

NM6103: Quantitative Research Methods

Instructor: Subhayan Mukerjee (mukerjee@nus.edu.sg)

AY2025-26 Sem II

COURSE DESCRIPTION

NM6103 provides a comprehensive foundation in the logic, design, and execution of quantitative social science research. Moving beyond simple data analysis, this course focuses on **quantitative reasoning**: the ability to transform complex social phenomena into measurable data and defensible claims.

The course is structured around three pillars: [The Logic of Scientific Inquiry](#), which covers the theoretical foundations of the scientific method, measurement, and causality; [Methods in Practice](#), which explores specific quantitative methods such as surveys, digital trace data, experimental design, causal inference; and [Meta-Perspectives](#), where we discuss ethics, the future of the field, paying particular attention to the implications of Generative AI. Through hands-on activities, group projects and critical “post-mortems” of published quantitative social science research, students will develop the technical rigour and ethical oversight required to produce and critique high-quality social research in the digital age.

REQUIRED TEXTBOOKS

- De Mesquita, E. B., & Fowler, A. (2021). *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.
- Salganik, M. J. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press.
- Ritchie, S. (2020). *Science fictions: Exposing fraud, bias, negligence and hype in science*. Random House.

EVALUATION COMPONENTS

Component	Weight	Due Date
Class Participation [details]	20%	Weekly
Response to Readings [details]	10%	Weekly
Lead a Class Discussion [details]	10%	Weeks 7–10
“Dream” Study Registered Report [details]	15%	Week 11
Peer Review [details]	5%	Week 13
Group Project (Total 40%)		
– Proposal [details]	10%	Week 7
– Final Project [details]	25%	Week 13
– Peer Feedback [details]	5%	Reading Week

PART I: THE LOGIC OF SCIENTIFIC INQUIRY

Week 1: The Theoretical Foundations of Quantitative Methods

Required Readings:

- Abend, G. (2008). The meaning of ‘theory’. *Sociological Theory*, 26(2), 173-199.
 - De Mesquita, E. B., & Fowler, A. (2021). *Thinking clearly with data*:
 - Chapter 1: Thinking Clearly in a data-driven world.
 - Godfrey-Smith, P. (2009). *Theory and reality*:
 - Chapter 1: Introduction.
 - Chapter 4: Popper: Conjecture and Refutation.
-

Week 2: Explanations and Mechanisms

Required Readings:

- Elster, J. (1989). *Nuts and bolts for the social sciences*:
 - Chapter 1: Explanation.
 - Chapter 2: Mechanism.
- Babbie, E. R. (2007). *The practice of social research*:
 - Chapter 4: Research Design.

Activity:

- Introduction to R and RStudio.
-

Week 3: Concepts and Measurements

Required Readings:

- Lauderdale, B. (2022). *Pragmatic Social Measurement*:
 - Chapter 1: What is Measurement?
- Babbie, E. R. (2007). *The practice of social research*:
 - Chapter 5: Conceptualization, Operationalization, and Measurement.
 - Chapter 6: Indexes, Scales, and Typologies.
- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529-546.

Activity:

- Validity and Reliability analysis.
-

Week 4: Finding the Signal

Required Readings:

- De Mesquita, E. B., & Fowler, A. (2021). *Thinking clearly with data:*
 - Chapter 2: Correlation: what is it and what is it good for?
 - Chapter 4: Correlation requires Variation.
 - Chapter 5: Regression for Describing and Forecasting.
 - Chapter 6: Samples, Uncertainty, and Statistical Inference

Activity:

- Understanding test statistics and statistical significance.
-

Week 5: The Question of Causality

Required Readings:

- De Mesquita, E. B., & Fowler, A. (2021). *Thinking clearly with data:*
 - Chapter 3: Causation: what is it and what is it good for?
 - Chapter 9: Why Correlation doesn't imply Causation.
 - Chapter 10: Controlling for Confounders.

Activity:

- Everything about linear models (and more).
-

Week 6: No Class [Lunar New Year]

Recess Period — No Readings Assigned

PART II: METHODS IN PRACTICE

Week 7: Doing Surveys

Required Readings:

- Salganik, M. J. (2019). *Bit by bit*:
 - Chapter 3: Asking Questions.
- Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74(5), 849-879.
- Parry, D. A., et al. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, 5(11), 1535-1547.

Readings for in-class presentation:

- Beyens, I., et al. (2024). Social media use and adolescents' well-being: Developing a typology of person-specific effect patterns. *Communication Research*, 51(6), 691-716.
- Hornik, R., et al. (2021). Association of COVID-19 misinformation with face mask wearing and social distancing in a nationally representative US sample. *Health Communication*, 36(1), 6-14.
- Altay, S., et al. (2024). News can help! The impact of news media and digital platforms on awareness of and belief in misinformation. *The International Journal of Press/Politics*, 29(2), 459-484.
- Reif, A., et al. (2025). The public trust in science scale: a multilevel and multidimensional approach. *Science Communication*, 47(5), 670-701.
- Hanitzsch, T., et al. (2024). Publish and perish: mental health among communication and media scholars. *Journal of Communication*, 74(6), 429-442.

Recommended Readings:

- Hillygus, D. S. (2011). The evolution of election polling in the United States. *Public Opinion Quarterly*, 75(5), 962-981.
- Jamieson, K. H., et al. (2023). Protecting the integrity of survey research. *Proceedings of the National Academy of Sciences Nexus*, 2(3), pgad049.

Activity:

- Survey weighting simulation.
-

Week 8: Analysing “Found” Data

Required Readings:

- Salganik, M. J. (2019). *Bit by bit*:
 - Chapter 3: Observing Behavior.
- Lazer, D., et al. (2021). Meaningful measures of human society in the twenty-first century. *Nature*, 595(7866), 189-196.
- boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.

Readings for in-class presentation:

- Allen, J., et al. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539.
- Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2), 405-450.
- King, G., et al. (2013). How censorship in China silences collective expression. *American Political Science Review*, 107(2), 326-343.
- Guilbeault, D., et al. (2024). Online images amplify gender bias. *Nature*, 626(8001), 1049-1055.
- Sun, L., et al. (2024). Smiling women pitching down: auditing representational and presentational gender biases in image-generative AI. *Journal of Computer-Mediated Communication*, 29(1), zmad045.
- Fischer, S., et al. (2020). Auditing local news presence on Google News. *Nature Human Behaviour*, 4(12), 1236-1244.

Recommended Readings:

- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665-668.
- Hargittai, E. (2020). Potential biases in big data. *Social Science Computer Review*, 38(1), 10-24.
- Boeschoten, L., et al. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, 4(2), 388-423.

Activity:

- Confounders, fixed effects and random effects.
-

Week 9: Designing Experiments

Required Readings:

- Salganik, M. J. (2019). *Bit by bit*:
 - Chapter 4: Running Experiments.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics*:
 - Chapter 2: The Experimental Ideal.

Readings for in-class presentation:

- Levy, R. E. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review*, 111(3), 831-870.
- Salganik, M. J., et al. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762), 854-856.
- Badrinathan, S. (2021). Educative interventions to combat misinformation. *American Political Science Review*, 115(4), 1325-1341.
- Van de Rijt, A., et al. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, 111(19), 6934-6939.
- Bail, C. A., et al. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 9216-9221.

- Mukerjee, S., & Yang, T. (2021). Choosing to avoid? Selective exposure and avoidance on social media. *Political Communication*, 38(3), 222-240.
- De Vreese, C. H., et al. (2017). Linking survey and media content data: Opportunities, considerations, and pitfalls. *Communication Methods and Measures*, 11(4), 221-244.

Recommended Readings:

- Munger, K. (2019). The limited value of non-replicable field experiments. *Social Media + Society*, 5(3).
- Montgomery, J. M., et al. (2018). How conditioning on posttreatment variables can ruin your experiment. *American Journal of Political Science*, 62(3), 760-775.
- Mutz, D. C. (2011). *Population-based survey experiments*. Princeton University Press.
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. WV Norton.

Activity:

- Moderation and heterogeneity of treatment effects.

Week 10: Inferring Causality

Required Readings:

- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics*:
 - Chapter 1: Questions about Questions.
- Huntington-Klein, N. (2021). *The Effect*:
 - Chapter 5: Identification.
 - Chapter 5: Causal Diagrams.

Readings for in-class presentation:

- Branas, C. C., et al. (2011). A difference-in-differences analysis of greening vacant urban space. *American Journal of Epidemiology*, 174(11), 1296-1306.
- Mukerjee, S., et al. (2023). Metrics in action: how social media metrics shape news production on Facebook. *Journal of Communication*, 73(3), 260-272.
- Jaidka, K., et al. (2019). Brevity is the soul of Twitter. *Journal of Communication*, 69(4), 345-372.
- Athey, S., et al. (2021). The impact of aggregators on internet news consumption. *National Bureau of Economic Research*.
- Yang, T., et al. (2021). Policing the digital divide. *Journal of Communication*, 71(4), 572-597.
- Wang, J., et al. (2022). Global evidence of expressed sentiment alterations during the COVID-19 pandemic. *Nature Human Behaviour*, 6(3), 349-358.
- Arendt, F., et al. (2023). News for life: improving the quality of journalistic news reporting to prevent suicides. *Journal of Communication*, 73(1), 73-85.

Recommended Readings:

- Huntington-Klein, N. (2021). *The effect: An introduction to research design and causality*. Chapman and Hall/CRC.

- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics*. Princeton University Press.
- Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge University Press.
- Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press.
- Sekhon, J. S., & Tiitunik, R. (2012). When natural experiments are neither natural nor experiments. *American Political Science Review*, 106(1), 35-57.
- Brand, J. E., et al. (2023). Recent developments in causal inference and machine learning. *Annual Review of Sociology*, 49(1), 81-110.

Activity:

- Matching.

PART III: META-PERSPECTIVES

Week 11: Ethics of Quantitative Research

Required Reading:

- Ritchie, S. (2020). *Science fictions: Exposing fraud, bias, negligence and hype in science*. Random House.
- Salganik, M. J. (2019). *Bit by bit*:
– Chapter 6: Ethics.

Recommended Readings:

- Munafò, M. R., et al. (2017). A manifesto for reproducible science. *Nature human behaviour*, 1(1), 0021.
- Camerer, C. F., et al. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, 2(9), 637-644.
- Nosek, B. A. et al. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11), 2600-2606.
- Miller, J. (2009). What is the probability of replicating a statistically significant effect?. *Psychonomic Bulletin & Review*, 16(4), 617-640.

Activity:

- Why research findings are hard to replicate (even when you are being honest!).
-

Week 12: Quantitative Research in the Age of Gen AI

Required Readings:

- Bail, C. A. (2024). Can Generative AI improve social science? *Proceedings of the National Academy of Sciences*, 121(21).
- Davidson, T., & Karell, D. (2025). Integrating Generative Artificial Intelligence into Social Science Research. *Sociological Methods & Research*.

- Messing, S., & Tucker, J. (2026). The train has left the station: Agentic AI and the future of social science research. *Brookings Institution*. [[link](#)]
- Garimella, K. (2026) Using AI Agents like Claude Code. [[YouTube Playlist](#)]

Recommended Readings:

- Kovanovic, V. (2025). What the hyperproduction of AI slop is doing to science. *The Conversation*. [[link](#)]
- Blau, W., et al. (2024). Protecting scientific integrity in an age of generative AI. *Proceedings of the National Academy of Sciences*, 121(22).
- Karp, D. (2026) Can AI Replace Social Science Researchers? [[link](#)]

Activity:

- Using agentic AI for doing quant research.
-

Week 13: Where do we go from here?

Required Readings:

- Salganik, M. J. (2019). *Bit by bit*:
– Chapter 7: The Future.
- Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour*, 1(1), 0015.
- Traberger, C. S., et al. (2026). AI is turning research into a scientific monoculture. *Communications Psychology*, 4(1), 37.

Activity:

- Final Presentations.

DETAILS OF EVALUATION COMPONENTS

Class Participation (20%)

Your active engagement is central to the success of this seminar. Participation means more than speaking up – it includes attentive listening, asking thoughtful questions, contributing to discussions, and building on your peers' ideas. Aim to bring curiosity, respect, and critical reflection to each session

Response to Readings (10%)

Each week, post a short reflection on the relevant Canvas discussion thread. Your response should go beyond a superficial summary of the readings: highlight a key idea, question, or critique from the readings that resonated with you, and connect it to broader seminar themes or your own perspective. Aim to spark dialogue by raising points that others can build on.

Leading a Class Discussion (10%)

As the discussion leader, your task is to lead a technical “post-mortem” of the assigned research paper. This will take place during the weeks’ dedicated to Surveys, Found Data, Experiments, and Causal Inference.

- **Theory:** Provide a very brief overview of the theoretical framework and the research question. This is necessary for context but should not be the main focus of your presentation.
- **Sampling, Data, and Measurement:** This is a primary focus. You must critically evaluate the data source, the sampling frame, and how the authors operationalized their variables. Be prepared to discuss construct validity and potential data quality issues.
- **Methods Design and Strategy :** This is a primary focus. Clearly explain the analytical strategy (e.g., how the authors attempt to isolate statistical effects). Is the research causal? If yes, how sure are you about the causal claim? You should be able to explain the logic of the design and why this specific strategy was chosen over others.
- **Statistical Inference:** Evaluate the appropriateness of the statistical models used. Discuss the interpretation of the coefficients, the handling of standard errors, and whether the conclusions drawn are actually supported by the numerical results.
- **Limitations:** Identify the most significant threats to internal or external validity. Consider the trade-offs the authors made and how these limitations might color the interpretation of the findings

“Dream” Study Registered Report (15%)

You have to design your “dream” quantitative study – that is based on your specific research interests. You have a virtually infinite budget (let’s say SGD 1 million). However, your design must be grounded in scientific reality and practical feasibility.

You will submit a “registered report”. A registered report is basically a data collection and analysis plan – that is peer-reviewed before data collection. This is a practice being embraced by many reputed journals to ensure that researchers don’t cherry pick significant findings or engage in p-hacking. To know more about registered reports see [this](#).

Your submission must comprise the following:

- **Problem Statement & Theoretical Framework:** What is the specific research gap you are filling? Why does this matter to the field?
- **Methodological Framework:** Detail your chosen approach (e.g., surveys, analysis of found data, experiments, causal Inference). You are allowed to combine multiple approaches if you need to.
- **Data Collection & Sampling:** Describe your data acquisition strategy. How will you ensure a representative and high-quality sample? Is your survey multi-wave? If you are using found data, what is the time of data collection? If you are using experiments, how will you recruit participants? Note that your sample doesn’t need to be “truly” representative, but it shouldn’t be a lazy sample like NUS undergrads.
- **Analytical Strategy & Rigor:** Which statistical tests or computational models will you employ? How will you ensure validity and reliability? Include supplementary materials such as survey questionnaires (if you are doing a survey) or list of keywords (for data collection if you are scraping social media data) or anything else that you feel will be important for assessing your study.
- **Ethical Considerations & Practical Limitations:** Discuss the ethical safeguards required and identify the inherent methodological limitations that even an infinite budget cannot solve.

For instance, don't think like an economist and say "let's do an experiment where we starve people in some districts and give unlimited food to people in other districts to see whether food abundance leads to better mental health" (I'm only partially joking about that example).

Your research design will be "peer reviewed" by another person in this class. To maintain the facade of anonymity, your name should not appear in your submission (in reality, they will probably get to know upon reading the submission who wrote it). For uniformity, please follow the APA 7th edition. formatting guidelines for preparing your manuscript. You can download a Microsoft Word template from Canvas or find a LaTeX template on Overleaf [here](#). Word limit is 1500 words, excluding references, tables, appendices, supplementary material. Please submit a PDF file.

Peer Review (5%)

You will evaluate one peer's Registered Report as a reviewer. Your task is to provide a high-level technical critique of the proposal's methodological soundness and statistical logic. Please do not make judgement about "reject" or "accept" in your review :)

A good peer review should have the following components:

- **Feasibility:** Is the data collection strategy feasible (even with the SGD 1 million budget cap)?
- **Internal & Statistical Validity:** Scrutinise the design for confounding variables, selection bias, or measurement error. Are the proposed statistical tests appropriate for the data?
- **Ecological & External Validity:** Evaluate if the "infinite budget" scale translates to real-world insights or if the findings remain "lab-bound."
- **Methodological "Stress Test":** Propose a specific failure (e.g., 40% attrition or platform changes) and analyse how it would impact the study's conclusions.
- **Technical Refinement:** Provide one high-level recommendation to improve precision (e.g., an alternative scale or a specific robustness check) without changing the core research question.
- **Ethics:** Is the design ethical? Are subjects in experiments aware that they are in an experiment? If there is deception involved, is it warranted? Deception by itself is not necessarily unethical in quant research, as long as it is well justified. Are there any other ethical lapses that the researcher missed?

Submit a PDF in any format. Ensure you don't go beyond 1000 words. There is no "minimum" word count, which is how real-world peer reviews are, but remember a good review is detailed and constructive. You are well within your right to send a two line peer review, but that also means you won't get high marks from me (and in the real world, this means you will be black listed as a future peer reviewer by the journal).

Group Project Proposal (10%)

For this submission, you will design and propose a quantitative research study on a social science topic of your choice. Your goal is to demonstrate a sophisticated understanding of quantitative research design, variable measurement, and statistical inference. While this is a proposal-style paper, it should be written with the rigor and structure of a peer-reviewed journal article. A suggested structure is as follows:

- **Introduction:** Identify the core social science problem or phenomenon you are investigating. Clearly state the research problem and justify why it is significant to both academia and society.
- **Theoretical Framework:** Explain the specific social science theories that ground your study. Detail how these theories help you derive your hypotheses or research questions. A good hypothesis should be specific, concise, and empirically testable

- **Methods:** This is the most critical section of your project. You must describe your research design with enough detail that another researcher could replicate your logic. You must address:
- **Design & Data:** Is the study descriptive, explanatory, causal, or a mix of both? Specify the data source (e.g., primary survey, secondary longitudinal data like the GSS, or found data like social media traces).
 - **Measurement:** Define your independent and dependent variables. How are they operationalised?
 - **Validity & Reliability:** How do you ensure your measures are valid (representing the construct intended) and reliable (consistent over time or across items)?
 - **Data Transformation:** Describe any recoding, scaling, or handling of missing data.
 - **Analysis Plan:** Specify the statistical tests (e.g., OLS regression, ANOVA, Chi-square) and how you will achieve statistical inference (p-values, confidence intervals, etc.).
- **Discussion:**
 - **Anticipated Findings:** What results do you expect to find based on your framework?
 - **Theoretical Implications:** How would these findings support, challenge, or refine the existing theories discussed in your framework?
 - **Limitations:** Acknowledge potential constraints such as sampling bias or measurement error.

The paper should be around 2000 words (excluding references, figure captions, tables). Please do not overshoot the word limit by more than a couple of hundred words. Please follow the APA 7th edition formatting guidelines for preparing your manuscript. You can download a Microsoft Word template from Canvas or find a LaTeX template on Overleaf [here](#).

Final Group Project (25%)

The final project will extend the initial proposal, incorporating my feedback and carrying the study through to completion. The deliverable will be a polished research paper. Its overall structure will mirror that of the proposal, with the addition of a dedicated Results section. Other components – such as the Introduction, Theoretical Framework, Methods, and Discussion – may be revised to reflect feedback, integrate findings, or account for adjustments made during the research process when contingencies prevented strict adherence to the original plan. Evaluation will be focused on:

- The extent to which feedback on the proposal has been effectively integrated into the final project
- The rigour of methodological implementation, including (but not limited to) data collection procedures, any preprocessing steps, the choice and application of statistical tests, the clarity of results reporting, and the interpretation of statistical findings
- The depth and coherence of the discussion of results
- The identification and articulation of limitations that emerged upon completion of the study

Please follow the APA 7th edition formatting guidelines for preparing your manuscript. You can download a Microsoft Word template from Canvas or find a LaTeX template on Overleaf [here](#). Submit a PDF of no more than 5000 words (excluding references, tables, captions, appendices, supplementary information).

Peer Feedback (5%)

You will be required to complete a survey where you will evaluate the contribution of your peers in the group project.