Stefan Müller, PhD
Assistant Professor and Ad Astra Fellow
School of Politics and International Relations
University College Dublin
Belfield, Dublin 4, Ireland
✉ stefan.mueller@ucd.ie
🖥 https://muellerstefan.net

Autumn Trimester 2022

Introduction to Statistics (POL40950)

Draft – Version: September 12, 2022

Latest version at: https://muellerstefan.net/teaching/2022-autumn-introstats.pdf

---

Term: Autumn Trimester 2022
Time: Lecture & Lab: Mon, 12:00–13:45
Location: QUI-009 (12:00); G5-DAE (13:00)
Credits: 10.0
Format: Lectures; lab work

Module Coordinator: Stefan Müller, PhD
stefan.mueller@ucd.ie
https://muellerstefan.net
Office: Newman Building, G312
Office hours: Tue, 14:30–15:30; Thu, 9:30–10:30
(sign up here)

Introduction

Welcome to Introduction to Statistics! In this module, you will learn about concepts such as measurement, variables, statistical data and get equipped to answer a social science research question using linear statistical models and the statistical programming language R.

Do you want to know whether more informed voters are more likely to have liberal values? Are political parties responsive to the issue priorities by voters? Are democracies less likely to initiate a war? Do high tax rates lead to higher levels of corruption? Can voters accurately predict the government that will be formed after an election? Answering such questions usually requires the analysis of data – information about people, parties, communication, firms, or nations.

There are many other statistical tools available to social scientists, but regression analysis is by far the most common. A thorough understanding of this method is required to read or write quantitative social science papers and research reports. Therefore, the course will mainly focus on regression analysis – including model specification (which variables to include in a model?) and statistical inference (how do I know whether my findings hold for cases beyond my sample?).

By the end of this module, you will have gained a basic understanding of statistics and the so-called frequentist approach of hypothesis testing. The lab sessions and two homework assignments will familiarise you with the R statistical programming language and prepare you two write a quantitative research paper.

The core textbook for the course is Ismay and Kim (2020), which is freely available online at https://moderndive.com.[1] This book takes a modern, data science approach to regression analysis. The differences between data science and more typical quantitative social science will be discussed in class, particularly in the context of model specification. In addition, the textbook by Llaudet and Imai (Forthcoming)[2] provides intuitive examples and explanations for most of the concepts discussed in this module.

We will read parts of the following textbooks for the applied components of this course, such as data import, data wrangling, and data visualisation. Both books will also help you with your homework assignments and the technical elements of your course paper.

- H. Wickham and G. Grolemund (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Sebastopol: O'Reilly. URL: https://r4ds.had.co.nz.

- K. Healy (2019). *Data Visualization: A Practical Introduction*. Princeton: Princeton University Press. URL: https://socviz.co.

We will work extensively with the R statistical programming language. The three books mentioned above (Ismay and Kim 2020; Wickham and Grolemund 2017; Healy 2019) provide intuitive examples and the corresponding R code (based on the `tidyverse` approach). In addition to these books, I recommend the following literature for introductions to statistical methods, regression, causal inference, and R:

- **Basic grasp of statistics and quantitative methods**:

  - D. Spiegelhalter (2020). *The Art of Statistics: Learning from Data*. London/New York: Penguin Books.

- **Research design and causal inference**:

  - N. Huntington-Klein (2021). *The Effect: An Introduction to Research Design and Causality*. Boca Raton: CRC Press. URL: https://theeffectbook.net.

  - S. Cunningham (2021). *Causal Inference: The Mixtape*. New Haven: Yale University Press. URL: https://mixtape.scunning.com.

- **R and regression analysis**:

  - D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf.

  - E. G. Larsen and Z. Fazekas (2021). *Quantitative Politics with R*. URL: http://qpolr.com.

- **Data visualisation**: C. O. Wilke (2019). *Fundamentals of Data Visualization: A Primer On Making Informative and Compelling Figures*. Sebastopol: O'Reilly. URL: https://clauswilke.com/dataviz/.

- **Quarto** (an open-source scientific and technical publishing system): URL: https://quarto.org/docs/get-started/hello/rstudio.html.

- **Online Lectures and Tutorials**: several scholars publish excellent videos on statistics, theoretical concepts, coding, and regression analysis. Below are selected channels that may be helpful for a recap of concepts and applications:

  - **Introduction to R for Economists** (N. Huntington-Klein): [YouTube Link]

---

[1] C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press. URL: https://moderndive.com.

[2] E. Llaudet and K. Imai (Forthcoming). *Data Analysis for Social Science: A Friendly Introduction*. Princeton: Princeton University Press. A link to the book was kindly provided by the authors. The URL will be distributed via Brightspace.

– **Quantitative Social Science Methods I** (G. King): [YouTube Link]

Learning Outcomes

1. basic understanding of working with R and RStudio

2. being able to wrangle, summarise, describe, and visualise statistical data

3. basic understanding of (frequentist) statistical inference

4. basic understanding of executing and interpreting multiple regression

5. preliminary understanding of logistic regression

Software

We will use the free statistical programming language R and the integrated development environment RStudio. Make sure to install both programs before the start of the module. Please update the programs to the latest versions if you have previously worked with R and RStudio. Ismay and Kim (2020: ch. 1) show how to install the required software.

- Download R for Mac, Windows, or Linux at: https://cloud.r-project.org

- Download RStudio (Desktop Open Source License [free]) at: https://rstudio.com/products/rstudio/download

Approaches to Teaching and Learning

The sessions consist of lectures and labs each week. The lectures are the first part of each session (12:00–12:45, QUI-009) and focus on the fundamental aspects of statistical inference and the interpretation of these methods and examples.

In the lab sessions (13:00–13:45, G5-DAE), students will be provided with clear instructions and solve problems related to data wrangling, visualisation, and statistical methods. The homework assignments are structured so that they gradually lead up to a comprehensive regression analysis and associated social science paper, putting the technical material of the class into practice.

I will distribute several short feedback surveys during the term in which you can ask questions and make improvement suggestions. We will communicate via Slack, a business communication platform, to clarify questions about the course content, homeworks, code, and examples. I will share the link to the Slack workspace on Brightspace. Students should check the Slack workspace at least three times per week. Questions about R code should be posted in the `#coding` channel. In previous modules, Slack facilitated communication and collaboration throughout the module.

Overview of Assessment

- Homework assignment (Week 4): 25%

- Homework assignment (Week 7): 25%

- Course paper (12 December 2022): 50%

Expectations and Grading

Students submit two **homework assignments** during the term (at the end of Week 4 and Week 7). Each homework counts toward 25% of the final grade. The homeworks will be distributed via Brightspace 14 days before the submission deadline as a Quarto file. We will discuss how to create and compile Quarto files in the first two weeks of the module. Students fill in the answers and solutions in the same Quarto file, rename it to `hw_01/02_surname_firstname.qmd`, knit it as an `html` file, and submit it via Brightspace. Only knitted `html` files will be accepted. More details on the homeworks will be provided throughout the module. Table 1 shows the grade conversion scheme that will be applied to the homework assignments.

Table 1: Grade conversion scheme for homeworks

Homeworks	UCD Grade
97–100%	A+
94–96%	A
91–93%	A–
88-90%	B+
85–87%	B
83–84%	B
80–82%	C+
77–79%	C
74–76%	C–
71–73%	D+
68–70%	D
65–67%	E+
54–64%	E
44–53%	E–
0–32%	F

Students also submit a **course paper** which counts toward 50% of the final grade. The research paper is a written analysis consisting of 4,000 words (including bibliography, captions, and footnotes). Students are required to answer a research question using quantitative methods and regression analysis. Students are free to answer questions from all fields of social science but must justify their choice and the relevance of the question. The course paper must address the following aspects: research gap and relevance; theory and expectations (based on previous research); data and methodological approach; results; conclusion and outlook. The course paper must be submitted via Brightspace as a `pdf` document by **Monday, 12 December 2022 (8pm)**. Detailed instructions on the research paper, the presentation, and the in-class discussion will be provided in class and on Brightspace.

For information on academic writing, I recommend the following two sources:

- P. Dunleavy (2014). *How to Write Paragraphs in Research Texts (Articles, Books and PhDs)*. URL: https://medium.com/advice-and-help-in-authoring-a-phd-or-non-fiction/how-to-write-paragraphs-80781e2f3054.

- S. B. Heard (2016). *The Scientist's Guide to Writing: How to Write More Easily and Effectively Throughout Your Scientific Career*. Princeton: Princeton University Press.

If you require information on proper citation style, please refer to the guidelines of the American Political Science Association:

- APSA Committee on Publications (2018). *Style Manual for Political Science (Revised 2018 Version)*. URL: https://connect.apsanet.org/stylemanual/.

Table 2: Student effort hours

Student effort type	Hours
Lectures	12
Computer Aided Lab	12
Autonomous Student Learning	200
Total	**224**

Plagiarism

Although this should be obvious, plagiarism – copying someone else's text without acknowledgement or beyond 'fair use' quantities – is not allowed. Plagiarism is an issue we take very serious here in UCD. Please familiarize yourself with the definition of plagiarism on UCD's website[3] and make sure not to engage in it.

Extenuating Circumstances

In the case that a student will not be able to meet an assessment deadline or will be absent from the course for an extended period of time, and this is known in advance, they should consult the UCD policies on extenuating circumstances found here: https://www.ucd.ie/students/studentdesk/extenuatingcircumstances/.

It is important that in such cases you make the issue known to the lecturer as soon as possible. The sooner that the lecturer is made aware of the situation, the more likely it is that you can be accommodated.

Late Submission Policy

If a student submits an assignment late, the following penalties will be applied:

- Coursework received at any time within two weeks of the due date will be graded, but a penalty will apply.
    - Coursework submitted at any time up to one week after the due date will have the grade awarded reduced by two grade points (for example, from $B-$ to C).
    - Coursework submitted more than one week but up to two weeks after the due date will have the grade reduced by four grade points (for example, from $B-$ to $D+$). Where a student finds they have missed a deadline for submission, they should be advised that they may use the remainder of the week to improve their submission without additional penalty.
- Coursework received more than two weeks after the due date will not be accepted. Regulations regarding extenuating circumstances apply.

Office Hours

We can meet either during my office hour on Monday from 14:30–15:30 (in person (Room G312, Newman Building or online via Zoom) and Thursday from 09:30–10:30 (online via Zoom). Please sign up for a meeting at https://calendly.com/mueller-ucd/office-hours.

---

[3]https://libguides.ucd.ie/academicintegrity.

Questions and Problems

In this module, we will discuss concepts, methods, and software you might not have heard of before. I know that parts of this module could be challenging and I will assist you as best as I can. In addition to the lectures and lab sessions, I offer weekly office hours only for participants of this module. The office hours will take place in person or via Zoom. I will share the link and password to the virtual room in the first lecture and post it on Brightspace.

If you struggle to solve problems relating to R or RStudio, please follow the steps outlined below before contacting me. At least one other person likely faced the same problem before or received the same error message.

1. Use the 'Search' function in the online books of the recommended textbooks (Ismay and Kim 2020; Wickham and Grolemund 2017; Healy 2019) and look up keywords related to your problem or the function that causes a problem. For questions about concepts, I recommend consulting the Glossary of Statistical Terms.

2. Try to summarise the problem in your own words and then google this summary. If the problem relates to R, add `rstats` to your search query. For example: `how to import csv file in rstats`. I am almost certain that you find a solution to most of your questions.

3. If your R code returns an error, I would advise you to Google the text of the error message. For example: `Error:  Can't subset columns that don't exist.`

⟶ If steps 1–3 still do not solve your problem or question, please send a message on Slack. I and your classmates are happy to help!

Syllabus Modification Rights

I reserve the right to reasonably alter the elements of the syllabus at any time by adjusting the reading list to keep pace with the course schedule. Moreover, I may change the content of specific sessions, depending on the participants' prior knowledge and research interests. If I make adjustments, I will send an email to all seminar participants and upload the revised syllabus to Brightspace.

Course Structure

Week 1: Accessing and Visualising Data I (12 September 2022)

What is quantitative political science? What are data? What is a variable?

Important: Please install R[4] and RStudio[5] before the start of the first lecture. Ismay and Kim (2020: ch. 1) provide detailed instructions on how to install the required software.

Mandatory Readings

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 1.

- D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf: ch. 3.

Optional

- J. Bryan (2019). *Data Wrangling, Exploration, and Analysis with R*. URL: https://stat545.com: ch. 2.

- D. Spiegelhalter (2020). *The Art of Statistics: Learning from Data*. London/New York: Penguin Books: ch. 2.

Week 2: Accessing and Visualising Data II (19 September 2022)

What is the unit of analysis. What are the different levels of measurement. How to summarise a dataset. How to describe your variables graphically. How to look at a distribution.

Mandatory Readings

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 2.

- D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf: ch. 4.

Optional

- J. Bryan (2019). *Data Wrangling, Exploration, and Analysis with R*. URL: https://stat545.com: ch. 2.

---

[4]Download R for Mac, Windows, or Linux at: https://cloud.r-project.org
[5]Download RStudio (Desktop Open Source License) at: https://rstudio.com/products/rstudio/download

Week 3: Descriptive Statistics I (26 September 2022)

How to describe your variables numerically, including the mean, mode, median, variance, and standard deviation.

Mandatory Readings

- K. Healy (2019). *Data Visualization: A Practical Introduction*. Princeton: Princeton University Press: ch. 1.

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 3–4.

Optional

- D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf: ch. 5.

- G. Wilson, J. Bryan, K. Cranston, J. Kitzes, L. Nederbragt, and T. K. Teal (2017). "Good Enough Practices in Scientific Computing". *PLoS Computational Biology* 13 (6): e1005510.

- E. G. Larsen and Z. Fazekas (2021). *Quantitative Politics with R*: ch. 3.

Week 4: Descriptive Statistics II (3 October 2022)

How to describe relations between variables graphically, including bar charts, scatterplots, and boxplots. Discussion of covariance and correlation.

Mandatory Readings

- E. Llaudet and K. Imai (Forthcoming). *Data Analysis for Social Science: A Friendly Introduction*. Princeton: Princeton University Press: ch. 3.5.

- A. McCauley and A. Ruggeri (2020). "From Questions and Puzzles to Research Project". *The SAGE Handbook of Research Methods in Political Science and International Relations*. Ed. by L. Curini and R. Franzese. Thousand Oaks: Sage: 26–43.

- G. Firebaugh (2008). *Seven Rules for Social Research*. Princeton: Princeton University Press: ch. 1.

Optional

- D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf: ch. 5.7.

Week 5: Simple Regression (10 October 2022)

Descriptive univariate linear regression models – how to look at the relation between two continuous variables.

Mandatory Readings

- E. Llaudet and K. Imai (Forthcoming). *Data Analysis for Social Science: A Friendly Introduction*. Princeton: Princeton University Press: ch. 4.

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 5.

Optional

- D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf: ch. 15.

- P. M. Kellstedt and G. D. Whitten (2018). *The Fundamentals of Political Science Research*. 3rd edition. Cambridge: Cambridge University Press: ch. 9.

Week 6: Multiple Regression (17 October 2022)

How to perform and interpret regression models with more than one independent variable. How to think about the difference between prediction and causal inference? Some discussion of model specification.

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 6.

- P. M. Kellstedt and G. D. Whitten (2018). *The Fundamentals of Political Science Research*. 3rd edition. Cambridge: Cambridge University Press: ch. 10.

Optional

- E. G. Larsen and Z. Fazekas (2021). *Quantitative Politics with R*: ch. 11.

- M. A. Hardy (1993). *Regression with Dummy Variables*. Newbury Park, CA: SAGE Publications.

Week 7: Multiple Regression – Categorical Independent Variables and Interactions (24 October 2022)

Categorical independent variables in multiple regression. Modeling interaction effects in multiple regression.

Mandatory Readings

- J. Fox (2015). *Applied Regresssion Analysis and Generalized Linear Models*. 3rd edition. Los Angeles: SAGE: ch. 7.

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 6.1.2.

Optional

- J. Hainmueller, J. Mummolo, and Y. Xu (2019). "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice". *Political Analysis* 27 (2): 163–192.

- W. R. Clark, M. J. Gilligan, and M. Golder (2006). "A Simple Multivariate Test for Asymmetric Hypotheses". *Political Analysis* 14 (3): 311–331.

Week 8: Bank Holiday: OPTIONAL – Virtual Q&A Session (31 October 2022)

The lecture and lab session are cancelled due to the Bank Holiday on 31 October. However, I will facilitate a virtual Q&A session, dealing with questions focusing on questions relating to theory, coding, or your final research projects. This session will take place on Zoom. We will discuss a date and time for this meeting in Week 2.

Week 9: Sampling Distributions and Central Limit Theorem (7 November 2022)

What are probabilities and probability distributions? Introduction to the normal distribution. What is statistical inference? Introduction to sampling methods. What is the Central Limit Theorem?

Mandatory Readings

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 8.

- P. M. Kellstedt and G. D. Whitten (2018). *The Fundamentals of Political Science Research*. 3rd edition. Cambridge: Cambridge University Press: ch. 7.

Optional

- E. Llaudet and K. Imai (Forthcoming). *Data Analysis for Social Science: A Friendly Introduction*. Princeton: Princeton University Press: ch. 6.

- R. Hoekstra, R. D. Morey, J. N. Rouder, and E.-J. Wagenmakers (2014). "Robust Misinterpretation of Confidence Intervals". *Psychonomic Bulletin & Review* 21 (5): 1157–1164.

Week 10: Hypothesis Tests and Confidence Intervals (14 November 2022)

What are hypothesis tests and confidence intervals? How to think of statistical inference in multiple regression analysis.

- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 9–10.

- R. L. Wasserstein and N. A. Lazar (2016). "The ASA's Statement on p-Values: Context, Process, and Purpose". *The American Statistician* 70 (2): 129–133.

Optional

- D. Spiegelhalter (2020). *The Art of Statistics: Learning from Data*. London/New York: Penguin Books: ch. 7.

Week 11: Multiple Regression – Diagnostics and Model Fit (21 November 2022)

How to think about model fit in the contexts of prediction and causal inference. Statistical versus modelling considerations in model specification. Common problems in regression analysis (and hints at solutions).

- D. Navarro (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf: ch. 15.9.
- P. M. Kellstedt and G. D. Whitten (2018). *The Fundamentals of Political Science Research*. 3rd edition. Cambridge: Cambridge University Press: ch. 11.

Week 12: Advanced Regression Models (28 November 2022)

Regression analysis when the dependent variable is binary – e.g. explaining whether or not a citizens turns out to vote on election day. Introduction to logistic regression. Intuition behind panel regression models

- P. M. Kellstedt and G. D. Whitten (2018). *The Fundamentals of Political Science Research*. 3rd edition. Cambridge: Cambridge University Press: ch. 12.
- P. Roback and J. Legler (2020). *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*. Boca Raton: CRC Press: ch. 6.

Optional

- J. Gill and M. Torres (2019). *Generalized Linear Models: A Unified Approach*. Newbury Park, CA: SAGE Publications.

Week 13: OPTIONAL – Reporting Data, Methods, and Results (5 December 2022)

How to present and interpret regression results. How to structure a quantitative research paper. How to convince the reader of the robustness of your results.

Mandatory Readings

- G. King (2006). "Publication, Publication". *PS: Political Science & Politics* 39 (1): 119–125.
- C. Ismay and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press: ch. 11.

Optional

- K. Healy (2019). *Data Visualization: A Practical Introduction*. Princeton: Princeton University Press: ch. 4–5.
- D. Spiegelhalter (2020). *The Art of Statistics: Learning from Data*. London/New York: Penguin Books: ch. 12.

References

APSA Committee on Publications (2018). *Style Manual for Political Science (Revised 2018 Version)*. URL: https://connect.apsanet.org/stylemanual/.

Bryan, J. (2019). *Data Wrangling, Exploration, and Analysis with R*. URL: https://stat545.com.

Clark, W. R., M. J. Gilligan, and M. Golder (2006). "A Simple Multivariate Test for Asymmetric Hypotheses". *Political Analysis* 14 (3): 311–331.

Cunningham, S. (2021). *Causal Inference: The Mixtape*. New Haven: Yale University Press.

Dunleavy, P. (2014). *How to Write Paragraphs in Research Texts (Articles, Books and PhDs)*. URL: https://medium.com/advice-and-help-in-authoring-a-phd-or-non-fiction/how-to-write-paragraphs-80781e2f3054.

Firebaugh, G. (2008). *Seven Rules for Social Research*. Princeton: Princeton University Press.

Fox, J. (2015). *Applied Regresssion Analysis and Generalized Linear Models*. 3rd edition. Los Angeles: SAGE.

Gill, J. and M. Torres (2019). *Generalized Linear Models: A Unified Approach*. Newbury Park, CA: SAGE Publications.

Hainmueller, J., J. Mummolo, and Y. Xu (2019). "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice". *Political Analysis* 27 (2): 163–192.

Hardy, M. A. (1993). *Regression with Dummy Variables*. Newbury Park, CA: SAGE Publications.

Healy, K. (2019). *Data Visualization: A Practical Introduction*. Princeton: Princeton University Press.

Heard, S. B. (2016). *The Scientist's Guide to Writing: How to Write More Easily and Effectively Throughout Your Scientific Career*. Princeton: Princeton University Press.

Hoekstra, R., R. D. Morey, J. N. Rouder, and E.-J. Wagenmakers (2014). "Robust Misinterpretation of Confidence Intervals". *Psychonomic Bulletin & Review* 21 (5): 1157–1164.

Huntington-Klein, N. (2021). *The Effect: An Introduction to Research Design and Causality*. Boca Raton: CRC Press.

Ismay, C. and A. Y. Kim (2020). *Statistical Inference via Data Science: A ModernDive into R and the tidyverse*. Boca Raton: CRC Press.

Kellstedt, P. M. and G. D. Whitten (2018). *The Fundamentals of Political Science Research*. 3rd edition. Cambridge: Cambridge University Press.

King, G. (2006). "Publication, Publication". *PS: Political Science & Politics* 39 (1): 119–125.

Larsen, E. G. and Z. Fazekas (2021). *Quantitative Politics with R*.

Llaudet, E. and K. Imai (Forthcoming). *Data Analysis for Social Science: A Friendly Introduction*. Princeton: Princeton University Press.

McCauley, A. and A. Ruggeri (2020). "From Questions and Puzzles to Research Project". *The SAGE Handbook of Research Methods in Political Science and International Relations*. Ed. by L. Curini and R. Franzese. Thousand Oaks: Sage: 26–43.

Navarro, D. (2022). *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners*. Version 0.6. URL: https://learningstatisticswithr.com/lsr-0.6.pdf.

Roback, P. and J. Legler (2020). *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*. Boca Raton: CRC Press.

Spiegelhalter, D. (2020). *The Art of Statistics: Learning from Data*. London/New York: Penguin Books.

Wasserstein, R. L. and N. A. Lazar (2016). "The ASA's Statement on p-Values: Context, Process, and Purpose". *The American Statistician* 70 (2): 129–133.

Wickham, H. and G. Grolemund (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Sebastopol: O'Reilly.

Wilke, C. O. (2019). *Fundamentals of Data Visualization: A Primer On Making Informative and Compelling Figures*. Sebastopol: O'Reilly.

Wilson, G., J. Bryan, K. Cranston, J. Kitzes, L. Nederbragt, and T. K. Teal (2017). "Good Enough Practices in Scientific Computing". *PLoS Computational Biology* 13 (6): e1005510.