Project Title: Mapping Social Determinants of Health in Utah: A Data-Driven Approach for Public Health and Clinical Decision-Making

Group 4 Names (emails, UIDs):
Sunho Im (u1526695@utah.edu, u1526695)
Joshua Kawasaki (J.K.Kawasaki@utah.edu, u1424902)
Kristin Knippenberg (k.knippenberg@utah.edu, u0243722)
Mahony Reategui Rivera (mahony.reategui@utah.edu, u1474537)
L Weaver (u1558626@utah.edu, u1558626)
Github Site: https://github.com/lwynholds/BMI6016_SDOH

Research question: What are the disparities in health outcomes in Utah, based on the intersection of air quality with other environmental health factors and SDOH?

Background and Motivation

The concept of Social Determinants of Health (SDOH) has been widely recognized in health research for approximately 25 years, following the early efforts by the World Health Organization (WHO) to contextualize systemic and population-level health disparities worldwide. In 2003, WHO introduced a comprehensive framework for SDOH, categorizing key social factors influencing health, including the social gradient, stress, early life, social exclusion, work, unemployment, social support, addiction, food, and transportation.(The World Health Organization (WHO), 2003) Since then, numerous SDOH frameworks have been developed at national and regional levels, reflecting these determinants' complex and multifaceted nature. While these frameworks aim to address population health holistically, their generalizability usually depends somewhat on context and population.(Braveman et al., 2011; Braveman & Gottlieb, 2014; Marmot, 2005; Marmot et al., 2012; Marmot et al., 2008)

In the United States, SDOH are central to public health and policy, and as such, are measured, tracked, and studied in detail by the Centers for Disease Control and Prevention (CDC) in cooperation with other organizations that conduct national survey data, such as Health and Human Services (HHS), the US Census and their American Community Survey (ACS), the Agency for Health Research and Quality (AHRQ), and the National Institutes of Health (NIH). A substantial portion of SDOH data is derived from national surveys, census records, and geographic datasets, providing valuable insights into social and environmental conditions that impact health outcomes. However, integrating these diverse data sources into actionable tools for clinical and public health applications remains a significant challenge.(Cantor et al., 2018)

Given the vast scope of SDOH research, identifying specific applications that yield tangible benefits can be overwhelming. SDOH influence health trajectories from birth to death, shaping risks for chronic diseases, accidents, and overall well-being. We focused on leveraging community-level rather than self-reported data to develop a practical and clinically relevant tool. In 2018, Cantor et al. published their Factors Affecting Communities and Enabling Targeted Services (FACETS) model, an open-data architecture, designed for New York City to

standardize and compile SDOH-related data at the census-tract level.(Cantor et al., 2018) FACETS facilitates the integration of community-level determinants with patient health records, enhancing decision-making in care planning.

Building on this approach, we propose developing a FACETS-based model tailored to the state of Utah, with special focus on air quality as a primary exposure. By integrating state-specific SDOH data, including environmental data, into a standardized framework, this initiative aims to provide a comprehensive tool for assessing community-level health determinants, ultimately supporting more informed public health interventions and clinical decision-making. Air quality, environmental health factors, and SDOH all influence health outcomes.(McRoy et al., 2017) However, their combined effects, particularly at the intersection of these factors, are less understood. Using this model, we hope to explore the disparities in health outcomes based on the intersection of air quality with other environmental health factors and SDOH.

Project Objectives

We propose to standardize Utah's SDOH data at the census tract (a "small, relatively permanent statistical subdivisions of a county" ** https://www.census.gov/programs-surveys/geography/about/glossary.html) level by integrating multiple sources of publicly available and state-specific data. Our objectives include:

- Standardizing Utah SDOH data at the census-tract level using:
    - ACS and US Census Bureau, CDC, Environmental Protection Agency (EPA), and US Department of Agriculture Food Access Research Atlas (USDA) data (data sources from the original FACETS study that are also relevant to Utah)
    - State datasets from the Utah Department of Health and Human Services
    - State SDOH variables not otherwise specified that are particular or pertinent to Utah (e.g., urban/rural, electoral participation)
    - Temporal, categorical, or spatial crosswalks, where necessary
- Creating a database of census tracts, within which an individual patient's address can be mapped to a Utah FACETS characterization of SDOH context
- Characterizing disparities in health outcomes based on the intersection of air quality with other environmental health factors and SDOH.

This project will help give context to any individual patient's health status and assist patients and providers in deciding how best to improve that patient's health, given the available resources.

Data

SDOH	Measure	URL to Source
Economic Stability	Employment , Food insecurity, Urban/rural classification, Employment, Housing instability, Poverty	USDA
		ACS
Education Access and Quality	Enrollment in higher education, High school graduation, Language and literacy, Early childhood development and education	ACS

Health Care Access and Quality	Health insurance status, Access to health services, Access to primary care, Health literacy, Health Care Services	ACS
		AHRQ
Neighborhood and Built Environment	Crime and violence, Access to Foods That Support, Healthy Dietary Patterns, Respiratory Hazard Index, Walkability Index, Quality of Housing	Utah Gov
		EPA
		EPA Walkability Index
Social and Community Context	Civic Participation, Discrimination, Incarceration, Social Cohesion	TBA
Other Indices	Social Vulnerability Index	CDC-SVI
	Election results	Utah Gov

See also (Office of Disease Prevention and Health Promotion, 2024)

Data Processing

Many of the datasets will be “analysis-ready” and will contain survey counts and percentages, means, and standard errors; however, careful consideration will be required to avoid aggregation errors, such as mismatching sample populations in the numerator and denominator for calculations. The FACETS study “highlights reverse geocoding latitude and longitude values to census tracts” for specific data types, which we will implement to ensure spatial alignment. To accomplish all this, we will use SQL and Python for data wrangling and integration, while geospatial data processing will be conducted using QGIS or ArcGIS.

A key component of our processing workflow will be the validation of data harmonization across sources. This involves assessing the temporal alignment of datasets to prevent inconsistencies due to varying collection periods and ensuring that categorical data maintains uniform classifications across different sources. In cases where direct mapping between datasets is not feasible, we will develop or source conversion tools and crosswalks to facilitate integration.

Design

The design of this project will focus on structuring SDOH data in a manner that allows for census-tract-level analysis while preserving the integrity of Utah-specific variables. The primary data types will include counts, percentages, means, and/or standard errors, ensuring flexibility in statistical interpretation. Individual tables will be designed based on the original data sources, with minimal transformation required for analysis-ready datasets. Nonetheless, we expect to encounter the most complexity in joining and combining these differing exposures to evaluate health outcomes.

A crucial design consideration will be the selection of Utah-specific SDOH indicators that can meaningfully contribute to the model. These indicators must be translatable to census-tract-level data, ensuring they can be effectively linked to individual patient addresses without compromising data granularity. The usability of the final dataset will be a guiding principle in our design approach, prioritizing clarity and ease of integration with healthcare applications.

Must-Have Features
All data elements must be mappable to census tracts for the Utah model to function effectively. If necessary, we will seek collaboration with other researchers for conversion tools or crosswalks to facilitate this process. In cases where such tools are unavailable, we will explore alternative data sources that offer similar granularity and coverage. Ensuring census-tract-level translation will be a non-negotiable feature of our dataset, as it directly impacts the accuracy and applicability of our model.

Optional Features
While the primary objective of this project is to create a standardized and actionable dataset, additional features may enhance its utility. One potential enhancement is the development of geocoded, interactive data layers that allow users to visualize SDOH trends spatially. While our initial interface may be relatively more basic, we will aim to provide a formula or look-up tool that can retrieve a FACETS profile based on patient address. This would enable users to quickly assess the SDOH context for specific locations, supporting both research and clinical decision-making.

By adhering to these methodological principles, we aim to develop a robust and scalable Utah FACETS model that is similar to the original FACETS methodology, while addressing the unique characteristics of the State of Utah's population and healthcare landscape.

Project Schedule

Week (Sat-Sun) or Due Date	Goals/Deliverables
Feb 16	SUBMIT PROJECT PROPOSAL (3-4 pages)
Feb 16-22	Identifying key attributes from databases and ensuring census-track mapping
Feb 23-Mar 1	Conceptual model design, variable mapping
Mar 2-8	Data cleaning/wrangling
Mar 7	MEET W/INSTRUCTORS FOR FEEDBACK #1
Mar 9-15	Physical database design and setup
Mar 16	INTERMEDIATE PRESENTATION AND PEER FEEDBACK. **Presentation follows proposal outline and current state of implementation. (**A relatively informal discussion in class of what you've been working on, to get feedback from your classmates, suggestions for data integrations, etc.)
Mar 16-22	Model implementation, exploration, and testing
Mar 23-29	Model validation
Mar 30	PROJECT UPDATE SUBMISSION VIA GITHUB (Instructions TBA)
Mar 30-Apr 5	Documentation of process
Apr 4	MEET W/INSTRUCTORS FOR FEEDBACK #2
Apr 6-12	Prepare analysis and results
Apr 13-19	Finalize report/presentation
Apr 21	FINAL PRESENTATION AND PEER FEEDBACK (So far, instructions are just like Intermediate on Mar 16)
Apr 22	FINAL PROJECT DUE

References

Social Determinants of Health (SDOH) Database (https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html

Behavioral Risk Factor Surveillance System (https://www.cms.gov/About-CMS/Agency-Information/OMH/resource-center/hcps-and-researchers/data-tools/sgm-clearinghouse/brfss

Nanda (National Neighborhood Data Archive) (https://nanda.isr.umich.edu

Braveman, P., Egerter, S., & Williams, D. R. (2011). The social determinants of health: coming of age. *Annu Rev Public Health*, *32*, 381-398. https://doi.org/10.1146/annurev-publhealth-031210-101218

Braveman, P., & Gottlieb, L. (2014). The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep*, *129 Suppl 2*(Suppl 2), 19-31. https://doi.org/10.1177/00333549141291S206

Canaday, F. T., Georas, S. N., & Croft, D. P. (2024). Examining the impact of air pollution, climate change, and social determinants of health on asthma and environmental justice. Curr Opin Pulm Med, 30(3), 276-280. https://doi.org/10.1097/MCP.0000000000001065

Cantor, M. N., Chandras, R., & Pulgarin, C. (2018). FACETS: using open data to measure community social determinants of health. *J Am Med Inform Assoc*, *25*(4), 419-422. https://doi.org/10.1093/jamia/ocx117

Cook, Q., Argenio, K., & Lovinsky-Desir, S. (2021). The impact of environmental injustice and social determinants of health on the role of air pollution in asthma and allergic disease in the United States. J Allergy Clin Immunol, 148(5), 1089-1101 e1085. https://doi.org/10.1016/j.jaci.2021.09.018

Gao, C., Sanchez, K. M., & Lovinsky-Desir, S. (2023). Structural and Social Determinants of Inequitable Environmental Exposures in the United States. Clin Chest Med, 44(3), 451-467. https://doi.org/10.1016/j.ccm.2023.03.002

Marmot, M. (2005). Social determinants of health inequalities. *Lancet*, *365*(9464), 1099-1104. https://doi.org/10.1016/S0140-6736(05)71146-6

Marmot, M., Allen, J., Bell, R., Bloomer, E., Goldblatt, P., Consortium for the European Review of Social Determinants of, H., & the Health, D. (2012). WHO European review of social determinants of health and the health divide. *Lancet*, *380*(9846), 1011-1029. https://doi.org/10.1016/S0140-6736(12)61228-8

Marmot, M., Friel, S., Bell, R., Houweling, T. A., Taylor, S., & Commission on Social Determinants of, H. (2008). Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet*, *372*(9650), 1661-1669. https://doi.org/10.1016/S0140-6736(08)61690-6

McRoy, L., Rust, G., & Xu, J. (2017). Factors Associated with Asthma ED Visit Rates among Medicaid-enrolled Children: A Structural Equation Modeling Approach. *AIMS Med Sci*, *4*(1), 71-82. https://doi.org/10.3934/medsci.2017.1.71

Office of Disease Prevention and Health Promotion. (2024, Jan 1). *Social Determinants of Health Literature Summaries*. US Department of Health and Human Services. Retrieved February 16 from https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries#block-sdohinfographics

The World Health Organization (WHO). (2003). *Social Determinants of Health: The Solid Facts, 2nd edition*. Retrieved from https://iris.who.int/handle/10665/326568

Zhu, G., Wen, Y., Cao, K., He, S., & Wang, T. (2024). A review of common statistical methods for dealing with multiple pollutant mixtures and multiple exposures. Front Public Health, 12, 1377685. https://doi.org/10.3389/fpubh.2024.1377685