Claremont Graduate University

Fall 2026

PP482 ADVANCED QUANTITATIVE RESEARCH METHODS

Course Instructor

Robert Klitgaard, University Professor

Office: Burkle 204

E-mail: robert.klitgaard@cgu.edu

Office hours: Tuesdays 9:00–11:45 am or by appointment.

Class Schedule

Course meetings: August 31 through December 14, 2026. No class September 7 (Labor Day).

Meeting day, time: Mondays, 1:00–3:50 pm.

Format: In person. Burkle 12.

Course Description

This is part of a four-course graduate sequence in statistical methods. It covers exploratory data analysis and multivariate statistical models such as regression analysis, data reduction techniques, panel data analysis, and difference-in-differences techniques.

The fall 2026 version aspires to broaden and deepen the experience.

Broaden, by enabling students from many professions and disciplines to succeed with the help of AI tutors customized to each student's field, background, and interests.

Deepen, by conveying how to think like a statistician.

First, "thinking like a statistician" means looking hard at data: where it comes from, what its variables mean, what patterns and exceptions it seems to contain. PP482 begins with exploratory data analysis, which help us look at and listen to the data before cranking up our statistical machinery. "The greatest value of a picture is when it forces us to notice what we never expected to see" –John Tukey. "Although we often hear that data speak for themselves, their voices can be soft and sly." –Frederick Mosteller

Second, "thinking like a statistician" means having "the maturity to realize that models are to be used, but not to be believed" –Henri Theil. Such "maturity" used to be mathematical, requiring years of preparation to plumb the depths of multivariate statistical models. Now, with the help of increasingly powerful AI platforms, students of many backgrounds can appreciate the strengths and weaknesses of advanced statistical tools—and indeed, get

the AI's help in figuring out which models and estimation techniques are suitable for what questions asked of their particular data.

Third, "thinking like a statistician" means judiciously applying statistical findings to the real world. What can be said about what works, where, for whom, and under what conditions? Statisticians understand that external validity is neither guaranteed ("my results apply everywhere") nor precluded ("every context is different, so we can't conclude anything about anyplace else"). AI can help us discern the limits and possibilities of inference and extrapolation.

Each class will teach tools, apply them to real examples, and enable students with AI help to apply them to datasets tied to their own interests.

Please bring your laptop to class.

Student Learning Outcomes

Think of learning goals using this matrix:

	Know	Do	See (like a statistician)
Without AI			
With AI			

We will fill this matrix in together over the semester — and your examinations will visit every cell.

By the end of this course, both *with* and *without* AI alongside, students will be able to:

1. Examine data before modeling. Explore any dataset — one variable at a time, then two — and diagnose its provenance, anomalies, and pathologies before estimating anything.
2. Build, choose, and translate models. Estimate and interpret multiple regressions. Choose among estimators by knowing what each assumes and how each fails. Explain any coefficient to a decision-maker in plain English.
3. Judge causal claims. Understand the logic and limits of randomized trials, natural experiments and regression discontinuity, difference-in-differences, and panel methods — including how changing the method can change the question being answered.
4. Hunt exceptions responsibly. Find the unusually effective performers hiding in residuals, distinguish real exceptions from noise, and come home empty-handed when that is the honest verdict.
5. Judge how far findings travel. Assess what works where, for whom, under what conditions — and what travels (often the ideas and variable lists) when the coefficients do not.

6. Interview and analyze data with an AI partner. Lead the consult, document the collaboration, and judge the AI's work—whether sound, incomplete, or wrong.

How to Read the Weekly Learning Priorities

This course covers a lot. You are not expected to master everything we encounter at the same depth. Each week the material is divided into three layers:

OWN: The central idea or habit of mind. You should be able to explain it in your own words, recognize when it matters, and carry it into later weeks.

USE: The main statistical tool or procedure. With notes and AI alongside, you should be able to apply it, understand what it is doing, interpret the result, and know its principal assumptions and failure modes.

RECOGNIZE: Important ideas, methods, or extensions that you should understand when you encounter them, but need not yet command on your own.

The categories move. By December, much of what began under Recognize should have migrated toward Use, and the most important lessons should simply be things you own. If a week feels crowded, return to the hierarchy: own the idea, use the tool, recognize the rest.

Two Roles for AI

In PP482 you will use AI in two different roles.

The Tutor knows what you are trying to learn. It helps you climb the week's ladder, rehearse, and review. ChatGPT can serve as your Tutor; another capable AI in a dedicated course conversation can too.

The Colleague approaches an analysis fresh, without the PP482 map. Use a new conversation with no course materials supplied for Full Consults, naïve analyses, and other exercises where you want an independent statistical response.

These are two roles, not necessarily two AIs. You may use the same platform for both; keep the conversations separate. Neither role is assumed to be right. Your job is to judge the work.

All the prompts mentioned below appear in "Your Package of Prompts." PP482 Package of Prompts.docx

"The Interrogator's Dozen" contains twelve questions to put to any analysis. The Interrogator's Dozen.docx

Prerequisites

PP481 or equivalent, or permission of the professor. Students are welcome from throughout CGU and the Claremont Colleges, with the permission of the professor.

Course Requirements

- Weekly “units”: 30 percent of the final grade.
- Class participation: 10 percent.
- Team mini-case (weeks 11–12): 10 percent.
- Midterm examination, October 19: 20 percent. Open book, open notes, AI in tow.
- Final examination, December 14: 30 percent. Open book, open notes, AI in tow.

There is no term paper and no textbook.

The Weekly Unit

Every week, you will practice answering three questions about a real analysis. 1. How would you explain this result to the proverbial mayor? 2. What cautions apply before extrapolating it? 3. Nonetheless, what does it suggest about the questions to ask experts and practitioners, and about candidates for success worth studying?

Use about two pages to provide three short answers to those questions, as well as one figure, one table, and one line naming the question you posed that mattered most (the follow-up you asked the AI that most changed the analysis).

Send each weekly unit by email to robert.klitgaard@cgu.edu — not through Canvas — by 5:00 pm on the Sunday before class. This applies every week; the assignments below simply say “Submit.” Please use the fixed subject line “PP482 Unit [number] — [Your last name]” (for example, “PP482 Unit 3 — Ramírez”).

Each week, please run the TUTOR prompt on the previous week’s OWN-USE-RECOGNIZE ladder.

> *Paste the week’s OWN / USE / RECOGNIZE ladder from the syllabus.*
>
> You are my statistics tutor for a graduate course. Here is last week’s learning ladder from PP482:
> OWN: [paste]
> USE: [paste]
> RECOGNIZE: [paste]
> Start by finding out what I already know: ask me three diagnostic questions, one at a time. Then tutor me up the ladder at my pace—quiz me, coach me, and make me practice with small examples from my own field or area of interest, which is [your field or interest]. Do not do the rungs for me: when I ask you to just give an answer, give me a hint instead. End by testing me on the two rungs I was weakest on.

Books and Software

No purchases. Three free references serve as our shelf for checking ourselves and our AI:

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and Other Stories*. Cambridge: Cambridge University Press. Free PDF and examples: https://avehtari.github.io/ROS-Examples/

Cunningham, Scott. 2026. *Causal Inference: The Remix*. New Haven: Yale University Press. Free online: https://mixtape.scunning.com

Huntington-Klein, Nick. 2025. *The Effect: An Introduction to Research Design and Causality*. 2nd ed. Boca Raton: Chapman & Hall/CRC Press. Free online: https://theeffectbook.net

Claremont Graduate University provides access to ChatGPT Edu for the duration of this course. You may instead use your own version of ChatGPT or another capable AI platform (Claude, Gemini, Grok, et al.). If your CGU access ends with the course, save any files or conversations you want to keep separately.

We will also use Gemini Notebook (formerly NotebookLM). We will have a Gemini Notebook for this course and practice using it.

Software: R is our common statistical language, building on PP481, but command syntax is not an end in itself. You may work directly with data through AI, ask AI to recommend methods, or have it generate and explain R code. In every case, you remain responsible for understanding what was done—the variables used, observations excluded, transformations made, estimator chosen, assumptions invoked, and what the result means.

Schedule of Classes

Readings marked * are optional.

August 31. First Look, Then Think, Then Model

Introductions and overview. Where do good hypotheses come from? Not from statistics but from mastery of your field. Statistics then gives hypotheses a profile: the sizes of effects under various conditions, and what can and cannot be concluded from the exceptions.

In class: The portrait of an expert statistical analyst, drafted together. John Snow and the 1854 cholera map — a man who knew his field's canon and defied its consensus, and won with a map and shoe leather. We will also meet the course's emblem: one research question whose answer runs from 0.04 to 2.9 depending on the statistical method. By week 11 you will be able to read it.

Readings:

Robert Klitgaard. 2026. *Graduate School, Augmented*. Ch. 1, "Think Like an Expert." https://gradschoolaugmented.com(30 minutes)

David A. Freedman. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology* 21: 291–313. (45 minutes); "Your Package of Prompts" PP482 Package of Prompts.docx provides guided-reading prompt — let your AI help with the technical

passages) https://www.scribd.com/document/637411335/Freedman-shoe-leather-statistical-model

Browse the UCLA John Snow site: https://www.ph.ucla.edu/epi/snow.html (30 minutes)

Do before class:

1. Run the prompt "Inside Tips" PP482 Package of Prompts.docx for applied statistics in public policy. Post on Canvas the two most surprising things the AI said experts notice or avoid.
2. Introduce yourself on Canvas: your field, your research interest, and your history with statistics — affection, dread, or both.

OWN: Good hypotheses come from knowing your field; statistics helps profile them, test them, and reveal their limits.

USE: Ask expert questions of an empirical claim: what do I notice, what should I doubt, and what evidence would change my mind?

RECOGNIZE: John Snow's larger lesson: models matter, but maps, field knowledge, and shoe leather can matter more.

September 14. Meet the Data Before You Master Them

Exploratory data analysis: one variable at a time, then two. Distributions, outliers, transformations. The question every dataset must answer first — where did these numbers come from? Two datasets: the *World Happiness Report* and a century of technology adoption in eight countries.

Robust estimation begins here: mean versus median versus trimmed mean when one wild value contaminates the data. The mean receives the course's first estimator biography: what it assumes, when it is best, how it fails. Estimator Biographies, part 1.docx

This week also sets a standing rule of the course, from John Tukey: exploratory analysis suggests clues; confirmatory analysis renders verdicts. Humans see patterns in noise, and a pattern found by looking is a hypothesis, not a finding. Verdicts require a pre-specified test or fresh data — and it is honorable to report that a clue did not survive.

In class: exploratory data analysis on the *World Happiness Report* and a century of technology adoption — including some curves that cannot be true; finding them is your job. Robust means; the first "estimator biography" (the mean). Estimator Biographies, part 1.docx

Readings:

Klitgaard. 2026. *Graduate School, Augmented*. Ch. 2, "Question Authority — All of It." https://gradschoolaugmented.com (30 minutes)

Browse our free reference shelf and bookmark it: Gelman, Hill & Vehtari, *Regression and Other Stories* (free PDF); Cunningham, *Causal Inference: The Remix* (free online); Huntington-Klein, *The Effect* (free online). (30 minutes)

John W. Tukey. 1980. “We Need Both Exploratory and Confirmatory.” *The American Statistician* 34(1): 23–25. Tukey-NeedExploratoryConfirmatory-1980.pdf

Assignment:

1. Run “Full Consult” PP482 Package of Prompts.docx on the technology-adoption file PP482_data_tech_adoption_8countries.csv with a fresh Colleague conversation: give it the data and the naïve prompt—what can I conclude? what next?—then earn your keep with follow-up questions (45–60 minutes). Save the transcript.
2. Run the lab “Break the Mean” (20 minutes): one wild value, watch the mean leave town.

First weekly deliverable, on the technology adoption data: one figure, one table, three short answers (the mayor, the cautions, the nonetheless). Due by email before 5:00 pm on Sunday, Sept. 13.

OWN: Look before you model. Patterns discovered while exploring are clues to test, not verdicts to announce.

USE: Read distributions and plots; compare mean, median, and trimmed mean; and explain why one summary may be more trustworthy than another.

RECOGNIZE: Transformations and robust estimation as responses to data whose shape or outliers make ordinary summaries misleading.

September 21. Where the Numbers Come From: Who Leads World Science?

In class: shares versus counts (the US share fell while US output rose 56 percent — write the senator’s sentence and the university president’s sentence); counting conventions; the house dataset (Quality of Government) introduced.

Readings:

Fedderke, “The Structural Parameters of Scientific Knowledge Production” (2026), Section 1 and Table 1 only. Fedderke_KPF.pdf

Browse the QoG Standard codebook: https://www.gu.se/en/quality-government/qog-data (30 minutes) Choose two countries you care about; look them up.

Do before class (nothing to submit):

1. Run “Provenance Interview” PP482 Package of Prompts.docx on the world-science file: who counted, how, and which convention could flip a conclusion? PP482_data_world_science_table1.csv
2. Write your three best why-questions about these data. Park them for November 23, when the author himself can answer some of them.
3. Choose two countries you care about; look up data about them in the QoG.

4. Write the Estimator Biography for the interquartile range (IQR). Use the template of the Estimator Biographies, part 1.docx
5. Use your Tutor prompt on last week's (that is Sept. 14) OWN, USE, RECOGNIZE.

   You are my statistics tutor for a graduate course. Here is last week's learning ladder from PP482:
   OWN: [paste]
   USE: [paste]
   RECOGNIZE: [paste]
   Start by finding out what I already know: ask me three diagnostic questions, one at a time. Then tutor me up the ladder at my pace—quiz me, coach me, and make me practice with small examples from my own field or area of interest, which is [your field or interest]. Do not do the rungs for me: when I ask you to just give an answer, give me a hint instead. End by testing me on the two rungs I was weakest on.

Assignment (submit before 5:00 pm Sunday Sept. 27)

Weekly unit, on the science data. Use the "Full Consult" prompt from Your Package of Prompts PP482 Package of Prompts.docx to analyze these data: PP482_data_world_science_table1.csv

1. How would you explain the result to the proverbial mayor? 2. What cautions apply before extrapolating it? 3. Nonetheless, what does it suggest about the questions to ask experts and practitioners, and about candidates for success worth studying?

Use about two pages to provide three short answers to those questions, as well as one figure, one table, and one line naming the question you posed that mattered most (the follow-up you asked the AI that most changed the analysis).

This week's OWN-USE-RECOGNIZE:

OWN: Provenance before inference: who counted, how they counted, and who is missing can change the story.

USE: Move between shares and counts and interrogate counting conventions before drawing substantive conclusions.

RECOGNIZE: A codebook is part of the data, not clerical background; definitions can create or erase apparent trends.

September 28. Advanced Seeing: Principal Components, Factors, and the Flinch

The World Justice Project gives eight different names to eight aspects of the rule of law. Countries, however, may not vary along eight independent lines. We begin with an intuitive question: when many measures march together, what common direction best describes the march? Principal components analysis compresses the observed variation. Factor analysis asks a different and bolder question: could one or more unobserved forces have generated the correlations, while each indicator retains something unique?

We will use both as lenses, never as metaphysical X-rays. A first component may reveal institutional quality, a broad development gradient, or a halo of countries that tend to receive good marks. The buried variance may be more interesting: it is where Singapore-like combinations, institutional tradeoffs, and exceptional cases can hide. We will learn the econometrician's flinch as well as the tool: components and factors depend on the variables, sample, scaling, model, and rotation; elegant scores do not identify causes.

In class, using the same QoG Basic Cross-Section: PCA without matrix algebra; the correlation heatmap, scree plot, loadings, scores, and the information lost by compression; Component 2 as the place to look for a different country story; factor analysis, uniqueness, and rotation; why econometricians flinch; and our first residual hunt—countries cleaner or less clean than their income predicts, with Estonia, Georgia, Rwanda, and Chile treated as questions rather than verdicts.

Readings:

World Justice Project. Rule of Law Index 2025, methodology section. (30 minutes) https://worldjusticeproject.org/rule-of-law-index/downloads/Index-Methodology-2025.pdf

Robert Klitgaard. 2012. Review of Bo Rothstein, The Quality of Government. *Perspectives on Politics* 10(3). (20 minutes)

Robert Klitgaard. 2021. *The Culture and Development Manifesto*. New York: Oxford University Press, ch. 9. (1 hour)

Do before class:

1. Download the QoG Basic Cross-Section dataset, Version Jan26, in CSV format from https://www.gu.se/en/quality-government/qog-data/data-downloads and upload that file to your AI. Run "Prepare the PCA Evidence" in the PP482 Package of Prompts. Use the eight World Justice Project scores, a common complete-case sample, and standardized variables. Keep the correlation display, scree plot, first two sets of loadings, and the countries at both ends of Component 2.

Then run "Name the Factor." Bring to class one page containing: the percentage of variance explained by Components 1 and 2; the two loading columns; three honest names for Component 1, including one deflationary name; the countries at both ends of Component 2; and one question the second component makes you want to ask. Bring the AI transcript as well. Nothing from this exercise is submitted before class.

Stop there, but keep the QoG file.

3. Use your Tutor prompt on last week's (September 21) OWN, USE, RECOGNIZE.

Assignment (submit before 5:00 pm Sunday, October 4):

Weekly unit, on the governance block. Use no more than two pages, plus your AI transcript. Include one figure and one table. Answer: (1) How would you explain the main PCA result to the proverbial mayor? (2) What cautions apply before treating Component 1 as a real thing, much less a cause? (3) Nonetheless, what did Component 2, the factor analysis, or the residual hunt suggest we should investigate next? Name one country worth a closer look,

give two rival explanations for its unusual position, and name one fact that would help distinguish them.

Add three sentences comparing PCA and factor analysis: the question each asks, what each does with variance unique to an indicator, and one reason an econometrician might flinch. End with one line naming the follow-up question you asked the AI that most changed your interpretation.

This week's OWN-USE-RECOGNIZE:

OWN: PCA compresses the variation we observed; factor analysis proposes a model for the common variation. Neither method discovers the one true name of a hidden thing.

USE: Read a correlation display, scree plot, loadings, scores, and uniqueness; compare a PCA with a factor solution; inspect the second dimension and regression residuals for country stories the summary hides.

RECOGNIZE: High correlations may reflect genuine common structure, shared methods and reputations, or a broader development gradient. Rotation, naming, and causal interpretation require judgment beyond the output.

Module 2: The Hard Questions

October 5. The Hard Question: Outcomes = f(Treatment, Everything Else)

Confounding and the counterfactual, taught twice with one logic. First, the hormone-replacement-therapy (HRT) reversal: how observational studies of millions of women showed large protective effects that a randomized trial then overturned. Second, the same trap in modern dress: early adopters of any technology — HRT included — are richer, healthier, and better connected, so the "benefit" in observational data is often a portrait of who adopts early. The week's hardest question is what we call "the mayor's question": in this case, what do you tell the patients?

We also learn to draw the problem: a first, gentle use of causal diagrams. One arrow from "who adopts early" into treatment, another into outcome — the confound, visible as a picture. For students who fear algebra, drawing the arrows is the most merciful formalism ever invented, and it upgrades every conversation with the mayor.

Multiple regression receives its full estimator biography this week: the assumptions of OLS, ranked with the load-bearing one first, each broken in the lab — simulate the violation, watch what happens to the estimate, and learn what (if anything) can be done. Exogeneity, linearity, independence, homoskedasticity: not a list of assumptions to memorize but a set of failure modes to recognize on sight.

In class: the hormone-therapy reversal; early adopters and why their good outcomes mislead; drawing the confound (first causal diagrams); OLS built and broken live — least squares as the mean with predictors, "holding constant" shown in pictures, each assumption violated in simulation.

Readings:

Gary Taubes. 2007. "Do We Really Know What Makes Us Healthy?" *The New York Times Magazine*, Sept. 16. Link (1 hour)

Regression and Other Stories, Ch. 1 ("Overview") and the first half of Ch. 18 on causal inference basics. (1 hour)

Causal Inference: The Remix, Ch. 3 ("Directed Acyclic Graphs"), opening sections. Link (45 minutes)

* Lawrence H. Summers. 1991. "The Scientific Illusion in Empirical Macroeconomics." *The Scandinavian Journal of Economics* 93(2): 129–148.

Do before class:

1. Run "Draw My DAG" PP482 Package of Prompts.docx for a treatment-outcome pair from your own area of interest (30 minutes).
2. Run "Steelman the Confounder" PP482 Package of Prompts.docx on the HRT cohort (Canvas): let the AI build the story that explains away the benefit — then check it against the data.
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: weekly unit, on the HRT cohort.

This week's OWN-USE-RECOGNIZE:

OWN: Every causal claim depends on a counterfactual, and confounding can manufacture a persuasive effect where none exists.

USE: Draw a causal diagram, fit and interpret multiple regression, and use simulation to see how a confounder can bias the estimate.

RECOGNIZE: OLS assumptions as distinct failure modes—especially exogeneity, but also linearity, independence, and homoskedasticity—rather than a list to memorize.

October 12. Exceptional Performers — and the Anticlimax

The residual hunt. Using real California school data, we regress achievement on poverty, rank the residuals, and study the districts that beat their circumstances — the unusually effective schools, revisited half a century after that literature began. Then, from the same files, the anticlimax: plot residuals against enrollment and watch small schools dominate both tails. Some exceptional performers are real; some are arithmetic. Telling them apart is the skilled verdict — and it is honorable to come home empty-handed.

Estimator craft on the same data: the residual funnel is heteroskedasticity made visible, and robust standard errors are the modern one-line answer (note the distinction, forever: robust estimators fix the estimate; robust standard errors fix the uncertainty). We re-run the regression robustly—does the outlier list change? Then we let the AI resample and rerank the schools many times and watch the supposed "top" school bounce around: an intuitive

introduction to bootstrapping and the instability of rankings. We also meet shrinkage. What a skilled analyst does about noisy small samples is borrow strength, not delete schools.

Two more pieces of regression craft, on the same body. Goodness of fit, interrogated: our model "explains 77 percent of the variance." What does that number buy the mayor, what does it not, and why is a high R^2 neither necessary for a good causal estimate nor sufficient for a useful one? And interactions: "what works for whom" is, statistically, an interaction term. We estimate one (does poverty's effect differ by English-learner share?), interpret it with marginal-effects plots rather than coefficient-squinting — including the notoriously slippery continuous-by-continuous case — and catalogue the classic misreadings.

In class: the unusually effective schools, revisited, with an anticlimax. Robust standard errors (the funnel is heteroskedasticity); interactions ("what works for whom," formalized); R^2 interrogated; the regression-table craft session. Midterm briefing.

Readings:

Howard Wainer. 2007. "The Most Dangerous Equation." *American Scientist* 95(3): 249–256. Link (45 minutes)

* Robert Klitgaard and George Hall. 1975. "Are There Unusually Effective Schools?" *Journal of Human Resources* 10(1): 90–106. https://doi.org/10.2307/145121 Available in this April 1973 RAND preprint: https://files.eric.ed.gov/fulltext/ED085402.pdf

Do before class:

1. With the provided script and your AI, rank the school residuals. Choose your "unusually effective" district and write three sentences on why you would want to visit it.
2. Then run "Noise or Excellence?" — and record honestly whether your district survived. PP482 Package of Prompts.docx
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: weekly unit, on the schools data.

OWN: Extreme performers may reflect real excellence or statistical noise. Small samples deserve special suspicion.

USE: Residual plots, sample size, and robust standard errors to diagnose what is happening; marginal-effects plots to interpret interactions.

RECOGNIZE: Bootstrapping as a way to examine the stability of rankings; shrinkage as a way to temper noisy extremes; the limits of R^2.

October 19. Midterm Examination

Five stations around one data patient: chart read (screen off); pencil markup of a printed analysis (screen off); the full consult (AI live); transcript audit (AI live); your signed briefing to the mayor. Graded on process. To rehearse: run "Mock Station" once during the prior week.

OWN: Process before answer: look first, ask where the data came from, identify the comparison, and translate honestly.

USE: Move between screen-off judgment, AI-assisted analysis, transcript audit, and a concise briefing to the mayor.

RECOGNIZE: A polished answer can still be wrong if the process missed the pathology or accepted the AI's first plausible story.

Module 3. The Causal Ladder

October 26. Randomization, and How It Gets Muddy

The 1954 Salk polio vaccine trial — 1.8 million children, and two research designs inside one historic experiment, one randomized and one not, with instructively different answers. Why randomization works, and what muddies it in the field: attrition, spillovers, refusal, ethics. The randomized trial as a triumph with boundaries, not a gold idol.

We open the module with the story of where all these methods came from — two tributaries, the Princeton labor economists and the Harvard statisticians, feeding one river. Knowing the history is part of mastering the field, and it explains why causal inference can feel "like walking in a third of the way into a science fiction movie."

In class: the 1954 Salk trial — two designs, 1.8 million children, instructively different answers; attrition, spillovers, refusal, and which way each biases; the field's origin story (two tributaries, one river).

Readings:

Causal Inference: The Remix, Ch. 2 ("Which Causal Inference?"). Link (1 hour)

Marcia Meldrum. 1998. "'A calculated risk': the Salk polio vaccine field trials of 1954." *BMJ* 317: 1233–1236. Link (30 minutes)

The Effect, Ch. 4 ("Describing Relationships") and Ch. 10 ("Treatment Effects"). (2 hours)

Do before class:

1. Run the lab "Muddy the Trial": take the clean design, add attrition, then spillovers, then refusal; note the direction each pushes. PP482 Package of Prompts.docx

Submit: weekly unit, on the Salk data.

OWN: Randomization earns comparability on average; field frictions can muddy that clean design in predictable ways.

USE: Diagnose attrition, spillovers, and refusal, and state which way each can push the estimated treatment effect.

RECOGNIZE: Exchangeability, ethics, and the field's two intellectual tributaries as part of understanding what randomized evidence can and cannot buy.

November 2. Natural Experiments and Discontinuity

Maimonides' rule: an 800-year-old religious edict creates a modern natural experiment on class size in Israeli schools. Regression discontinuity: what it identifies, and for whom — the effect is local, true near the threshold, which is itself a lesson in honest translation to the mayor.

Readings:

Joshua Angrist and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114(2): 533–575. Read with your AI as guide: aim for the design and Figures, not the algebra. (1½ hours)

Causal Inference: The Remix, Ch. 6 ("Regression Discontinuity"), first half. Link (1 hour)

Do before class:

1. Run "RDD Picture Quiz" until you can call four of five plots correctly. PP482 Package of Prompts.docx
2. Write the mayor's sentence stating exactly whom the effect applies to — and also write the tempting overly general sentence you must not give the mayor.
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: weekly unit, on the class-size data.

OWN: A discontinuity can identify a credible effect near a threshold; that local result must be translated locally.

USE: Read regression-discontinuity plots critically and distinguish a genuine jump from a smooth curve or threshold manipulation.

RECOGNIZE: Natural experiments as opportunities created by rules or institutions rather than by the researcher.

November 9. Difference-in-Differences — and a Field That Moved

Parallel trends, event-study plots—and a cautionary tale about the field itself: around 2020, econometricians discovered that the workhorse difference-in-differences estimator is biased when treatment rolls out at different times. Textbooks printed before 2021 teach methods now known to be flawed. Your AI may know the newer literature—or may still offer stale advice with perfect confidence. Your job is to date-stamp and check it against *The Effect*.

Readings:

The Effect, Ch. 18 ("Difference-in-Differences"), including the section on staggered rollout. (1.5 hours)

Causal Inference: The Remix, Ch. 10 ("Complex Diff-in-Diff Designs"), opening sections. Here is the post-2020 literature in teachable form. Note disagreements with *The Effect*, and take any disagreement to your AI. Link (45 minutes)

* Baker, Andrew, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2026. "Difference-in-Differences Designs: A Practitioner's Guide." *Journal of Economic Literature* 64 (2): 498–557.DOI: 10.1257/jel.20251650. Baker-et-al-2026-difference-in-differences-designs-a-practitioner-s-guide.pdf

Do before class:

1. Run "Date-Stamp" PP482 Package of Prompts.docx with a fresh Colleague: ask what changed in DiD after 2020 and what a 2019 textbook gets wrong or leaves incomplete. Verify one claim against the shelf.
2. Run "Audit This DiD" PP482 Package of Prompts.docx with a fresh Colleague on the provided analysis; judge whether its advice is current and whether the estimator matches the design.
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: weekly unit, on the DiD patient.

OWN: Difference-in-differences lives or dies by its comparison: parallel trends matter, and staggered treatment can make old two-way fixed-effects answers misleading.

USE: Audit a DiD by identifying treatment timing and controls, reading the event-study plot, and checking whether the estimator matches the design.

RECOGNIZE: Goodman-Bacon decomposition and the post-2020 family of alternatives for staggered treatment timing.

November 16. Keystone I: The Estimand Trap — and the Ladder, Built by You

We return to the world-science data, now with the full toolkit. Before class, give a fresh Colleague the panel and ask it to estimate the effect of human capital on scientific output. Do not coach it toward a particular estimator. Bring what it does.

In class: we compare the answers. If your Colleague used fixed effects, ask what variation that estimator is using. If it rejected fixed effects or proposed alternatives, make it explain why. Then we read, at last, the emblem from week 1: the same elasticity running from 0.04 to 2.9 across methods. When the method changes, the estimand may change.

Then the sandbox: teams re-estimate the human-capital effect with different estimators — fixed effects, its cross-sectionally corrected cousins, the pooled variants — each team steering its AI, and post their answers on the board. The spread from 0.04 to 2.9 assembles before your eyes, out of your own runs. The most important figure of the course stops being something you were shown and becomes something you made.

Readings:

Fedderke 2026, Sections 5.2–5.3 and Tables 9, 11, and 12, with the guided-reading prompt provided. (1.5 hours) Fedderke_KPF.pdf

Do before class:

1. Give a fresh Colleague the world-science panel (Canvas) and ask it to estimate the effect of human capital on scientific output. Bring its answer and its reasoning. Do not steer it toward a particular estimator.
2. Reread your estimator biographies in the field guide — you will need them at the ladder.
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: nothing this week.

OWN: Changing the estimator can change the question. Before comparing coefficients, ask whether the estimand changed.

USE: Compare estimates from different panel approaches and explain, in plain English, what variation each estimator is using.

RECOGNIZE: A coefficient near zero under fixed effects need not contradict a large cross-sectional relationship; the methods may be answering different questions.

November 23. Keystone II: The Residual Hunt (Guest: Prof. Johannes Fedderke)

Prof. Fedderke joins by Zoom to present the substantive stakes of the world-science study and his conclusions — and to answer, at last, some of the why-questions you parked in September. Then the residual hunt: teams claim countries from his tables of exceptional performers — Japan, Israel, China, Türkiye — and begin qualitative mini-cases in class: what institutional story explains the residual, what evidence would discriminate it from rivals, and what should we see in five years if it is true. His residuals, your shoe leather.

Instrumental variables arrive as the cautionary coda: the settler-mortality story and its reply — how a celebrated instrument was smelled out, a decade later.

Readings:

Fedderke 2026, Tables 9 and 12, revisited with fresh eyes: read the country notes as fifteen invitations. (30 minutes) Fedderke_KPF.pdf

Do before class:

1. Rank your three preferred countries for the team mini-case.
2. For your first choice, draft two rival institutional stories, each with one piece of evidence that would discriminate it.
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: nothing this week — the team memo is due Monday morning, November 30.

OWN: A residual is an invitation to investigate, not an explanation of why a case is exceptional.

USE: Turn an exceptional case into a research plan: rival stories, discriminating evidence, and a prediction that could prove you wrong.

RECOGNIZE: Instrumental variables through the settler-mortality debate: ingenious identification can remain vulnerable to provenance and measurement challenges.

Module 4. Does It Travel?

November 30. External Validity. Logit and the Mini-Case Verdicts

How far do findings travel? When diffusion closes gaps (Kenya's M-PESA) and when it does not (broadband). When the coefficient doesn't travel, the ideas often do: the conditions to examine, the candidates to visit, and the questions to carry. Team mini-case presentations. And a warning for quants and quals: the seductive anecdote cons as readily as the starred regression.

A new estimator for a new kind of outcome: adoption is yes-or-no, and OLS is the wrong tool. Maximum likelihood, taught as "which parameter values make our data least surprising," and logistic regression as its child — estimated on real survey microdata (who adopts mobile money, given income, gender, education, rural/urban). MLE's biography joins the field guide: powerful under its assumptions, brittle when the model is misspecified.

A twenty-minute cameo appearance from an actor that belongs in this course by right: synthetic control, the statisticians' answer to the single case study. One treated state; a weighted "donor pool" of untreated states builds its counterfactual twin. For a course devoted to marrying the case method to statistics, this is a thesis exhibit — try to take the con out of both econometrics and the case method, in one method.

In class: external validity — when diffusion closes gaps (M-PESA) and when it doesn't; maximum likelihood and logistic regression on real adoption microdata; team mini-case presentations, with Prof. Fedderke's written reactions read aloud; a 20-minute cameo on synthetic control

Readings:

World Bank Global Findex 2025, overview chapter on mobile money and financial inclusion. Link (45 minutes)

Regression and Other Stories, Ch. 13 ("Logistic regression"), first half. (1 hour)

* *Causal Inference: The Remix*, Ch. 11 ("Synthetic Control"), opening sections. Link

Do before class:

1. Team mini-case memo, due Monday, November 30, at noon — Thanksgiving comes first; the memo can wait a bit. Two pages: the institutional story behind your country's residual, the evidence that would discriminate it from rivals, and what we should see in five years if it is true.

2. Run “MLE From Scratch” (30 minutes): a coin, then a logit, by simulation. PP482 Package of Prompts.docx
3. Use your Tutor prompt on last week’s OWN, USE, RECOGNIZE.

Submit: the team memo, Monday noon (this replaces the weekly unit).

OWN: A finding can be sound where it was estimated and still travel badly. Ask what features of the setting govern whether it travels.

USE: Fit and translate logistic regression in probabilities for concrete people or cases, and explain why a binary outcome calls for a different model.

RECOGNIZE: Maximum likelihood as the logic behind logit, and synthetic control as a statistical approach to constructing a comparison for a single treated case.

December 7. Grand Rounds — and the Crescendo

First, Grand Rounds, as in a teaching hospital: six data patients around the room, each carrying one disease from this course — a confound, a hidden nonlinearity, the funnel, a staggered rollout, selection, measurement error. Pairs rotate bed to bed: diagnose, prescribe, or refer to design. Everything that can go wrong with regression, and what to do about it — one week before the final examination, which this session happens to rehearse.

After the break, we face the crescendo: prediction itself. The world-science study projects to 2047 — conditional forecasts, carefully labeled as such. We ask about the condition most likely to fail: the production function itself. If AI augments every researcher, the returns to human capital are not what they were; if machines read and recombine the whole stock of knowledge, the compounding parameter may lurch. Live evidence in the room: federal data now tracks firm AI adoption every two weeks, and the adoption gradient looks just like every gradient we have studied — large, rich, information-sector firms first. You are sitting on the curve you are analyzing, using the tool whose diffusion you are measuring — and the course’s own AI improved while the semester ran. Students sketch scenarios under which the 2047 projections break, and name the early indicators to watch. (Students of PP338 will recognize the method.)

Review of the course.

Readings:

Census Bureau, Business Trends and Outlook Survey — AI supplement, latest release. Link (30 minutes)

Anthropic Economic Index, latest report. Link (30 minutes)

Fedderke 2026, Section 5’s forward projections. (30 minutes) Fedderke_KPF.pdf

Do before class:

1. Run “Break the Forecast”: two scenarios that break the 2047 projections, each with its earliest observable indicator (one page). PP482 Package of Prompts.docx

2. Reread your own field-guide entries; bring the one you are proudest of, and the one you would now rewrite.
3. Use your Tutor prompt on last week's OWN, USE, RECOGNIZE.

Submit: the scenarios page.

OWN: Every extrapolation is a bet that the underlying machine stays enough the same; good forecasts name what would make that bet fail.

USE: Diagnose familiar pathologies in unfamiliar data and build break-the-forecast scenarios with early indicators to watch.

RECOGNIZE: AI adoption as a live example of structural change: it may alter both the inputs to knowledge production and the production function itself.

December 14. Final Examination

The same five-station format as the midterm, with several things wrong at once. Open book, open notes, AI in tow. Designed for three hours; you may take four. You will leave with something rarer than a grade: a field guide your class wrote, and a set of reflexes no future estimator — human or machine — will easily bluff.

OWN: The course's habits should now travel with you: interrogate the data, the model, the AI, and your own hopes.

USE: Diagnose a data patient with several problems at once and decide what can be estimated, what must be qualified, and what should be investigated next.

RECOGNIZE: Knowing when the evidence cannot answer the question is part of expert statistical judgment, not a failure of technique.

University Policies

CGU institutional policies apply, including academic integrity and accommodations; see https://bulletin.cgu.edu and https://cgu.edu/registrar. Your work must be your own; your collaborations — with classmates and with AI — must be transparently documented, which in this course is not a burden but the method.

University Resources for Students

Access to university resources and policies can be found in Canvas using the left navigation menu.

- **Student Resources**
- **Requirements & Responsibilities**
- **Center for Writing & Rhetoric**
- **The Claremont Colleges Library**

Technical Resources

- Office of Information Technology (OIT)
- Canvas Guides
- Microsoft 365 Training
- **(?) Help** Button (accessible through left navigation bar in Canvas)