EuDARe
CHALLENGES TO DEMOCRACY AND REPRESENTATION IN THE EU

Funded by the European Union

Jean Monnet Chair EuDARe

Challenges to Democracy and Representation in the EU

Department of Political and Social Sciences, University of Catania

Research Design and Data Analysis in Political Science

Master's Programme in Global Politics and Euro-Mediterranean Relations (LM-62); optional for the Master's Programme in Data Science

Academic year 2026-27

Syllabus

Syllabus as distributed to students.

The course is taught under the Jean Monnet Chair EuDARe (Erasmus+ Jean Monnet Chair, project 101127735). This cover sheet was added by the Chair to the archive copy published on eudare.eu.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Research Design and Data Analysis in Political Science

Academic year 2026–2027, first semester

Marcello Carammia
Jean Monnet Chair EuDARe
Associate Professor of Political Science
Department of Political and Social Sciences
University of Catania
marcello.carammia@unict.it

Second year, first semester. Elective course in the base curriculum of the Master Programme in Global Politics and Euro-Mediterranean Relations (LM-62). Optional course for students of the Master Programme in Data Science.

Course overview

This course covers the methods for designing political and social science research. We discuss how to design a research project, from asking good research questions to selecting data sources, formulating and testing hypotheses, and evaluating findings. Throughout, we highlight the differences between designing research to explore problems, to develop theories, and to make forecasts.

Each student chooses a topic and develops a research project during the course. This is not a primarily technical course. The emphasis is on the ability to design, carry out and evaluate research, more than on analytical techniques or statistical modelling.

The course is attended by students of the Master Programme in Global Politics and Euro-Mediterranean Relations and by students of the Master Programme in Data Science. Most of the course works well for both groups. For some laboratory sessions we organise separate activities: political science students follow training sessions on R, while data science students analyse their own data with methods they already know, or work on the design of their project. With data science students we focus on how data science and new data sources can help to understand political and social problems. In past editions the two groups have learnt from each other.

1. Expected learning outcomes (RA)

Knowledge and understanding By the end of the course, students can identify the theories, hypotheses and methods used in empirical political science research. They know the main research designs – experimental, large-N, comparative, case study and mixed designs – and the logic of causal inference. They understand how large and new data sources can contribute to the study of political and social problems.

Applying knowledge and understanding Students apply different methods to political science research questions. They design a research project of their own, with a research question, a theory, testable hypotheses, and a plan for data collection and analysis. They draw causal diagrams (DAGs) to represent a theory, and they carry out basic data analysis in R during the laboratory sessions.

Making judgements Students analyse data to measure concepts, make comparisons and draw inferences. They assess which design and method suit a given research question, and they evaluate the design of published research. The "research in focus" discussions of published studies and the laboratory exercises build this capacity.

Communication skills Students communicate political science concepts, theories and methods in writing, in the research design proposal. They present their project, its findings and its implications in front of an audience, in the three student workshops and in the oral exam.

Learning skills Students learn to recognise the most suitable method for a research question, including methods based on large data sources, and to read methodological and applied literature on their own. The research design proposal, developed individually over the full term, and the individual reading before each session build the capacity to continue studying and doing research independently.

2. Prerequisites (PR)

No formal prerequisites. The course starts from the basics of research design and methods. A general background in political or social science is useful but not indispensable. Basic notions of statistics and data analysis are useful for the individual research project, and are also not indispensable: the laboratory sessions introduce the tools that students need.

3. Course contents (CN)

1. **Foundations** (sessions 1-5): the scientific study of politics; research questions and the literature review; theory; concepts, operationalisation, measurement and description; explanation and causality. First student workshop on research questions.
2. **Causal diagrams** (sessions 6-7, practical): DAGs of published research; the DAG of each student's own project.
3. **Research designs** (sessions 8-12): experimental designs; large-N designs; comparative designs and case selection; case studies, within-case analysis, mixed and nested designs. Second student workshop on theory, hypotheses and causal diagram.

4. **Data analysis and student projects** (sessions 13-16): data analysis with R in three laboratory afternoons, with separate tracks for political science and data science students; third student workshop with the presentation of the research designs.

The detailed session-by-session programme is in section 9.

4. Assessment method (MV)

Assessment covers the work done during the semester and the final exam. Students work on an individual research design proposal over the full term, present it in two workshops, submit it as a written paper, and discuss it at the final exam.

Assessment method

- **Type of assessment:** continuous assessment, a written paper, and a final exam with a written and an oral part.
- **Structure and components (with weight):**
  - *Class participation and workshop presentations – 30%*: participation in class discussions and in the "research in focus" sessions, and the three presentations of the student's own project in the student workshops: research question (10 November); theory, hypotheses and causal diagram (30 November, written deliverable of one to two pages due on 27 November); research design (15 January).
  - *Final paper – 40%*: an original research design proposal of 3,000–5,000 words, built from the three workshop deliverables and the comments received on them. The paper has a fixed structure: introduction and research question; literature and theoretical framework; hypotheses; concepts and operationalisation; data; research design; limitations and expected findings. It must contain the causal diagram built in class, a data note produced in the R laboratory on the data the student proposes to use (for data science students, their own analysis), and a one-page revision memo stating what changed after each workshop and why. The design section is at least half of the paper. A separate document gives detailed instructions and the grading rubric. The paper is submitted one week before the exam date.
  - *Final exam – 30%*: a **written** part (10%) in which the student outlines a research design on a topic assigned at the exam, chosen among three – research question, theory and hypotheses, concepts and measurement, design and data, main threats to inference – in 45 minutes, without materials or devices; and an **oral** part (20%) of about 20 minutes, which discusses the student's research design proposal and the key topics of the course.
- **Intermediate assessment:** there is no midterm test. Participation and the workshop presentations run through the semester and count for 30% of the final grade, as detailed above.
- **Assessment criteria:** grading follows the programme's shared grading grid below (18–30 and *lode*), based on knowledge of the topics, capacity to apply and discuss them critically, and clarity of exposition. In this course, particular weight is given to the coherence between research question, theory, hypotheses and design in the student's proposal, to the ability to justify design choices, and to the correspondence between the paper and the work presented in class. A paper that the student cannot defend at the oral exam does not pass the paper component.
- **Data science students:** attendance is not compulsory for students of the Master in Data Science, who choose one of two options by e-mail to the instructor by 9 November. *Option 1 – attending the practical*

activities: the student attends at least 70% of the hours of the causal-diagram practicals (16 and 17 November), the three workshops (10 and 30 November, 15 January) and the R laboratory (11–13 January); attending the lectures is recommended. Assessment is the same as above. *Option 2 – non-attending:* final paper 50% (same structure; the student prepares the causal diagram and the data note alone; no revision memo), written design outline 20%, oral exam 30% on the paper and on all the reference texts. Non-attending students may send a one-page research question by 13 November and a sketch of theory, hypotheses and causal diagram by 4 December, and receive written comments.

- **Language:** English.
- **Use of AI tools:** AI tools may be used for language editing, exploring sources, and even for supporting the generation of contents. Their use has to be acknowledged and explained in an AI disclaimer, attached to the final paper. Undeclared use will be treated as academic misconduct, and passive copying of AI contents will be treated as plagiarism. Students' command of the contents of the course and their own work will be assessed at an oral exam.

Learning assessment may also be carried out on-line, should the conditions require it.

To ensure equal opportunities and in compliance with current laws, interested students may request a personal interview in order to plan any compensatory and/or dispensatory measures based on educational objectives and specific needs. Students can also contact the CInAP (Centro per l'integrazione Attiva e Partecipata – Servizi per le Disabilità e/o i DSA) referring teacher within their department (https://www.cinap.unict.it/content/referenti).

Grading scale

In Italy, grades for individual courses range from 18 (minimum pass) to 30 (maximum). The highest grade may be awarded with *lode* (cum laude, distinction) for outstanding performance. The scale is not directly convertible to percentages or GPA, but it is roughly comparable to "sufficient–excellent" ranges used internationally.

Band	Level	Descriptor
0–17	Failed	The student does not demonstrate knowledge of the basic concepts of the course and/or has not completed the required share of work.
18–20	Pass – Sufficient	The student demonstrates a basic knowledge of the topics of the course but shows significant difficulties in discussing or applying them, with limited capacity to engage critically.
21–24	Fair – Discrete	The student demonstrates an adequate knowledge of the topics of the course and can engage with materials, cases, or discussions in a simple way, though often requiring guidance.
25–27	Good	The student demonstrates a solid knowledge of the topics of the course and can address assignments, case analyses, or discussions independently, with only minor inaccuracies.
28–30	Very Good – Excellent	The student demonstrates comprehensive knowledge of the topics of the course, engaging autonomously with readings, cases, and debates, and making appropriate connections across themes, with very few errors.

Band	Level	Descriptor
30 e lode	Distinction	The student demonstrates outstanding mastery of the course topics, addressing assignments and discussions with originality, critical insight, and virtually no errors.

5. Example exam questions (ED)

- *Written part – design outline on an assigned topic (example topic):* "Does exposure to social media increase political polarisation?" Outline a research design: state a research question, a theory and one or two testable hypotheses; define and operationalise the key concepts; indicate the data and the cases or units; choose a design (experimental, large-N, comparative or case study) and justify it; name the main threats to inference and how the design addresses them.
- *Written part – example topic:* "Why do some countries adopt stricter asylum policies than others?" Same task.
- *Written part – example topic:* "Do cash transfers reduce support for populist parties?" Same task.
- *Oral, on the research design proposal:* State your research question and your main hypothesis. How do you measure the key concepts, and what are the limits of these measures?
- *Oral, on the research design proposal:* Why did you choose this design and these cases or data? Which alternative design did you consider, and what would it have allowed you to say that your design does not?

6. Texts and readings (TD)

Reference texts

1. Toshkov, D. (2016). *Research Design in Political Science*. London: Palgrave. This is the main text of the course.
2. Kellstedt, P. M. & Whitten, G. D. (2018). *The Fundamentals of Political Science Research* (3rd edition). Cambridge: Cambridge University Press.
3. Huntington-Klein, N. (2022). *The Effect. An Introduction to Research Design and Causality*. Boca Raton: Chapman and Hall/CRC. Free online version at https://theeffectbook.net
4. Cunningham, S. (2021). *Causal Inference: The Mixtape*. New Haven: Yale University Press. Free online version at https://mixtape.scunning.com
5. Knopf, J. W. (2006). Doing a Literature Review. *PS: Political Science & Politics*, 39(1), 127–132. https://doi.org/10.1017/S1049096506060264
6. Llaudet, E. & Imai, K. (2022). *Data Analysis for Social Science: A Friendly and Practical Introduction*. Princeton: Princeton University Press. Used in the R laboratory.

Students should not be intimidated by the amount of readings. Before each session, students normally read one chapter on the general topic of the session (see section 9). One more reading is normally a piece of published research – the *research in focus* – that we use to see how general questions are addressed in applied research.

Recommended readings

- Bergstrom, C. T. & West, J. D. (2020). *Calling Bullshit: The Art of Skepticism in a Data-Driven World*. New York: Random House.
- Curini, L. & Franzese, R. (eds.) (2020). *The SAGE Handbook of Research Methods in Political Science and International Relations*. London: Sage.
- Gerring, J. (2012). *Social Science Methodology: A Unified Framework* (2nd edition). Cambridge: Cambridge University Press.
- King, G., Keohane, R. O. & Verba, S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Lowndes, V., Marsh, D. & Stoker, G. (eds.) (2017). *Theory and Methods in Political Science* (4th edition). Basingstoke: Palgrave Macmillan.

7. Course delivery (MS)

The course has 30 hours of lectures and 12 hours of laboratory, in 16 sessions of two, three or four hours between 26 October 2026 and 15 January 2027. The laboratory hours are distributed over the course, as detailed below and in section 9.

Lectures (erogative teaching). In each lecture session the instructor presents the topic, following the reference text, and keeps the presentation as interactive as possible. Lectures build knowledge and understanding of research design.

Interactive activities. Three formats involve students actively.

1. *Research in focus.* In most lecture sessions we discuss one or two published studies, for their substance (the question they address and their findings) and for their research design (the data and the design used). This activity builds the capacity to make judgements on published research.
2. *Student workshops.* Three sessions include students' presentations of their own research design proposal: research question (10 November), theory, hypotheses and causal diagram (30 November), research design (15 January). Students receive comments from the instructor and from the other students and use them to develop the final paper. The workshops build communication skills and the capacity to apply the course contents to an original project.
3. *Class work.* Short exercises inside the lecture sessions, for example on variable and case selection in a large-N design.

Laboratory (12 hours). The laboratory hours are hours of assisted practical work:

- *Causal diagrams (5 hours, 16 and 17 November).* Students draw Directed Acyclic Graphs (DAGs) to represent theories, first on published research and then on their own project, which they build in class and stress-test with their peers.
- *Workshop 2 (1 hour, 30 November).* Presentation and discussion of each student's theory, hypotheses and causal diagram.
- *Data analysis with R (6 hours, within the three four-hour sessions of 11, 12 and 13 January).* Political science students follow a guided introduction to R with the tidyverse (data handling and graphs with ggplot2) and to data analysis, with applied exercises on real political science datasets. Data science students, who already have these skills, analyse the data for their own proposal, or replicate and extend a published

study, and report to the class on the last day. The instructor divides the time between the two groups. Students bring their own laptops; installation instructions are sent in December.

Students prepare for each session by completing the assigned readings, and they work on their research design proposal over the full term.

If the course is delivered in blended or remote mode, appropriate adjustments may be made to the above, in order to ensure consistency with the syllabus.

8. Attendance (MF)

Attendance is compulsory for at least 70% of the teaching hours of the course (lectures and laboratory), with the partial exemptions provided by the Glopem academic regulations for working students, athletes, students in situations of difficulty and students with disabilities.

Students of the Master in Data Science, for whom attendance is not compulsory, choose between attending the practical activities and a non-attending programme, as described in section 4.

9. Detailed programme (PRG)

The course runs from 26 October 2026 to 15 January 2027, with 30 hours of lectures and 12 hours of laboratory. Classes start at 16:00; the length of each session is given in the table. Sessions marked “practical” or “laboratory” carry the laboratory hours. Numbers in the “Readings” column refer to the reference texts in section 6; further readings are from the recommended list.

Date	Hours	Type	Topic	Readings	Further readings
Mon 26 Oct	2	Lecture	**Introduction to the scientific study of politics, and to the course.** What political (or social) science means. Approaching politics scientifically. Contents, structure and activities of the course; the research design proposal.	[1] ch. 1; [2] ch. 1	
Mon 2 Nov	2	Lecture	**Research questions, types of research, and the literature review.** Types of research questions. Why good research questions matter, and how to formulate them. Finding, selecting and organising the literature.	[1] ch. 2; [5]	Clark (2020); McCauley & Ruggeri (2020)
Tue 3 Nov	2	Lecture	**Theory.** The function of theory in social science, and the difference from theories in natural science. Paradigms, frameworks, theories and models. Developing and assessing theories. Hypotheses.	[1] ch. 3; [2] ch. 2	
Mon 9 Nov	2	Lecture	**Concepts, operationalisation, measurement and description.** The role of concepts and the problem of concept definition. Operationalis-	[1] ch. 4–5	Gerring (2012) ch. 5, 7

Date	Hours	Type	Topic	Readings	Further readings
			ing concepts. Measurement strategies and descriptive inference.		
Tue 10 Nov	3	Lecture (2h) and workshop (1h)	**Explanation and causal relations.** Types of explanation. Notions of causality and causal inference. **First student workshop:** presentation of the research questions.	[1] ch. 6; [3] ch. 6	[4] ch. 3; King, Keohane & Verba (1994), selected chapters
Mon 16 Nov	2	Practical (2h)	**Causal diagrams (I).** Reading and drawing DAGs on published research.	[3] ch. 7	
Tue 17 Nov	3	Practical (3h)	**Causal diagrams (II).** Each student builds the DAG of their own project from a one-page theory sketch, and peers stress-test it.	[3] ch. 8–9	
Mon 23 Nov	2	Lecture	**Experimental research designs.** Goals and logic of experimental research. The design of experiments. Randomised controlled trials and quasi-experiments. Analysis and limitations. Experiments in political science and public policy research.	[1] ch. 7	
Tue 24 Nov	2	Lecture	**Large-N research designs (I): logic and strategies.** Conditions and strategies for causal inference: natural experiments, instrumental variables, mediation analysis, conditioning.	[1] ch. 8 (pp. 200–227)	[4] ch. 4–10; [3] ch. 14–20
Mon 30 Nov	3	Lecture with class work (2h) and workshop (1h)	**Large-N research designs (II): designs, estimation and pitfalls.** Common designs: time series, cross-sectional, panel and multilevel designs. Estimating causal effects; uncertainty and statistical significance. Variable and case selection; levels of analysis; measurement error and missing data. Class work on a large-N design. **Second student workshop:** theory, hypotheses and causal diagram (written deliverable due 27 November).	[1] ch. 8 (pp. 227–257)	[3] ch. 18
Tue 1 Dec	2	Lecture	**Comparative designs and case selection.** Logic and types of small-N comparative research. Most similar and most different systems designs. Qualitative comparative analysis. Use and limitations.	[1] ch. 9	King, Keohane & Verba (1994), selected chapters
Mon 21 Dec	2	Lecture	**Case studies, within-case analysis, mixed and nested designs.** Selecting evidence to observe. Conducting case study research. Selecting and using cases in mixed and nested designs. Use and limitations.	[1] ch. 10–11	King, Keohane & Verba (1994), selected chapters

Date	Hours	Type	Topic	Readings	Further readings
Mon 11 Jan	4	Lecture (2h) and laboratory (2h)	**Data analysis with R (I).** Political science students: RStudio, objects, data frames, reading data, first graphs with ggplot2. Data science students: own data or replication of a published study.	[6] ch. 1	
Tue 12 Jan	4	Lecture (2h) and laboratory (2h)	**Data analysis with R (II).** Political science students: analysing an experiment (difference in means), including the class's own data; graphs of one and two variables. Data science students: own data or replication.	[6] ch. 2–3	
Wed 13 Jan	4	Lecture (2h) and laboratory (2h)	**Data analysis with R (III).** Political science students: simple regression; the data note for the research design proposal. Data science students: short report to the class.	[6] ch. 4	
Fri 15 Jan	3	Workshop	**Third student workshop.** Presentation of the research designs. Communicating research and structuring a research proposal.	[1] ch. 12	

An additional two-hour session may be added at the end of the course to complete the student presentations. The date is announced in class.