Australian National University

POLS8042
Quantitative Research for Political Science

Semester 2, 2024

Course Convener:	Dr. Thiago Nascimento da Silva
E-mail:	Thiago.Silva@anu.edu.au
Website:	http://thiagosilvaphd.com
Office Hours:	Tuesdays, 9:30 a.m. - 10:30 a.m (or by appointment).
Room:	3.67. RSSS Building (third floor).
Mode of Delivery:	In-person.
Class Time:	Tuesdays, 2:00 p.m. - 4:00 p.m.
Classroom:	Lectorial 1.21.
Address:	RSSS building, First Floor, 146 Ellery Crescent, Acton, 2601.
More Info:	https://programsandcourses.anu.edu.au/2024/course/POLS8042

Course description:

This course will introduce students to the fundamental concepts and tools of quantitative methodology in political science. The content covers a range of topics, including descriptive statistics, basic probability, statistical inference and regression analysis. Additionally, students will have the opportunity to gain practical experience in statistical computing and data analysis by utilising the statistical software R and applying the knowledge they have acquired in class to their own research. A strong background in mathematics is not a prerequisite for this course. All that is required is a willingness to actively participate and learn.

Learning Outcomes:

Upon successful completion, students will have the knowledge and skills to:

1. Understand the foundations of quantitative analysis that are commonly employed across the discipline;
2. Critically evaluate the complexity of contemporary politics from the perspective of solid research design and empirical analysis;
3. Generate and visualise descriptive and inferential statistics for political phenomena using statistical programming software, and;
4. Apply relevant concepts and quantitative techniques to complete a research project and prepare a presentation suitable for delivery at a graduate-level political science conference.

Reading Requirements:

For each class, there will be compulsory readings and optional readings. The latter provide opportunities to better understand the topic, but are not compulsory.

Required Textbooks:

Students are required to obtain copies of the following books:

- Kellstedt, Paul and Guy D. Whitten. 2018. *The Fundamentals of Political Science Research.* Cambridge: Cambridge University Press. [Henceforth: KW 2018.]
- Imai, Kosuke and Nora Webb Williams. 2022. *Quantitative Social Science: An Introduction in Tidyverse*. Princeton: Princeton University Press. [Henceforth: IW 2022.]

Recommended Textbooks:

Some lectures will be based on book chapters or academic articles, which will be provided on the Wattle website by the convener.

To gain a deeper understanding of the fundamental concepts covered in this course, students can enhance their confidence in statistics and the application of mathematics in political science by exploring supplementary readings. Below, you'll find a list of suggested additional references (though not exhaustive) that may prove valuable in this context.

- Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. Mostly harmless econometrics: An empiricist's companion. Princeton: Princeton University Press.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2014. Mastering 'metrics: The path from cause to effect. Princeton: Princeton University Press.

- Bueno de Mesquita, Ethan and Anthony Fowler. 2021. Thinking Clearly About Data: A Guide to Quantitative Reasoning and Analysis. Princeton: Princeton University Press.
- Cunningham, Scott. 2021. Causal Inference: The Mixtape. New Haven: Yale University Press.
- Evans, Merran, et al. 2011. Statistical Distributions. Hoboken: John Wiley & Sons.
- Gill, Jeff. 2006. Essential Mathematics for Political and Social Research. Cambridge: Cambridge University Press.
- Greene, William H. 2018. Econometric Analysis. New York: Pearson.
- Hansen, Bruce E. 2022. Econometrics. Princeton: Princeton University Press.
- Imai, Kosuke. 2018: Quantitative Social Science: An Introduction. Princeton: Princeton University Press.
- Long, J. Scott. 1997. Regression Models for Categorical and Limited Dependent Variables. Newbury Park: Sage.
- Moore, Will and David A. Siegel. 2013. A Mathematics Course for Political and Social Research. Princeton: Princeton University Press.
- Pearl, Judea and Dana Mackenzie. 2018. The Book of Why: The New Science of Cause and Effect. New York: Basic Books.
- Pearl, Judea, Madelyn Glymour and Nicholas P. Jewell. 2019. Causal Inference in Statistics: A Primer. Hoboken: John Wiley & Sons.
- Wooldridge, Jeffrey M. 2019. Introductory Econometrics: A Modern Approach. Boston: Cengage Learning.

In this course, students will also be introduced to statistical computing and data analysis using the free statistical software R. It is recommended to download RStudio, a free user interface for R. We will be using RStudio during lectures as it provides an easier way to interact with the R environment.

While no prior knowledge of R is required, there are many great resources and tutorials available online that will be useful for learning R. The following free tutorials can be particularly helpful:

- "An Introduction to R" (2023) by W. N. Venables, D. M. Smith and the R Core Team.

- "simpleR: Using R for Introductory Statistics" (2002) by John Verzani.

For those that prefer textbooks and "how-to-do" books, I list some suggestions below:

- Adler, Joseph. 2010. R in a Nutshell by Joseph Adler. O'Reilley Media.
- Braun, W. John, and Duncan J. Murdoch. 2021. A first course in statistical programming with R. Cambridge: Cambridge University Press.
- Bruce, Peter, Andrew Bruce and Peter Gedeck. 2020. *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python*. Sebastopol: O'Reilly Media.
- Grolemund, Garrett. 2014. Hands-on programming with R: Write your own functions and simulations. Sebastopol: O'Reilly Media.
- Maindonald, John and W. John Braun. 2010. Data Analysis and Graphics Using R.
- Wickham, Hadley, Mine Çetinkaya-Rundeland and Garrett Grolemund. 2023. R for data science: Import, tidy, transform, visualise, and model data. Sebastopol: O'Reilly Media. [Second Edition.] We will use several chapters from this book. An online (and free) version of the second edition of this book is available at `https://r4ds.hadley.nz`. The source of the book is also available at the book's GitHub repository at `https://oreil.ly/Q8z_O`.

It is also a good idea to start building good programming habits.

- "R Style: An Rchaeological Commentary" (2022) by Paul E. Johnson is a useful text that provides do's and don'ts for programming in R.
- "The tidyverse Style Guide" (2023) by Hadley Wickham is highly recommended. Available at `https://style.tidyverse.org`.

Course Schedule:

Week	Date	Class Topic
1	23 July	Introduction: Research Design, R and RStudio
2	30 July	Research Questions and Exploratory Data Analysis
3	6 August	Causality in Experimental Research Design
4	13 August	Causality in Observational Studies
5	20 August	Measurement and Types of Data
6	27 August	Probability, Statistical Distributions and Hypothesis Testing
		Mid-Semester Break
7	17 September	OLS I: Introduction to Regression Analysis
8	24 September	OLS II: Multiple Regression - Model Specification and Estimation
9	1 October	OLS III: Interactions
10	8 October	Regression Discontinuity Design
11	15 October	An Introduction to MLE: Logistic Regression
12	22 October	An Introduction to Time Series Analysis

We will be using Wattle to display eventual changes to the course schedule, grades, give feedback, and make announcements: `https://wattlecourses.anu.edu.au/course/view.php?id=39132`.

Assessment Requirements:

Assessment Task	Value	Due Date	Learning Outcomes
1. Lecture Participation and Weekly Readings	10%	N/A	1, 2
2. List of Exercises 1	10%	Week 3, 14H	1, 2, 3
3. List of Exercises 2	10%	Week 6, 14H	1, 2, 3
4. List of Exercises 3	10%	Week 9, 14H	1, 2, 3
5. List of Exercises 4	10%	Week 12, 14H	1, 2, 3
6. Midterm Exam	20%	Week 7, 14H	1, 2, 3
7. Research Paper	30%	12 November, 14H	1, 2, 3, 4

Assessment Details:

- *Assessment Task 1: Lecture Participation and Weekly Readings (10%).* Active participation in the seminar is crucial for student learning and will be evaluated based on a combination of attendance and engagement in course discussions. It is expected that students attend all classes and course activities. Additionally, students must complete the assigned compulsory readings before each class and come prepared to actively participate in discussions.

  The students' participation may encompass various activities, such as contributing in overall discussions, collaborating in group problem-solving, delivering short presentations, and engaging in "Try it out" pop-up quizzes, surveys, or assessments during class.

- *Assessment Tasks 2-5: Assignments (40% total; 10% each).* The material for this course is best learned through practice. Therefore, there will be four sets of exercises throughout the semester. These exercises will test students' understanding of the concepts covered in class and their abilities in various cognitive domains, including comprehension and application of knowledge and problem-solving skills.

- *Assessment Task 6: Midterm Exam (20%).* The midterm exam will assess students' general understanding of aspects of empirical analysis covered in the first part of the course. The midterm exam will consist of multiple-choice exercises from course materials such as slides, lectures and mandatory readings. The midterm exam duration will be 1 hour for the 6-unit option and 2 hours for the 12-unit option.

- *Assessment Task 7: Research Paper (30%)*. Students are required to submit a political science research essay using quantitative methods.

  The research paper will evolve from a research paper proposal that the students will have to submit between Week 6 and Week 9 of the course, and associated feedback received from peers and the convenor between Week 10 and Week 12 of the course.

  The research paper will require students to apply the concepts and methods covered in the course to their main research interest and will consist of a 3000-word document (for the 6-unit option) and a 6000-word document (for the 12-unit option).

For additional information regarding the assessments, please refer to the course's Wattle website. We will utilize Wattle to communicate any modifications to the course schedule, grades, provide feedback, and make announcements. You can access the Wattle website at: .

Grading Policy:

The grading scale for this class follows the standard grade distribution provided in http://policies.anu.edu.au/policies/assessment_of_student_learning/policy

HD (80-100%): Work of exceptional quality, which demonstrates comprehensive understanding of the subject matter, mastery of relevant skills, sophisticated or original critical and conceptual analysis, and outstanding quality in clarity, precision and presentation of work.

D (70-79%): Work of superior quality, which demonstrates a thorough knowledge and understanding of the subject matter, proficiency in relevant skills, and analytical and conceptual ability of a high order.

C (60-69%): Work of good quality, which displays a good understanding of the subject matter and a sound grasp of relevant skills.

P (50-59%): Work of satisfactory quality, which displays an adequate understanding of most of the subject matter and a sufficient grasp of relevant skills.

N (0-49%): Work which is incomplete or displays an inadequate understanding of the subject matter or an inadequate grasp of relevant skills.

Class and University Policies

Assessment Requirements: The ANU is using *Turnitin* to enhance student citation and referencing techniques, and to assess assignment submissions as a component of the University's approach to managing Academic Integrity. For additional information regarding *Turnitin* please visit the Academic Skills website. In rare cases where online submission using *Turnitin* software is not technically possible; or where not using *Turnitin* software has been justified by the Course Convener and approved by the Associate Dean (Education) on the basis of the teaching model being employed; students shall submit assessment online via 'Wattle' outside of *Turnitin*, or failing that in hard copy, or through a combination of submission methods as approved by the Associate Dean (Education). The submission method is detailed below.

Online Submission: You will be required to electronically sign a declaration as part of the submission of your assignment. Please keep a copy of the assignment for your records. Unless an exemption has been approved by the Associate Dean (Education) submission must be through *Turnitin*.

Extensions and Penalties: Extensions and late submission of assessment pieces are covered by the Student Assessment (Coursework) Policy and Procedure. Extensions may be granted for assessment pieces that are not examinations or take-home examinations. If you need an extension, you must request an extension in writing on or before the due date. If you have documented and appropriate medical evidence that demonstrates you were not able to request an extension on or before the due date, you may be able to request it after the due date.

Late Submission: Late submission of assessment tasks without an extension are penalised at the rate of 5% of the possible marks available per working day or part thereof. Late submission of assessment tasks is not accepted after 10 working days after the due date, or on or after the date specified in the course outline for the return of the assessment item. Late submission is not accepted for take-home examinations.

Educational Policies: ANU has educational policies, procedures and guidelines , which are designed to ensure that staff and students are aware of the University's academic standards, and implement them. Students are expected to have read the Academic Integrity Rule before the commencement of their course. Other key policies and guidelines include:

- Academic Integrity Policy and Procedure
- Student Assessment (Coursework) Policy and Procedure
- Special Assessment Consideration Guideline and General Information

- Student Surveys and Evaluations
- Deferred Examinations
- Student Complaint Resolution Policy and Procedure
- Code of Practice for Teaching and Learning

Mark Moderation: Marks that are allocated during Semester are to be considered provisional until formalised by the College examiners meeting at the end of each Semester. If appropriate, some moderation of marks might be applied prior to final results being released.

Distribution of Grades: Academic Quality Assurance Committee monitors the performance of students, including attrition, further study and employment rates and grade distribution, and College reports on quality assurance processes for assessment activities, including alignment with national and international disciplinary and interdisciplinary standards, as well as qualification type learning outcomes. Since first semester 1994, ANU uses a grading scale for all courses. This grading scale is used by all academic areas of the University.

Referencing Requirements: Accepted academic practice for referencing sources that you use in presentations can be found via the links on the Wattle site, under the file named "ANU and College Policies, Program Information, Student Support Services and Assessment." Alternatively, you can seek help through the Students Learning Development website. The Academic Skills website has information to assist you with your writing and assessments. The website includes information about Academic Integrity including referencing requirements for different disciplines. There is also information on Plagiarism and different ways to use source material.

Privacy Notice: The ANU has made a number of third party, online, databases available for students to use. Use of each online database is conditional on student end users first agreeing to the database licensor's terms of service and/or privacy policy. Students should read these carefully. In some cases student end users will be required to register an account with the database licensor and submit personal information, including their: first name; last name; ANU email address; and other information. In cases where student end users are asked to submit 'content' to a database, such as an assignment or short answers, the database licensor may only use the student's 'content' in accordance with the terms of service – including any (copyright) licence the student grants to the database licensor. Any personal information or content a student submits may be stored by the licensor, potentially offshore, and will be used to process the database service in accordance with the licensors terms of service and/or privacy policy. If any student chooses not to agree to the database licensor's terms of service or privacy policy,

the student will not be able to access and use the database. In these circumstances students should contact their lecturer to enquire about alternative arrangements that are available.

Academic Integrity: Academic integrity is a core part of the ANU culture as a community of scholars. The University's students are an integral part of that community. The academic integrity principle commits all students to engage in academic work in ways that are consistent with, and actively support, academic integrity, and to uphold this commitment by behaving honestly, responsibly and ethically, and with respect and fairness, in scholarly practice.

The University expects all staff and students to be familiar with the academic integrity principle, the Academic Integrity Rule 2021, the Policy: Student Academic Integrity and Procedure: Student Academic Integrity, and to uphold high standards of academic integrity to ensure the quality and value of our qualifications.

The Academic Integrity Rule 2021 is a legal document that the University uses to promote academic integrity, and manage breaches of the academic integrity principle. The Policy and Procedure support the Rule by outlining overarching principles, responsibilities and processes. The Academic Integrity Rule 2021 commences on 1 December 2021 and applies to courses commencing on or after that date, as well as to research conduct occurring on or after that date. Prior to this, the Academic Misconduct Rule 2015 applies.

The University commits to assisting all students to understand how to engage in academic work in ways that are consistent with, and actively support academic integrity. All coursework students must complete the online Academic Integrity Module (Epigeum), and Higher Degree Research (HDR) students are required to complete research integrity training. The Academic Integrity website provides information about services available to assist students with their assignments, examinations and other learning activities, as well as understanding and upholding academic integrity.

Feedback

Staff Feedback: Feedback from teaching staff to students will be available on an ad hoc basis by request, and promptly (no longer than two weeks) following the submission of assessments. All assessment feedback will be delivered via Wattle in the first instance.

Student Feedback: ANU is committed to the demonstration of educational excellence and regularly seeks feedback from students. One of the key formal ways students have to provide feedback is through Student Experience of Learning Support (SELS) surveys. The feedback given in these surveys is anonymous and provides the Colleges, University Education Committee and Academic Board with opportunities to recognise excellent teaching, and opportunities for improvement.

More information on student surveys at ANU and reports on the feedback provided on

ANU courses, is provided at: http://unistats.anu.edu.au/surveys/selt/students/ and http://unistats.anu.edu.au/surveys/selt/results/learning/

Word Limits (per CASS policy):

1. Written assessment in a 6-unit undergraduate course will normally be in the range of 4000 to 5000 words, including examinations, or equivalent. This word limit may be increased by 10% to accommodate the use of the Harvard, or other in-text, referencing system in which references are placed in the body of the main text.
2. Written assessment in a 6-unit postgraduate course will normally be in the range of 5000 to 6000 words, including examinations, or equivalent. This may be increased by 10% to accommodate the use of the Harvard, or other in-text, referencing system in which references are placed in the body of the main text.
3. Assessment in double-weighted courses should normally have double the word count.
4. Bibliographies and footnotes using the Oxford (or other end-note or footnote) system of referencing will not be included in the word count, though notes will be assessed to ensure conformity with footnoting and disciplinary conventions.
5. Appendices and tables will not be included in the word count.
6. There should be an appropriate relationship between the word length of a particular piece of assessment and its weighting in the total assessment regime.
7. Where an assignment exceeds the prescribed word length, no penalty will be applied for the initial 10% excess words. Thereafter a 10% penalty will apply.
8. Students will declare the word length, without appendices, bibliography and Oxford-style footnotes on the cover sheet submitted with their assignment.

Access and Inclusion: The convener of this course and the ANU's Office of Access and Inclusion (A&I) support students whose participation in academic studies is affected by disabilities, mental health conditions, chronic medical conditions and short term illnesses/conditions. A&I also supports carers, international under 18 students and elite athletes.

If you require additional support and/or an Education Access Plan, please register with A&I. EAPs are now managed through the Student Access Portal, which sends all EAPs to the relevant conveners. Students do not need to send EAPs to conveners or tutors.

Diversity Statement: The instructor of this course supports the Australian National University commitment to diversity, and welcomes individuals from any racial, ethnic, religious, age, gender, sexual orientation, class, disability, and nationality. In the spirit of this vital commitment, in this course each voice in the classroom has something of value to contribute to all discussions. Everyone is expected to respect the different experiences, beliefs and values expressed by fellow students and the instructor, and will engage in reasoned discussion that refrains from derogatory comments about other people, cultures, groups, or viewpoints. For more information, visit "equity-and-diversity" site https://services.anu.edu.au/business-units/human-resources-division/equity-and-diversity.

Laptop and phone policy: In order to ensure an active participation and to keep your attention on the important things (our class), please avoid distracting yourself through (unnecessary) electronic devices or applications. For further insights on the consequences of multitasking, I recommend the study by Bellur, Nowak and Hull (2015) (https://bit.ly/2GnyTf2). They found that in-class multitasking leads to significantly lower performance.

Support for Students: The University offers students support through several different services. You may contact the services listed below directly or seek advice from your Course Convener, Student Administrators, or your College and Course representatives (if applicable).

- ANU Health, safety & wellbeing for medical services, counselling, mental health and spiritual support.
- ANU Access & inclusion for students with a disability or ongoing or chronic illness.
- ANU Dean of Students for confidential, impartial advice and help to resolve problems between students and the academic or administrative areas of the University.
- ANU Academic Skills supports you make your own decisions about how you learn and manage your workload.
- ANU Counselling Centre promotes, supports and enhances mental health and wellbeing within the University student community.
- ANUSA supports and represents undergraduate and ANU College students.
- PARSA supports and represents postgraduate and research students.

Course Outline:

Week 0: Preparation

Before our first class, I want you to be aware of the organization of this course.

1. Please read this course guide carefully, as it lists the policies and goals of this course, the required readings, the assignments and deadlines, how grades will be determined, when exams will be given, and other important information of your interest.

2. Please complete the "Pre-semester survey" available in Wattle. In the survey, you will be asked to answer some questions about yourself, your previous experience with the use of math in political science research, and your expectations about this course.

3. Read: Ford, Paul. 2016. "What is code? If you don't know, you need to read this." In.: The Best American Magazine Writing 2016. Eds.: Sid Holt and The American Society of Magazine Editors. New York: Columbia University Press. https://www.degruyter.com/document/doi/10.7312/asme16957-004/html.

4. Take a look at Andrew Gelman's (Columbia University) blog on *Statistical Modeling, Causal Inference, and Social Science*: https://statmodeling.stat.columbia.edu.

Week 1

23 July. Week 1. Introduction: Research Design, R and RStudio

In this session, we will provide an overview of the seminar, including a brief course outline and address any organizational matters related to it.

We will cover introductory readings on the role of descriptive analysis, how to approach political science as a science (Kellstedt and Whitten 2018), and provide an introduction to statistical programming and data analysis using R (Imai and Williams 2022).

Throughout the lectures, we will utilize RStudio, a free user interface for R, as it offers a more user-friendly way to interact with the R environment.

Before our first class, please read the Tutorial: Getting Started with R and RStudio. Following the tutorial, you will be able to correctly install R and RStudio (in the right order: first R, then RStudio) and take the initial steps towards programming with R using RStudio.

Compulsory readings (read them in the order suggested):

- Kellstedt, Paul M., and Guy D. Whitten. 2018. "The Scientific Study of Politics." In.: The Fundamentals of Political Science Research. Cambridge: Cambridge University Press [Third Edition]. Chapter 1, pages 1-24.
- Imai, Kosuke and Nora Webb Williams. 2022. "Introduction to R and the tidyverse." In Quantitative Social Science: An Introduction in tidyverse. Princeton: Princeton University Press. Chapter 1, pages 1–38.

Optional Further readings:

* King, Gary, Robert O. Keohane, and Sidney Verba. 1994. "The Science in Social Science." In.: Designing Social Inquiry: Scientific Inference in Qualitative Research. Princeton: Princeton University Press. Chapter 1, pp. 3–33.
* King, Gary, Michael Tomz and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." AJPS 44(2): 341–355.
* Gujarati, Damodar N., and Dawn C. Porter. 2009. "Appendix B. Rudiments of Matrix Algebra." In.: *Basic Econometrics*. New York: McGraw-hill, pp. 838-848.
* Venables, W. N., D. M. Smith, and the *R* Core Team. 2023. *An Introduction to R*. Available at http://cran.r-project.org/doc/manuals/R-intro.pdf
* Verzani, John. 2002. *simpleR: Using R for Introductory Statistics*. Available at http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf
* Adler, Joseph. 2010. *R in a Nutshell.* Sebastopol: O'Reilly Media.

Week 2

30 July. Research Questions, Theoretical Models, and Exploratory Data Analysis

A good research question serves as the foundation for any successful research endeavour. However, coming up with interesting research questions can be a challenging task. In this session, we have readings to help us understand the process of generating (interesting) research questions and how to answer them by developing theoretical models.

In this lecture, we will also introduce Exploratory Data Analysis (EDA), a key step to summarize the main characteristics of your data. In EDA, we use simple graphs and charts to find patterns, spot unusual data points, and make sure you know your data before doing more detailed and advanced analysis.

Compulsory readings:

- Clark, William R. 2020. "Asking Interesting Questions." In.: The SAGE Handbook of Research Methods in Political Science and International Relations. Eds. Luigi Curini and Robert Franzese. Newbury Park: SAGE. Chapter 1, pages 7-25.
- Huntington-Klein, Nick. 2021. "Research Questions." In.: *The effect: An introduction to research design and causality*. Boca Raton: Chapman and Hall/CRC. Available at https://theeffectbook.net/ch-ResearchQuestions.html
- Kellstedt, Paul M., and Guy D. Whitten. 2018. "The Art of Theory Building." In.: The Fundamentals of Political Science Research. Cambridge: Cambridge University Press [Third Edition]. Chapter 2, pages 25–55.

Further readings:

* Maki, Daniel P., and Maynard Thompson. 1972. *Mathematical models and applications: With emphasis on the social life, and management sciences*. Englewood Cliffs: Prentice-Hall.
* Lave, Charles A. and James G. March. 1993. *An Introduction to Models in the Social Sciences*. New York: University Press of America.
* Morton, Rebecca. 1999. "What Makes a Model Formal?" In.: *Methods & Models: A Guide to the Empirical Analysis of Formal Models in Political Science*. Cambridge: Cambridge University Press. Chapter 2, pages 33–74.
* Bruce, Peter, Andrew Bruce and Peter Gedeck. 2020. "Exploratory Data Analysis." In.: Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python. Sebastopol: O'Reilly Media. Chapter 1, pages 1–46.
* Hullman, Jessica, and Andrew Gelman. 2021. "Designing for interactive exploratory data analysis requires theories of graphical inference." *Harvard Data Science Review* 3(3): 1–70.

Week 3

6 August. Week 3. Causality in Experimental Research Design

Statistical inference and causal inference are two fundamental approaches in political science research methods. Statistical inference involves drawing generalisations and making predictions about a population based on sample data, while causal inference seeks to establish cause-and-effect relationships between variables, allowing researchers to understand the causal mechanisms driving political outcomes.

In this lecture, we will define causality and explore it from the perspective of experimental research design, where researchers control the treatments and randomly assign participants to different groups (e.g., control and treatment groups). We will examine why this approach allows for greater confidence in establishing causality. Despite its advantages, we will also discuss the challenges of experimental research, including ethical concerns, feasibility issues, and the lack of external validity of its results.

Compulsory readings:

- Pearl, Judea and Dana Mackenzie. 2018. "The Ladder of Causation." In.: The Book of Why. Chapter 1, pp. 23-52.
- KW 2018. Chapters 3 and 4, pp: 56–103.
- IW 2022. "Causality." First Part of Chapter 2, pp: 38–64.
- Wickham, Hadley and Garrett Grolemund. 2023. *R for Data Science Import, Tidy, Transform, Visualize, and Model Data*. Sebastopol: O'Reilly Media. Chapter 1, "Data Visualization" and Chapter 2, "Workflow: Basics." Available at https://r4ds.hadley.nz.
- Wickham, Hadley. 2023. The tidyverse Style Guide. Availa https://style.tidyverse.org.

Optional further readings:

* Pearl, Judea. 2009. *Causality*. Cambridge: Cambridge University Press.
* Morton, Rebecca B. and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.
* Pearl, Judea, Madelyn Glymour and Nicholas P. Jewell. 2019. *Causal Inference in Statistics: A Primer*. Hoboken: John Wiley & Sons.
* Pearl, Judea and Dana Mackenzie. 2018. *The book of why: The new science of cause and effect*. New York: Basic Books.
* Huntington-Klein, Nick. 2021. *The effect: An introduction to research design and causality*. London: Chapman and Hall/CRC. Available at https://theeffectbook.net

Week 4.

13 August. Causality in Observational Studies

In this lecture, we will extend our discussion of causality to understand it in observational studies, which rely on naturally occurring data without any intervention or manipulation by researchers. While these studies are more practical and often used in political science, the lack of randomization in observational studies makes it more challenging to draw causal conclusions, and the potential for omitted variable bias poses a constant threat to the validity of results. Therefore, causal inference in observational studies requires controlling for confounding variables to minimize bias and approximate causality accurately.

Compulsory reading:

- IW 2022. "Causality." Second part of Chapter 2, pp: 65–87.
- Bueno de Mesquita, Ethan and Anthony Fowler. 2021. "Part I. Establishing a Common Language." In.: Thinking Clearly About Data: A Guide to Quantitative Reasoning and Analysis. Princeton: Princeton University Press. [Part I, pp: 11–52.]
- Cunningham, Scott. 2021. "Directed Acyclic Graphs." In.: Causal Inference: The Mixtape. New Haven: Yale University Press. [Chapter 3, pp: 96–118.]

Optional further readings:

* Rubin, Donald. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5): 688-701.
* Brady, Henry. 2003. "Models of Causal Inference." Paper presented at the Midwest Political Science Association Conference. University of California, Berkeley.
* Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.
* Cunningham, Scott. 2021. *Causal Inference: The Mixtape*. New Haven: Yale University Press.

Week 5.

20 August. Measurement and Types of Data

In this lecture, we will cover the fundamental aspects of measurement in political science research, emphasizing the importance of selecting appropriate variables and indicators to study political phenomena. Students will learn about different data types and their respective characteristics and use in research. Furthermore, the class will address various types of variables, such as continuous, nominal, ordinal, and categorical, and their implications for data analysis and interpretation in political science research. The students will also explore the data coding process and the significance of reliability and validity of their measures.

Compulsory readings:

- KW 2018. Chapters 5 and 6, pages: 104–142.
- IW 2022. Chapter 3, pages: 88–143.
- Wickham, Çetinkaya-Rundel and Grolemund. 2023. "Data Visualization." In.: R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. Sebastopol: O'Reilly. Chapter 1, pp: 1-31. Available at https://r4ds.hadley.nz/data-visualize

Optional further readings:

* Collier, David, and James E. Mahon. 1993. "Conceptual 'stretching' revisited: Adapting categories in comparative analysis." *American Political Science Review* 87(4): 845-855.
* Collier, David, and John Gerring. 2009. *Concepts and method in social science: the tradition of Giovanni Sartori*. London: Routledge.
* Munck, Gerardo L. 2009. *Measuring democracy: A bridge between scholarship and politics*. Baltimore: John Hopkins University Press.
* Schedler, A. 2012. "Judgment and Measurement in Political Science." *Perspectives on Politics* 10(1):21–36.
* Curini et al. 2020. "Part 3: Conceptualization and Measurement." In.: *The SAGE Handbook of Research Methods*. Pages: 331-384.
* Goertz, Gary. 2020. *Social science concepts and measurement: New and completely revised edition*. Princeton: Princeton University Press.

Week 6.

27 August. Week 6. Probability, Statistical Distributions and Hypothesis Testing

In this lecture, we will cover the fundamental concepts of probability, statistical distributions, and hypothesis testing. We will start with probability, which quantifies the likelihood of different outcomes and forms the basis for statistical distributions and hypothesis testing. Understanding probability helps us grasp the randomness in data and the behavior of sample averages.

Next, we will explore various statistical distributions, including the normal, binomial, and Poisson distributions. These distributions describe how data points are spread out and are crucial for modeling political data. Through simulation, we will understand the Central Limit Theorem, which states that the distribution of sample means approximates a normal distribution as the sample size gets larger, regardless of the population's distribution. This is a crucial assumption in regression analysis, which we will cover in the second half of this course.

Additionally, we will introduce the principles of hypothesis testing, focusing on formulating null and alternative hypotheses, selecting basic tests (e.g., difference of means), and interpreting p-values.

Compulsory readings:

- KW 2018. Chapters 7 and 8, pp.: 143–187.
- IW 2022. Chapter 6, pp.: 279-356.

Further readings:

* Wooldridge, Jeffrey. 2019. "Math Refresher B: Fundamentals of Probability." In.: Introductory Econometrics. Boston: Cengage, pp: 684-713.
* Cunningham, Scott. 2021. "Probability and Regression Review." In.: Causal Inference: The Mixtape. New Haven: Yale University Press. From Sections 2.1 to Section 2.12 only.
* Hansen, Bruce E. 2022. "Chapter 9: Hypotheses Testing." In *Econometrics*. Princeton: Princeton University Press, pp: 225–261.
* Forbes, Catherine, Merran Evans, Nicholas Hastings, and Brian Peacock. 2011. *Statistical Distributions*. Hoboken: John Wiley & Sons.

Week 7.

17 September. OLS I: Introduction to Regression Analysis

This lecture covers the fundamentals of ordinary least squares (OLS) regression analysis. Students will begin by learning about correlation, including graphical visualization techniques such as scatter plots. We will then introduce the population regression model, the sample regression model, and the concept of the regression line. The lecture will also cover the assumptions underlying OLS and the consequences of violating these assumptions. Emphasizing the importance of regression analysis in political science research, students will work with basic examples that include only a dependent variable and one independent variable. The focus will be on estimation, interpreting coefficients, and assessing model fit.

Compulsory reading:

- David Freeman, Robert Pisani and Roger Purves. 2007. "Chapter 8: Correlation" and "Chapter 9: More about Correlation." In.: Statistics. New York: W. W. Norton Company, pp.: 119-157.
- KW 2018. Chapter 9, pp.: 188–214.
- IW 2022. Chapter 4. Section 4.2 Linear Regression, pp. 162-187.

Optional further readings:

* Wooldridge, Jeffrey M. 2019. "The Simple Regression Model." In.: Introductory Econometrics: A Modern Approach. Boston: Cengage Learning. Chapter 2, pp.: 20–65.
* Gujarati, Damodar N., and Dawn C. Porter. 2009. "Chapter 2: Two-Variable Regression Analysis: Some Basic Ideas" and "Chapter 3: Two-Variable Regression Model: The Problem of Estimation." In.: *Basic Econometrics*. New York: McGraw-hill, pp. 34-95.
* Bueno de Mesquita, Ethan and Anthony Fowler. 2021. "Chapters 5 and 6." In.: *Thinking Clearly About Data: A Guide to Quantitative Reasoning and Analysis*. Princeton: Princeton University Press, pp.: 74-112.

Week 8.

24 September. OLS II: Multiple Regression - Model Specification and Estimation

This lecture introduces multiple regression analysis, where more than one independent variable is included in the model. Students will learn the principles of model specification, how to estimate and interpret multiple regression models, and the two key reasons why correlation does not imply causation: the potential presence of confounders and the possibility of reverse causation. The lecture will also cover the assessment of multicollinearity and the importance of proper model specification to ensure unbiased and consistent estimates.

Compulsory readings:

- Bueno de Mesquita, Ethan and Anthony Fowler. 2021. "Chapter 9: Why Correlation Doesn't Imply Causation" In.: *Thinking Clearly About Data: A Guide to Quantitative Reasoning and Analysis*. Princeton: Princeton University Press, pp.: 159-192.
- KW 2018. Chapters 10 and 11, pp.: 215–272.

Optional further readings:

* Gujarati, Damodar N., and Dawn C. Porter. 2009. "Chapter 7: Multiple Regression Analysis: The Problem of Estimation." In.: *Basic Econometrics*. New York: McGraw-hill, pp. 188-232.
* Greene, William H. 2018. "The Linear Regression Model." In.: Econometric Analysis. New York: Pearson. Chapter 2, pp.: 12-27.
* Wooldridge, Jeffrey M. 2019. "Chapters 3 and 4." In.: Introductory Econometrics: A Modern Approach. Boston: Cengage Learning. Chapter 3, pp.: 66–162.
* Bueno de Mesquita, Ethan and Anthony Fowler. 2021. "Chapter 10: Controlling for Confounders" In.: *Thinking Clearly About Data: A Guide to Quantitative Reasoning and Analysis*. Princeton: Princeton University Press, pp.: 159-192.

Week 9.

1 October. OLS III: Interaction Models

Most hypotheses in political science are conditional, meaning that the relationship between variables often depends on the context or level of another variable. Therefore, it is essential to understand interactive multiple regression models. The focus of this lecture will be on how to specify, estimate, and interpret models that include interaction terms between independent variables accurately. The lecture will cover the theoretical rationale for interactions and demonstrate their application in political science research.

Compulsory readings:

- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding interaction models: Improving empirical analyses." Political Analysis 14(1): 63–82.
- Berry, William, Matt Golder, and Daniel Milton. 2012. "Improving Tests of Theories Positing Interaction." *Journal of Politics* 74(3): 653–671.

Optional further readings:

* Franzese, Robert, and Cindy Kam. 2009. *Modeling and interpreting interactive hypotheses in regression analysis*. Ann Arbor: University of Michigan Press.
* Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models." *Political Analysis* 27: 163-192.
* Block Jr., Ray, Matt Golder and Sona N. Golder. 2023. "Evaluating Claims of Intersectionality." *The Journal of Politics* 85(3): 795–811.
* Clark, William Roberts and Matt Golder. 2024. *Interaction Models: Specification and Interpretation*. Cambridge: Cambridge University Press.

Week 10.

8 October. Regression Discontinuity Design

This lecture introduces the regression discontinuity design (RDD), a method to approximate causal inference in observational studies by using a cutoff or threshold in an assignment variable to identify causal effects. By comparing observations just above and just below this cutoff, RDD approximates a randomized experiment, allowing researchers to estimate the causal impact of an intervention or treatment. Students will learn the key principles and assumptions of RDD, how to implement the design, and how to interpret the results. The lecture will also include practical replications to illustrate the application of RDD in political science research.

Compulsory readings:

- Bueno de Mesquita, Ethan and Anthony Fowler. 2021. "Chapter 12: Regression Discontinuity Designs" In.: *Thinking Clearly About Data: A Guide to Quantitative Reasoning and Analysis*. Princeton: Princeton University Press, pp.: 243-265.
- Angrist, Joshua and J. Pischke. 2015. "Regression Discontinuity Designs" In.: Mastering Metrics: The Path from Cause to Effect. Chapter 4, pp. 147-177.
- IW 2022. Chapter 4. Section 4.3 Regression and Causation, pp. 188-215.

Optional further readings:

* Cattaneo, Idrobo and Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge: Cambridge University Press.
* Cattaneo, Idrobo and Titiunik. 2023. *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Cambridge: Cambridge University Press.

Week 11.

15 October. An Introduction to MLE: Logistic Regression

This lecture covers maximum likelihood estimation (MLE) with a focus on logistic regression. In political science, many phenomena of interest are dichotomous, meaning they occur or do not occur, or are present or absent. Students will learn how to model binary dependent variables using logistic regression, understand the concept of the likelihood function, and interpret coefficients and odds ratios. The lecture will also discuss the differences between OLS and logistic regression models.

Compulsory readings:

- KW 2018. First part of Chapter 12, pp.: 273–282.
- Pampel, Fred C. 2020. "Chapters 1 and 2." In.: *Logistic Regression: A Primer*. Thousand Oaks: Sage.

Optional further readings:

* Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks: Sage.
* Gujarati, Damodar N., and Dawn C. Porter. 2009. "Chapter 15: Qualitative Response Regression Models." In.: *Basic Econometrics*. New York: McGraw-Hill, pp. 541-590.

* Pampel Fred C. 2020. *Logistic Regression: A Primer*. Thousand Oaks: Sage.
* Hagenaars, Jacques A. P., Steffen Kühnel and Hans-Jürgen Andress. 2024. *Interpreting and Comparing Effects in Logistic, Probit and Logit Regression*. Thousand Oaks: Sage.

Week 12.
22 October. An Introduction to Time Series Analysis

Given that most political phenomena unfold over time and are influenced by past events, understanding time series analysis is crucial in political science research. This lecture provides a brief introduction to time series analysis, covering the characteristics of time series data, the importance of stationarity, and techniques to model temporal dependencies. The lecture will cover basic methods such as autoregressive (AR) models and moving average (MA) models, as well as introduce more advanced techniques like ARIMA models, to effectively analyze and interpret temporal data.

Compulsory readings:

- KW 2018. Second part of Chapter 12, pp.: 282–298.
- De Boef, Suzanna, and Luke Keele. 2008. "Taking time seriously." *American Journal of Political Science* 52(1): 184–200.

Optional further reading:

* Gujarati, Damodar N., and Dawn C. Porter. 2009. "Chapters 16 and 17" In.: *Basic Econometrics*. New York: McGraw-hill, pp. 591-672.
* Box-Steffensmeier, Janet M., John R. Freeman, Matthew P. Hitt, and Jon CW Pevehouse. 2014. *Time Series Analysis for the Social Sciences*. Cambridge: Cambridge University Press.

Changes to Syllabus

The convener reserves the right to update/modify/clarify the syllabus with advance notification.