

# QUANTITATIVE THEORY AND METHODOLOGY FOR THE SOCIAL SCIENCES

Columbia University  
GR5010, FALL 2022  
Thurs 6:10PM-8:00PM  
302 FAYERWEATHER / REMOTE

**Instructor:** Marco Morales  
**Email:** marco.morales@columbia.edu  
**Office:** 509E International Affairs Building  
**Office Hours:** by appointment

**TA:** Cindy Zhang  
**Email:** xz3106@tc.columbia.edu

## I. Overview

---

This course — one of the two foundational courses in the QMSS curriculum — is designed as an in-depth introduction to the social sciences and its methodologies. It is intended to give a broad overview so students can intelligently combine ideas in solving real-world problems.

We will focus on the logic and design of social research, beginning with some concepts and topics common to research across the social sciences. We will later move on to understanding the principles behind an array of methodologies used in the social sciences: causal inference, experimentation, observational studies, formal models, surveys, and applied machine-learning techniques. We will analyze their applications using cases drawn from the research literature.

The focus of this course is not on the techniques themselves — you will have ample opportunity to learn that deeply in other courses — but in understanding the logic behind the use of these tools to extract meaningful answers from their applications.

**Prerequisites:** it is assumed that you have had at least one semester of graduate-level statistics involving linear regression and analysis of variance. Some basic mathematics and algebra will also be assumed.

## II. Course Tools

---

The course will rely on a combination of reading materials, pre-recorded video lectures, and asynchronous written discussion to enhance your learning experience. We will use the following tools:

- **Canvas** will serve as the main repository for class materials, including all required readings, pre-recorded video lectures, class assignments instructions and submission;
- **Ed Discussion (within Canvas)** will serve as the primary means of written asynchronous communication where students can communicate, share ideas, post questions, and submit team exercises in writing;
- **Flip** will serve as the platform where students submit weekly video reading presentations; and
- **Zoom** will serve as the platform to hold remote live class sessions.

## III. Course Materials

---

All materials required for each session will be made available on a dedicated weekly module on **Canvas** seven days before each class. There, you will find **Required Readings**, **pre-recorded video lectures**, **Topic Readings**, **Thesis Proposal Deliverable instructions** and **class slides**. Please make sure to consume materials in the indicated order. Also, please note that:

**Required Readings** are mandatory and should be completed before each class. They are the theoretical foundation to discuss each week's topic.

**pre-recorded video lectures** are mandatory as well and should be completed before each class. We have little time and much material to cover. These videos contain the lecture portion of the class that you should consume on your own, so that — armed with that knowledge — we can focus on deeper discussions during time together in class.

**Topic Readings** are also mandatory and should be completed before each class. These are empirical applications of the methods we will discuss and analyze in class.

**Complementary Readings** are intended to serve as further (and future) references if you ever want to delve deeper on a particular topic, but are not a requirement to this class.

**Thesis Readings** two reference texts are intended to help you with the drafting and crafting of your Thesis Proposal, but will not be discussed in class. Read the assigned chapters in the suggested order/dates:

- Booth, W. C., Colomb, G. C., and Williams, J. M. (2016). *The Craft of Research*. University of Chicago Press, Chicago, IL, fourth edition.
- Weston, A. (2009). *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition.

#### IV. Course Dynamics

---

While students are expected to attend class live, we found that students benefited from opportunities for asynchronous learning, and so we will continue to utilize many of these tools. Specifically,

- (1) students will be able to communicate, share ideas and questions and post content via Ed Discussions.
- (2) students will be able to work collaboratively on small group presentations. Each group member will make part of the presentation, and the whole group will map out the parts and flow of the presentation. You will utilize Flip which allows you to record your part when you are ready.
- (3) TAs will hold weekly 1-hour recitation/lab sessions, where they will “live” lead students through labs and/or homework and/or lecture review. But these sessions will be recorded and uploaded to Courseworks for students who cannot make the recitation times or who would like to review the material multiple times. Additionally, you may not see TAs in class, but you are welcome to visit them during their office hours, whether virtually or in-person, depending on TA availability.

**Synchronous Participation vs. Asynchronous Participation:** This course is designed to have a combination of synchronous and asynchronous participation to enhance your learning experience. It is our strong expectation that you will participate synchronously when required so that you can benefit fully from your peers and the live instruction. That said, it is completely understandable that your circumstances may make that very difficult, at least on some occasions. Please alert us when that is the case. On those occasions, the synchronous portions can be done asynchronously as well. Likewise, assignments and some forms of participation can also be done asynchronously.

**Expectation of Regular Participation and Utilization of Course tools:** We will be monitoring student participation and completion of assignments using the corresponding tools throughout the semester. We want to make sure that students are consistently engaged, and if that becomes difficult, that students alert us to their situations.

**In preparation for each class:**, you should have (i) read, thought about and be prepared to discuss all Required and Topic Readings assigned for the week; (ii) watched the pre-recorded video lecture in Canvas and your classmates' Topic Reading video presentations in Flip; and (iii) posted questions or comments on Ed Discussion.

**During each live class:** we will devote the **first 30 mins of the class** to answer questions live, clarify concepts, and answer related questions where appropriate. We will spend the **remaining 90 minutes of the class** discussing assigned Topic Readings in 45min blocks. In some classes, we will spend 5-10 minutes carrying out class exercises.

**Thesis Proposal Requirement:** At the end of the course, you will turn in a 4-page **low-fidelity prototype of your thesis**. Throughout the semester, you will be required to turn in four **Deliverables** which are foundational to the prototype. Feedback will be provided on these deliverables where required. Deliverables should be **submitted through Canvas on the indicated dates/times**.

**Late Submission Policy:** Assignments, deliverables and your thesis proposal are expected to be submitted on the due date. For every day after the submission date, 10% of the maximum grade will be deducted from the score.

## **V. Course Requirements**

---

The final grade of the course will be based of your **fulfillment** of each of the following requirements:

**Class participation (10%):** Students are expected to have completed a number of activities before and during each class, including (i) regular submission of questions, answers and comments related to the material for the week through *Ed Discussion*; (ii) consistent class participation during our weekly session discussions.

**Topic presentations (20%):** Students will be assigned to teams that will submit 5-min video oververviews of a **Topic Reading**. These video presentations should be submitted 7 days before class.

**Thesis Proposal Deliverables (40%):** Students will submit pieces of their proposal four times during the semester. Make sure to submit each one through Canvas on the indicated dates/times.

**Low-fidelity thesis prototype (30%):** Students will submit a 4-page low fidelity prototype of their thesis at the end of the semester. Make sure to submit it through Canvas on the indicated date/time.

All written work must be original and produced exclusively for this class. You are expected to follow the University's guidelines for the submission of written work.

## VI. Course Outline

---

### WEEK 1: INTRODUCTION TO THE COURSE

*What this course is (and what it is not). Course overview. What is so unique about quantitative methods applied to the social sciences? Why do we need models to understand the world? Why is it useful to have statistical models in the social sciences?*

### WEEK 2: THE “SCIENCE” OF SOCIAL SCIENCE

*What is so scientific about social sciences? A look at “the method”. Inductive v deductive perspectives. Theories, hypothesis and falsifiability. Links to quantitative methods. Mechanisms.*

#### Required Readings:

- Gelman, A. (2011). Induction and deduction in Bayesian data analysis. *Rationality, Markets and Morals*, 2:67–78.
- Elster, J. (2007). *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge University Press, Cambridge, MA. [Ch 1-2]

#### Thesis Readings:

- Booth, W. C., Colomb, G. C., and Williams, J. M. (2016). *The Craft of Research*. University of Chicago Press, Chicago, IL, fourth edition. [Ch 3]
- Weston, A. (2009). *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition. [Ch I-II]

#### *Complementary Readings:*

- Popper, K. (2002[1935]). *The Logic of Scientific Discovery*. Routledge, New York, NY.
- Kuhn, T. S. (2012). *The Structure of Scientific Revolution*. University of Chicago Press, Chicago, IL, fourth edition.
- King, G., Keohane, R. O., and Verba, S. (1995). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, Princeton, NJ.
- Gelman, A. and Shalizi, C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):8–38.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 34:103–115.

**WEEKS 3 | 4: CAUSALITY AND CAUSAL INFERENCE (I & II)**

*Causes of effects or effects of causes? The search for causes: from Aristotle to Fisher. The fundamental problem of causal inference. The Neyman-Rubin model.*

**Required Readings:**

- Holland, P. W. (1980). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.
- Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modelling, decisions. *Journal of the American Statistical Association*, 100(469):322–331.

Thesis Readings:

- Booth, W. C., Colomb, G. C., and Williams, J. M. (2016). *The Craft of Research*. University of Chicago Press, Chicago, IL, fourth edition. [Ch 4]
- Weston, A. (2009). *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition. [Ch VI-VII]

Complementary Readings:

- Imbens, G. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge, MA.
- Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. Cambridge University Press, Cambridge, MA, second edition.
- Sekhon, J. S. (2004). Quality meets quantity: Case studies, conditional probability and counterfactuals. *Perspectives on Politics*, 2(2):281–293.
- Dawid, A. (2000). Causal inference without counterfactuals. *Journal of the American Statistical Association*, 95(450):407–448.
- Page, S. (2006). Path dependence. *Quarterly Journal of Political Science*, 1(1):87–115.

**[9/29] - Thesis Proposal Deliverable #1 (research topic) due.**

**WEEK 5: EXPERIMENTS AND RANDOMIZATION**

*Theoretical Foundations of Experiments. Statistical foundations of experiments. Taxonomy of randomized experiments. Randomized experiments as the golden standard for causal inference. Inference from randomized experiments.*

**Required Readings:**

- Druckman, J. N. and Green, D. (2021). A new era of experimental political science. In Druckman, J. N. and Green, D., editors, *Advances in Experimental Political Science*. Cambridge University Press, New York, NY.
- Angrist, J. D. and Pischke, J.-S. (2015). *Mastering ‘Metrics’: The Path from Cause to Effect*. Princeton University Press, Princeton, NJ. [Ch 1]

**Topic: Voter Turnout**

- De la O, A. (2013). Do conditional cash transfers affect electoral behavior? Evidence from a randomized experiments in Mexico. *American Journal of Political Science*, 57(1):1–14.
- Wantchekon, L. (2003). Clientelism and voting behavior: Evidence from a field experiment in Benin. *World Politics*, 55(3):399–422.

Thesis Readings:

- Booth, W. C., Colomb, G. C., and Williams, J. M. (2016). *The Craft of Research*. University of Chicago Press, Chicago, IL, fourth edition. [Ch 15]

Complementary Readings:

- Martel García, F. and Wantchekon, L. (2010). Theory, external validity, and experimental inference: Some conjectures. *Annals of the American Academy of Political Science*, 628(132-147).
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press, Princeton, NJ.
- Imbens, G. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge, MA.
- Morton, R. and Williams, K. (2010). *Experimental Political Science and the Study of Causality*. Cambridge University Press, Cambridge, MA.
- Salganik, M. J. (2018). *Bit by bit*. Princeton University Press, Princeton, NJ. [Ch 4]

**WEEK 6: OBSERVATIONAL STUDIES**

*How are they different from experiments? Understanding their limitations for causal inference. Statistical tools to cope with non-random assignment of treatments. Inference from observational studies.*

**Required Readings:**

- Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, 2(3):808–840.
- Przeworski, A. (2009). Is the science of Comparative Politics possible? In Boix, C. and Stokes, S. C., editors, *Oxford Handbook of Comparative Politics*. Oxford University Press, New York, NY.

**Topic: Selected statistical applications to observational studies**

- Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.
- Erikson, R. S. and Titiunik, R. (2015). Using regression discontinuity to uncover the personal incumbency advantage. *Quarterly Journal of Political Science*, 10(1):101–119.

*Complementary Readings:*

- Cochran, W. G. (2015[1972]). Observational studies. *Observational Studies*, 1(1):126–136.
- Rubin, D. B. (2006). *Matched sampling for causal effects*. Cambridge University Press, Cambridge, MA.
- Rosenbaum, P. R. (2002). *Observational Studies*. Springer, New York, NY.
- Imai, K., King, G., and Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the royal statistical society: series A (statistics in society)*, 171(2):481–502.
- Winship, C. and Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25(1):659–706.
- Sekhon, J. S. and Titiunik, R. (2012). When natural experiments are neither natural nor experiments. *American Political Science Review*, 106(1):35–57.

**[10/13] - Thesis Proposal Deliverable #2 (hypotheses) due.**

**WEEK 7: CONCEPTS, MEASUREMENT, AND MEASUREMENT ERROR**

*Research design and the research question. Measurements as a function of concepts. Theoretical consequences of measurement error. Statistical consequences of measurement error.*

**Required Readings:**

- Geddes, B. (1990). How the cases you choose affect the answers you get: Selection bias in comparative politics. *Political Analysis*, 2(1):131–150.
- Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *Journal of Economic Perspectives*, 15(4):57–67.

**Topic: Economic Perceptions**

- Michelitch, K., Morales, M. A., Owen, A., and Tucker, J. A. (2012). Looking to the future: Prospective economic voting in 2008 presidential elections. *Electoral Studies*, 31(4):838–851.

Thesis Readings:

- Booth, W. C., Colomb, G. C., and Williams, J. M. (2016). *The Craft of Research*. University of Chicago Press, Chicago, IL, fourth edition. [Ch 7-9, 12]

Complementary Readings:

- Alwin, D. F. (2007). *Margins of Error: A Study of Reliability in Survey Measurement*. John Wiley & Sons, Hoboken, NJ.
- Goertz, G. (2008). Concepts, theories, and numbers: A checklist for constructing, evaluating, and using concepts or quantitative measures. In Box-Steffensmeier, J., Brady, H. E., Collier, D., and Goertz, G., editors, *Oxford Handbook of Political Methodology*. Oxford University Press, New York, NY.
- Jackman, S. (2008). Measurement. In Box-Steffensmeier, J., Brady, H. E., Collier, D., and Goertz, G., editors, *Oxford Handbook of Political Methodology*. Oxford University Press, New York, NY.

**WEEK 8: IDENTIFICATION**

*Statistical Inference v Identification. Extrapolation. Selection. Endogeneity. The empirical - and theoretical - problems of endogeneity. Quantitative methods to address identification. Instrumental variables.*

**Required Readings:**

- Angrist, J. D. and Pischke, J.-S. (2015). *Mastering ‘Metrics’: The Path from Cause to Effect*. Princeton University Press, Princeton, NJ. [Ch 3]
- Morgan, S. L. and Winship, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, Cambridge, MA, second edition. [Ch 3]

**Topic: Selected statistical applications to address identification**

- Miguel, E., Satyanath, S., and Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4):725–753.
- Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

*Complementary Readings:*

- Manski, C. F. (1999). *Identification Problems in the Social Sciences*. Harvard University Press, Cambridge, MA.
- Aronow, P. M. and Miller, B. T. (2019). *Foundations of Agnostic Statistics*. Cambridge University Press, Cambridge, MA. [Ch 6-7]
- Keele, L. (2015). The discipline of identification. *PS: Political Science & Politics*, 48(1):102–105.
- Angrist, J. D. and Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives*, 15(4):69–85.
- Antonakis, J., Bendahan, S., Jacquart, P., and Lalive, R. (2014). Causality and endogeneity: Problems and solutions. In *Oxford Handbook of Leadership and Organizations*. Oxford University Press, New York, NY.
- Angrist, J. D., Imbens, G., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

**[10/27] - Thesis Proposal Deliverable #3 (data description) due.**

**WEEKS 9 | 10: SURVEY RESEARCH AND SURVEY METHODOLOGY (I & II)**

*What is wrong (and what is right) with polls? The Total Survey Error Paradigm. Respondent Selection Issues. Response Accuracy Issues. Survey Administration Issues.*

**Required Readings:**

- Bautista, R. (2012). An overlooked approach in survey research: Total Survey Error. In *Handbook of Survey Methodology for the Social Sciences*. Springer, New York, NY.
- Kennedy, C., Blumenthal, M., Clement, S., Clinton, J. D., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Saad, L., Witt, E., and Wlezien, C. (2017). An evaluation of 2016 election polls in the United States. AAPOR Ad Hoc Committee on 2016 Election Polling.

**Topic: Question Wording and Response Scales**

- Lundmark, S., Gilljam, M., and Dahlberg, S. (2015). Measuring generalized trust: An examination of question wording and the number of scale points. *Public Opinion Quarterly*, 80(1):26–43.

**Topic: Non-response and Data Quality**

- Fricker, S. and Tourangeau, R. (2010). Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly*, 74(5):934–955.

**Topic: Mode of Data Collection**

- Sakshaug, J. W., Yan, T., and Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly*, 74(5):907–933.

**Topic: Probability v Non-probability Samples**

- Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., and Wang, R. (2011). Comparing the accuracy of rdd telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75(4):709–747.

Thesis Readings:

- Booth, W. C., Colomb, G. C., and Williams, J. M. (2016). *The Craft of Research*. University of Chicago Press, Chicago, IL, fourth edition. [Ch 13-14, 16-17]
- Weston, A. (2009). *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition. [Ch VIII]

*Complementary Readings:*

- Weisberg, H. F. (2009). *The Total Survey Error Approach: A guide to the new science of survey research*. University of Chicago Press, Chicago, IL.
- Tourangeau, R., Rips, L. J., and Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge University Press, Cambridge, MA.
- Foster, I., Ghani, R., Jarmin, R. S., Kreuter, F., and Lane, J. (2016). *Big Data and Social Science: A practical guide to methods and tools*. CRC Press, New York, NY.
- De Leeuw, E. D., Hox, J., and Dillman, D. (2012). *International Handbook of Survey Methodology*. Routledge, New York, NY.
- Salganik, M. J. (2018). *Bit by bit*. Princeton University Press, Princeton, NJ. [Ch 3]

[11/10] - Thesis Proposal Deliverable #4 (methodological strategy) due.

**WEEK 11: MODELING HUMAN BEHAVIOR**

*Why Models? Can behavior be modeled? Applied Game Theory and the Rational Choice paradigm. Formal models to help explain collective behavior*

**Required Readings:**

- Page, S. E. (2018). *The model thinker*. Basic Books, New York, NY. [Ch 2-4]

**Topic: Legislative Behavior**

- Krehbiel, K. (1998). *Pivotal politics: A theory of US lawmaking*. University of Chicago Press, Chicago, IL. [Ch 2]

**Topic: Partisanship and partisan bias**

- Achen, C. H. (2002). Parental socialization and rational party identification. *Political Behavior*, 24(2):151–170.

*Complementary Readings:*

- Friedman, M. (2008). The methodology of positive economics. In Hausman, D. J., editor, *The Philosophy of Economics*. Cambridge University Press, Cambridge, MA.
- Becker, G. S. (1976). *The economic approach to human behavior*. University of Chicago Press, Chicago, IL.
- Sen, A. (1999). The possibility of social choice. *American Economic Review*, 89(3):349–378.

- Riker, W. H. (1995). The political psychology of rational choice theory. *Political Psychology*, 16(1):23–44.
- Arrow, K. J. (1994). Methodological individualism and social knowledge. *The American Economic Review*, 84(2):1–9.
- Morton, R. B. (1999). *Methods and models: A guide to the empirical analysis of formal models in political science*. Cambridge University Press, Cambridge, MA.
- Osborne, M. J. (2004). *An Introduction to Game Theory*. Oxford University Press, New York, NY.
- Kahneman, D. (2011). *Thinking Fast and Slow*. Farrar, Strauss and Groux, New York, NY.
- Chwe, M. S.-Y. (2014). *Jane Austen: Game Theorist*. Princeton University Press, Princeton, NJ.
- Bhattacharya, A. (2021). *The Man from the Future: the visionary life of John von Neumann*. W. W. Norton & Co, New York, NY. [Ch 6]

## WEEK 12: ACADEMIC HOLIDAY

## WEEKS 13 | 14: APPLIED MACHINE LEARNING (I & II)

*What do we understand by Machine Learning? ML for inference and prediction. ML applications to Social Science problems.*

### Required Readings:

- Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3):289–310.
- Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., and Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature*, 595(7866):181–188.

### Topic: Text as Data - Estimating Ideology from Texts

- Laver, M., Benoit, K., and Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2):311–331.
- Barberá, P. (2014). Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis*, 23(1):76–91.

**Topic: Machine Learning and Causal Inference**

- Bansak, K., Ferwerda, J., Hainmueller, J., Dillion, A., Hangqartner, D., Lawrence, D., and Weinstein, J. (2018). Improving refugee integration through data-driven algorithmic assignment. *Science*, 359:325–329.
- Green, D. and Kern, H. L. (2012). Modeling heterogenous treatment effects in survey experiments with bayesian additive regression trees. *Public Opinion Quarterly*, 76(3):491–511.

*Complementary Readings:*

- Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2):3–27.
- Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.
- Athey, S. (2019). The impact of machine learning on economics. In Agrawal, A., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press, Chicago, IL.
- Grimmer, J., Roberts, M. E., and Stewart, B. M. (2022). *Text as Data: a New Framework for Machine Learning and the Social Sciences*. Princeton University Press, Princeton, NJ.
- Kern, H. L., Stuart, E. A., Hill, J., and Green, D. (2016). Assessing methods for generalizing experimental impact estimates to target populations. *Journal of Research on Educational Effectiveness*, 9(1):103–127.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning*. Springer, New York, NY, second edition.
- Murphy, K. P. (2012). *Machine Learning: a probabilistic approach*. MIT Press, Cambridge, MA.
- Manning, C., Raghavan, P., and Schütze, H. (2009). *An Introduction to Information Retrieval*. Cambridge University Press, Cambridge, MA.

**[12/15] - LOW-FIDELITY THESIS PROTOTYPE DUE.**

### **Statement on Academic Integrity**

Columbia's intellectual community relies on academic integrity and responsibility as the cornerstone of its work. Graduate students are expected to exhibit the highest level of personal and academic honesty as they engage in scholarly discourse and research. In practical terms, you must be responsible for the full and accurate attribution of the ideas of others in all of your research papers and projects; you must be honest when taking your examinations; you must always submit your own work and not that of another student, scholar, or internet source. Graduate students are responsible for knowing and correctly utilizing referencing and bibliographical guidelines. When in doubt, consult your professor. Citation and plagiarism-prevention resources can be found at the GSAS page on Academic Integrity and Responsible Conduct of Research.

Failure to observe these rules of conduct will have serious academic consequences, up to and including dismissal from the university. If a faculty member suspects a breach of academic honesty, appropriate investigative and disciplinary action will be taken following the Dean's Discipline procedures.

### **Statement on Disability Accommodations**

If you have been certified by Disability Services (DS) to receive accommodations, please either bring your accommodation letter from DS to your professor's office hours to confirm your accommodation needs, or ask your liaison in GSAS to consult with your professor. If you believe that you may have a disability that requires accommodation, please contact **Disability Services** at 212-854-2388 or [disability@columbia.edu](mailto:disability@columbia.edu).

***Important:*** To request and receive an accommodation you must be certified by DS.