Computational approaches to health, especially approaches harnessing "big data" offer researchers emerging methods and novel data to understand social inequalities. Through data sources such as social media and smart city technology (e.g., streaming passenger sensor data from public buses), the public is generating trillions of data points and drawing the attention of researchers interested in classic subject areas such as understanding development of inequalities and newer research areas like the connection between online and offline social worlds. Big data has unique strengths and limitations that are magnified by the methods used to analyze big data in the social sciences, a methodology broadly known as computational social science. Computational social science employs computational approaches such as cryptography and machine learning algorithms to analyze, simulate, and model behavioral phenomena. Additionally, computational methods provide large-scale access to unstructured data types such as text, images, and audio that previously eluded social science research. This dissertation explores the substantive and methodological contributions and limitations of blending big data, computational social science, and synthetic data in two health domains. Synthetic datasets are used to maintain data privacy and simulate data entry uncertainty by leveraging multiple imputation and most importantly for these studies, intentionally inaccurate (e.g., not the actual collected value in the original data set) data. The first study uses Twitter and merges image and geolocation data to assess demographic variations in physical activity attitudes. Chapter one demonstrates how an integrated data gathering and refinement approach can produce a high quality social media data set. Sentiment polarity findings indicate that racial minorities (especially women) discuss physical activity as positive and often more positive than whites challenging conventional hypotheses of race and physical activity attitudes. The second and third chapter use synthetic data based on Homeless Management Information Systems (HMIS) administrative data from a county comprising one major metropolitan area. HMIS data was collected from multiple service providers in federally funded housing programs such as transitional housing, emergency shelters, rapid-rehousing, support services only, etc. The second chapter’s substantive aim is to identify family typologies of homelessness service use. Identifying typologies of service use is a goal of homelessness research to align homeless interventions with users and identifying family typologies have become a research focus over the last twenty years. The substantive goal for the third chapter is to examine how family characteristics such as household structure and parents’ race interact with housing program interventions to influence homelessness exit pathways. Researchers have disputed how influential housing program and family characteristics are for homelessness duration or persistence in homeless cycles making targeted interventions difficult. The methodological contribution of this chapter is to use random forest classifiers to predict exit pathways across different family characteristic constellations. This dissertation engages multiple computational social science approaches with synthesized administrative and social media data. Strategies and results reveal the potential and pitfalls from leveraging emerging methodology and data to investigate health disparities.
People Involved: Kivan Polimis
Adviser(s): Kyle Crowder, Hedwig (Hedy) Lee, Emilio Zagheni
Status of Research: Completed/published
Research Type: Graduate Dissertations
Related Fields: Applied Sociology, Big Data, Computational Methods, Demography, Health
Health Disparities, Homelessness, Public Opinion, Race and Ethnicity

Department of Sociology · University of Washington · 211 Savery Hall, Box 353340 · Seattle, WA 98195-3340
Main Office: (206) 543-5882 · Student Services: (206) 543-5396 · Fax (206) 543-2516 · Email: uwsoc@uw.edu

Copyright © 2015-2020 University of Washington · Privacy · Terms · Site Map · Contact Us

Source URL: https://soc.washington.edu/research/graduate/developing-computational-approaches-investigate-health-inequalities