PLSC 503: "Multivariate Analysis for Political Research"

Spring 2025

Professor Christopher Zorn
Department of Political Science
Pennsylvania State University
E-mail: zorn@psu.edu
Mondays, 9:00 a.m. - 12:00 p.m. ET
Willard 370

Course Description

This is the second (full) course in quantitative methods in Penn State's political science Ph.D. program. The course introduces students to regression models for the analysis of quantitative data, and provides a basis of knowledge for more advanced statistical methods. It will also have a substantial programming/computation focus. The course assumes basic math literacy, including familiarity with probability theory, properties of estimators, rudimentary calculus, and linear algebra, as well as mastery of the basic statistics taught in PLSC 502. The bulk of the course will focus on regression models for continuous response variables, and will include discussions of the mathematical bases for such models, their estimation and interpretation, model assumptions and techniques for addressing violations of those assumptions, model diagnostics, and topics related to model specification and functional forms. The course will also address a range of other topics, including missing data, regularization, validation and cross-validation, and tools for robust inference (jackknife, bootstrap, randomization, and so forth), as well as maximum likelihood and generalized linear models (logit, Poisson regression, etc.). In short, this is designed to be a "modern regression" course.

Note that all course materials (including this syllabus, slides, notes, data, computer code, homework exercises, etc.) are or will be available on a dedicated Github repo, which can be found at https://github.com/PrisonRodeo/PLSC503-2025-git. Throughout this syllabus, hot links are in Penn State Blue.

Texts

"Required":

The scare quotes around "required" are there because these texts are not, actually, required; the topics we'll cover from them can be found in lots of places, including for free on the web (e.g., here and here).

Weisberg, Sanford. 2014. *Applied Linear Regression*, 4th Ed. New York: Wiley. (See also *ALR*'s Wiley page and Weisberg's homepage.)

Faraway, Julian J. 2016. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression*, 2nd Ed. London: Chapman & Hall.

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and Other Stories*. New York: Cambridge University Press. It's terrific, covers a *lot* of ground, and has a good webpage. It's also *very* Bayesian (neither good nor bad, just the truth) and very reliant on Stan (and `rstan`).

Additional readings as necessary, all of which will be available via JSTOR™or on the course github repo.

The Weisberg text will be the source of most listed readings in the course. It's expensive, but a good reference, and is available both used and in an e-book version. Faraway is better for the latter part of the course, and may also be used in PLSC 504. There are many first editions of Faraway available used; either version will work for this course. Gelman-Hill-Ventari (hereinafter "*GHV*") will be referenced periodically; it's perhaps good to have an IRL copy, but there's a PDF of it in the "Readings" folder of the Github repository.

Strongly Recommended:

Fox, John. 2015. *Applied Regression Analysis and Generalized Linear Models*, Third Edition. Thousand Oaks, CA: Sage Publications. Nice to have if you can get it cheaply; previous versions of this course used this as its main text.

Fox, John, and Sanford Weisberg. 2019. *An R and S-Plus Companion to Applied Regression*, Third Edition. Thousand Oaks, CA: Sage Publications. A companion to the Fox text, for R users.

Nagler, Jonathan. 1996. "Coding Style and Good Computing Practices." *The Political Methodologist* 6(2):2-8. See also this blog post from 2015. The original article is old, and also contains words to live by.

Shalizi, Cosma. 2021. *Advanced Data Analysis from an Elementary Point of View*. This is an unpublished draft of a textbook that Cosma Shalizi (CMU) has been working on for roughly a decade. It's very comprehensive, super smart, often funny, and worth keeping around.

Other Good Linear Regression Texts:

Blackwell, Matthew. 2023. *A User's Guide to Statistical Inference and Regression*. A Quarto book.

Chatterjee, Samprit, and Ali S. Hadi. 2013. *Regression Analysis by Example*, 5th Ed. New York: Wiley.

Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken. 2002. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Ed. Lawrence Erlbaum.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Harrell Jr., Frank E. 2015. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. New York: Springer.

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. 2012. *Introduction to Linear Regression Analysis*, 5th Ed. New York: Wiley.

Seber, George A.F., and Alan J. Lee. 2003. *Linear Regression Analysis*, 2nd Ed. New York: Wiley.

Books on "Econometrics":

Don't bother.

Books on R (a few from a very long list; many of these are available electronically via the Penn State Libraries)

Crawley, Michael J. 2014. *Statistics: An Introduction Using R*, 2nd Ed. New York: Wiley.

Everitt, Brian S., Torsten Hothorn. 2014. *A Handbook of Statistical Analyses Using R*, 3rd Ed. Boca Raton, FL: Chapman & Hall.

Mailund, Thomas. 2022. *Beginning Data Science in R: Data Analysis, Visualization, and Modelling for the Data Scientist*. New York: APress.

Maindonald, John, and John Braun. 2013. *Data Analysis and Graphics Using R: An Example-Based Approach*, 3rd Ed. New York: Cambridge University Press.

Murrell, Paul. 2019. *R Graphics*, 3rd Ed. Boca Raton, FL: Chapman & Hall. (Website is here).

Navarro, Danielle. 2021. *Learning Statistics With R* (and the `lsr` R package on CRAN).

A Few Other R Resources (mostly online)

R Reference Card 2.0 (also in Chinese).

The R Language: A Short Companion.

Robert Kabakoff's Quick-R (really excellent).

Owen, W. J. 2010. The R Guide.

Peng, Roger P. 2022. R Programming for Data Science.

Phillips, Nathaniel D. 2018. YaRrr! The Pirate's Guide to R.

Ricci, Vito. 2005. Fitting Distributions With R.

Ricci, Vito. 2005. R Functions For Regression Analysis.

Santana, Julio Sergio, and Efrain Mateos Farfan. 2014. *El Arte de Programar en R: Un Lenguaje Para la Estadistica*.

Shupinov, Alexay. 2019. *Visual Statistics. UseR!*.

Wickham, Hadley, and Garrett Grolemund. 2017. R for Data Science. O'Reilly. (Note: This is not an R book, it's a Tidyverse book, and the Tidyverse is not R. That doesn't make it a bad resource, just a very limited one. For my take on the Tidyverse, see pp. 6-7 of the syllabus for PLSC 502.)

The Methods Preceptor

Morrgan Herlihy is the methods preceptor for PLSC 503. She is a Political Science Ph.D. candidate who studies American Politics and social data analytics, with a particular focus on courts and judicial politics. She will serve as a "first line of defense" in the course: she can assist you with course material, software and programming issues, and other matters related to the course work. He can be reached via e-mail at `mth5492 [at] psu [dot] edu`.

Grading

Grading will be based on a total of 1000 points, divided as follows:

- Homework exercises: Ten worth 50 points each.
- A final project, worth 500 points.

Details for the homework assignments and the final project will be announced in class. An overview of expectations for the final project is available on the court Github repository. Also, note that all homework exercises and the final paper should be submitted electronically, as PDF files. Note that **only PDF files will be accepted**, without exception. If you do not yet know how to create a PDF file from a document you've created, please go learn; it is not difficult.

Some Other Useful Resources

The **Inter-University Consortium for Political and Social Research** (ICPSR), at the University of Michigan, maintains an extensive archive of data in the social and behavioral sciences. Much of it is accessible via their homepage (`http://www.icpsr.umich.edu`).

The **Political Methodology Section** of the American Political Science Association was created to provide APSA members with an interest in political methodology with a forum in which to meet and discuss ideas. The section publishes a quarterly newsletter (*The Political Methodologist*), a quarterly journal on political methodology (*Political Analysis*), conducts a discussion list on topics relating to political methodology, and maintains an extensive electronic archive of papers, accessible via their homepage (at `https://polmeth.org/`).

The **Comprehensive R Archive Network** (CRAN) (`http://cran.r-project.org/`) is the place to go for downloads, packages, and documentation for the R statistical language. Similarly, the **Stata**^{TM} homepage (`http://www.stata.com`) is a valuable resource for questions about Stata statistical software.

Obligatory Statement on Academic Integrity

Academic integrity is the pursuit of scholarly activity in an open, honest and responsible manner. Academic integrity is a basic guiding principle for all academic activity at The Pennsylvania State University, and all members of the University community are expected to act in accordance with this principle. Consistent with this expectation, the University's Code of Conduct states that all students should act with personal integrity, respect other students' dignity, rights and property, and help create and maintain an environment in which all can succeed through the fruits of their efforts.

Academic integrity includes a commitment by all members of the University community not to engage in or tolerate acts of falsification, misrepresentation or deception. Such acts of dishonesty

violate the fundamental ethical principles of the University community and compromise the worth of work completed by others.

In cases of any violation of academic integrity it is the policy of the Department of Political Science to follow procedures established by the College of the Liberal Arts. More information on academic integrity and procedures followed for violation can be found here.

Obligatory Statement on Accommodations for Disabilities

Penn State welcomes students with disabilities into the University's educational programs. Every Penn State campus has an office for students with disabilities. Student Disability Resources (SDR) website provides contact information for every Penn State campus (here). For further information, please visit the Student Disability Resources website (here).

In order to receive consideration for reasonable accommodations, you must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation: See documentation guidelines here. If the documentation supports your request for reasonable accommodations, your campus disability services office will provide you with an accommodation letter. Please share this letter with your instructors and discuss the accommodations with them as early as possible. You must follow this process for every semester that you request accommodations.

Obligatory Statement on Counseling and Psychological Services

Many students at Penn State face personal challenges or have psychological needs that may interfere with their academic progress, social development, or emotional wellbeing. The university offers a variety of confidential services to help you through difficult times, including individual and group counseling, crisis intervention, consultations, online chats, and mental health screenings. These services are provided by staff who welcome all students and embrace a philosophy respectful of clients' cultural and religious backgrounds, and sensitive to differences in race, ability, gender identity and sexual orientation.

Counseling and Psychological Services at University Park (CAPS)
(http://studentaffairs.psu.edu/counseling/): 814-863-0395

Counseling and Psychological Services at Commonwealth Campuses
(http://senate.psu.edu/faculty/counseling-services-at-commonwealth-campuses/)

Penn State Crisis Line (24 hours / 7 days/week): 877-229-6400. Crisis Text Line (24 hours / 7 days/week): Text LIONS to 741741.

Obligatory Statement on Educational Equity and Reporting Bias

Penn State takes great pride to foster a diverse and inclusive environment for students, faculty, and staff. Consistent with University Policy AD29, students who believe they have experienced or observed a hate crime, an act of intolerance, discrimination, or harassment that occurs at Penn State are urged to report these incidents as outlined on the University's Report Bias webpage (http://equity.psu.edu/reportbias/).

Obligatory Statement on Religious Observances

The Religious and Spiritual Observances Calendar is compiled by the Center for Spiritual and Ethical Development in consultation with campus and community religious leaders. It specifies those holy days of the major world religions for which observance may require students to depart from their normal routine at the University. Please note that only those holy days which occur when Penn State classes are in session are listed. This is not, therefore, an exhaustive list of all major holy days in each religious tradition.

Non-Obligatory Statement on Generative AI, Large Language Models, etc.

You're undoubtedly well aware of the existence of large language models (LLMs) – e.g., ChatGPT – and other artificial intelligence (AI) tools for language / image creation. Having been described as everything from making everyone their own version of Tony Stark to a Lovecraftian shoggoth (and most things in between), LLMs are currently creating a sometimes-depressing, sometimes-hilarious panic among faculty in legacy academic disciplines and programs. Most of that panic revolves around the use of LLMs to "cheat," in the traditional sense: to create work that deceptively gives the impression that the student knows something they do not. Beyond its intrinsically duplicitous nature, such use in a conventional classroom setting gives rise to concerns about equity and (potentially) devalues the experience / credential for other class members.

The other side of the LLM equation is that they are powerful tools for augmenting learning and creating new knowledge. Experience suggests that, in line with other technological advances (the printing press, personal computers, search engines, etc.), it is wiser to adapt to LLMs than to attempt to limit or ban their use. This is especially true in a course like this one, where (a) LLMs are particularly useful tools for learning technical skills (e.g., the R programming language) and (b) the long-term, repeated nature of graduate school creates disincentives for "cheating" in a conventional sense. Accordingly, enrollees in PLSC 503 are welcome to use generative AI tools, such as ChatGPT, to assist them with their work in the course. In doing so, it is important to remember that such AI tools are capable of making errors, and that it is each student's responsibility to verify the information they receive from the such a tool. In addition, any information obtained from a generative AI source must be noted/cited in the student's work, just as they would cite any other source.

Course Schedule

Reading the things assigned here each week is a good idea. If you miss a reading or two now and again, no worries; some weeks have more to read than others, and sometimes life intervenes.

Linear Regression

- **January 13**: *Introduction + Review of Bivariate Linear Regression*
  Readings (for background):
  - Weisberg: Preface + Chapters 1 and 2 + Appendices A.1 - A.4.
  - *GHV*, Chapter 7.
  - Alley, Joshua. 2021. "An Open Collection of Political Science Research with OLS Models and Cross-Sectional Data." *Political Methodologist* blog, September 8, 2021.
  - Roberts, Margaret E. 2018. "What is Political Methodology?" *PS: Political Science & Politics* 51:597-601.
  - Lewis-Beck, Michael S., and Andrew Skalaban. 1990. "When to Use R-Squared." *Political Methodologist* 3(2):11-12.
  - King, Gary. 1990. "When Not to Use R-Squared." *Political Methodologist* 3(2):9-11.
  - Luskin, Robert C. 1991. "R-Squared Encore." *Political Methodologist* 4(1):21-23.

- **January 20**: **NO CLASS** – Martin Luther King, Jr. Holiday

- **January 27**: *Multivariate Regression: Estimation and Inference*
  Readings:
  - Weisberg, Chapter 3 (pp. 51-68) and 6 (pp. 133-150) and Appendix A.8.
  - Berk, Richard, Lawrence Brown, Andreas Buja, Edward George, Emil Pitkin, Kai Zhang, and Linda Zhao. 2014. "Misspecified Mean Function Regression: Making Good Use of Regression Models That Are Wrong." *Sociological Methods & Research* 43:422-451.
  - Carlin, John B., and Margarita Moreno-Betancur. 2023. "On the Uses and Abuses of Regression Models: A Call for Reform of Statistical Practice and Teaching." Working paper: Murdoch Children's Research Institute.
  - Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. "Using Graphs Instead of Tables in Political Science." *Perspectives on Politics* 5:755-771.
  - Wysocki, Anna C., Katherine M. Lawson and Mijke Rhemtulla. 2022. "Statistical Control Requires Causal Justification." *Advances in Methods and Practices in Psychological Science* 5:1-19.

Homework One due.

- **February 3**: *Practical Multivariate Linear Regression*
  Readings:
  - Berk, Richard. 2010. "What You Can and Can't Properly Do with Regression." *Journal of Quantitative Criminology* 26(4):481-487.
  - Gelman, Andrew. 2008. "Scaling Regression Inputs by Dividing by Two Standard Deviations." *Statistics in Medicine* 27:2865-2873.
  - Keele, Luke, Randolph T. Stevenson, and Felix Elwert. 2020. "The Causal Interpretation of Estimated Associations in Regression Models." *Political Science Research and Methods* 8:1-13.
  - Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58:1083-1091.
  - Spirling, Arthur, and Brandon M. Stewart. 2022. "What Good is a Regression? Inference to the Best Explanation and the Practice of Political Science Research." Working paper: Princeton University.
  - Westreich, Daniel, and Sander Greenland. 2013. "The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients." *American Journal of Epidemiology* 177:292-298.

- **February 10**: *Dichotomous Predictors, (Non-)Linearity, and Data Transformations*
  Readings:
  - Weisberg, Chapters Four (pp. 67-93), Five (pp. 98-123), and Eight (pp. 185-199).
  - *GHV*, Chapter 12.
  - Rittmann, Oliver, Marcel Neunhoeffer, and Thomas Gschwend. 2023. "How to Improve the Substantive Interpretation of Regression Results When the Dependent Variable is Logged." *Political Science Research and Methods* 14: forthcoming. DOI: `https://doi.org/10.1017/psrm.2023.29`
  - Recommended: Breiman and Friedman (1985), Chen and Roth (2023), Bellego et al. (2022), and/or Shadden and Zorn (2011).

  Homework Two due.

- **February 17**: *Variance Issues: Collinearity, "Robustness," and Regularization*
  Readings:

- Weisberg, Chapter 7 (pp. 156-179).
- Long, J. Scott, and Laurie H. Ervin. 2000. "Using Heteroscedasity-Consistent Standard Errors in the Linear Regression Model." *The American Statistician* 54:217-224.
- King, Gary, and Margaret E. Roberts. 2014. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What To Do About It." *Political Analysis* 22:1-21.
- Vu, Patrick. 2024. "Do Standard Error Corrections Exacerbate Publication Bias?" Working paper: Brown University.
- One or more online resources on regularization, e.g. here, here, here, here, and/or here.

Homework Three due.

- **February 24**: *Residuals, Outliers, and Diagnostics + Endogeneity / Simultaneity*
Readings:
  - Weisberg, Chapter 9 (pp. 204-226).
  - Kennedy (pp. 107-109; 180-191).
  - Baissa, Daniel, and Carlisle Rainey. 2020. "When BLUE Is Not Best: Non-Normal Errors and the Linear Model." *Political Science Research and Methods* 8:136-148.
  - Lal, Apoorva, Mackenzie William Lockhart, Yiqing Xu, and Ziwen Zu. 2021. "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice based on Over 60 Replicated Studies." Working paper, Stanford/UCSD.
  - Sovey, Allison J., and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55:188-200.

Homework Four due.

- **March 3**: *Variable Selection and Multiplicative Interactions*
Readings:
  - Weisberg, Chapter 10 (pp. 234-248).
  - Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26(November):797-833.
  - Brambor, Thomas, William R. Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14:63-82.

- Esarey, Justin, and Jane Lawrence Sumner. 2018. "Marginal Effects in Interaction Models: Determining and Controlling the False Positive Rate." *Comparative Political Studies* 51:1144-1176.
- Hainmueller, Jens, Jonathan Mummolo, and Xiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* 27:163-192.

Homework Five due.

- **March 10**: **NO CLASS** (Spring Break)

- **March 17**: *Bootstrapping, etc. + Missing Data Redux*
  Readings:
  - *GHV*, Chapter 17.
  - Harden, Jeffrey J. 2011. "A Bootstrap Method for Conducting Statistical Inference with Clustered Data." *State Politics & Policy Quarterly* 11:223-246.
  - Hesterberg, Tim C. 2015. "What Teachers Should Know About the Bootstrap: Resampling in the Undergraduate Statistics Curriculum." *The American Statistician* 69:371-386.
  - Lall, Ranjit. 2016. "How Multiple Imputation Makes a Difference." *Political Analysis* 24:414-433.
  - Mooney, Christopher Z. 1996. "Bootstrap Statistical Inference: Examples and Evaluations for Political Science." *American Journal of Political Science* 40:570-602.

Homework Six due.

Likelihood-Based Regression

- **March 24**: *Maximum Likelihood: Introduction + Optimization*
  Readings:
  - Weisberg, Appendix A.11.
  - *GHV*, Chapter 8.
  - Fox, Appendix D6, pp. 92-95.

- Some random web resources (among many):
  - The PSU Stat 415 MLE page
  - A Gentle Introduction to Maximum Likelihood Estimation, on *Towards Data Science* (with some minimal Pythion code)
  - Steele, Joel S. 20??. Examples of Maximum Likelihood Estimation and Optimization in R, and Regression with Nonlinear Transformations.

- **March 31**: *MLE: Inference and Testing*
  Readings:
  - Weisberg, Chapter 12 (pp. 270-279).
  - Buse, A. 1982. "The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note." *The American Statistician* 36(3):153-57.
  - Freedman, D. A. 2006. "On the So-Called 'Huber Sandwich Estimator' and 'Robust' Standard Errors." *The American Statistician* 60:299-302.

  Homework Seven due.

- **April 7**: *Models for Binary Outcomes*
  Readings:
  - Faraway (pp. 25-50).
  - *GHV*, Chapter 13.
  - Bloome, Deirdre, and Shannon Ang. 2022. "Is the Effect Larger in Group A or B? It Depends: Understanding Results From Nonlinear Probability Models." *Demography* 59:1459-1488.
  - Breen, Richard, Kristian Bernt Karlson, and Anders Holm. 2018. "Interpreting and Understanding Logits, Probits, and Other Nonlinear Probability Models." *Annual Review of Sociology* 44:39-54.
  - Chen, Kaicheng, Robert S. Martin, and Jeffrey M. Wooldridge. 2023. "Another Look at the Linear Probability Model and Nonlinear Index Models." Working paper: Michigan State University.
  - Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models." *American Journal of Political Science* 57:263-277.

- Karlson, Kristian Bernt, and Ben Jann. 2023. "Marginal Odds Ratios: What They Are, How to Compute Them, and Why Sociologists Might Want to Use Them." *Sociological Science* 10:332-347.
- Mize, Trenton D. 2019. "Best Practices for Estimating, Interpreting, and Presenting Nonlinear Interaction Effects." *Sociological Science* 6:81-117.
- Fun blog posts:
  - Kubinec, Robert. 2024. "Lost in Transformation: The Horror and Wonder of Logit." Blog post, April 9, 2024.
  - Westfall, Jake. 2018. "Logistic Regression Is Not Fucked." *Cookie Scientist* blog, March 18, 2018 (via the Wayback Machine).

- **April 14**: *Models for Nominal Responses*
  Readings:
  - Faraway, pp. 129-139.
  - Paolino, Philip. 2021. "Predicted Probabilities and Inference with Multinomial Logit." *Political Analysis* 29:416-421.

  Homework Eight due.

- **April 21**: *Models for Ordinal Responses*
  Readings:
  - Faraway, pp. 139-149.
  - *GHV*, pp. 273-278.
  - Kleinbaum, David G., and Mitchel Klein. 2010. *Logistic Regression: A Self-Learning Text*, 3rd Ed. New York: Springer. Chapter 13: Ordinal Logistic Regression.
  - Winship, Christopher, and Robert D. Mare. 1984. "Regression Models with Ordinal Variables." *American Sociological Review* 49(4):512-25.

  Homework Nine due.

- **April 28**: *Models for Event Counts*
  Readings:
  - Faraway, pp. 83-94.

- Gould, William. 2011. "Use Poisson Rather Than Regress; Tell a Friend." *The Stata Blog*, August 22, 2011.
- Wooldridge, Jeffrey. 2001. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press. Chapter 19 (pp. 645-659).

Homework Ten due.

- **May 7**: **Final Papers / Projects Due**

Now That You're Done With Your First Year...

- **...read the following things:**
  - Breiman, Leo. 2001. "Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)." *Statistical Science* 16(3):199-231.
  - Daoud, Adel, and Devdatt Dubhashi. 2023. "Statistical Modeling: The Three Cultures." *Harvard Data Science Review* Issue 5.1 (Winter 2023).
  - Molnar, Christopher. 2023. *Modeling Mindsets: The Many Cultures of Learning From Data*. MUCBOOK:Heidi Seibold.