GESIS Summer School in Survey Methodology 2023

Syllabus for course:
“Causal Inference with Directed Acyclic Graphs (DAGs)”

Lecturer: Paul Hünermund
Affiliation: Copenhagen Business School
Email: phu.si@cbs.dk
Homepage: www.p-hunermund.com

Date: 02 – 04 August 2023
Time: 09:30-11:30 | 13:00-15:00
Venue: Online via Zoom

About the Lecturer

Paul Hünermund is an Assistant Professor of Strategy and Innovation at Copenhagen Business School. He holds a Ph.D. in business economics from KU Leuven in Belgium, and a master’s degree in economics from the University of Mannheim. His work has been published in Journal of Management Studies, Research Policy, Harvard Business Review, International Journal of Industrial Organization, Economics Letters, and Advances in Methods and Practices in Psychological Science, among others. It has been covered by Süddeutsche Zeitung, Frankfurter Allgemeine Zeitung, Politiken, and Neue Zürcher Zeitung. Dr. Hünermund serves on the editorial board of the Journal of Causal Inference and on the executive team of the Technology and Innovation Management division at the Academy of Management.

Selected Publications:

Course Description:

This course offers an introduction into causal inference with directed acyclic graphs (DAGs). DAGs combine mathematical graph theory with statistical probability concepts and provide a powerful approach for causal modeling. Originally developed in the computer science and artificial intelligence field, they recently gained increasing traction also in other scientific disciplines (such as economics, political science, sociology, health sciences, and philosophy). DAGs allow to check the validity of causal statements based on intuitive graphical criteria, that do not require algebra. In addition, they open the possibility to completely automatize the causal inference task with the help of special identification algorithms. As an encompassing framework for causal reasoning, DAGs are becoming an essential tool for everyone interested in data science and machine learning. The course provides a good...
overview of the theoretical advances that have been made in the field of causal data science in the last thirty years. The focus lies on practical applications of the theory and students will be put into the position to apply the covered methodologies in their own research. In particular, common causal inference challenges such as backdoor adjustment, bad controls, instrumental variables, selection bias, and external validity will be discussed in one single framework. Hands-on examples using dedicated libraries in R will guide through the presented material. There are no prerequisites for participating, but a good working knowledge in basic statistics and R are a plus.

Keywords:
Causal diagrams, regression analysis, instrumental variables, selection bias, external validity

Course Prerequisites:
- Basic statistics
- Basic knowledge in R is helpful

Target Group:
Participants will find the course useful if:
- they plan to do quantitative analyses in their own research.
- they want to get a better conceptual understanding of causal inference.
- they are curious to learn new data science skills.
- they are interested in an introduction into the field of causal AI.

Course and Learning Objectives:
By the end of the course participants will:
- gain a better understanding of common causal inference problems.
- be able to draw better connections between a variety of quantitative methodologies.
- master a powerful formalism for causal modeling.
- have deeper insights into methodological approaches from the field of causal AI.
- acquire various practical tools for solving causal inference challenges in their own research.

Organizational Structure of the Course:
The class is organized in two times two hours sessions (with short breaks in between) per day. The afternoon sessions will comprise practice tutorials in R run by a TA (the lecturer will be available for questions).

Software and Hardware Requirements:
Practice exercises will be taught in R. To follow the hands-on examples, it is recommended that participants have R (https://cran.r-project.org/) and RStudio (https://www.rstudio.com/) installed on their laptops. Both programs are free and open source.
Day-to-day Schedule and Literature:

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
</table>
| 1   | ▪ Structural causal models  
      ▪ Causal diagrams  
      ▪ D-separation  
      ▪ R packages: dagitty, ggdag |
|     | **Compulsory reading:**  
|     | **Suggested reading:**  
| 2   | ▪ Backdoor criterion  
      ▪ Front-door criterion  
      ▪ Surrogate experiments  
      ▪ Z-identification |
|     | **Compulsory reading:**  
|     | **Suggested reading:**  
| 3   | ▪ Recovering from selection bias  
      ▪ Transportability  
      ▪ The data fusion paradigm  
      ▪ Identification algorithms  
      ▪ R package: causaleffect |
|     | **Compulsory reading:**  
|     | **Suggested reading:**  

Preparatory reading:

**Additional Recommended Literature:**