Soc 3580

Big Data on the Social World

Spring, 2024
Friday 2:00PM – 4:30PM
Statler 291

Cristobal Young
Department of Sociology
344 Uris Hall
cristobal.young@cornell.edu
Office Hours: by appointment

This course showcases frontier research that uses big data and graphical analysis to understand our social world. Topics include inequality and opportunity, success in higher education, the gender wage gap, taxing the rich, Chinese censorship, the spread of false news, online dating, and other issues relevant to contemporary society. Although this is not a statistical methods course, prior training in data science (e.g., CS1380/ORIE1380/STSCI1380) or quantitative methods for the social sciences is highly recommended. This course showcases frontier research using big data to understand our social world.

We focus on research that offers (1) engaging topics, which pulls us into a deeper reflection on the social world; (2) uses 'big' / administrative data; and (3) visualizes the analysis so that non-technical readers can see and understand how the evidence emerges. We discuss research design through applied cases, such as natural experiments and discontinuity analysis. A central theme of the course is, how do social scientists analyze big data?

Please note that the readings may be revised as the course unfolds. The assigned readings are available on Canvas, in which you should be automatically enrolled after registering for the course. Please come to lecture having read the assigned readings for that day.

Course requirements:

Class participation	15%
Memos on readings	25%
Mid-term exam	30%
In class final exam	30%

Class participation

As an upper-year seminar, class discussion will be an important part of the course. I will present on the readings, and then open the floor for discussion. To facilitate thoughtful discussion involving everyone, we will follow the main lectures with a think-pair-share format. (1) spend a minute writing down questions / comments about the lecture; (2) pair up with your neighbor and discuss each of your questions, formulating the best answers you can; (3) share what you discussed with the full class. This will allow you to test out and elaborate on your initial ideas in a low-stakes setting (with your neighbor) before having to "go public."

Memos on weekly readings

Students will write **five memos** based on the readings of the week – one memo every second week. Memos are two pages (max 750 words) and will focus on the empirical analysis in the paper. What evidence do the authors draw on? How does their analysis work: can you explain what the authors did in plain language? What do their findings tell us?

Do not spend time in your memo talking about how the reading made you feel, or what was surprising to you. Memos lose points when they narrate your experience of reading article. Focus on the content: efficiently summarize your key points, comments, arguments, and questions from the readings. You may use point form to highlight your key questions for class discussion. The memos are graded "excellent" (5), "good" (4), "satisfactory" (3) or "unsatisfactory" (0).

> *Please upload your memos to Canvas by 9pm the night before class.*
> ***Upload files in Word. Use 1.5 spacing (as I will print them).***

Exam

The final exam will be comprehensive and closed-book, with questions focusing on the broad themes of the course. The content will be drawn from specific questions previously discussed in class.

Class Schedule

Week 1. Jan 26. Introduction.

Week 2: Feb 2. Inequality and Social Mobility.

Before break:

Saez, Emmanuel. 2008. "Striking it Richer: The Evolution of Top Incomes in the United States." *Pathways*. (Updated)

Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*. Vol. 27(3): 79-102.

After break:

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*. Vol. 129(4): 1553-1623.

Start with the Non-Technical Summary.

Exercise: Find the place where you have lived the most in the US using the Racial Dot Map. Does this area seem to you more as racially integrated or segregated? Does the map fit with your lived experience of the area?

Week 3: Feb 9. Higher Education and Social Mobility

Haskins, Ron. 2008. "Education and Economic Mobility." Economic Mobility Project: Pew Charitable Trusts.

Leonhardt, David. 2017. "America's Great Working Class Colleges." *New York Times*. *Also*: Look up data for Cornell University.

Chetty, Raj, John Friedman, Emmanuel Saez, Nicholas Turner, Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States." *Quarterly Journal of Economics*. Vol 135(3): 1567–1633.

Week 4: Feb 16. Social Networks

Before break:

Young, Cristobal, and Chayoon Lim. 2014. "Time as a Network Good." *Sociological Science*.

After break:

Bailey, et al. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives*.

Optional further reading:
Bailey, et al. 2022. "Peer Effects in Product Adoption." *American Economic Journal: Applied Economics*, 14 (3): 488-526.

Week 5. Feb 23. Social Capital

Before break:

Chetty, Raj, et al. 2022. "Social capital I: measurement and associations with economic mobility." *Nature*.

> *Optional further reading:* Chetty, Raj, et al. 2022. "Social capital II: determinants of economic connectedness." *Nature*.

After break:

Young, Cristobal, Barum Park, Benjamin Cornwell, and Nan Feng. 2024. "Inequality in Social Capital: Evidence from Big Diverse Data." Working Paper.

Week 6: Mar 1. Covid-19 Pandemic

Before break:

Chetty, Raj, et al. 2020. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." NBER working paper #27431.

After break:

Polack, Fernando, et al. 2020. "Safety and Efficacy of the [Pfizer] BNT162b2 mRNA Covid-19 Vaccine." *New England Journal of Medicine*. Vol. 383:2603-2615

Dagan, Noa, et al. 2021. "BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting." *New England Journal of Medicine.*

Week 7: Mar 8. Natural Experiments.

"Mastering 'Metrics." Read to understand the graphical analysis. SKIP section 5.1, as it is a painfully boring example. Section 5.2 picks up directly from Chapter 4.

DeAngelo, Gregory, and Benjamin Hansen. 2014. "Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities." *American Economic Journal: Economic Policy*, 6 (2): 231-57.

Examples from:

Chetty, Raj, Nathaniel Hendren, Lawrence Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review*. Vol. 106(4): (pp. 855-902)

Carrell, Scott, Mark Hoekstra, James West. 2010. "Does drinking impair college performance? Evidence from a regression discontinuity approach." *Journal of Public Economics*.

Pierson, et al. 2020. "A large-scale analysis of racial disparities in police stops across the United States." *Nature Human Behaviour*. Vol. 4:736-45.

Week 8: Mar 15. In-class mid-term exam

Week 9: Mar 22. Taxing the Rich.

   Before break: First: watch this video

   Young, Cristobal. 2018. *Taxing the Rich: the Myth of Millionaire Tax Flight*. Stanford University Press. Chapters 1, 2, and 5.

   After break:

   Web-scraping data on elites

      Young (2018) on World Billionaires (chapter 3, p 43-54)

      2 short essays by Seth Stephens-Davidowitz: "The Geography of Fame" and "Just How Nepotistic Are We?"

Week 10. Mar 29. No Class.

Week 11: Apr 5. Spring Break.

Week 12: Apr 12. Gender Wage Gap

Before break:

Belkin, Lisa. "The Opt-Out Revolution." Pp. 332-36 in Grusky reader.

Correll, Shelley, Stephen Benard, and In Paik. 2007. "Getting a Job: Is There a Motherhood Penalty?" Pp 365-77 in Grusky reader.

Optional: "On Pay Gap, Millennial Women Near Parity – For Now." *Pew Research Center*. Read Introduction and Chapter 1. (Pages 1 to 25)

After break:
Kleven, Henrik, Camille Landais, and Jakob Søgaard. 2018. "Children and Gender Inequality: Evidence from Denmark." NBER Working Paper No. 24219.

--- 2019. "Child Penalties across Countries: Evidence and Explanations." AER. Pp. 121-26.

Optional further reading: "The Geography of Child Penalties and Gender Norms: Evidence from the United States." https://www.nber.org/papers/w30176

Week 13: Apr 19: Text Analysis – Language as Data.

* *How to make a spam filter*

Wu, Alice. 2018. "Gendered Language on the Economics Job Market Rumors Forum." *AEA Papers and Proceedings*, 108, Pp. 175-179.

After break:

Cheng, Justin, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2017. "Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions." CSCW 2017.

Week 14: Apr 26. Chinese Censorship

Gary King, Jennifer Pan, and Margaret E Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review*. Vol. 107, 2 (May), Pp. 1-18.

Week 15: May 3. **In class final exam**