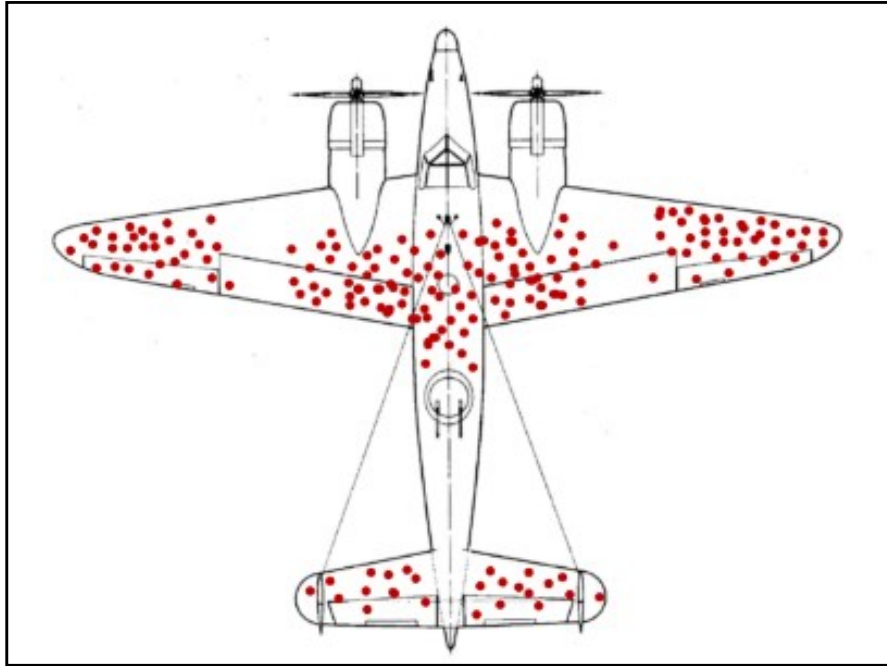
Correlation is not causation



A quote from Pearson

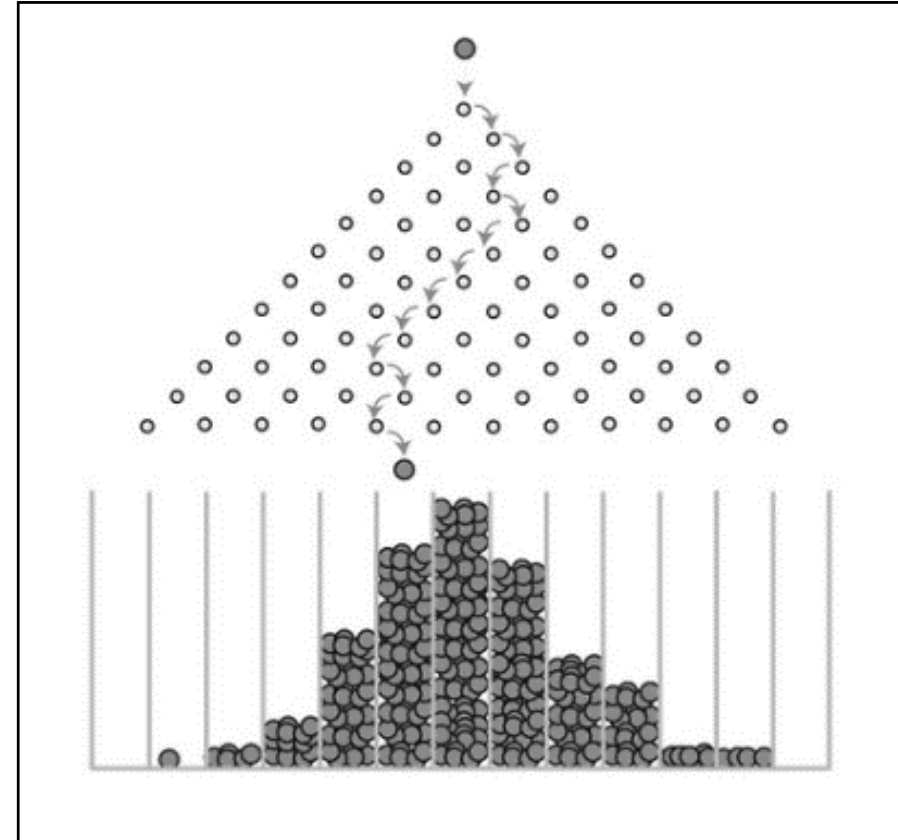
Beyond such discarded fundamentals as “matter” and “force” lies still another **fetish** amidst the **inscrutable arcana** of modern science, namely, the category of cause and effect.

The importance of knowing causality



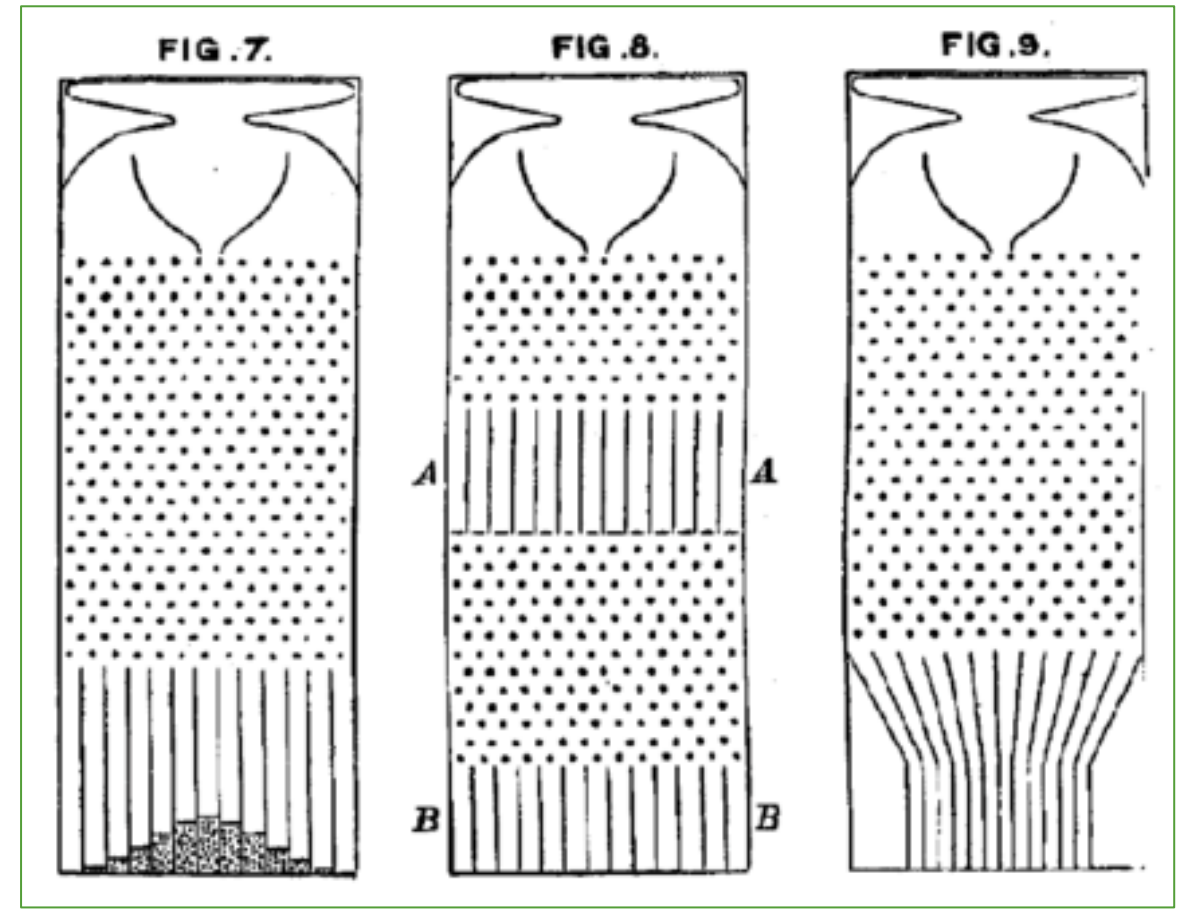
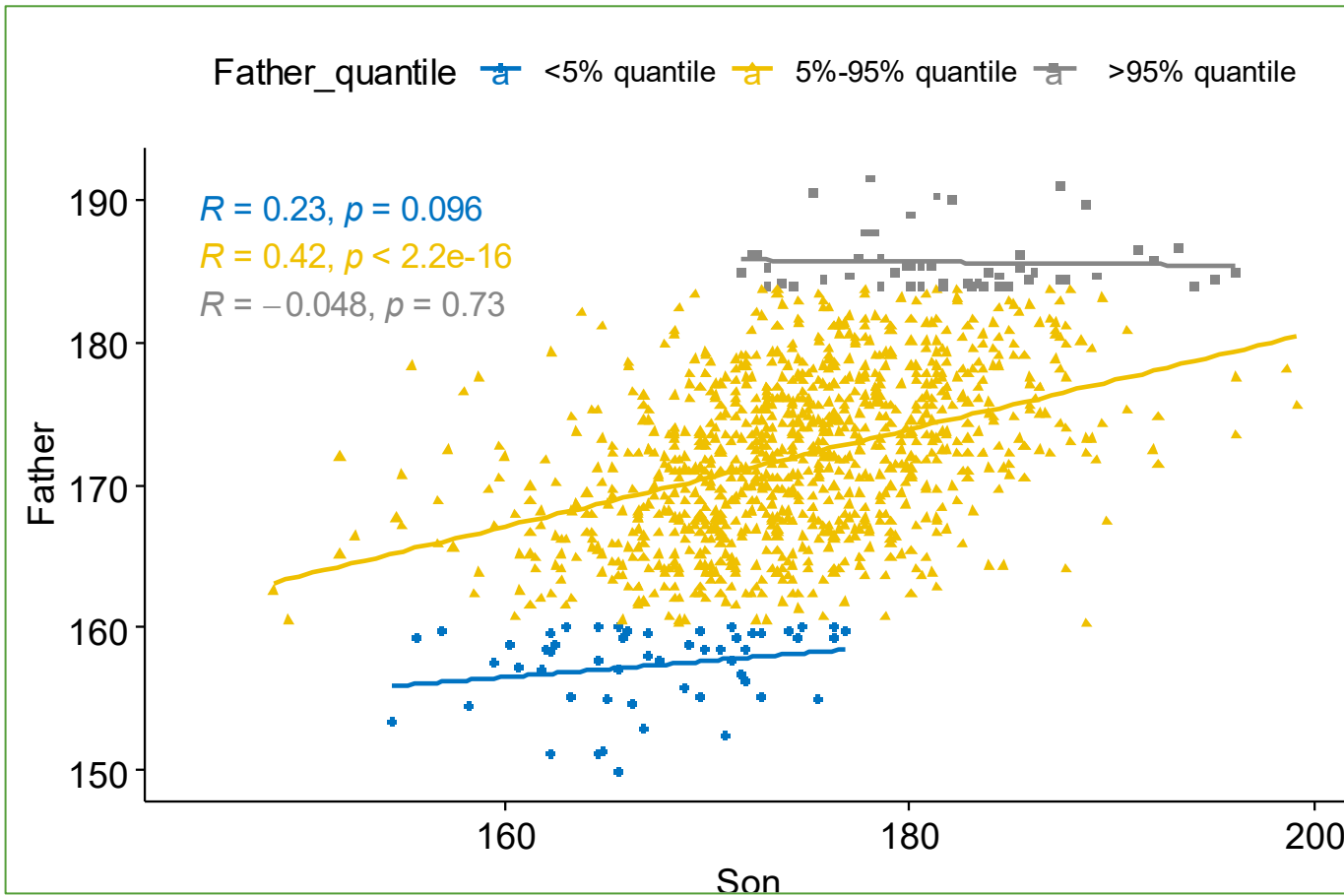
Wald, Abraham. (1943). *A Method of Estimating Plane Vulnerability Based on Damage of Survivors*. Statistical Research Group, Columbia University.

Survival Bias



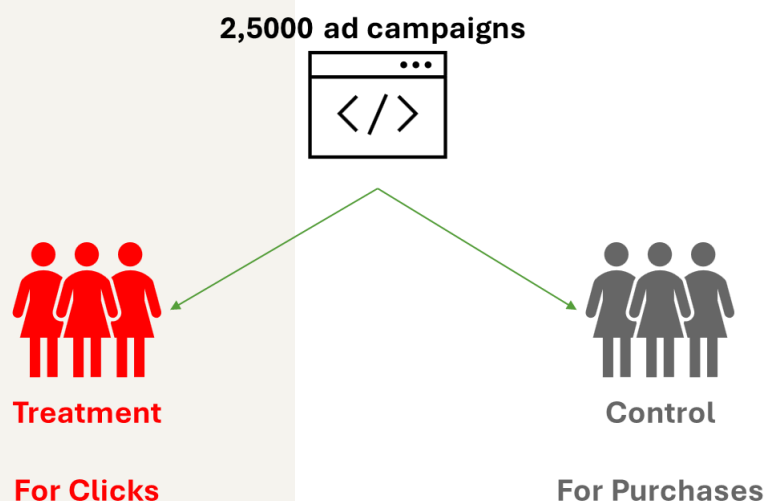
Galton, F. (1894). *Natural inheritance*. Macmillan and Company.

Regression to the Mean



The importance of knowing causality

- A double-layered Galton board to study heredity.
- Regression-to-the-mean.



We're standing up to Apple for small businesses everywhere

At Facebook, small business is at the core of our business. More than 10 million businesses use our advertising tools each month to find new customers, hire employees and engage with their communities.

Many in the small business community have shared concerns about Apple's forced software update, which will limit businesses' ability to run personalized ads and reach their customers effectively.

Forty-four percent of small to medium businesses started or increased their usage of personalized ads on social media during the pandemic, according to a new Deloitte study. Without personalized ads, Facebook data shows that the average small business advertiser stands to see a cut of over 60% in their sales for every dollar they spend.

While limiting how personalized ads can be used does impact larger companies like us, these changes will be devastating to small businesses, adding to the many challenges they face right now.

Small businesses deserve to be heard. We hear your concerns, and we stand with you.

Join us at fb.com/SpeakUpForSmall

FACEBOOK

More examples

Apple vs. Facebook

Allow "Facebook" to track your activity across other companies' apps and websites?

[Here, in addition to other screens, Facebook can explain why users should allow tracking.]

Ask App not to Track

Allow

TV CAMPAIGNS SUPPORTED
BY TWITTER SEE A

+16%
INCREASE IN ROI

The study's design

Studios tend to spend most of their advertising budgets around a film's theatrical release date. To get a comprehensive view of performance, we looked at nearly four years of paid theatrical history data — from December 2011 to June 2015. The research team used multivariate regression analysis (a process that measures and predicts the sales impact of various media channels) to understand the effects of changes in Twitter media for movies. Researchers also looked specifically at comedy and action genres in the US, as these categories were among the biggest box office performers.



Self-introduction

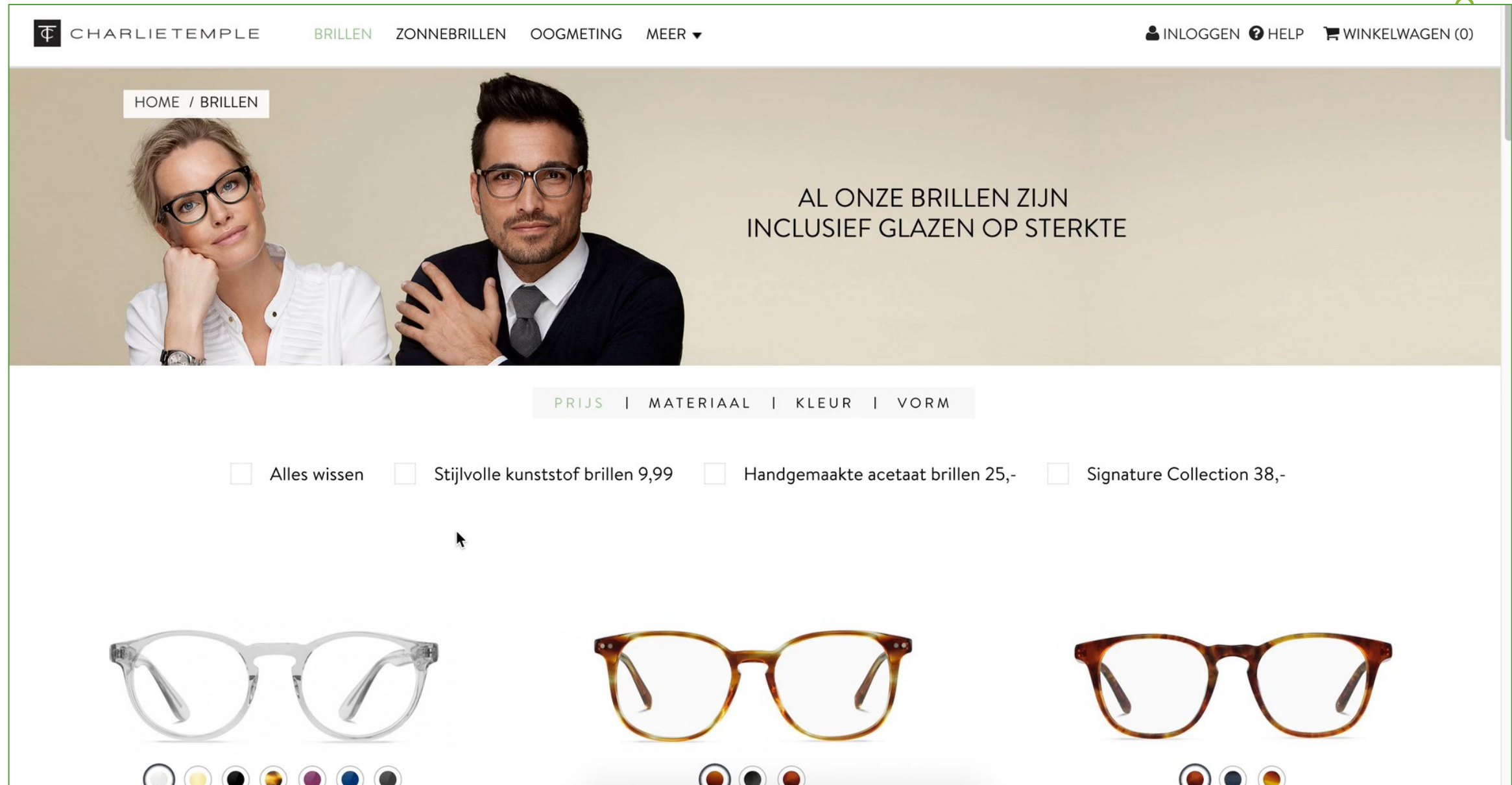
Xi Chen

Applying and developing causal inference methods for

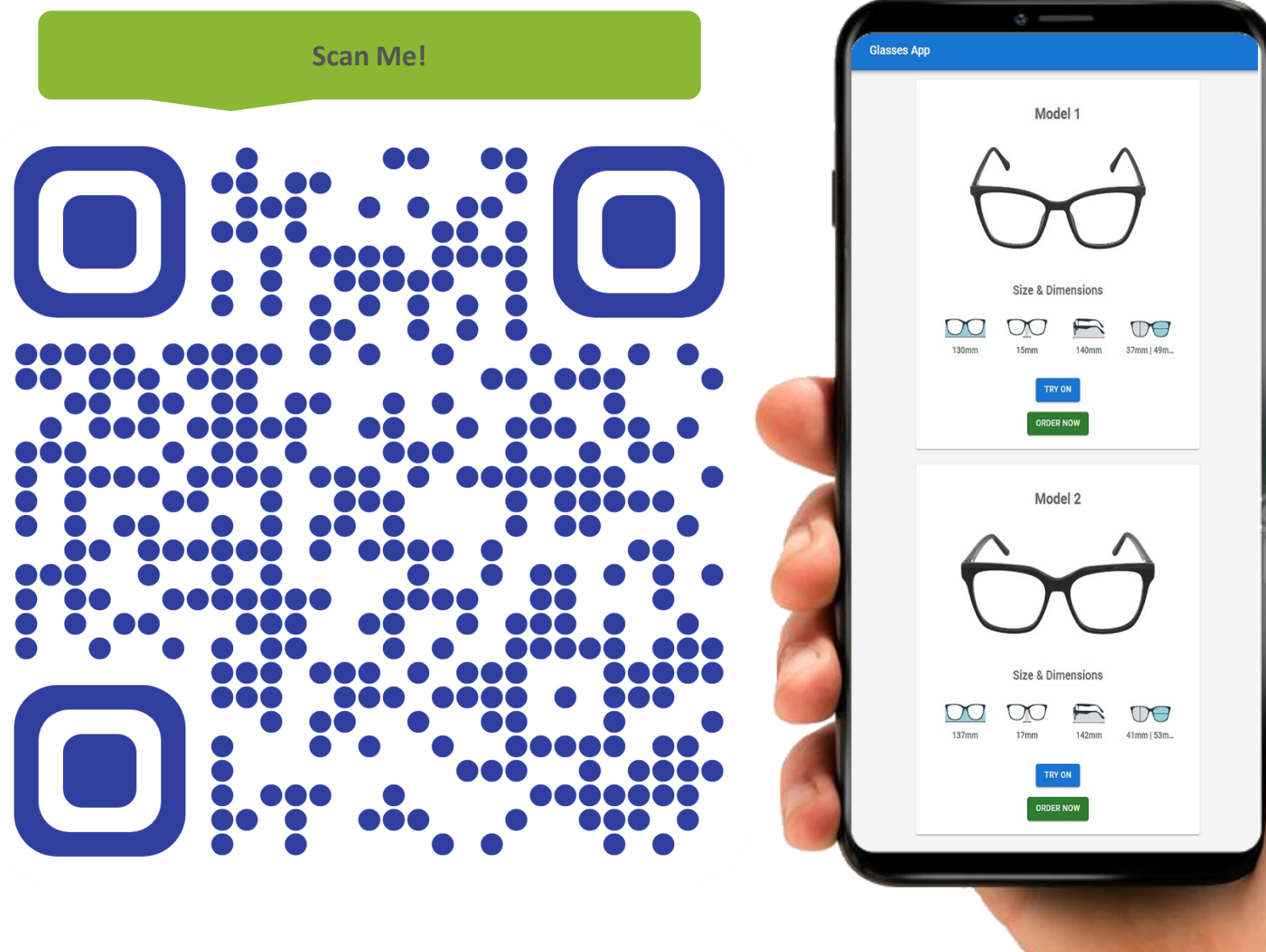
- digital marketing
- policy evaluation
- business research

rich experiences with data analysis
working with companies

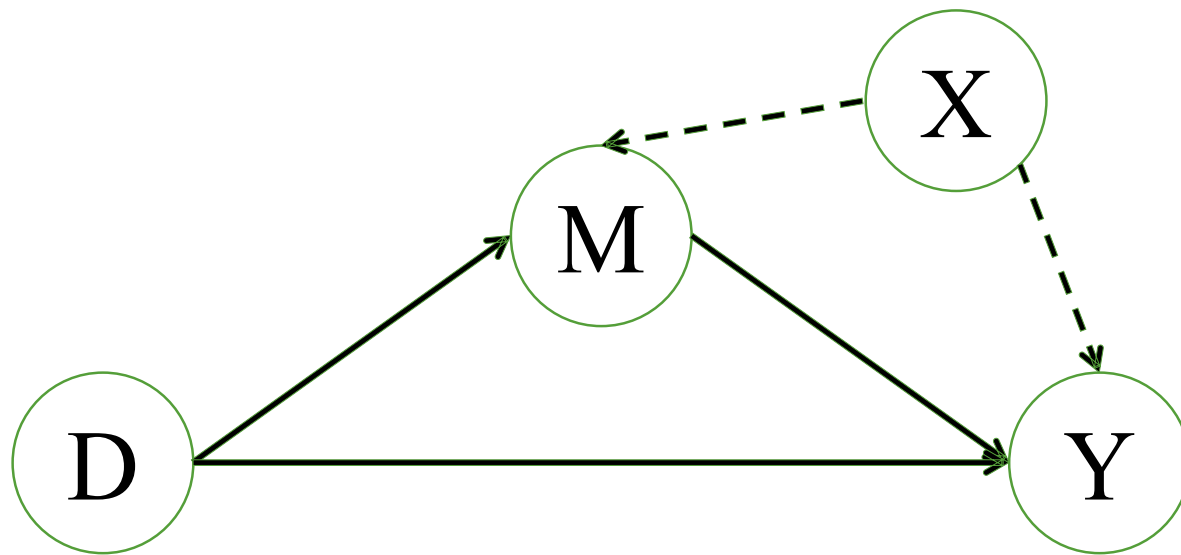
AR Virtual Fitting on Product Returns



AR Virtual Fitting on Product Returns



Boosting Conventional Mediation Analysis

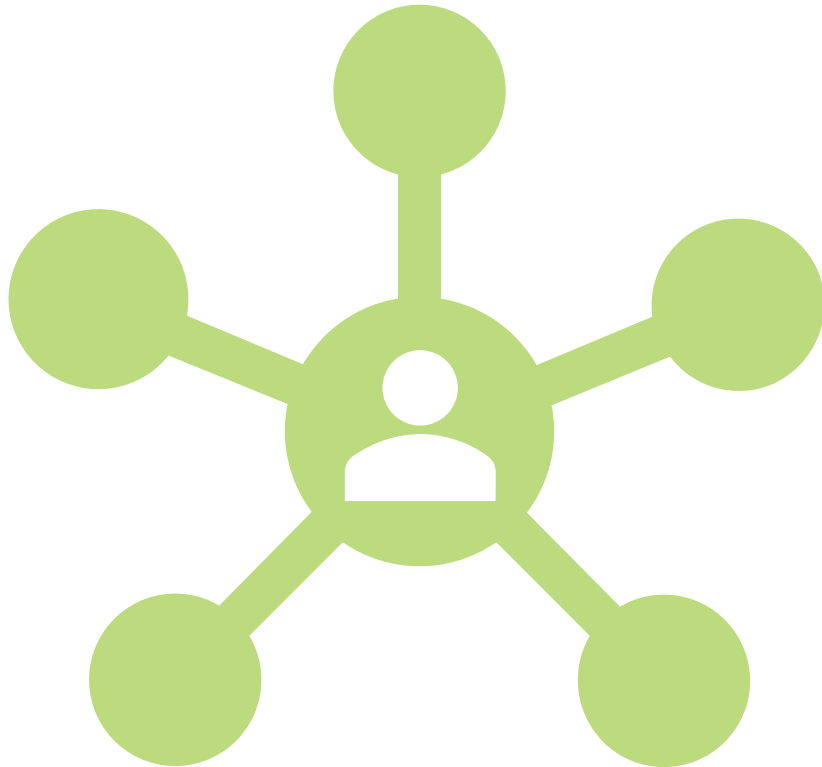


```
library(boostmed)
```

```
causal_med_civ(civ_example_data,  
               x = "D",  
               m = "M",  
               y = "Y")
```

```
## $strength_of_constructed_IV  
## Chi-square Stats      P-value  
##      432.7623          0.0000  
##  
## $exogeneity_of_constructed_IV  
## Chi-square Stats      P-value  
##      0.03431541        0.98298865  
##  
## $acme_IV  
## acme_constructed_IV    lb 95% CI    ub 95% C  
I  
##      0.5271170         0.4604242    0.597255  
2
```


To get acquainted...



- A short self-intro with
 - Your name
 - Your background



Ask anything about causality or
causal inference

21 Questions Game: causal inference ver.

No causation without correlation

Running a regression gives you effects

e.g. $\text{Sales} = b \cdot \text{Price}$

Adding more variables always gives you better estimates

e.g. $\text{Sales} = b \cdot \text{Price} + \text{Seasonality} + \dots$

21 Questions Game: causal inference ver.

"No causation without manipulation"

"Only experiments can establish causality"

"Causes are only those things that could, in principle, be treatments in experiment" (Holland 1986, p.954)

"Such questions ['no series of actions can be inferred from the description of the treatment'] have no causal answer within our framework" (Rubin 1978, p.39)

How experiments establish causation?

What's so unique about experiments?

21 Questions Game: causal inference ver.

A or B?

A. Causation can be inferred solely from data.

B. Causal inference always requires assumptions on top of data.

Any other questions?

Overview of the course

The thinking behind the course design

- The decision problem:

- **max** *LearningExperiences*

subject to: {
5 days and 25 hours
Compact Learning
A Complex Topic
Diversity of Our Group

Your expectations from the pre-survey

1. Understanding and applying causal inference in research of various fields.
2. Designing and assessing causal research.
3. Diving deeper into causal inference.
4. Causal inference tools for experimental data analysis or econometric analysis.
5. To apply causal inference in your own research projects.



The thinking behind the course design



To integrate different perspectives for problem-solving



To focus on the in-depth understanding

Only on essential technical details



To communicate “principles” of causal inference

For your critical thinking



To use blended learning methods

To address class diversity
To gain hands-on experiences
To reduce your prep efforts

First-principles based approach

We go back to the source!

1

Start from the basic problems

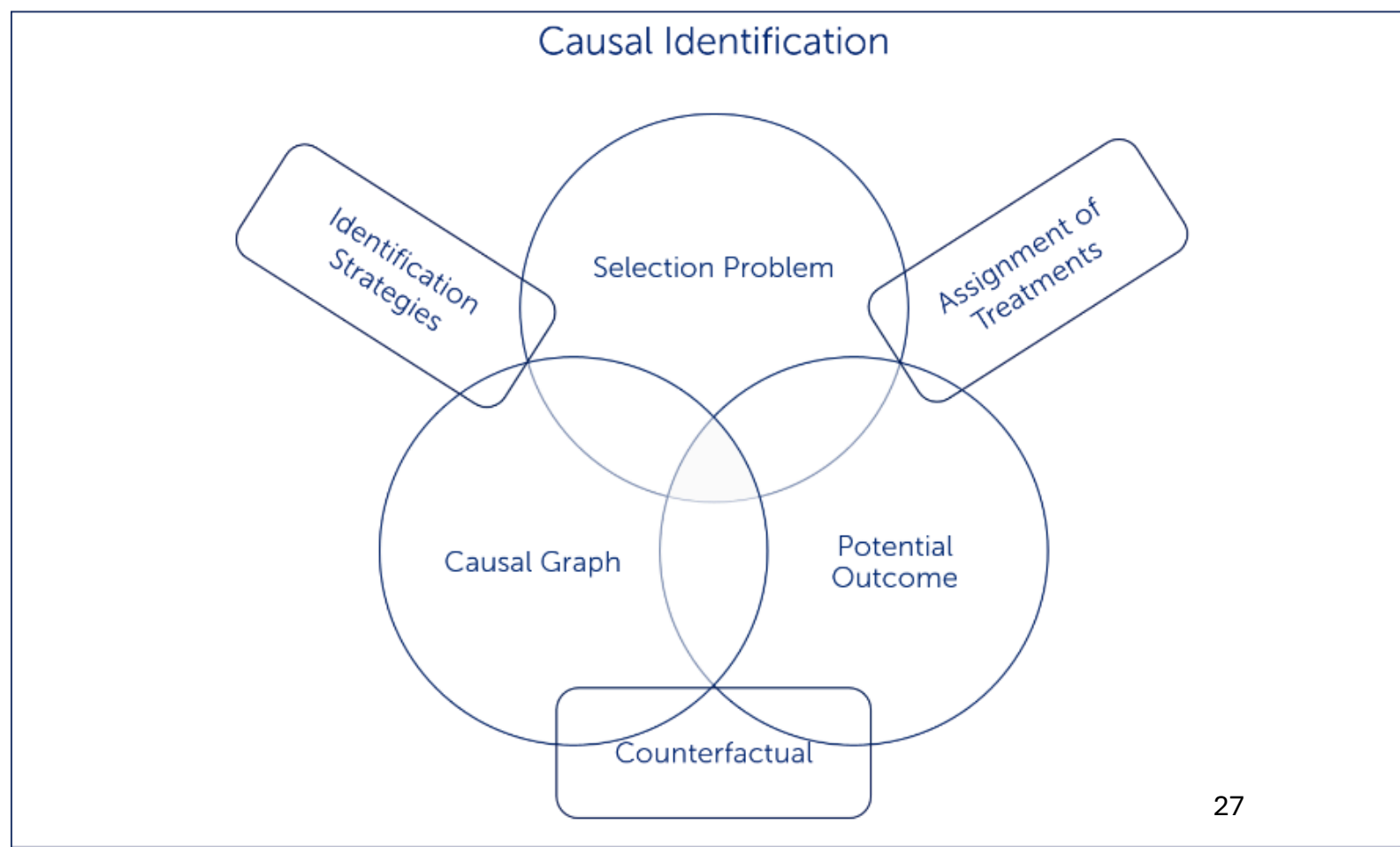
2

Break them down into fundamental blocks

3

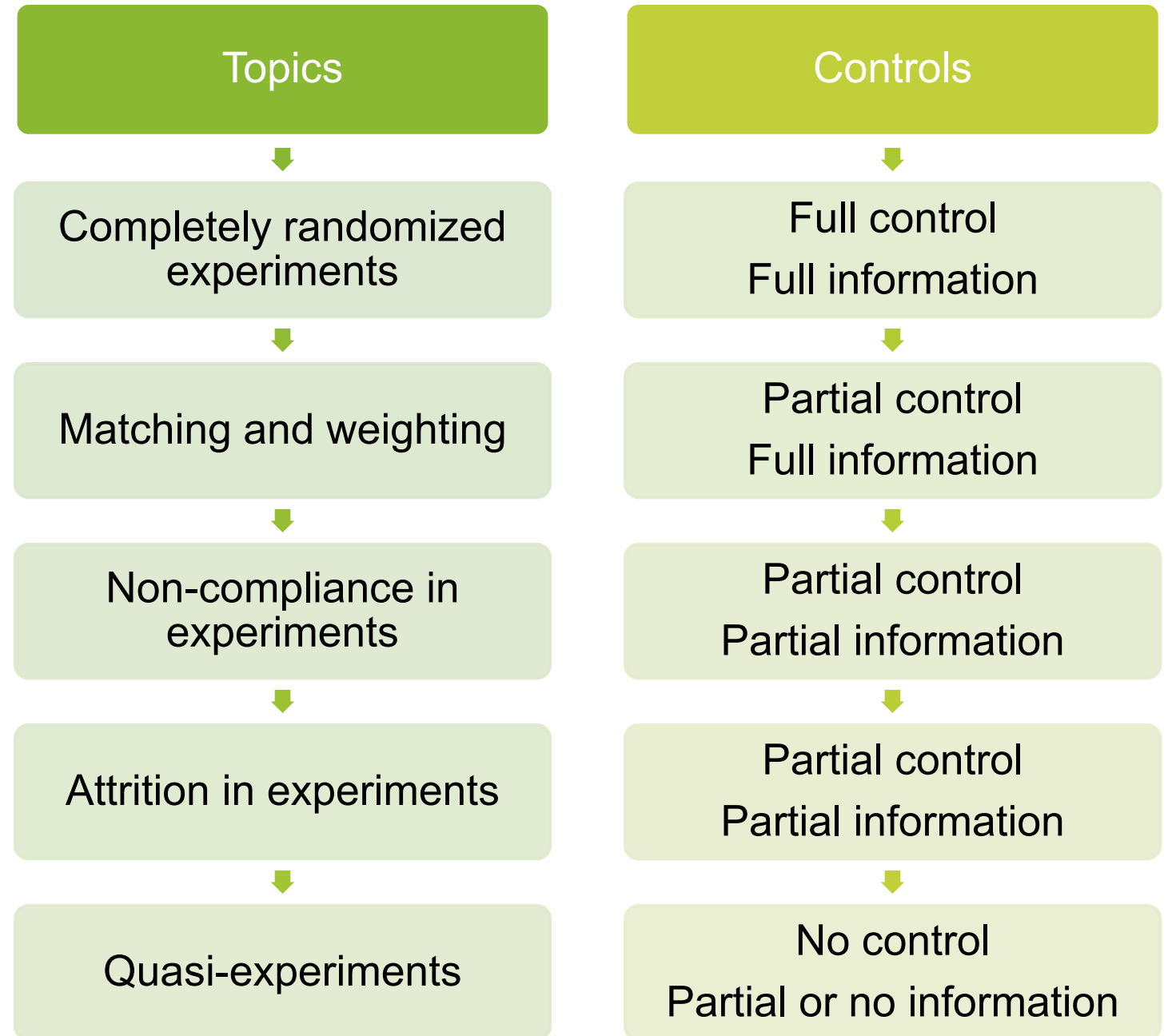
Derive the methods / solutions

A multi-disciplinary perspective

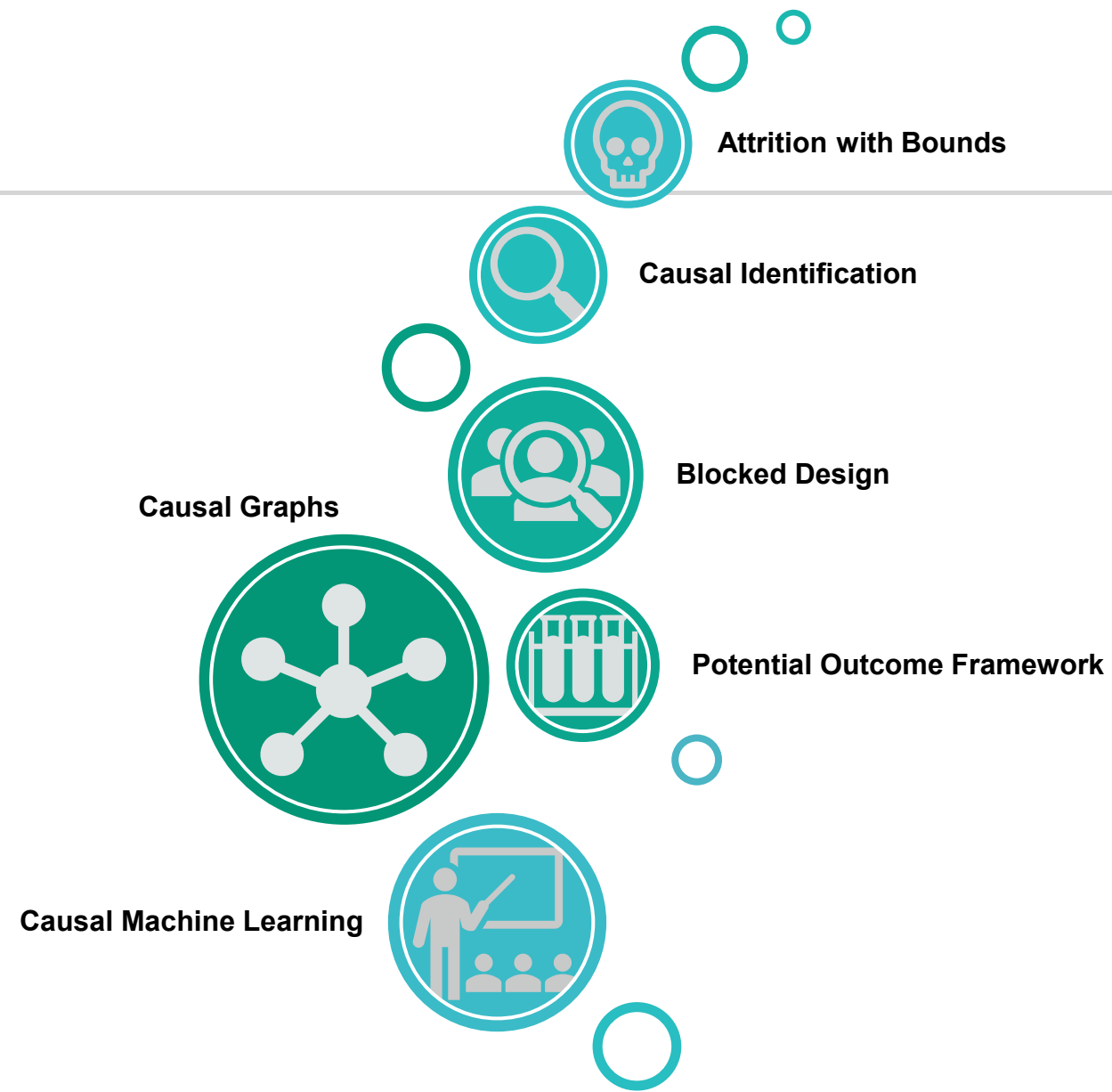


Statistical and Machine Learning Methods

Design logic of the course

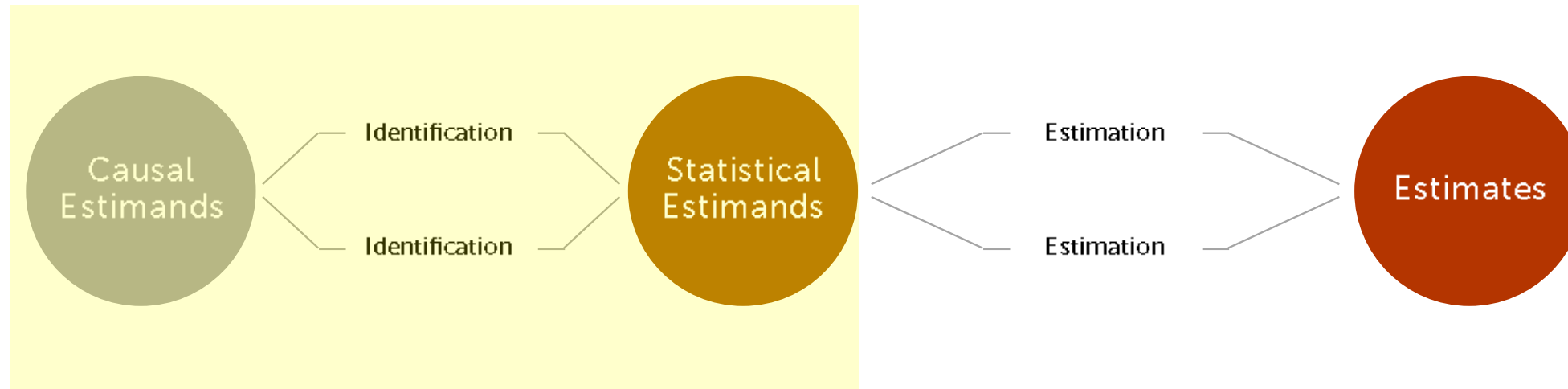


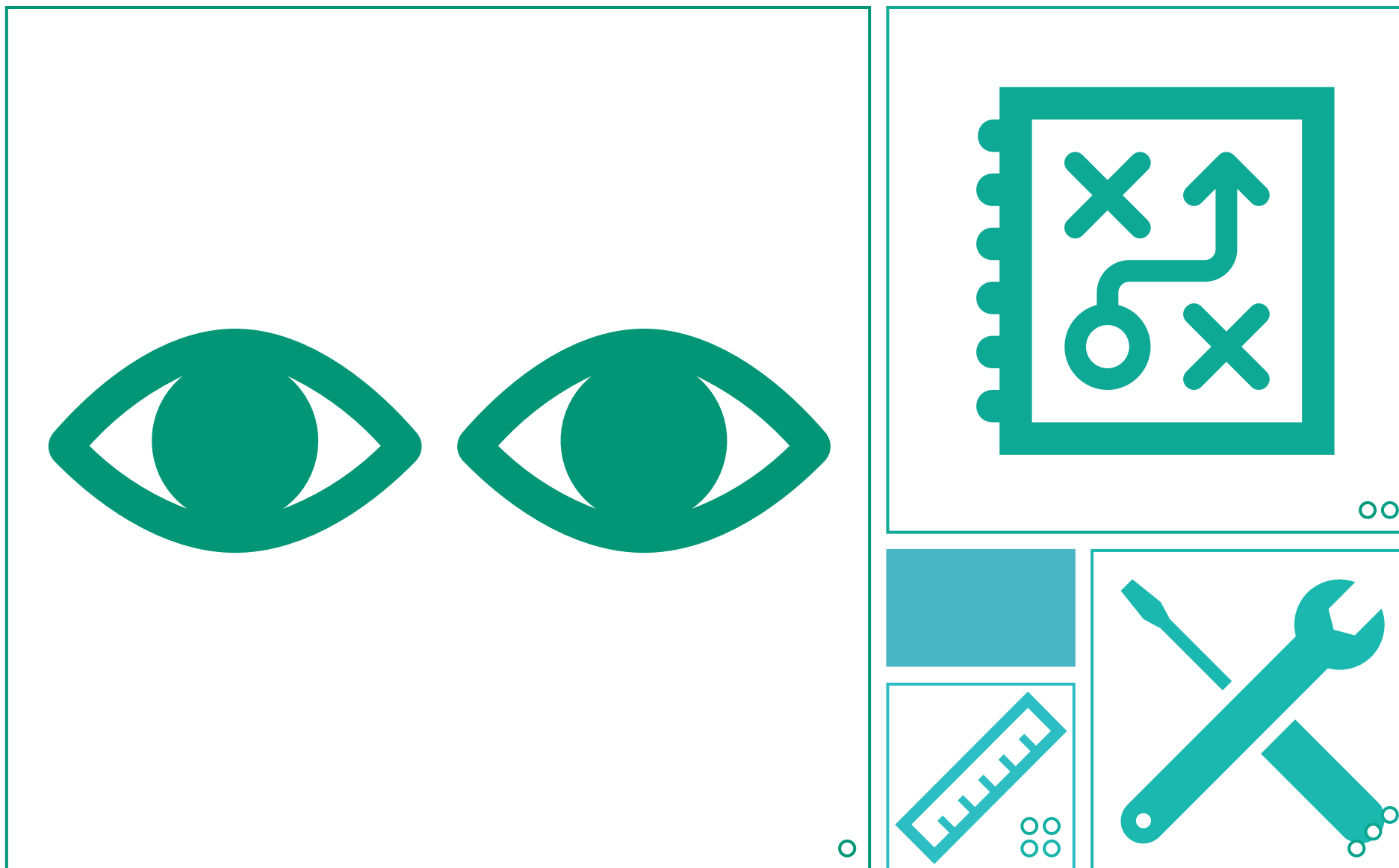
Some topics



Presentation of each topic

- The core of causal inference is causal identification – finding assumptions that help you learn causal relationships from data.



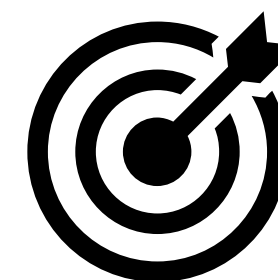


○ Recognize
causal problems

○○ Propose an identification
strategy | Data + Theory

○○○ Find or create an
appropriate method

○○○ Validate your
results truthfully



Blended learning Lectures

- Lectures
 - The basic concepts and understanding
 - Systematic way of implementation
 - Limitations of tools
- Principles
 - Principles of causal inference

Blended learning Discussions

- Discussion
 - Lecturing is “dull”
 - Of selected topics
 - Based on papers or cases
- Active participation is important
 - At least skim through the papers

Blended learning Cases

- Cases
 - To apply the models to solve real problems
 - Data- and model-driven
 - Step-to-step guide for doing projects

Course materials

Textbooks (see course manual)

*(**optional, reference only)*

- You do not have to read them if you do not feel like it.

Slides

(most important)

Rotterdam School of Management
Erasmus University

RSM
Erasmus

Causal Inference & Experimentation

Lecture 1 Introduction to the course

Dr. Xi Chen



RSM - a force for positive change

Course materials

Papers (for discussion, on Github)

*(**required)*

- Try to skim through them...



Extended readings

*(**optional, if you are interested)*

- At your leisure time...



Assessment

Integrative Assignment (three weeks)

- Online in the last week (data and questions on Github)
- Data analytics for causal inference problems
- 2 types of questions
 - Analytical & interpretation
 - Open conceptual or managerial questions
- Focus on problem-solving
 - If needed, pseudo-codes will be given.



Any other questions?