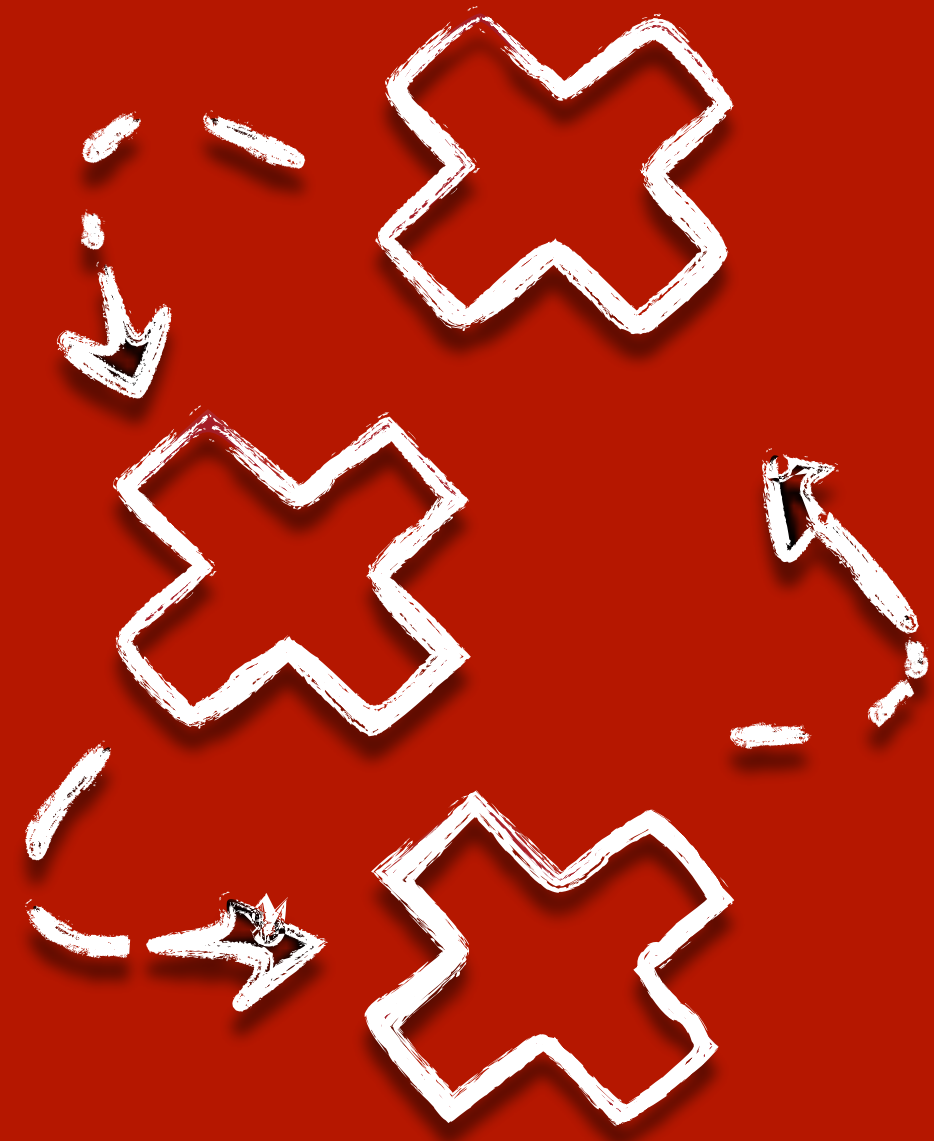


# Causal Data Science

Lecturer: Sara Magliacane

UvA - Spring 2023

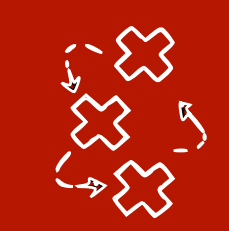


# Causal Data Science

## Lecture 1.1: Logistics

Lecturer: Sara Magliacane

UvA - Spring 2023



# Overview

- Lecturer (lectures):
  - Sara Magliacane, [s.magliacane@uva.nl](mailto:s.magliacane@uva.nl), <https://saramagliacane.github.io/>
- Teaching Assistants (tutorials):

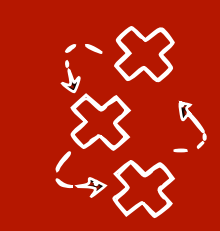


Mátyás Schubert, [m.schubert@uva.nl](mailto:m.schubert@uva.nl)



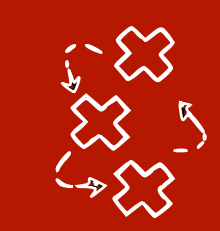
James Nevin, [j.g.nevin@uva.nl](mailto:j.g.nevin@uva.nl)

- Canvas site: <https://canvas.uva.nl/courses/36213>, Datanose [https://datanose.nl/#course\[110264\]](https://datanose.nl/#course[110264])



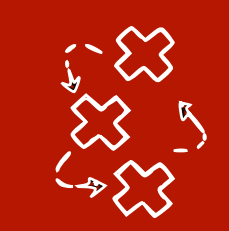
# Schedule

- Follow Datanose [https://datanose.nl/#course\[110264\]](https://datanose.nl/#course[110264])
- Two **lectures** weekly (Mondays 11-13 in SP G0.05 and Thursdays 9-11 in SP D1.116)
  - On Mondays I propose we have a break for lunch at 12:00-12:30



# Schedule

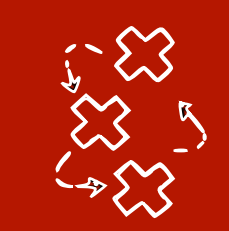
- Follow Datanose [https://datanose.nl/#course\[110264\]](https://datanose.nl/#course[110264])
- Two **lectures** weekly (Mondays 11-13 in SP G0.05 and Thursdays 9-11 in SP D1.116)
  - On Mondays I propose we have a break for lunch at 12:00-12:30
- One **lab session** weekly (Fridays 11-13 in SP F1.02)
  - The lab sessions will generally apply what we discussed in the lectures that week



# Schedule

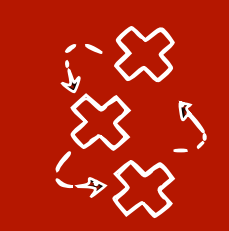
- Follow Datanose [https://datanose.nl/#course\[110264\]](https://datanose.nl/#course[110264])
- Two **lectures** weekly (Mondays 11-13 in SP G0.05 and Thursdays 9-11 in SP D1.116)
  - On Mondays I propose we have a break for lunch at 12:00-12:30
- One **lab session** weekly (Fridays 11-13 in SP F1.02)
  - The lab sessions will generally apply what we discussed in the lectures that week
    - Short (~15-30 minutes) tutorial presentation by the TAs
    - The rest is time to work on a **group project** applying the tutorial, TAs answer questions





# Group Projects

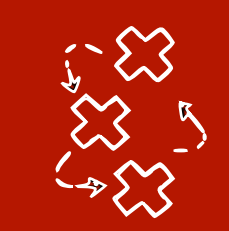
- Self-organise in groups of 3/4, to be decided by **17th February**:
  - At least one very comfortable programming in **Python**
  - If you don't have a group, post in **Canvas Discussions**



# Group Projects

- Self-organise in groups of 3/4, to be decided by **17th February**:
  - At least one very comfortable programming in **Python**
  - If you don't have a group, post in **Canvas Discussions**
- The project will be on a **dataset of your own choosing** (we can also suggest some datasets)
- You will apply the theoretical ideas from the lectures, and tools from the tutorials on this problem
- **Outcome:** Project presentation in week 12 (March 20-23) during usual lecture hours





# Tentative course outline - topics

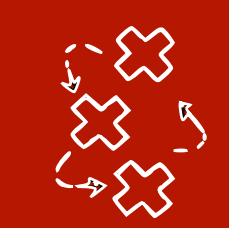
6/02/2023	Introduction
9/02/2023	Probability recap
13/02/2023	Graphical models, d-separation
16/02/2023	Causal graphs, Interventions, SCMs
20/02/2023	Covariate adjustment: backdoor criterion
23/02/2023	Covariate Frontdoor criterion, Instrumental variables
27/02/2023	Counterfactuals, potential outcomes, estimating causal effects 1
2/03/2023	Estimating causal effects 2 (matching, IPW)
6/03/2023	Constraint based structure learning
9/03/2023	Score based structure learning, restricted models
13/03/2023	Do-calculus, transportability, Joint Causal Inference
16/03/2023	Causality-inspired ML, recap of the course

Background on causal graphs

We know the causal graph, how do we estimate causal effects?

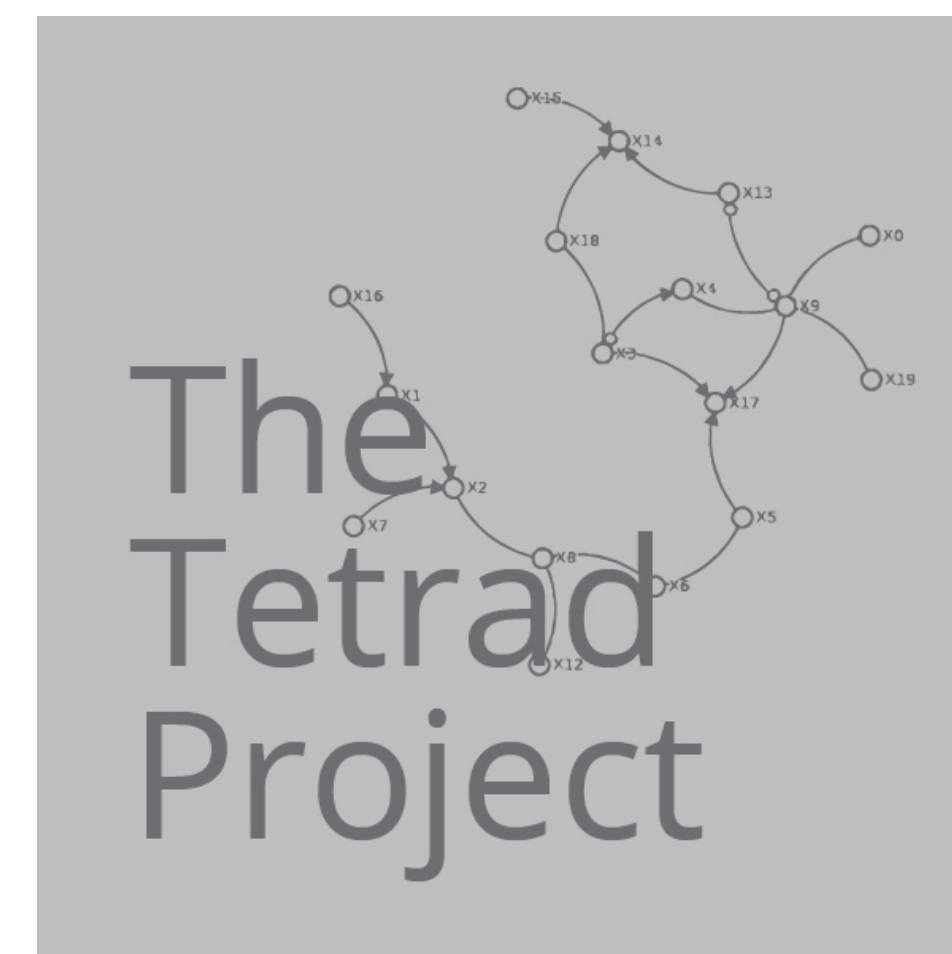
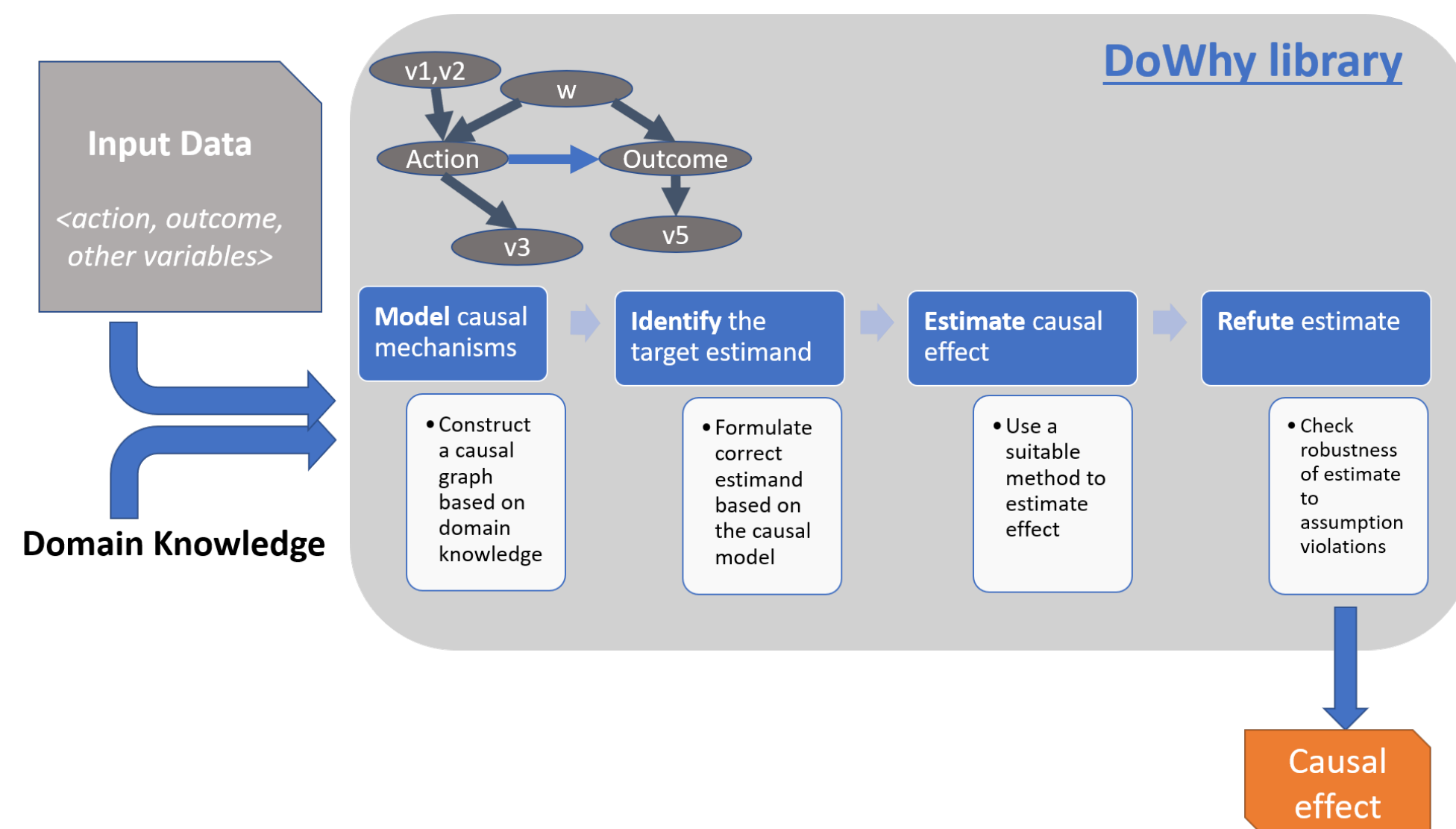
What happens if the graph is unknown?

Cutting edge research



# Lab sessions

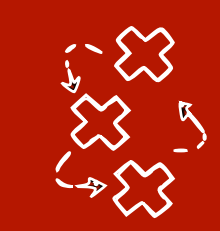
- **Programming language:** **Python**, only for experts also **R (no support)**
- Each week a short introduction to the tools
  - Jupyter notebook will be made available on Canvas after the tutorials
  - Ask questions during lab sessions or **through the Discussion tab in Canvas**



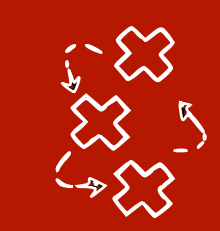


# Grading

- **Class participation/basic theoretical understanding (10%)** - Canvas quizzes during lectures
- **Project presentations (20%)** - same grade for all participants in the group
- **Written exam (70%)** - similar exercises to the Canvas quizzes and some open-ended questions to show understanding of the theory
- **The final grade is a weighted average of the individual grades, passing grade is  $\geq 5.5$**
- See also <https://canvas.uva.nl/courses/36213/pages/assessment-slash-grading>



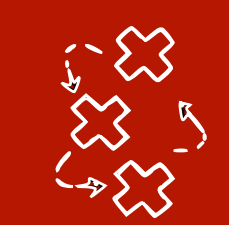
6/02/2023	Introduction
9/02/2023	Probability recap
10/02/2023	<a href="#">Tutorial 1: numpy, pandas, matplotlib, setting up projects</a>
13/02/2023	Graphical models, d-separation
16/02/2023	Causal graphs, Interventions, SCMs
17/02/2023	<a href="#">Tutorial 2: intro to dowhy, correlations and dependence testing</a>
20/02/2023	Covariate adjustment: backdoor criterion
23/02/2023	Covariate Frontdoor criterion, Instrumental variables
24/02/2023	<a href="#">Tutorial 3: Backdoor estimation</a>
27/02/2023	Counterfactuals, potential outcomes, estimating causal effects 1
2/03/2023	Estimating causal effects 2 (matching, IPW)
3/03/2023	<a href="#">Tutorial 4: Frontdoor estimation, IVs, estimation</a>
6/03/2023	Constraint based structure learning
9/03/2023	Score based structure learning, restricted models
10/03/2023	<a href="#">Tutorial 5: Constraint-based causal discovery</a>
13/03/2023	Do-calculus, transportability, Joint Causal Inference
16/03/2023	Causality-inspired ML, recap of the course
17/03/2023	<a href="#">Tutorial 6: Score-based causal discovery</a>
20/03/2023	<b>Presentations of projects</b>
23/03/2023	<b>Presentations of projects</b>
24/03/2023	<a href="#">[Optional] open hours pre-exam</a>
31/03/2023	<b>Exam!</b>
2/06/2023	<b>Resit exam</b>



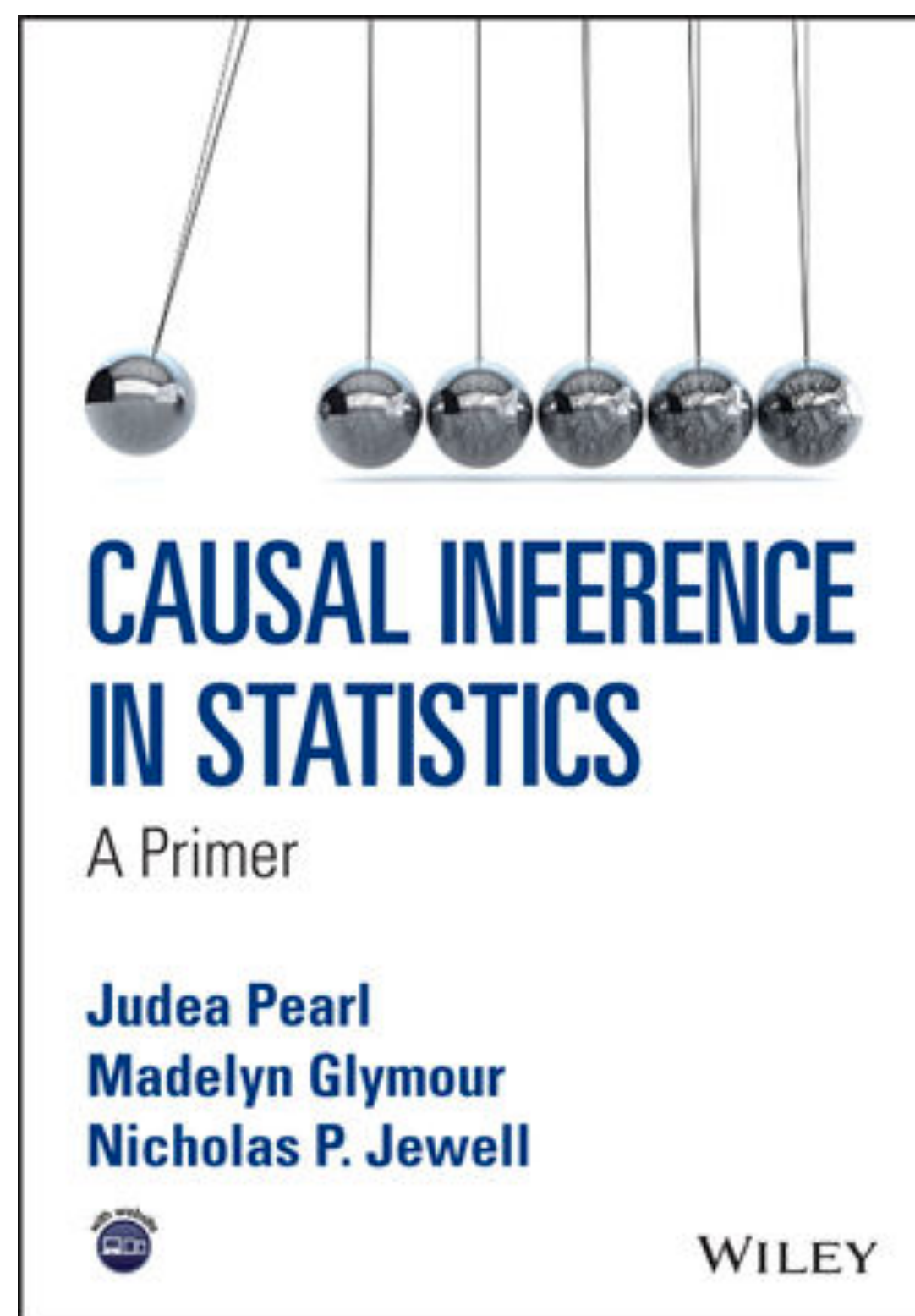
# Grading project presentations

- See also <https://canvas.uva.nl/courses/36213/pages/assessment-slash-grading>
- Project presentation grading (each presentation will be ~15 minutes):
  - Motivation, description of dataset, causal questions, assumptions (10% grade)
  - Exploratory Data Analysis, testing correlations/independences (15% grade, [follow Tutorial 1&2](#))
  - Identify estimands for backdoor, frontdoor criterion and IVs (20% grade, [follow Tutorial 3&4](#))
  - Apply different methods to estimated the causal effects (10% grade, [follow Tutorial 4](#))
  - Causal discovery, constraint-based and score-based (20% grade, [follow Tutorials 5&6](#))
  - Validation and sensitivity analysis, discussion (20% grade, [follow Tutorial 6](#))

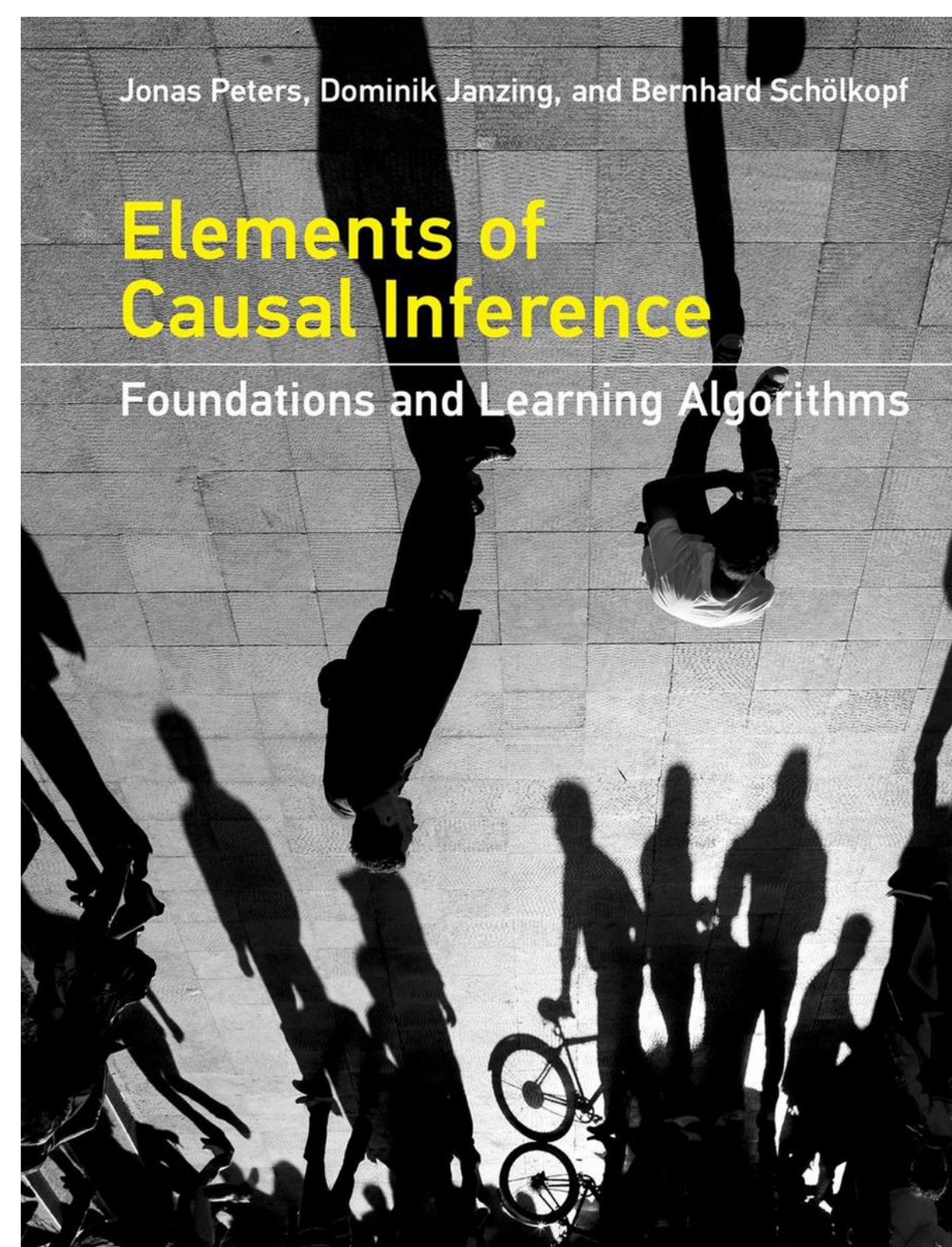




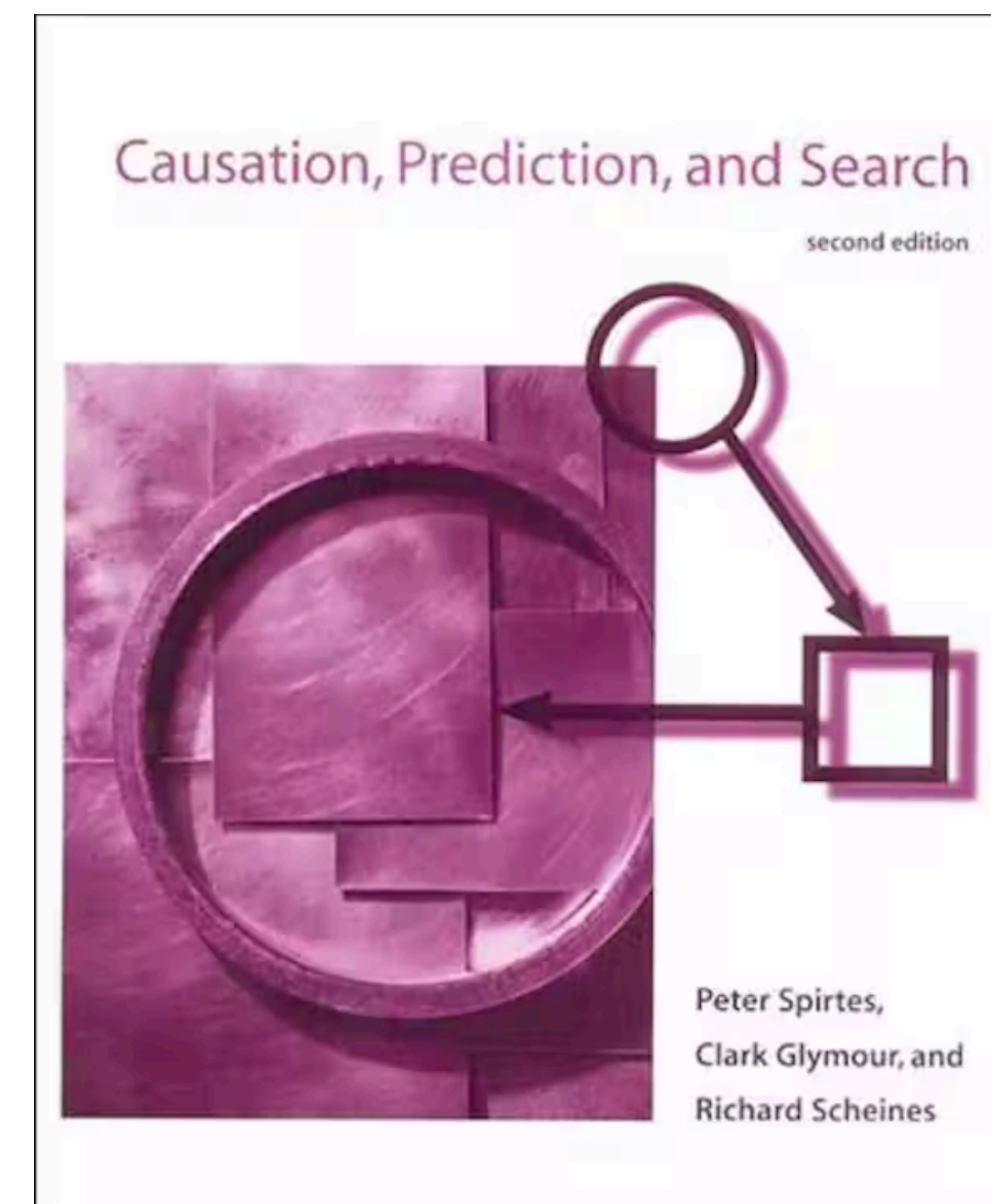
# Some foundational textbooks



<http://bayes.cs.ucla.edu/PRIMER/>



[http://web.math.ku.dk/~peters/jonas\\_files/ElementsOfCausalInference.pdf](http://web.math.ku.dk/~peters/jonas_files/ElementsOfCausalInference.pdf)

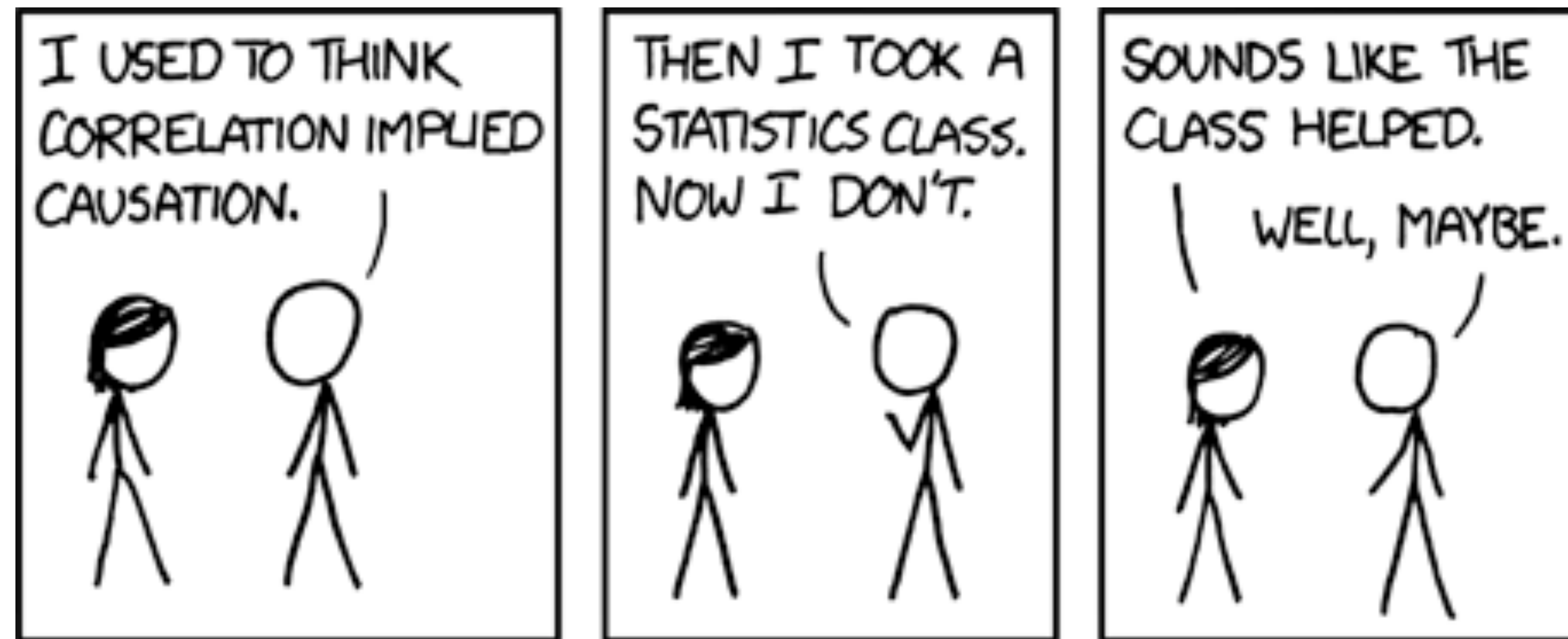


[https://www.researchgate.net/publication/242448131\\_Causation\\_Prediction\\_and\\_Search](https://www.researchgate.net/publication/242448131_Causation_Prediction_and_Search)



# Questions??

- If you have any follow-up question use the [Discussions tab in Canvas](#)



<https://xkcd.com/552/>