

Syllabus

Instructor

 [Jan Pfänder](#) |  Institut Jean Nicod, ENS-PSL |  jan.pfander@psl.eu | [Bluesky](#)

Course Description and Objectives

In science, but also in politics, in court, in airport-bestsellers, or at your lunch table, claims are made based on data and science. In this course, you will learn to think critically about data, and whether it actually supports the claim it is meant to. A less fancy title for this class would have been “Introduction to Research Design”.

This class is hands-on, meaning that you will learn how to design a study, organize a project, analyze data, present results and read scientific literature.

The class is structured into four main blocks.

1. *Data Analysis* - You will develop the competences to understand and analyze data
2. *Statistics* - You will learn basic statistical concepts.
3. *Causal Inference* - You will get an understanding of how different research designs do or do not permit causal claims.
4. *Science* - You will get an idea of how science works, including how evidence accumulates and how scientific publishing works

Schedule

An up-to date schedule can be found on the [course website](#). This is also the place where all course content, including slides, assignments and additional resources will be posted.

Language

Although the class will be held in English, you can ask questions in French at any time.

Participation

This class will be highly interactive. You are strongly encouraged to participate in lectures and interrupt at any time. This means asking deep and challenging questions, but also asking simple clarification questions like “I’m just not getting this, please explain it in some new way” or “I’m lost, can you remind me why we’re talking about this?”. Don’t blame yourself if you don’t get something. Teachers are also just humans—for me to help you, I need you to communicate with me.

A lot of what you do in this course will involve writing code, and coding is a skill that is best learned by doing. Therefore, as much as possible, you will be working on a variety of tasks and activities throughout each lecture.

Requirements

You are expected to bring a laptop, tablet, or Chromebook (or whatever else you use and can have R and RStudio run on it) to each class so that you can participate in the in-class exercises. Please make sure your device is fully charged before you come to class, as the number of outlets in the classroom will not be sufficient to accommodate everyone.

If this presents a problem for you, please let me know via email or in person and we will figure something out.

Evaluation

This class is not graded—you either pass or fail. This is meant to alleviate pressure for high performance. However, you are expected to prepare for class by completing assigned readings and problem sets, attend all lectures, and meaningfully contribute to in-class exercises and discussion.

Assignments

To pass this class, you will need to have completed **all assignments**. These include

- Weekly problem sets
- A pre-registration for a study
- A group presentation on a scientific paper

It is ok if you get things wrong, but for you to pass, I need to see that you at least tried.

Attendance

Since being present and doing your homework is the only requirement to pass, I will very strictly apply the Règlement de la scolarité et des études:

“Si une étudiante ou un étudiant est signalé absent trois fois dans une UE de 24h ou plus, au cours d’un même semestre, sans justification valable, il ou elle obtient la note finale de zéro à cette UE.”

I will check presence during every session and report each absence. You can miss two classes, but starting with the third missed class, you will need to justify each absence. All justifications are to be presented to the secretary and will be validated by them

Credit

The content of this course is inspired by and draws on material from openly and freely available courses from brilliant teachers, including

- Andrew Heiss’ [“Program Evaluation for Public Service”](#)
- Louis Sirugue’s [“Introduction to Econometrics & Programming”](#)
- Mine Çetinkaya-Rundel’s [“STA 199: Introduction to Data Science and Statistical Thinking”](#) and [“STA 101 - Data Analysis and Statistical Inference”](#)
- Megan Hall’s workshop on [“Keeping it tidy”](#)
- Sam Shanny-Csik’s [Data Visualization & Communication](#)

Readings Assigned Throughout the Semester

- Bergstrom, Carl T., and Jevin D. West. 2020. *Calling bullshit: the art of skepticism in a data-driven world*. First edition. New York: Random House.
- Cohen, Jessica, and Pascaline Dupas. 2010. “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment*.” *The Quarterly Journal of Economics* 125 (1): 1–45. <https://doi.org/10.1162/qjec.2010.125.1.1>.
- Dattani, Saloni. 2022. “Why Randomized Controlled Trials Matter and the Procedures That Strengthen Them.” *Our World in Data*, February. <https://ourworldindata.org/randomized-controlled-trials>.
- Gould, Elliot, Hannah S. Fraser, Timothy H. Parker, Shinichi Nakagawa, Simon C. Griffith, Peter A. Vesik, Fiona Fidler, et al. 2025. “Same Data, Different Analysts: Variation in Effect Sizes Due to Analytical Decisions in Ecology and Evolutionary Biology.” *BMC Biology* 23 (1): 35. <https://doi.org/10.1186/s12915-024-02101-x>.

- Harrer, Mathias, Pim Cuijpers, Furukawa Toshi A, and David D Ebert. 2021. *Doing Meta-Analysis with r: A Hands-on Guide*. 1st ed. Boca Raton, FL; London: Chapman & Hall/CRC Press.
- Lakens, Daniël. 2022. *Improving Your Statistical Inferences*. Zenodo. https://lakens.github.io/statistical_inferences/.
- Open Science Collaboration. 2015. “Estimating the Reproducibility of Psychological Science.” *Science* 349 (6251). <https://doi.org/10.1126/science.aac4716>.
- Rohrer, Julia M. 2018. “Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data.” *Advances in Methods and Practices in Psychological Science* 1 (1): 27–42. <https://doi.org/10.1177/2515245917745629>.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. “False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant.” *Psychological Science* 22 (11): 1359–66. <https://doi.org/10.1177/0956797611417632>.
- Thaler, Richard. 1980. “Toward a Positive Theory of Consumer Choice.” *Journal of Economic Behavior & Organization* 1 (1): 39–60.
- Turner, Erick H., Annette M. Matthews, Eftihia Linardatos, Robert A. Tell, and Robert Rosenthal. 2008. “Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy.” *New England Journal of Medicine* 358 (3): 252–60. <https://doi.org/10.1056/NEJMs065779>.
- Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Golemund. 2023. *R for data science: import, tidy, transform, visualize, and model data*. 2nd edition. Beijing Boston Farnham Sebastopol Tokyo: O’Reilly. <https://r4ds.hadley.nz/>.