Tutorial

Health Atlas: Tutorial of a Visualization Tool and Data Resource for Place-Based Social and Structural Determinants of Health

Debora L Oh[1], MSc, PhD; Kathryn E Kemper-McIsaac[1], MPH; Dan Meltzer[1], MA, MPH; Eric Brelsford[2], MS; Kelsey Taylor[2], MSc; Dhananjay Vinay Dixit[2], BA; Andrea Nickerson[1], MS; Riya Desai[3], MS, MPH, MD; Salma Shariff-Marco[1], MPH, PhD; Courtney Lyles[4], PhD; Scarlett Gomez[1], MPH, PhD; Mark J Pletcher[1], MPH, MD; Mindy C DeRouen[1,5], MPH, PhD

[1]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

[2]Stamen Design, San Francisco, CA, United States

[3]Kaiser Permanente Oakland Medical Center, Oakland, CA, United States

[4]Center for Healthcare Policy and Research, University of California, Davis, Sacramento, CA, United States

[5]Department of Public Health Sciences, College of Health, Education, and Social Transformation, New Mexico State University, Las Cruces, NM, United States

Corresponding Author:

Debora L Oh, MSc, PhD
Department of Epidemiology and Biostatistics, University of California, San Francisco
550 16th St., Floor 2
San Francisco, CA 94143
United States
Phone: 1 415-476-2300
Email: debora.oh@ucsf.edu

Abstract

Places have a significant impact on health. Thus, examining place-based structural and social determinants can help inform effective public health interventions. HealthAtlas.ucsf.edu provides a single online platform to explore multiple domains of place-based data. The initial California-only version of Health Atlas was launched in April 2020, and a national version was launched in October 2024. An artificial intelligence–assisted search function was integrated in January 2026. Health Atlas includes data on over 200 variables across 5 topical domains: demographics, socioeconomic, neighborhood, environment, and health and health care. The data were either (1) obtained from the American Community Survey, CDC PLACES, and other public sources or (2) obtained through collaboration with research partners. Users can visualize and aggregate data for 7 geographic levels: census tract, zip code tabulation area, county, congressional district, core-based statistical area, public use microdata area, and state. Most variables are available across all 50 states in the United States, the District of Columbia, and Puerto Rico. Users can explore Health Atlas to better understand the relationship between selected variables via maps, histograms, scatterplots, and summary plots. Users can select custom areas to suit specific needs. Customizable datasets can be downloaded for further use. Health Atlas provides freely available, user-friendly data and tools to support researchers, community organizations, government entities, and other public health professionals. We envision that Health Atlas will contribute to evidence-based, community-based initiatives; impactful health equity research; and effective public health programs.

J Med Internet Res 2026;28:e89065; doi: 10.2196/89065

Keywords: demography; socioeconomic factors; neighborhood characteristics; environment; social environment; public health; social determinants of health

Introduction

When it comes to health, place matters. Where we live, work, and play influences health in both positive and negative ways. Places have a significant impact on health outcomes independent of individual-level factors [1]. Examining place-based social determinants of health, such as access to food, housing, air quality, and social support, as well

as structural determinants, including the historical context, governance, policies, and norms that shape local environments, can help inform effective public health interventions [2,3].

While the concept of place can be subjective, smaller areas are often defined by census tract or zip code boundaries, and larger areas are often defined by counties, congressional districts, or states. Using these definitions, publicly accessible place-based data can be leveraged to identify factors underlying the health of a population. Prior studies have shown that neighborhood socioeconomic status, measured at the census tract-level, is associated with increased rates of heart disease, cancer, and all-cause mortality, even after individual-level socioeconomic status is taken into account [4-6]. Other studies show that zip codes in which children grow up influence their future income, educational attainment, fertility, and marriage outcomes [7]. Analyses of the impact of state Medicaid expansion status on comorbidity burden show the influence of state-level policy on health inequities [8].

Over the last decade, the availability of publicly accessible place-based data has greatly expanded [9-11]. However, the growth of data resources has been segmented and nonuniform. Many websites developed by individual governmental organizations or research groups offer domain-specific data, such as social environment, climate, health outcomes, or housing [12-15]. In addition, new or expanded place-based measures have been developed that are not yet accessible through online platforms [16,17]. For some domains (eg, rurality [18]), multiple measures exist but often are not available in one place to facilitate comparisons or assess suitability for particular applications. Furthermore, the availability of national data has become inconsistent and uncertain due to cuts to federal funding for biomedical research and shifting priorities at federal agencies [19-21]. The result is an extensive but evolving, complex, and fragile landscape of place-based measures.

UCSF Health Atlas [22] provides a single online platform for place-based data to help users explore place-based factors that influence health. It includes a curated repository of publicly available population estimates as well as expert-driven composite measures. Most data are available across all 50 states in the United States, the District of Columbia, and Puerto Rico.

The development of Health Atlas was initiated by the University of California, San Francisco (UCSF) Population Health and Health Equity program and led by the UCSF Population Health Data Initiative. The website was created in collaboration with Stamen Design, and its content was developed with expertise from the UCSF Department of Epidemiology and Biostatistics, UCSF Disparities Research Environment and Multi-omics Lab, and Plain Language Health. Health Atlas development began in 2019, and the initial California-only version was launched in April 2020 [23]. Over time, a mobile-enabled version, Census 2020 geographies, and new domains of data were added. In fall 2024, the Health Atlas team launched a national version that included national maps and data [24]. An artificial intelligence (AI)–assisted search function was integrated in January 2026.

Data Assets

Health Atlas supports data visualization for 7 geographic levels: census tract, zip code tabulation area, county, congressional district, core-based statistical area, public use microdata area, and state. While not all data are available at every geographic level, most data are provided at the census tract level, the smallest unit of geography in Health Atlas (representing areas of approximately 4000 residents). Health Atlas provides data based on 2010 Census and 2020 Census boundaries; variables are merged on a geographic identifier, such as the Census Tract Federal Information Processing Series code [25], and transformed into flat files based on geographic granularity. Providing multiple census geographies allows researchers to select, download, and merge data relevant to their time period of interest.

To date, Health Atlas includes data on 206 variables (also referred to as “characteristics”) across 5 topical domains: demographics, socioeconomic, neighborhood, environment, and health and health care (Table 1). Domains were modeled after the Office of Disease Prevention and Health Promotion’s social determinants of health conceptual framework, which includes domains of health care access and quality, education access and quality, social and community context, economic stability, and neighborhood and built environment. Variables were selected for inclusion based on this framework, as well as feedback from expert scholars and potential users from academic, medical, and community settings. Data sources and variables are vetted for inclusion based on their relevance to place-based social determinants of health, accessibility (eg, from a public source), national scope, availability at the geographic boundaries represented in Health Atlas, and sustainability (eg, feasibility to update over time). Included data are curated primarily from public sources including the American Community Survey (ACS) [26] (eg, demographics, socioeconomics), US Centers for Disease Control and Prevention (CDC) PLACES: Local Data for Better Health [12] (eg, health and health care), and US Environmental Protection Agency (EPA) Environmental Justice Screen [27] (eg, environmental exposures).

Table 1. Health Atlas variable domains, subdomains, and sources (2025 version).[a]

Domain and subcategory	Sources
Demographic	
• Age • Sex	• ACS[b] 5-year estimates [28]

Domain and subcategory	Sources
• Race and ethnicity	
Socioeconomic	
• Socioeconomic indices	• Public Health Alliance of Southern California [29] • Boston University: Institute for Equity in Child Opportunity & Healthy Development [30] • CDC[c]: ATSDR[d] [31] • Custom measures generated by DREAM[e] Lab [16] and HEAN[f] [32]
• Poverty and income	• CDC PLACES [12] • NCI[g] [33] • USDA[h] Economic Research Service [34] • HUD[i] Comprehensive Housing Affordability Strategy [35] • ACS 5-year estimates [28] • ACS 5-year estimates [28]
• Income inequality	
• Employment	
• Education	
• Digital access	
Neighborhood	
• Total population	• ACS 5-year estimates [28]
• Built environment	• USDA Food Access Research Atlas [36] • University of Michigan: NaNDA[j] [37] • EPA[k] EJ[l] Screen [27]
• Housing	• HUD Comprehensive Housing Affordability Strategy [35] • ACS 5-Year Estimates [28] • CDC PLACES [12]
• Language and ethnic enclaves	• Custom measures generated by UCSF[m] DREAM Lab [16] • ACS 5-Year Estimates [28]
• Rurality	• USDA Economic Research Service [34] • US Census Bureau [38]
• Segregation	• Custom measures generated by HEAN [32] • NCI [33]
• Social environment	• CDC PLACES [12] • ACS 5-Year Estimates [28]
• Structural racism	• University of Michigan: Institute for Social Research [39]
• Transit	• CDC PLACES [12] • ACS 5-Year Estimates [28] • University of Michigan: NaNDA [37] • EPA EJ Screen [27]
Environment	
• Air pollution	• EPA EJ Screen [27] • Custom measures generated by UC[n] CCHE[o] [17]
• Heat	• Custom measures generated by UC CCHE [17]
• Extreme precipitation	• Custom measures generated by UC CCHE [17]
• Environmental indices	• EPA Public Health and Environmental Systems Division [40] • CDC: ATSDR [31] • FEMA[p] [13]
Health and Healthcare	
• Health outcomes	• CDC PLACES [12] • CDC NVSS[q] [41]
• Health care access	• ACS 5-year estimates [28]
• Health care use	• CDC PLACES [12]
• Health behaviors	• CDC PLACES [12]
• Disability	• ACS 5-year estimates [28] • CDC PLACES [12]
• Mental health	• CDC PLACES [12]

[a]Additional details for each variable are provided in Multimedia Appendices 1 and 2.
[b]ACS: American Community Survey.

[c]CDC: US Centers for Disease Control and Prevention.
[d]ATSDR: Agency for Toxic Substances and Disease Registry.
[e]DREAM: Disparities Research Environment and Multi-omics.
[f]HEAN: Health Excellence Action Network.
[g]NCI: National Cancer Institute.
[h]USDA: US Department of Agriculture.
[i]HUD: US Department of Housing and Urban Development.
[j]NaNDA: National Neighborhood Data Archive.
[k]EPA: US Environmental Protection Agency.
[l]EJ: Environmental Justice.
[m]UCSF: University of California, San Francisco.
[n]UC: University of California.
[o]CCHE: Center for Climate, Health, and Equity.
[p]FEMA: Federal Emergency Management Agency.
[q]NVSS: National Vital Statistics System.

The data are selected based on the most recent version available for a given geographic boundary. For example, the 2015-2019 ACS variables (eg, for race/ethnicity) are available for Census 2010 geographies because they are the most recent variables available for that geography; the Census 2020 boundary is currently associated with the 2019-2023 ACS variables. Health Atlas aims to update data annually following the release of new data from the Census and to search for additional updates to other variables at the same time. Health Atlas does not maintain longitudinal data, so when new data are released at an existing geographic boundary, it replaces the old version.

Full descriptions, data sources, available years, and geographies for each variable are provided on the Health Atlas and in Multimedia Appendices 1 and 2.

Variable Types

American Community Survey Variables

Five-year estimates from the ACS are downloaded from the tidycensus application programming interface package in R through a batch download. The Health Atlas includes many variables directly generated by and reported from the ACS. For example, "Overcrowding" is pulled directly from the Selected Housing Characteristics table. Others, such as "Seniors living alone," are calculated internally based on ACS estimates:

$$\frac{\text{Male householder over 65 living alone} + \text{Female householder over 65 living alone}}{\text{Total population 65 years and over}}$$

Data From Government Sites

Other publicly available variables, such as those from the EPA Environmental Justice Screen and CDC PLACES, are downloaded directly as .csv files from government websites.

Index Variables

Health Atlas also includes selected index variables generated for Health Atlas by the UCSF Disparities Research Environment and Multi-omics Lab research group and the national Health Excellence Action Network. These include indices measuring socioeconomic status (ie, neighborhood socioeconomic status [nSES] [5], the neighborhood deprivation index [42], and racialized income segregation [also known as the Index of Concentration at the Extremes] [43]), indices measuring racial or ethnic segregation (ie, the local exposure and isolation index [LEx/Is] [44] and the location quotient [LQ]) [44], and indices measuring ethnic enclave status (ie, Asian American enclave and Hispanic enclave) [45,46]. These census tract-level indices are generated using ACS data and are presented as quintiles. Quintiles of nSES and ethnic enclave status are state-specific, based on the distribution of index scores in each state. All other indices are scaled to the United States.

Historic Redlining

Historic redlining data were added to Health Atlas through a partnership with the UCSF Environmental Research and Translation for Health Center to highlight the effects of structural racism on neighborhoods tracing back to discriminatory home loan practices implemented by the Home Owners' Loan Corporation (HOLC) in the 1930s. The HOLC created a ranking system for neighborhoods, now known as redlining, that categorized them by grade from A ("Best") to D ("Hazardous") based on data, including the quality of housing, recent values of sales and rentals, and the racial and ethnic composition of a neighborhood's residents [29]. Health Atlas reports historical redlining using HOLC mortgage security risk maps that have been digitized by the Mapping Inequality Project [47] and overlaid with 2010 and 2020 Census tracts for selected US cities in ArcGIS by the University of Michigan, Institute for Social Research [48]. In Health Atlas, the color scheme from the original HOLC risk grades was implemented, where Green indicates Grade A, Blue indicates Grade B, Yellow indicates Grade C, and Red indicates Grade D.

Climate Variables

The Health Atlas also includes multiple derived variables related to climate change in the state of California from the University of California Center for Climate, Health and Equity. Variables were derived using component data from the EPA Air Quality System and the Gridded Surface Meteorological data [49,50]. These data were added to contextualize air pollution, heat, and precipitation at the neighborhood level to inform local-level climate adaptation strategies. Example variables include average wildfire particulate matter 2.5, average heat index, number of days

>95th percentile for maximum temperature, or maximum precipitation.

Data Infrastructure

Data are partitioned by geographic level, census year, and variable and stored in Amazon Web Services S3 buckets, a public cloud storage resource [51]. This type of data organization (commonly called "hive partitioning") is performed by using the DuckDB [52] and Polars [53] software packages. Storing the data statically reduces dependence on the server, making the platform more resilient.

Maps are generated using Mapbox [54] maps, which include data from OpenStreetMap [55]. Data descriptions are updated in Airtable [56], a cloud-based data storage platform, and mapped to the public-facing website. Metadata about each variable are loaded from Airtable into Health Atlas. When a user selects a variable and geography, the platform pulls only the relevant data from the S3 bucket and can load the values for the selection independently.

We aimed to meet Web Content Accessibility Guidelines level AA compliance for the interface with high text-to-background contrast ratios, minimum text sizes, and minimal decorative text. Maps and data visualizations are not strictly accounted for under Web Content Accessibility Guidelines, but we tested common points of concern, such as colorblind-friendly color palettes and figure-to-background ratios, to improve legibility during the design process.

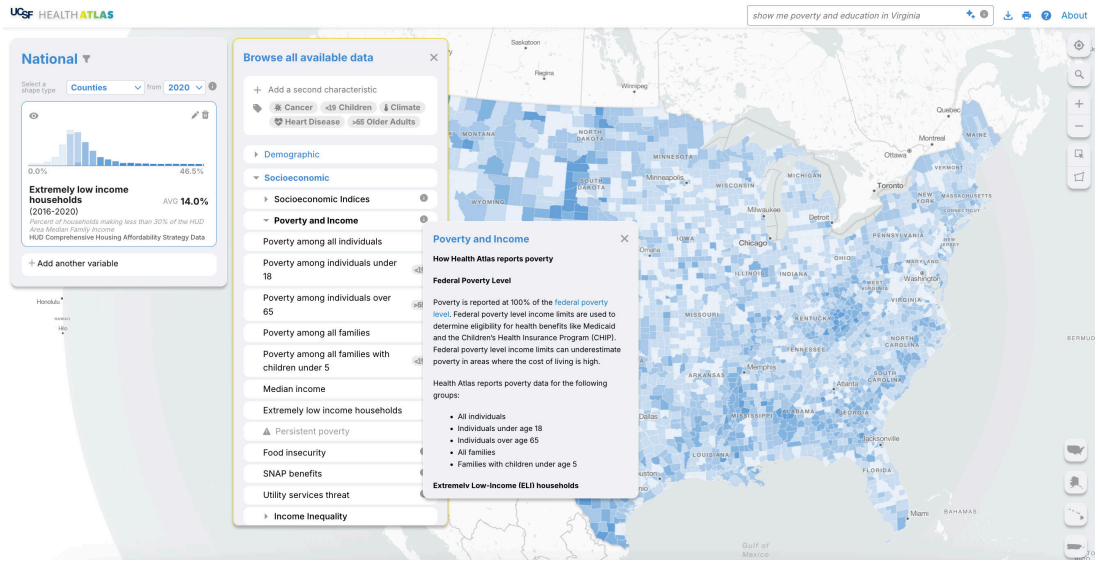
An AI-assisted search function is integrated into Health Atlas to help users select their data of interest more efficiently. Suggestions are provided based on natural language processing via large language models and are limited to variables available in the Health Atlas.

Tutorial

Selecting a Variable

To select a variable, users can click on the name of the currently visible variable in the left-side panel. Another panel will pop up to browse all available data grouped by domain (demographic, socioeconomic, neighborhood, environment, or health and health care) and by tags (cancer, children, climate, heart disease, or older adults). If a variable is not available for a certain geography, it will be grayed out. If users attempt to select a geography or variable combination that is not valid, a pop-up panel will appear to provide an option to switch to the appropriate geographic boundary. Information on each variable, including source year, can be found in data descriptions by clicking the "i" icon next to the variable name. Figure 1 displays a pop-up window with the data description for Poverty and Income variables.

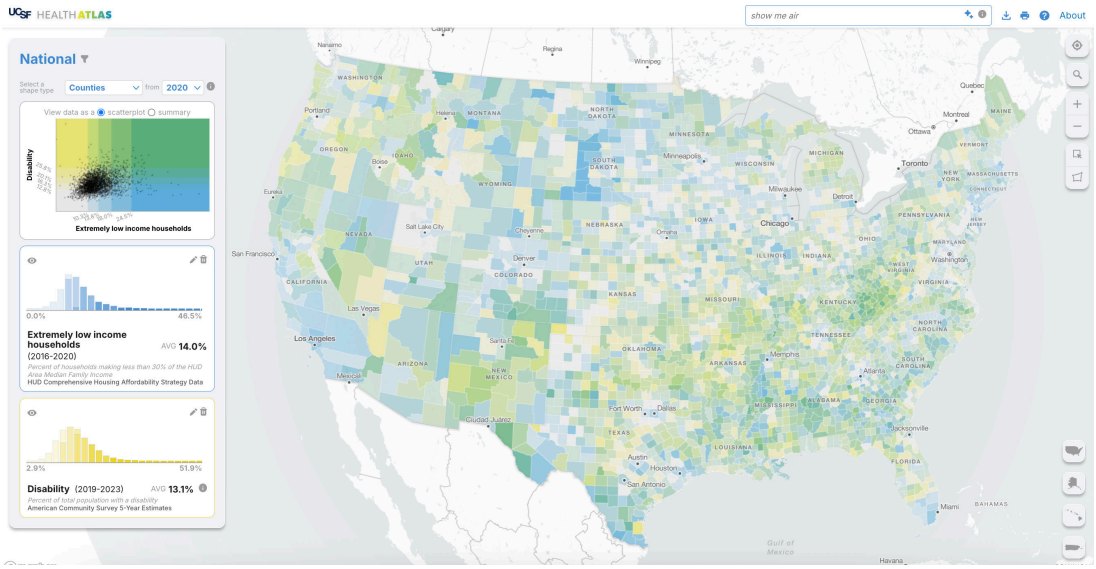
Figure 1. Health Atlas map showing the percentage of households that are extremely low-income across the United States at the 2020 Census tract-level. The selection panel shows a data description of Poverty and Income variables.



Visualizing Data With Maps

Data are visualized in univariate and bivariate choropleth maps. Figure 2 displays the Health Atlas with 2 variables selected for 2020 census tracts: the percent of households that are extremely low-income and the percent of people with a disability. The color of the first variable (extremely low-income) ranges from light to dark blue, while the second variable (disability) ranges from light to dark yellow, and the data with the highest values for both are depicted in green. When zoomed out to approximately the continental United States, maps are displayed in an Albers projection; when zoomed into a state or county-level scale, maps are displayed in a Mercator projection. With inset maps in the bottom right corner of the main map, users can elect to view the contiguous United States, Alaska, Hawai'i, or Puerto Rico. Users can easily navigate to noncontiguous areas (and back) via the insets.

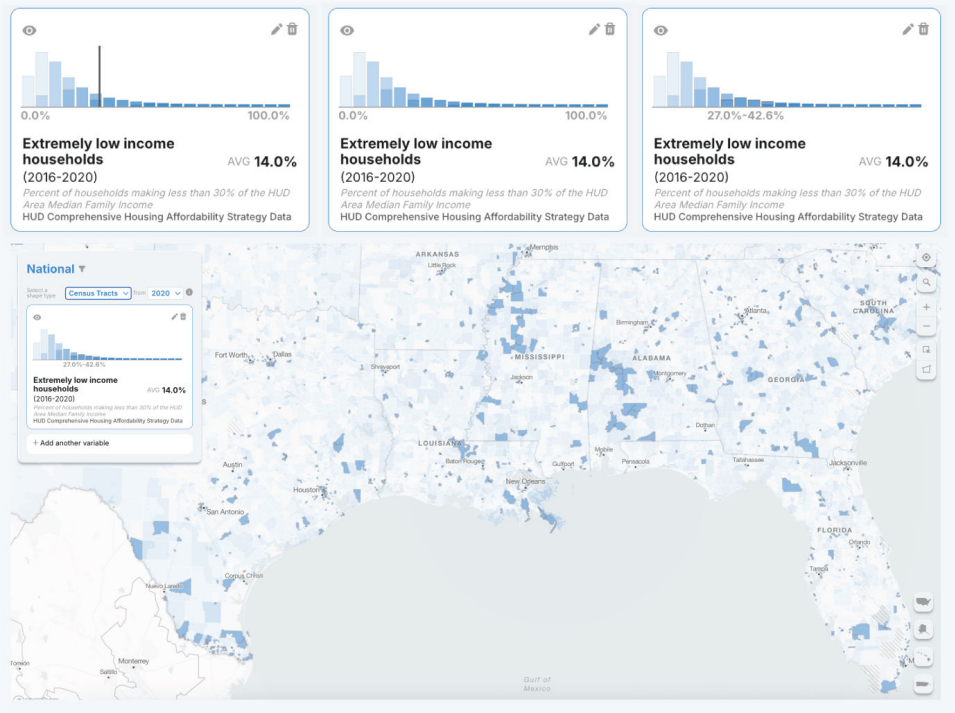
Figure 2. Health Atlas map of the percentage of households that are extremely low-income and the percentage of people with a disability across the United States at the 2020 Census tract-level.



Using Histograms

Histograms act as a map legend and allow users to identify areas with similar values of the chosen variable; mousing over a geographic area shows where that unit lies in the overall distribution of the histogram by displaying a vertical line corresponding to variable values for that area (Figure 3). Histograms display data distributions using 20 bins for continuous variables, 5 bars for quintiles, 4 bars for some categorical variables (eg, rural or urban and historical redlining), and 2 bars for binary values. The map in Figure 3 shows how mousing over a choropleth bin in the histogram highlights geographic areas with values falling in that bin.

Figure 3. Three different states of a histogram on Health Atlas visualizing the distribution of census tracts according to the percentage of households that are extremely low-income. HUD: US Department of Housing and Urban Development.



The bins of the histograms use equal intervals, but the colors of the bars follow natural breaks, which allow users to get a clear sense of the contours of the data. Natural breaks partition the data based on gaps in the data distribution [57] and are particularly effective for mapping variables with nonuniform distributions, which applies to most Health Atlas data from ACS, CDC PLACES, and other public sources. Choropleth bins are determined using *ckmeans* from the simple-statistics JS library by Tom MacWright [58]. A few Health Atlas variables are divided into equal intervals since they are displayed as quintiles (eg, nSES, ethnic enclaves) or scored as relative measures and ranked between 0 and 1 or 0

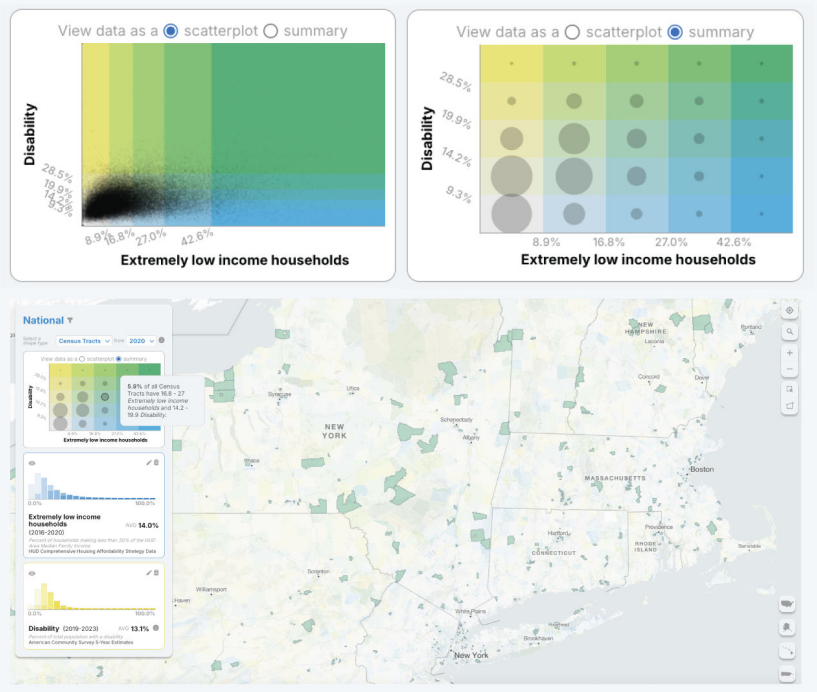
and 100 (eg, CDC Environmental Justice Index and Federal Emergency Management Agency National Risk Index, respectively).

Using Scatterplots and Summary Plots

Like histograms, scatterplots serve as a map legend and a way to see the texture of the data. Scatterplots display how 2 variables relate to each other, with each dot representing a geographic unit (eg, each state, county, etc, depending on the selected level of aggregation). Scatterplots can help users descriptively visualize whether there is a linear, nonlinear, or null relationship between 2 selected variables on the map. Mousing over a geographic unit on the map outlines the point corresponding to that area on the scatterplot, so users can see the bivariate relationship of 2 selected variables for that geographic area.

The summary view is helpful for geographies such as census tracts, which generate thousands of points on the scatterplot. In addition, mousing over a bubble on the summary plot allows users to identify geographic areas that are similar in the relationship between 2 variables. Figure 4 shows that mousing over the center bubble on the summary plot highlights on the map the 5.9% of census tracts across the United States where 16.8% to 27% of households have extremely low income and 14.2% to 19.9% of residents have a disability.

Figure 4. Bivariate distributions of extremely low-income households and disability across the United States at the 2020 Census tract level in the Health Atlas, using a scatterplot and summary plot.

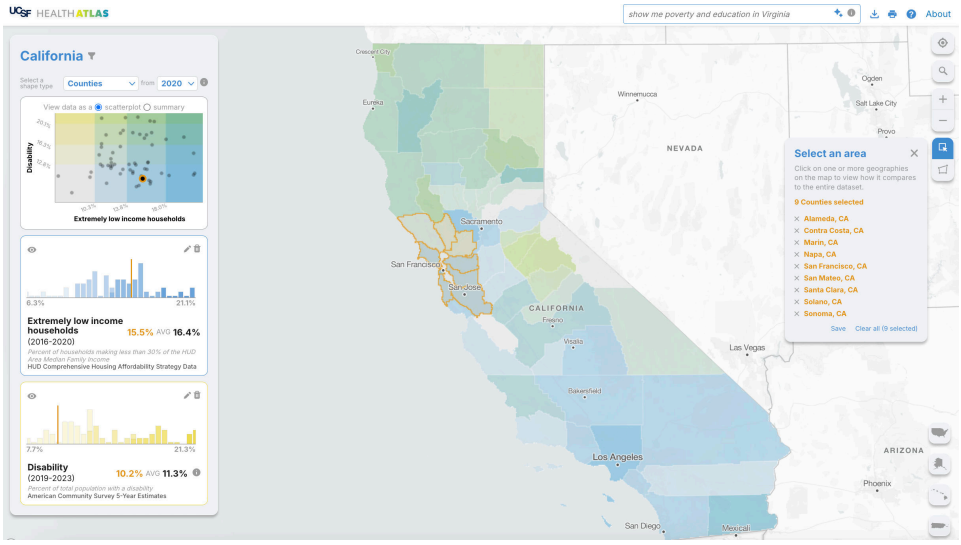


Custom Area Selection

Users can select a custom area by using the "Select an area" or "Draw a shape" tools on the right side of the map. Once a custom area has been selected, weighted average values for the selected area are shown in orange in the histogram or histograms, scatterplot, and summary visualization. Figure 5 shows the selection of 9 San Francisco Bay Area counties outlined in orange on a map filtered for the state of California. The relative position of the bivariate relationship of the selected variables in the selected area is indicated with an orange point on the scatter and summary plots.

Raw data for population denominators at the census tract level are used to generate weighted averages for custom selection areas. For example, the value for "under 5 years old" for a joint selection of census tracts is calculated by summing each census tract value and dividing it by the total population. Meanwhile, the value for "poverty among individuals over 65" is calculated by summing each census tract value within the selected tracts and dividing it by the population size over 65. Some variables, such as indices or climate data, are not aggregated based on a population denominator. The Health Atlas data dictionary describes which population denominators, if applicable, are used to calculate weighted averages (Multimedia Appendix 1).

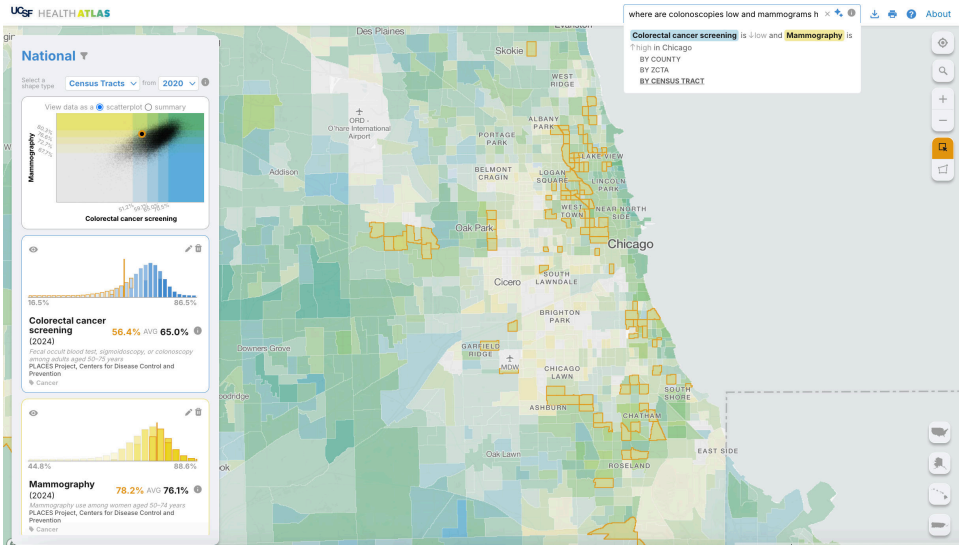
Figure 5. Health Atlas custom selection of 9 San Francisco Bay Area counties on a map filtered for the state of California showing the percentage of households that are extremely low-income and the percentage of residents that have a disability at the county level.



AI-Assisted Search

Users can use the AI-assisted search function at the top of the website to guide variable selection. The search bar can process queries (eg, "show me poverty and education in Virginia") and return variable and geography selections. The AI-assisted search function can also help users to zoom to specific areas of the map, filter to specific states, view variable suggestions, select available geographic boundaries, and identify regions with high or low values (Figure 6).

Figure 6. Health Atlas map of AI–assisted search identifying a map based on the query "Where are colonoscopies low and mammograms high in Chicago?"

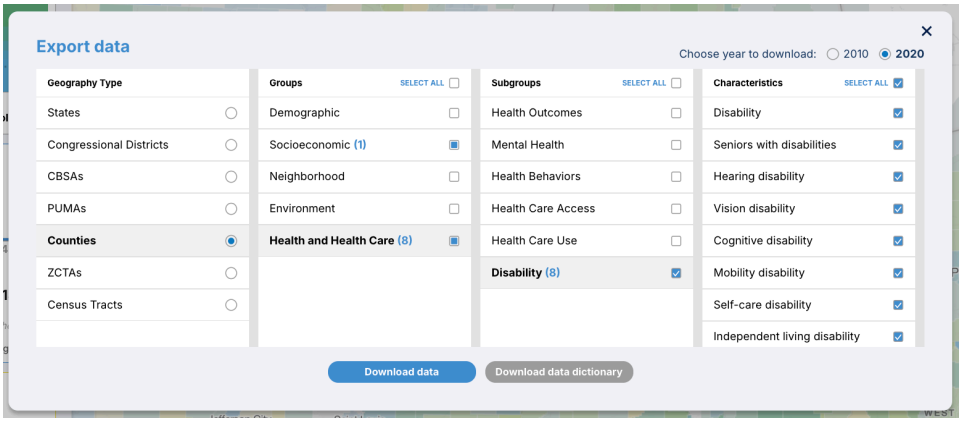


Data Export

Users can download Health Atlas data by clicking the download icon at the top right corner of the website. Data downloads will reflect the current parameters on the map and can be adjusted before download. Custom data exports allow users to download all data, or alternatively, data for a selected area and selected variables. Figure 7 shows an example of 1 user's specifications to download data for a custom area selection for 10 socioeconomic variables and 8 health and health care (disability) variables at the county-level. The full data dictionary is also available for download.

Data are provided with geographic identifiers, facilitating the merging of exported data with study-specific datasets. A data dictionary is also available for download and includes the following information for each variable: name, brief description, categorization, units (eg, percent or quintile), data source, denominator for aggregation (if applicable), and available geographic units (Multimedia Appendices 1 and 2). Maps and visualizations can be exported as .png images for integration into presentations, reports, or proposals.

Figure 7. Health Atlas custom data download for socioeconomic variables and health and health care variables at the county-level. CBSA: core-based statistical area; ZCTA: zip code tabulation area.



Discussion

Health Atlas is a publicly available resource with over 200 place-based variables that can support a wide array of users interested in public health. Researchers can download data for analysis and maps for presentations or reports. Policymakers can gain insights into custom regions of interest. Community members can visualize local health and environmental factors for advocacy. Educators and learners can use mapping features to explore a variety of topics. By including national-level data, Health Atlas expands opportunities for collaboration by facilitating multistate comparisons and partnerships. Furthermore, it provides access to data that have not been readily available previously, such as segregation and ethnic enclave measures.

Health Atlas is aligned with contemporary conceptual frameworks that emphasize that health inequities largely stem from structural and social determinants of health across multiple levels of influence, including place [1,59]. Data hosted on Health Atlas, therefore, include measures of structural and social determinants of health across domains, such as the built and social environment, environmental and climate stressors, and health care access and use.

Health Atlas provides place-based variables across multiple domains free of cost to the user, curated from publicly available sources and/or research groups that collectively have not been previously collated or easily accessed (Table 1) [16,17,32]. It also provides comprehensive data downloads and data visualizations across domains and geographic units (eg, census tract, county, state). Providing granular levels of data, such as census tracts, is key to meaningfully capturing neighborhood-level exposures. Providing data at multiple levels, including nonnested geographic units such as congressional districts and public use microdata areas, provides users the flexibility to use the data for different applications such as advocacy, policy, and administration. Other resources, such as PolicyMap [60] and SocialExplorer [61], include a wider set of similar data but require institutional licenses for access to the data, whereas Health Atlas data are publicly available. Tools such as ArcGIS [62] can also provide maps with more sophisticated features, including spatial analysis or the ability to layer study data in a secure environment; however, access to these tools typically requires licenses. While some licensed access additionally allows for a platform in which users can upload and incorporate their own data into maps, Health Atlas allows users to download a customizable dataset to append to their data.

The ability to gain insights from data is key and should not be limited to those with advanced training. Health Atlas provides easy-to-access maps, histograms, scatterplots, and summary plots to allow users to better understand the distribution of the data. Data descriptions and links to original sources are included for each variable. Through partnership development, Health Atlas makes new measures available to the public (eg, segregation, climate). For these measures, more detailed descriptions of how variables were created are provided, with citations to relevant published articles.

Health Atlas has remained nimble in response to user input, new partnerships, and changing external circumstances. The site has evolved to include tags for filtering variables related to selected topics, additional variable domains and categories, and a function to easily export map images. The Health Atlas scientific team also has back-end capabilities to update data and data descriptions as needed.

Since Health Atlas was first released in April 2020, it has been used for a variety of public health applications. During the COVID-19 pandemic, Health Atlas allowed users to visualize COVID-19 data alongside place-based data to make evidence-based decisions for organizations and individuals [63]. It also helped researchers and policymakers better understand how different communities were being impacted by the pandemic, and how social factors, such as neighborhood-level English proficiency and disability, contributed to disparities in COVID-19 outcomes [64,65]. The University of California Davis Comprehensive Cancer Center is currently using Health Atlas to identify areas with low breast cancer screening to plan locations for their mobile mammogram bus program (J. Dang, personal communication, October 30, 2025).

Health Atlas has also been used to link place-based data to geocoded patient addresses from individual-level datasets or electronic health records, providing insight into multilevel

determinants of health. At UCSF, a project linking nSES from Health Atlas to electronic health records data found that older adults living in low SES neighborhoods had lower odds of advance care planning [66], a process by which people communicate their preferences for future medical care. Another project used nSES from Health Atlas to adjust for confounding at the neighborhood level and provide a clearer picture of the direct benefits of a digital health intervention among patients participating in cardiac rehabilitation [67]. Additional research used geospatial methods to plot hotspots of patients with serious mental illness and characterize neighborhoods based on racial or ethnic composition and socioeconomic status using data accessed through Health Atlas [68]. These examples of multilevel research studies demonstrate the intersection of individual and place-based factors and provide context and nuance for health professionals and policymakers to better address structural and social determinants of health.

While Health Atlas can help to quickly visualize geographic differences within the framework of an ecological study, users should avoid attributing place-based factors or associations to individuals residing within those places (ie, ecological fallacy) [69]. The interpretation of descriptive results should include consideration of the time periods over which measures were derived (source years) and possible confounding factors. Moreover, place-based data cannot replace local expertise on how residents interact with places and how these various attributes influence their health.

Data available through Health Atlas are pulled primarily from publicly available sources. If the reliability and availability of emerging data change, the tool may not accurately represent the population and the places where they reside. In some cases, data were not available for certain areas or years (eg, PLACES Health-Related Social Needs, which were not derived for all states by the CDC) not relevant for certain regions (eg, ethnic enclaves, which were only generated for states with sufficient representation of Asian American or Hispanic residents), or were not generated for all states due to the complexity of the data (eg, modeled climate data such as heat). Furthermore, the use of geographic boundaries such as census tracts can limit data presentation and analysis for Health Atlas [59]. While the availability of multiple geographic units allows for flexibility, the results of an analysis may be dependent on the size, shape, or orientation of the chosen geographic unit (ie, Modifiable Areal Unit Problem) [70].

Health Atlas is a scientifically robust data repository that represents an important contribution to the landscape of place-based data available to the public. We envision that the national Health Atlas will contribute to evidence-based community-based initiatives, impactful health equity research, and effective public health programs. Future directions for Health Atlas include further integration of AI, improved accessibility, incorporation of longitudinal data, and automated reports.

Acknowledgments

We thank the following individuals for their contribution to this project: Kirsten Bibbins-Domingo, Indira D'Souza, Jessica Fields, Jennifer Reihm (Pierce), Logan Williams, Alec Burch, Jeff Klein, Bernardo Loureiro, Eric Rodenbeck, Stuart Gansky, Kim Rhoads, Edwin Charlebois, Hyelee Kim, Nancy Cheng, Grace Ma, Peggy Reynolds, Sapna Thottathil, Sigal Maya, Tarik Benmarhnia, Sherry Weiser, and Arianne Teherani.

No generative AI tools were used at any stage in the preparation of this manuscript; however, AI is integrated directly into the Health Atlas tool. Figure 6 was generated by taking a screenshot of the Health Atlas website after typing "Where are colonoscopies low and mammograms high in Chicago?" into the AI-assisted search bar.

Funding

Initial project funding was provided by the University of California, San Francisco (UCSF) School of Medicine Dean's Office. Additional funding was provided by the UCSF Population Health Data Initiative (residual class settlement funds in the matter of April Krueger v. Wyeth, Inc., Case No. 03-cv-2496, US District Court, SD of Calif.); Health Excellence Action Network (U24MD017250); UCSF Environmental Research and Translation for Health Center (P30ES030284); UC Center for Climate, Health and Equity; Temple University Center for Asian Health (R01 MD017521); and Smiles of Hope (R56DE032675).

Data Availability

Data can be downloaded from UCSF Health Atlas [22].

Authors' Contributions

Conceptualization: DLO, KEK-M, SS-M, CL, SG, MCD
Data curation: DLO, KEK-M, DM, RD, SS-M, SG, MCD
Funding acquisition: SG, MJP, CL
Methodology: DLO, KEK-M, DM, EB, KT, RD, SS-M, DVD, CL, SG, MCD
Project administration: DLO, KEK-M, AN, DVD, MCD
Software: EB, KT
Supervision: DLO, DVD, MJP, MCD
Visualization: DLO, KEK-M, DM, EB, KT, MCD
Writing – original draft: DLO, KEK-M, AN, MCD
Writing – review & editing: DLO, KEK-M, DM, EB, KT, AN, RD, SS-M, CL, SG, MJP, MCD

Conflicts of Interest

None declared.

Multimedia Appendix 1

Health Atlas data dictionary variable categories, subcategories, names, denominators, units, and sources.
[XLSX File (Microsoft Excel File), 26 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Health Atlas data dictionary variable categories, subcategories, names, export availability, geographies, and Census years.
[XLSX File (Microsoft Excel File), 30 KB-Multimedia Appendix 2]

References

1. Closing the gap in a generation: health equity through action on the social determinants of health. World Health Organization; 2008. URL: https://iris.who.int/server/api/core/bitstreams/cb08095c-55c8-484e-bff6-0e9c78fd38dd/content [Accessed 2026-06-12]
2. Kershaw KN, Magnani JW, Diez Roux AV, et al. Neighborhoods and cardiovascular health: a scientific statement from the American Heart Association. Circ Cardiovasc Qual Outcomes. Jan 2024;17(1):e000124. [doi: 10.1161/HCQ.0000000000000124] [Medline: 38073532]
3. Chen X, Lee C, Huang H. Neighborhood built environment associated with cognition and dementia risk among older adults: a systematic literature review. Soc Sci Med. Jan 2022;292:114560. [doi: 10.1016/j.socscimed.2021.114560] [Medline: 34776284]
4. Bird CE, Seeman T, Escarce JJ, et al. Neighbourhood socioeconomic status and biological "wear and tear" in a nationally representative sample of US adults. J Epidemiol Community Health. Oct 2010;64(10):860-865. [doi: 10.1136/jech.2008.084814] [Medline: 19759056]
5. Gomez SL, Glaser SL, McClure LA, et al. The California Neighborhoods Data System: a new resource for examining the impact of neighborhood characteristics on cancer incidence and outcomes in populations. Cancer Causes Control. Apr 2011;22(4):631-647. [doi: 10.1007/s10552-011-9736-5] [Medline: 21318584]
6. Bosma H, van de Mheen HD, Borsboom GJ, Mackenbach JP. Neighborhood socioeconomic status and all-cause mortality. Am J Epidemiol. Feb 15, 2001;153(4):363-371. [doi: 10.1093/aje/153.4.363] [Medline: 11207154]
7. Chetty R, Hendren N. The impacts of neighborhoods on intergenerational mobility I: childhood exposure effects. Q J Econ. Aug 1, 2018;133(3):1107-1162. [doi: 10.1093/qje/qjy007]
8. Akinyemiju T, Jha M, Moore JX, Pisu M. Disparities in the prevalence of comorbidities among US adults by state Medicaid expansion status. Prev Med. Jul 2016;88:196-202. [doi: 10.1016/j.ypmed.2016.04.009] [Medline: 27095325]
9. Chishtie J, Bielska IA, Barrera A, et al. Interactive visualization applications in population health and health services research: systematic scoping review. J Med Internet Res. Feb 18, 2022;24(2):e27534. [doi: 10.2196/27534] [Medline: 35179499]
10. Dixon BE, Dearth S, Duszynski TJ, Grannis SJ. Dashboards are trendy, visible components of data management in public health: sustaining their use after the pandemic requires a broader view. Am J Public Health. Jun 2022;112(6):900-903. [doi: 10.2105/AJPH.2022.306849] [Medline: 35446601]
11. Thorpe LE, Gourevitch MN. Data dashboards for advancing health and equity: proving their promise? Am J Public Health. Jun 2022;112(6):889-892. [doi: 10.2105/AJPH.2022.306847] [Medline: 35446603]
12. PLACES: local data for better health. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/places/index.html [Accessed 2025-10-02]
13. National Risk Index. Federal Emergency Management Agency (FEMA). URL: https://hazards.fema.gov/nri/ [Accessed 2025-10-02]
14. EJI technical documentation. Agency for Toxic Substances and Disease Registry (ATSDR), Centers for Disease Control and Prevention (CDC). Mar 12, 2024. URL: https://www.atsdr.cdc.gov/place-health/php/eji/eji-technical-documentation.html [Accessed 2025-11-05]
15. Data catalog. US Department of Housing and Urban Development (HUD). URL: https://data.hud.gov/datasets [Accessed 2025-11-12]
16. DREAM Lab. URL: https://dreamlab.ucsf.edu/ [Accessed 2025-10-02]
17. Climate Health. URL: https://climatehealth.ucsf.edu/ [Accessed 2025-10-02]
18. Long JC, Delamater PL, Holmes GM. Which definition of rurality should I use?: The relative performance of 8 federal rural definitions in identifying rural-urban disparities Med Care. 2021;59(Suppl 5):S413-S419. [doi: 10.1097/MLR.0000000000001612]

19. Stone W, Huang P. Some federal health websites restored, others still down, after data purge. National Public Radio (NPR). Feb 6, 2025. URL: https://www.npr.org/sections/shots-health-news/2025/02/06/nx-s1-5288113/cdc-website-health-data-trump [Accessed 2025-10-02]
20. Cox C, Rae M, Kates J, Wager E, Ortaliza J, Dawson L. A look at federal health data taken offline. KFF. Feb 2, 2025. URL: https://www.kff.org/hiv-aids/a-look-at-federal-health-data-taken-offline/ [Accessed 2025-11-06]
21. Brady J. Far more environmental data is being deleted in Trump's second term than before. National Public Radio (NPR). Aug 8, 2025. URL: https://www.npr.org/2025/08/08/nx-s1-5495338/climate-change-environment-websites-trump [Accessed 2025-10-02]
22. UCSF Health Atlas. URL: https://healthatlas.ucsf.edu/?v%5B0%5D=pct_eli_all&v%5B1%5D=pct_disabled&f%5B0%5D=national&g=county&zoomgeo=true&y=2020&c%5B0%5D=-99.6126&c%5B1%5D=44.5714&z=2.81&s= [Accessed 2026-06-12]
23. Chiu LW. Population health and health equity is excited to announce the launch of new website. UCSF Population Health and Health Equity. Apr 3, 2020. URL: https://pophealth.ucsf.edu/news/population-health-and-health-equity-excited-announce-launch-new-website [Accessed 2025-10-02]
24. Colliver V. UCSF Health Atlas expands its reach with nationwide mapping tool. UCSF (University of California San Francisco). Oct 10, 2024. URL: https://www.ucsf.edu/news/2024/10/428646/ucsf-health-atlas-expands-its-reach-nationwide-mapping-tool [Accessed 2025-10-02]
25. Understanding geographic identifiers (GEOIDs). United States Census Bureau. URL: https://www.census.gov/programs-surveys/geography/guidance/geo-identifiers.html [Accessed 2025-06-05]
26. American Community Survey (ACS). United States Census Bureau. URL: https://www.census.gov/programs-surveys/acs.html [Accessed 2025-10-02]
27. EJSCREEN: environmental justice screening and mapping tool. United States Environmental Protection Agency (EPA). URL: https://19january2021snapshot.epa.gov/ejscreen [Accessed 2025-10-02]
28. American community survey 5-year data (2009-2023). United States Census Bureau. URL: https://www.census.gov/data/developers/data-sets/acs-5year.html [Accessed 2025-10-02]
29. Public Health Alliance. URL: https://www.thepublichealthalliance.org/ [Accessed 2025-10-02]
30. Institute for Equity in Child Opportunity & Healthy Development (IECOHD). Boston University, School of Social Work. URL: https://www.bu.edu/ssw/research/centers/iecohd/ [Accessed 2025-10-02]
31. Agency for Toxic Substances and Disease Registry (ATSDR). URL: https://www.atsdr.cdc.gov/index.html [Accessed 2025-10-02]
32. Health Excellence Action Network. URL: https://health-excellence-action.org/homepage [Accessed 2025-10-02]
33. National Cancer Institute. URL: https://www.cancer.gov/ [Accessed 2025-10-02]
34. Economic Research Service, US Department of Agriculture (USDA). URL: https://www.ers.usda.gov/ [Accessed 2025-10-02]
35. Consolidated Planning and Comprehensive Housing Affordability Strategy (CHAS) data. HUD USER. URL: https://www.huduser.gov/portal/datasets/cp.html [Accessed 2025-10-02]
36. Food Access Research Atlas. Economic Research Service, US Department of Agriculture (USDA). URL: https://www.ers.usda.gov/data-products/food-access-research-atlas [Accessed 2025-10-02]
37. National Neighborhood Data Archive (NaNDA). URL: https://nanda.isr.umich.edu/ [Accessed 2025-10-02]
38. United States Census Bureau. URL: https://www.census.gov/ [Accessed 2025-11-05]
39. University of Michigan, Institute for Social Research. URL: https://isr.umich.edu/ [Accessed 2025-10-02]
40. About the Public Health and Environmental Systems Division (PHESD). United States Environmental Protection Agency (EPA). URL: https://www.epa.gov/aboutepa/about-public-health-and-environmental-systems-division-phesd [Accessed 2025-10-02]
41. National Vital Statistics System (NVSS). National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC). 2025. URL: https://www.cdc.gov/nchs/nvss/index.htm [Accessed 2025-10-02]
42. NDI: Neighborhood Deprivation Index. The Comprehensive R Archive Network (CRAN). URL: https://cran.r-project.org/web/packages/ndi/index.html [Accessed 2026-06-12]
43. Krieger N, Waterman PD, Spasojevic J, Li W, Maduro G, Van Wye G. Public health monitoring of privilege and deprivation with the index of concentration at the extremes. Am J Public Health. Feb 2016;106(2):256-263. [doi: 10.2105/AJPH.2015.302955] [Medline: 26691119]
44. Bemanian A, Beyer KMM. Measures matter: the local exposure/isolation (LEx/Is) metrics and relationships between local-level segregation and breast cancer survival. Cancer Epidemiol Biomark Prev. Apr 1, 2017;26(4):516-524. [doi: 10.1158/1055-9965.EPI-16-0926]

45. Guan A, Pruitt SL, Henry KA, et al. Asian American enclaves and healthcare accessibility: an ecologic study across five states. Am J Prev Med. Dec 2023;65(6):1015-1025. [doi: 10.1016/j.amepre.2023.07.001] [Medline: 37429388]
46. Guan A, Shariff-Marco S, Henry KA, et al. Latino enclaves and healthcare accessibility: an ecologic study across five states. J Gen Intern Med. Mar 2025;40(4):739-748. [doi: 10.1007/s11606-024-08974-6] [Medline: 39285075]
47. Mapping inequality: redlining in new deal America. Digital Scholarship Lab, University of Richmond. URL: https://dsl.richmond.edu/panorama/redlining/ [Accessed 2025-10-01]
48. Historic redlining scores for 2010 and 2020 US census tracts. openICPSR. URL: https://www.openicpsr.org/openicpsr/project/141121/version/V2/view [Accessed 2026-06-12]
49. Aguilera R, Luo N, Basu R, et al. A novel ensemble-based statistical approach to estimate daily wildfire-specific $PM_{2.5}$ in California (2006-2020). Environ Int. Jan 2023;171:107719. [doi: 10.1016/j.envint.2022.107719] [Medline: 36592523]
50. GRIDMET: University of Idaho Gridded Surface Meteorological Dataset. Earth Engine Data Catalog. URL: https://developers.google.com/earth-engine/datasets/catalog/IDAHO_EPSCOR_GRIDMET [Accessed 2025-10-01]
51. Amazon S3. AWS (Amazon Web Services). URL: https://aws.amazon.com/s3/ [Accessed 2026-04-10]
52. DuckDB. URL: https://duckdb.org/ [Accessed 2025-10-02]
53. Polars. URL: https://www.pola.rs/ [Accessed 2025-10-02]
54. Mapbox. URL: https://www.mapbox.com [Accessed 2025-11-05]
55. OpenStreetMap. URL: https://www.openstreetmap.org/ [Accessed 2025-11-06]
56. AI app building. Airtable. URL: https://www.airtable.com/lp/ai-psu-plp [Accessed 2025-11-05]
57. Natural breaks classification. Esri (GIS Dictionary). URL: https://support.esri.com/en-us/gis-dictionary/natural-breaks-classification [Accessed 2025-10-02]
58. Simple Statistics. URL: https://simple-statistics.github.io/ [Accessed 2025-12-04]
59. Dasgupta N, Kapadia F. The future of the public health data dashboard. Am J Public Health. Jun 2022;112(6):886-888. [doi: 10.2105/AJPH.2022.306871] [Medline: 35613427]
60. PolicyMap. URL: https://www.policymap.com/ [Accessed 2026-04-16]
61. Social Explorer. URL: https://www.socialexplorer.com/my-home [Accessed 2026-04-16]
62. ArcGIS. URL: https://www.arcgis.com/index.html [Accessed 2026-04-16]
63. Walsh B. To aid coronavirus fight, the times releases database of California cases. Los Angeles Times. Apr 6, 2020. URL: https://www.latimes.com/california/story/2020-04-06/coronavirus-fight-la-times-releases-its-california-cases-database [Accessed 2025-10-02]
64. Oh DL, Kemper KE, Meltzer D, Canchola AJ, Bibbins-Domingo K, Lyles CR. Neighborhood-level COVID vaccination and booster disparities: a population-level analysis across California. SSM Popul Health. Jun 2023;22:101366. [doi: 10.1016/j.ssmph.2023.101366] [Medline: 36873265]
65. Oh DL, Meltzer D, Wang K, et al. Neighborhood factors associated with COVID-19 cases in California. J Racial Ethn Health Disparities. Dec 2023;10(6):2653-2662. [doi: 10.1007/s40615-022-01443-y] [Medline: 36376642]
66. Nouri S, Lyles CR, Rubinsky AD, et al. Evaluation of neighborhood socioeconomic characteristics and advance care planning among older adults. JAMA Netw Open. Dec 1, 2020;3(12):e2029063. [doi: 10.1001/jamanetworkopen.2020.29063] [Medline: 33301019]
67. Oclaman JM, Murray ML, Grandis DJ, Beatty AL. The association between mobile app use and change in functional capacity among cardiac rehabilitation participants: cohort study. JMIR Cardio. May 15, 2023;7(1):e44433. [doi: 10.2196/44433] [Medline: 37184917]
68. Iturralde E, Rubinsky AD, Nguyen KH, Anderson C, Lyles CR, Mangurian C. Serious mental illness, glycemic control, and neighborhood factors within an urban diabetes cohort. Schizophr Bull. Apr 30, 2024;50(3):653-662. [doi: 10.1093/schbul/sbad122] [Medline: 37597839]
69. Robinson WS. Ecological correlations and the behavior of individuals. Int J Epidemiol. Apr 2009;38(2):337-341. [doi: 10.1093/ije/dyn357] [Medline: 19179346]
70. Buzzelli M. Modifiable areal unit problem. In: Kobayashi A, editor. International Encyclopedia of Human Geography. Vol . 2019:2. 169-173. [doi: 10.1016/B978-0-08-102295-5.10406-8]

Abbreviations

ACS: American Community Survey
AI: artificial intelligence
CDC: US Centers for Disease Control and Prevention
EPA: US Environmental Protection Agency
HOLC: Home Owners' Loan Corporation
NCI: National Cancer Institute

nSES: neighborhood socioeconomic status
UCSF: University of California, San Francisco

Edited by Amaryllis Mavragani; peer-reviewed by Beatriz Manzor-Mitrzyk, Lei Xiong, Nien-Tzu Chang; submitted 05.Dec.2025; final revised version received 29.Apr.2026; accepted 03.Jun.2026; published 09.Jul.2026

Please cite as:
Oh DL, Kemper-McIsaac KE, Meltzer D, Brelsford E, Taylor K, Dixit DV, Nickerson A, Desai R, Shariff-Marco S, Lyles C, Gomez S, Pletcher MJ, DeRouen MC
Health Atlas: Tutorial of a Visualization Tool and Data Resource for Place-Based Social and Structural Determinants of Health
J Med Internet Res 2026;28:e89065
URL: https://www.jmir.org/2026/1/e89065
doi: 10.2196/89065

© Debora L Oh, Kathryn E Kemper-McIsaac, Dan Meltzer, Eric Brelsford, Kelsey Taylor, Dhananjay Vinay Dixit, Andrea Nickerson, Riya Desai, Salma Shariff-Marco, Courtney Lyles, Scarlett Gomez, Mark J Pletcher, Mindy C DeRouen. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 09.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.