

Soc 596: Computational Social Science

Matthew J. Salganik
Department of Sociology
Princeton University

01-04-Preview for next class (Big Data)



“Big data”:

- ▶ three Vs: Volume, Velocity, Variety
- ▶ not collected for the purposes of academic research
- ▶ government administrative data and business administrative data

Bit by Bit: Social Research in the Digital Age, Chapter 2:

Observing behavior.

— Farber. 2015. Why you Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers. *Quarterly Journal of Economics*.

— Helft. 2008. Google Uses Searches to Track Flu's Spread. *New York Times*.

Ginsberg et al. 2008. Detecting influenza epidemics using search engine query data. *Nature*.

Goel et al. 2010. Predicting consumer behavior with Web search. *PNAS*.

Lazer et al. 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science*.

Yang et al. 2015. Accurate estimation of influenza epidemics using Google search data via ARGO. *PNAS*.

Google Correlate: The Comic Book.

Spend 2 minutes playing with Google Trends.

Spend 2 minutes playing with Google Correlate.

— Mas and Moretti. 2009. Peers at Work. *American Economic Review*.

- ▶ Sunday at midnight: upload research proposal, annotating the chapter using hypothes.is
- ▶ Monday at midnight: give feedback on research proposal, upload feedback on chapter
- ▶ Tuesday, 2 - 5: class
- ▶ Wednesday, 2 - 5: lab (optional)