Congressional District Health Dashboard Technical Document

Updated April 28, 2026

SECTION 1: Introduction

Document Mission

This document is written for individuals interested in the technical details of the Dashboard. It provides information on which data sources, sub-tables, variables, and formulas were used to operationalize all Dashboard metrics and explains the rationale for analytic decisions.

Users are invited to contact the Dashboard (info@CDhealthdashboard.org) with general feedback or questions not addressed below.

Metric Overview

The Dashboard presents measures in one of three different formats: percentage, rate, or index. The type of measure is determined by the data that are analyzed to derive each estimate. Most metrics are aggregated to the congressional district level from census tract or county data; measures are also available by demographic subgroups or at the tract level if the underlying data allow for such disaggregation. For more information years of data available for each metric, see the Data Dictionary, Appendix A.

Domain	Metric Name	Metric Description	Data Source	Tract Estimates	Demographic Subgroups	Geography Aggregation
Clinical Care	Dental Care	Percentage of adults who reported visiting a dentist in the past year	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
	Designated Primary Care Shortage Area	Percentage of population living in a designed primary care shortage area	Health Resources & Services Administration	T	Not Available	From Tract
	Medicaid Enrollment	Percentage of total population enrolled in Medicaid in the past quarter	American Community Survey and Centers for Medicare & Medicaid Services	F	Not Available	tract & state
	Prenatal Care	Percentage of births for which prenatal care began in the first trimester	Natality Data, National Vital Statistics System, National Center for Health Statistics	F	Race/Ethnicity	From County
	Routine Checkup, 18+	Percentage of adults who reported visiting a doctor for a routine checkup in the past year	PLACES Project, Centers for Disease Control and Behavioral Risk	T	Not Available	From Tract

			Factor Surveillance System			
	Uninsured	Percentage of population ≤64 years old without health insurance	American Community Survey, U.S. Census Bureau	T	Age, Sex, Race/Ethnicity	From Tract
Health Behavior	Binge Drinking	Percentage of adults who reported binge drinking in the past 30 days	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
	Physical Inactivity	Percentage of adults who reported no leisure-time physical activity in the past 30 days	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
	Smoking	Percentage of adults who reported current cigarette smoking	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
	Teen Births	Births per 1,000 females aged 15-19	Natality Data, National Vital Statistics System, National Center for Health Statistics	F	Race/Ethnicity	From County
Health Outcomes	Breast Cancer Deaths	Deaths due to breast cancer per 100,000 females	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	F	Race/Ethnicity	From County
	Cardiovascular Disease Deaths	Deaths due to cardiovascular disease per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for	F	Sex, Race/Ethnicity	From County

		Health Statistics			
Colorectal Cancer Deaths	Deaths due to colorectal cancer per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	F	Sex, Race/Ethnicity	From County
Diabetes	Percentage of adults who reported having diabetes	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
Firearm Homicides	Deaths due to firearm homicide per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	F	Sex, Race/Ethnicity	From County
Firearm Suicides	Deaths due to firearm suicide per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	F	Sex, Race/Ethnicity	From County
Frequent Mental Distress	Percentage of adults who reported experiencing ≥14 days of poor mental health in the past 30 days	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
Frequent Physical Distress	Percentage of adults who reported experiencing ≥14 days of poor physical health in the past 30 days	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
High Blood Pressure	Percentage of adults who reported having	PLACES Project, Centers for	T	Not Available	From Tract

	high blood pressure	Disease Control and Behavioral Risk Factor Surveillance System			
Independent Living Difficulty	Percentage of adults who reported difficulty doing errands alone because of a physical, mental, or emotional condition	American Community Survey, U.S. Census Bureau	T	Age	From Tract
Life Expectancy	Average life expectancy at birth, in years	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	F	Sex, Race/Ethnicity	From County
Life Expectancy - Census Tract-Level, 2015	Average life expectancy at birth, in years	U.S. Small-Area Life Expectancy Estimates Project (USALEEP)	T	Not Available	From Tract
Low Birthweight	Percentage of live births with low birthweight (<2500 grams)	Natality Data, National Vital Statistics System, National Center for Health Statistics	F	Race/Ethnicity	From County
Obesity	Percentage of adults with a body mass index (BMI) ≥30 kg/m², based on their reported weight and height	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
Opioid Overdose Deaths	Deaths due to opioid overdose per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	F	Sex, Race/Ethnicity	From County
Premature Deaths (All Causes)	Years of potential life lost before age 75	Multiple Cause of Death Data, National Vital	F	Sex, Race/Ethnicity	From County

		per 100,000 people	Statistics System, National Center for Health Statistics			
Physical Environment	Air Pollution - Ozone	Average daily maximum concentration (parts per billion) of ground-level ozone throughout a month	George Mason University	T	Not Available	From Tract
	Air Pollution - Particulate Matter	Average daily concentration ($\mu g/m^3$) of fine particulate matter (PM2.5) per cubic meter of air throughout a month	George Mason University	T	Not Available	From Tract
	Housing with Potential Lead Risk	Percentage of housing stock with potential elevated lead risk	American Community Survey, U.S. Census Bureau	T	Not Available	From Tract
	Lead Exposure Risk Index	Index (1-10) reflecting poverty-adjusted risk of housing-based lead exposure	American Community Survey, U.S. Census Bureau	T	Not Available	From Tract
Social and Economic Factors	Broadband Connection	Percentage of households with high speed broadband internet connection (cable, fiber optic, DSL)	American Community Survey, U.S. Census Bureau	T	Not Available	From Tract
	Children in Poverty	Percentage of children living in households ≤100% of the federal poverty level	American Community Survey, U.S. Census Bureau	T	Race/Ethnicity	From Tract
	Chronic Absenteeism	Percentage of public school students who miss 10% or more school days in an academic year. Note: this metric is at the congressional district (not school district) level.	National Center for Education Statistics, U.S. Department of Education	F	Sex, Race/Ethnicity	-999

Food Insecurity	Percentage of adults who reported that during the past year they did not have enough food and did not have money to get more	PLACES Project, Centers for Disease Control and Behavioral Risk Factor Surveillance System	T	Not Available	From Tract
High School Completion	Percentage of adults ≥25 years old with a high school diploma or equivalent, or higher degree	American Community Survey, U.S. Census Bureau	T	Age, Sex, Race/Ethnicity	From Tract
Income Inequality	Index (-100 to +100) reflecting households with income at the extremes of the national income distribution (the top or bottom 20%)	American Community Survey, U.S. Census Bureau	T	Not Available	From Tract
Neighborhood Racial/Ethnic Segregation	Index (0-100) reflecting the geographic clustering of racial/ethnic groups across the area	American Community Survey, U.S. Census Bureau	F	Not Available	From Tract
Racial/Ethnic Diversity	Index (0-100) reflecting how evenly distributed the population is across the racial/ethnic groups living in this area	American Community Survey, U.S. Census Bureau	T	Not Available	From Tract
Rent Burden	Percentage of households where ≥30% of income is spent on rent	American Community Survey, U.S. Census Bureau	T	Not Available	From Tract
SNAP Participation	Percentage of households participating in Supplemental Nutrition Assistance Program (SNAP) in the past quarter	American Community Survey and U.S. Department of Agriculture Food and Nutrition Service	F	Not Available	tract & state
Unemployment	Percentage of population ≥16 years old who are unemployed but seeking work	American Community Survey, U.S. Census Bureau	T	Age, Sex, Race/Ethnicity	From Tract

	Youth Not in Work or School	Percentage of youth aged 16-19 who are neither working nor in school, also known as "disconnected youth"	American Community Survey, U.S. Census Bureau	T	Sex	From Tract

Metric Selection Criteria

The following metric inclusion criteria were used to compile accurate, consistent, and comparable data across 5 overarching domains:

- Rigorous methods underlying the original data collection
- Data available to the Dashboard analytic team
- Evidence of importance and validity in academic literature
- Metrics that are amenable to intervention
- Time lag between the Dashboard release and data collection ≤ 5 years
- Updated regularly, preferably at least every 2 years
- Balanced across the 5 domains (clinical care, health behaviors, health outcomes, physical environment and social and economic factors)
- When possible:
    - Aligned with other existent population health reporting frameworks (e.g., County Health Rankings & Roadmaps, Vital Signs, Culture of Health)
    - Disaggregated by census tracts or demographics
    - Available for all desired geographies

Dashboard Team

Avalon Aragon, MPH	Program Coordinator
Jacqueline Betro, MPA	Assistant Director
Amelia Bohan	Program Coordinator
Samantha Breslin, MPA	Program Supervisor
Chelsea Brown	Grants Manager
Marc N. Gourevitch, MD, MPH	Primary Investigator
Yoomin Hwang	Senior Project Coordinator
Aminata Kaba	Research Analyst
Neil Kleiman, PhD	Co-Primary Investigator (City Policy/Partnerships)
Taylor Lampe, MPH	Senior Data Analyst
Yuruo Li, PhD, MPH	Lead Data Analyst
Liang, PharmD, MSPH	Data Analyst
Hedy Ludwig, MPH	Data Analyst
Isabel Nelson, MPH	Lead Data Analyst
Annie Robinson, MS	Senior Program Coordinator
Eileen Shea, MPH	Data Analyst
Ben Spoer, PhD, MPH	Program Director
Jay Stadelman, MPH	Data Analyst
Lorna E. Thorpe, PhD	Co-Primary Investigator (Methods)
Hannah Wade, MPH	Senior Population Health Policy Analyst

Updates to Technical Documentation

This technical document is updated iteratively as needed. The date of the most recent update is noted on its first page and footer.

Please see the Appendix for an outline of changes made to each version of this document.

Feedback or Errors

Users are encouraged to contact the Dashboard with comments or questions regarding www.congressionaldistricthealthdashboard.org and any documents available for download from it, including this Technical Document, at info@CDhealthdashboard.org.

Downloading Dashboard Data

Users should note that much of the data outlined in this document is available for free download at www.congressionaldistricthealthdashboard.org/data-access.

Users should consult the Data Dictionary, available at www.congressionaldistricthealthdashboard.org/data-access, for more detail.

Please contact the Dashboard at info@CDhealthdashboard.org with any questions or concerns.

Citing Dashboard Data and Technical Document

The Dashboard should be cited when the data or graphics are used, including in published presentations, articles, research, blogs, policy documents, and other print or digital media.

We encourage use of Dashboard data and visualizations, and suggest the following citation:

> Department of Population Health, NYU Langone Health. Congressional District Health Dashboard. www.congressionaldistricthealthdashboard.org. Accessed [INSERT DATE OF ACCESS].

To cite our Technical Document, we suggest the following:

> Dashboard Team. *Congressional District Health Dashboard Technical Document*. New York: Congressional District Health Dashboard; [YEAR]. Available at www.congressionaldistricthealthdashboard.org/technical-documentation. Accessed [INSERT DATE OF ACCESS].

SECTION 2: Congressional District Overview

Introduction to Congressional Districts

Congressional districts are regions designated by state governments intended to proportionally represent the state population in the House of Representatives, the lower house of Congress. Though the process of redistricting varies by state, all are required by United States federal law to redraw their district lines for the election immediately after each Decennial Census. Public Law (PL) 94-171 (December 1975) requires that the Census Bureau provides block level data to states within one year of the Census day, from which states will build their respective congressional districts.[1]

PL 94-171 also guides the reapportionment of seats (i.e. congressional districts) across states[1]. First, the national "ideal population size" for a district is calculated by dividing the recent Decennial Census national population by 435 (i.e. the number of congressional districts designated in the House of Representatives). Each state is mandated at least 1 district. From there, the number of districts in each state is determined by dividing state populations by the "ideal population size" and incorporating the "Equal Proportion Method" from the Census.[2]

The Dashboard currently provides metric data for the 119th Congress. The 119th Congress is the session of Congress elected November 2024 and inaugurated in January 2025.

Nonvoting Delegates

Metrics presented on the Congressional District Health Dashboard are intended to be national in scope, but unfortunately, for a supermajority of our metrics, data are unavailable for "nonvoting bodies" that send nonvoting delegates to Congress. For this reason, the Congressional District Health Dashboard has elected to not include these bodies at this time (see list below). One exception is the District of Columbia (DC), which is available in most national data sources, and therefore the Dashboard has elected to include the DC nonvoting district on the website.

These bodies include:

Current nonvoting bodies designating delegates to the House of Representatives, in alphabetical order:

- American Samoa
- The Cherokee Nation *(designate, awaiting confirmation)*
- Guam
- the Northern Mariana Islands
- Puerto Rico

Current nonvoting bodies not currently designating a delegate to the United States Congress:

- The Choctaw Nation

Note: the Congressional District Health Dashboard is exploring ways to expand data access for these districts. Please reach out to us if you have any suggestions.

SECTION 3: Analytic Decisions

Data Disclaimer

Estimates presented in the Dashboard are subject to the same limitations as those inherent in source datasets. We identify the most likely sources of bias as necessary for each metric, and users should consult the data sources to understand potential biases more fully.

Aggregating to Congressional District Estimates: Methods and Approaches

The Congressional District Health Dashboard calculates most congressional district estimates by aggregating census tract or county level estimates ("source geographies") to the congressional district level, as these are the geographies for which data are most widely available and nationally comprehensive. You can see which metrics are derived from which geography under each metric-specific section, and in the metric table in the introduction. We use census tract data whenever possible, as they are smaller geographies and better nest within congressional districts. This makes tracts more likely to generate accurate congressional estimates, especially for smaller demographic subgroups.

Our method is conceptually similar to a dasymetric approach,[3] in which population distributions from underlying geographies (in this case, census blocks) are used to derive population-weighted estimates by aggregating from the source geography.[4,5]

Creating Population Crosswalks

We acquire 119th congress block-CD equivalency files from the Census [6], and access 2020 block population counts from tables in the 2020 Census Demographic and Housing Characteristics File (DHC) using the variables below. We combine these population counts with the block-CD equivalency files and sum to create tract-CD and county-CD population crosswalks. These population crosswalks represent the population count of each source geography (tract or county) that overlaps with different congressional districts for various weights of interest.

Population Weight	**Variable(s)**	**Variable Definition(s)**
Adult (18+)	P10_001N	*Total Population 18 Years and Over*
Asian (Hispanic-inclusive)	P12D_001N P12E_001N	*Asian alone* *Native Hawaiian and Other Pacific Islander alone*
Asian (non-Hispanic)	P9_008N P9_009N	*Asian alone, not Hispanic or Latino* *Native Hawaiian and Other Pacific Islander alone, not Hispanic or Latino*
Black (Hispanic-inclusive)	P12B_001N	*Black or African American alone*
Black (non-Hispanic)	P9_006N	*Black or African American alone, not Hispanic or Latino*
Hispanic	P9_002N	*Hispanic or Latino*
Housing Units	H1_001N	*Total Housing Units*
Other (Hispanic-inclusive)	P12C_001N P12F_001N P12G_001N	*American Indian and Alaska Native alone* *Some Other Race alone* *Two or more races*
Other (non-Hispanic)	P9_007N P9_010N P9_011N	*American Indian and Alaska Native alone, not Hispanic or Latino* *Some Other Race alone, not Hispanic or Latino* *Two or more races, not Hispanic or Latino*
Renter Households	H4_004N	*Renter Occupied Housing Units*
Total	P9_001N	*Total Population*
Total Households	P16_001N	*Total Households*
White (non-Hispanic)	P9_005N	*White alone, not Hispanic or Latino*

Choosing Population Weights

When choosing which population weight(s) to use for each metric, the Dashboard attempts to match the underlying population sample defined by the data source. For example, measures assessed at the household-level (i.e. Broadband Connection) utilize a Total Households weight.

Within a metric, total population, sex, and age demographic subgroups use the same population weights. Race/ethnicity demographic subgroups use race-specific population weights, matching the Hispanic inclusion/exclusion definitions of the data source. All aggregated metrics and their associated weight(s) are documented in the following table. Please email info@CDhealthdashboard.org with any questions about our weight selection.

Metric Name	**Population Weight(s)**
Air Pollution - Ozone	Total
Air Pollution - Particulate Matter	Total
Binge Drinking	Adult (18+)
Breast Cancer Deaths	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Broadband Connection	Total Households
Cardiovascular Disease Deaths	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Children in Poverty	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Colorectal Cancer Deaths	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Dental Care	Adult (18+)
Designated Primary Care Shortage Area	Total
Diabetes	Adult (18+)
Firearm Homicides	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Firearm Suicides	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Food Insecurity	Adult (18+)
Frequent Mental Distress	Adult (18+)
Frequent Physical Distress	Adult (18+)
High Blood Pressure	Adult (18+)
High School Completion	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Housing with Potential Lead Risk	Housing Units
Income Inequality	Total Households
Independent Living Difficulty	Total
Lead Exposure Risk Index	Housing Units
Life Expectancy	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Low Birthweight	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Medicaid Enrollment	Total

Metric Name	Population Weight(s)
Neighborhood Racial/Ethnic Segregation	Hispanic, White (non-Hispanic), Black (non-Hispanic), Asian (non-Hispanic), Other (non-Hispanic)
Obesity	Adult (18+)
Opioid Overdose Deaths	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Physical Inactivity	Adult (18+)
Premature Deaths (All Causes)	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Prenatal Care	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Racial/Ethnic Diversity	Hispanic, White (non-Hispanic), Black (non-Hispanic), Asian (non-Hispanic), Other (non-Hispanic)
Rent Burden	Renter Households
Routine Checkup, 18+	Adult (18+)
SNAP Participation	Total Households
Smoking	Adult (18+)
Teen Births	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Unemployment	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Uninsured	Total, Hispanic, White (non-Hispanic), Black (Hispanic-inclusive), Asian (Hispanic-inclusive), Other (Hispanic-inclusive)
Youth Not in Work or School	Total

Population Crosswalk Modification for Estimates in 2010 Census Vintage

Some metrics presented on the Dashboard are only available in 2010 Census vintages for certain data periods. To create population crosswalks between 2010 census tracts or counties and 119th congressional districts (which are in 2020 vintages), we incorporated a 2010 to 2020 block interpolation weight obtained from IPUMS's National Historic Geographic Information Systems (NHGIS) geographic crosswalks.[7] This weight represents the expected proportion of the 2010 block's population and housing units located in each 2020 block. Metrics, data periods, and estimate types using this modified method are documented below.

Metric Name	Data Periods with 2010 Census Vintages	Applicable Estimate Types
Binge Drinking	2019, 2020, 2021	Total
Breast Cancer Deaths	2020, 2021	Race/Ethnicity
Cardiovascular Disease Deaths	2020, 2021	Sex, Race/Ethnicity
Colorectal Cancer Deaths	2020, 2021	Sex, Race/Ethnicity
Dental Care	2018, 2020	Total
Diabetes	2019, 2020, 2021	Total
Firearm Homicides	2020, 2021	Sex, Race/Ethnicity
Firearm Suicides	2020, 2021	Sex, Race/Ethnicity
Frequent Mental Distress	2019, 2020, 2021	Total

Metric Name	Data Periods with 2010 Census Vintages	Applicable Estimate Types
Frequent Physical Distress	2019, 2020, 2021	Total
High Blood Pressure	2017, 2019	Total
Life Expectancy	2019, 2020, 2021	Total (2019 only), Sex, Race/Ethnicity
Life Expectancy - Census Tract-Level, 2015	2015	Total
Low Birthweight	2020, 2021	Race/Ethnicity
Obesity	2019, 2020, 2021	Total
Opioid Overdose Deaths	2020, 2021	Sex, Race/Ethnicity
Physical Inactivity	2019, 2020, 2021	Total
Premature Deaths (All Causes)	2020, 2021	Sex, Race/Ethnicity
Prenatal Care	2020, 2021	Race/Ethnicity
Routine Checkup, 18+	2019, 2020, 2021	Total
Smoking	2019, 2020, 2021	Total
Teen Births	2020, 2021	Race/Ethnicity

Formula

To derive congressional district estimates, we assigned estimates from source geographies (tract or county) to their overlapping congressional districts using the aforementioned population crosswalks. We then created a population weight (*P)* for the appropriate demographic subgroup by dividing the overlapping population count by the full congressional district population count. Population counts from source geographies with missing estimates were dropped from the calculation.

We multiplied this population weight by the source geography estimate (tract or county), then summed all weighted estimates to calculate the final derived congressional district estimate (see Equation). This method is applied to rate and percentage metric calculations. Unless otherwise noted in metric-specific sections, the Dashboard calculates the full metric at the source geography and then aggregates to congressional districts. Some metrics require combining multiple variables to calculate the estimate. If one variable is missing (NA) then we exclude it from the source geography estimate calculation.

$$Est_{derived,\ CD} = \sum est_{source\ geo} * P_{(source\ geo\ population\ in\ CD\ |\ CD\ population)}$$

When using county source geographies, there are some instances in denser urban areas where the congressional district boundary is smaller than the county boundary. In these cases the population weight = 1, and the congressional district estimate is identical to the source geography (county) estimate.

Formula Modification for Count Data

The Dashboard sometimes calculates derived congressional district *count* estimates, which requires a modification to the population weight formula indicated above. The weight in this instance represents the proportion of the full source geography population (tract or county) contained in the area of that geography *overlapping* with the congressional district. This adjustment is to properly reflect the non-proportional count estimate.

Confidence Intervals for Aggregated Estimates

The Dashboard does not release confidence intervals (CIs) for any aggregated estimates, due to substantial imprecision in calculated margins of error (MOEs), which is induced in using a standard sum of squares approach.

Censoring/Flagging Aggregated Estimates with Missing Contributing Data

The Dashboard team censors or flags congressional district derived estimates that are missing sufficient contributing data, which can impact estimate accuracy. Please email info@CDhealthdashboard.org for more information about our censorship development process.

Criteria differ by source geography. County to congressional district estimates are censored and flagged with more stringent criteria because counties (as compared to tracts) overlap more poorly with each congressional district. Therefore, each missing county estimate has greater impact on the full congressional district estimate. For metric estimates that require combining multiple subgroup/component variables, populations for subgroup/component variables that are missing may not contribute to censoring or flagging.

Criteria for Censoring and Flagging Derived Estimates *% Population Missing From Source Geography Data*		
	Censor	**Flag**
Tract	>25%	> 10% and < 25%
County	>10%	> 0% and < 10%

Censored estimates are removed from the website and downloadable data. Flagged estimates are noted in "Data Tips" on the website, or in downloadable data. Users should consult the Downloadable Data Codebook, available at www.congressionaldistricthealthdashboard.org/data-access, for more detail.

Note that these criteria differ from the criteria for metrics calculated from the National Vital Statistics System. Please see that data source section for more details.

Validating Estimates

The Dashboard team completed extensive analyses to validate our methods and analytic decisions for deriving congressional district estimates. Please email us at info@CDhealthdashboard.org to learn more.

At-large District Codes + Estimates

Some states send only one representative to the House of Representatives because their populations do not meet the "ideal population size." These states are considered "at-large" districts. The Congressional District Health Dashboard made the decision to use "[state FIPS code]01" to designate these districts, as opposed to the Census Bureau's designation of "[state FIPS code]00", for internal consistency.

Because at-large districts share boundaries with their state, the Dashboard presents state estimates for these at-large districts. See "State + National Estimates" below for more information on calculating state estimates.

Census Tract Estimates

Census tract estimates are provided on the website and for download for select metrics. Demographic subgroup estimates are not provided for census tracts due to the small population counts. Users should

note that some census tract estimates may be unstable due to low population count and sampling bias. Interested users can access downloadable data and census tract confidence intervals (when available) to better assess estimate reliability. See www.congressionaldistricthealthdashboard.org/data-access for more detail.

Census Tract-CD Assignment

Census tracts in 2020 geographies were assigned to 119 congressional districts using the 2020 block-CD equivalency file detailed in section "Creating Block Equivalency Files". Census tracts in 2010 geographies were assigned to congressional districts by transforming 2010 blocks into 2020 blocks using the 2010 to 2020 block interpolation weights obtained from IPUMS's National Historic Geographic Information Systems (NHGIS) geographic crosswalks.[7] Then, the 2020 block-CD equivalency file was used.

The Census infrequently modifies Census tract IDs/FIPS codes between Decennial Censuses to fix errors and account for county geography changes.[8] To maintain consistency across the website, the Dashboard standardizes all data sources' FIPS codes back into original Decennial Census vintages. Interested users can email info@CDhealthdashboard.org for more information.

National + State Estimates

National estimates on the Dashboard represent the unweighted average of congressional district estimates by metric and year for the total population only. Estimates for identical or similar metrics that use the nation as a sampling frame may produce different estimates. Average estimates are calculated after censoring criteria are applied. See the section "Censoring/Flagging Estimates with Missing Contributing Data" for more details.

State estimates on the Dashboard may represent either:

- the unweighted average of congressional district estimates by metric and year for total population
- an aggregation of tract-level total population estimates, using the same method as congressional districts outlined in "SECTION 3: Analytic Decisions"
- total population estimates calculated directly from the data source using a state sampling frame

See "Geography-Specific Notes" under each metric section to learn more.

Confidence Intervals

Confidence intervals (CIs), also known as confidence limits, provide a measure of the variation around a given estimate of a population value. For consistency, this document exclusively uses the term confidence intervals.

Confidence intervals are provided in downloadable data for census tracts and states, where available. As noted in section "Confidence Intervals for Aggregated Estimates", confidence intervals are not calculated for aggregated estimates.

Confidence Interval Calculation

Dashboard CIs are reported at the 90% level. Ninety-five percent CIs are most commonly reported in the scientific literature. However, the Dashboard reports 90% CIs for a number of reasons. Most notably, the Census Bureau recommends calculation of 90% CIs when using American Community Survey data.[9] The Dashboard opted to construct 90% CIs from standard errors where necessary to ensure consistency

between measures. There are a number of formulas for deriving CIs; selection depends on properties of the underlying data source. See Section 4 below for specifics on the formula used.

Confidence intervals for percentages were manually restricted to minimum 0 and maximum 100 when raw values exceeded these bounds. As a rule, CIs were not calculated for the Dashboard's index values because indices reflect a weighted composite of measures that are then scaled, making CI calculation relatively complicated and less meaningful.

Metric Subgroup Race/Ethnicity Categories

Where possible, the Dashboard disaggregates metrics by the following demographic groups: Asian (Asian or Native Hawaiian or Pacific Islander (NHOPI)); Black/African American; Hispanic/Latino; white (not Hispanic or Latino); and other (some other race, 2 or more races, or American Indian/Alaska Native (AIAN)).[10] Federal guidelines for reporting data by demographics[10] mandate separate categories for AIAN and NHOPI. However, the geographic areas used to generate Dashboard estimates generally lack large enough populations for reporting stable estimates for these groups. The Dashboard therefore combines NHOPI with Asian and AIAN with "other race" and two or more races, as data availability allows. See the metric-specific sections for more details.

Health Snapshots - District Facts

Demographic Data

Congressional district demographic data on the Health Snapshots - District Facts pages are calculated by aggregating ACS tract-level 5-year estimates via the method described in the section "Getting to Congressional District level Data: Methods and Approaches." Variables selection was informed by consultations with demographic experts and community partners (see more in a blog from our city website). *Asian*, *American Indian and Alaska Native*, and *Native Hawaiian or Other Pacific Islander*. subgroups reflect persons who either identify as this race alone or in combination with other races or ethnicities, allowing for greater representation that aligns with group identities. This may result in totals that exceed 100%. See the below table for demographic group name and variable details.

Demographic Group	Variable(s)	Variable Definitions
Total population	DP05_0001	*Estimate!!SEX AND AGE!!Total population*
American Indian and Alaska Native	DP05_0085	*Estimate!!Race alone or in combination with one or more other races!!Total population!!American Indian and Alaska Native*
Asian	B02018_001	*Estimate!!Total Groups Tallied: (ASIAN ALONE OR IN ANY COMBINATION BY SELECTED GROUPS)*
Black or African American	DP05_0045	*Estimate!!RACE!!Total population!!One race!!Black or African American*
Hispanic	B03001_003	*Estimate!!Total:!!Hispanic or Latino:*
Native Hawaiian or Other Pacific Islander	B02019_001	*Estimate!!Total Groups Tallied: (NATIVE HAWAIIAN AND OTHER PACIFIC ISLANDER ALONE OR IN ANY COMBINATION BY SELECTED GROUPS)*
Other race alone	DP05_0074	*Estimate!!RACE!!Total population!!One race!!Some Other Race*
White, non-Hispanic	DP05_0096	*Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!White alone*
Two or more races	DP05_0035	*Estimate!!RACE!!Total population!!Two or More Races*
Age 0-17	DP05_0019	*Estimate!!SEX AND AGE!!Total population!!Under 18 years*

Age 18-64	DP05_0001 - DP05_0024 - DP05_0019	*See other rows*
Age 65+	DP05_0024	*Estimate!!SEX AND AGE!!Total population!!65 years and over*

District Density Scale

The District Density Scale on the Health Snapshots – District Facts page measures the extent to which a district is urban, suburban, or rural. You can learn more about the methods and technical decisions behind this feature on its technical overview page.

Website CD + Tract Maps

119 congressional district and 2020 Census tract website maps were created by combining the Dashboard-created 2020 national block-CD equivalency file with 2020 block spatial files. Block shapes were dissolved into congressional districts, or 2020 tracts in congressional districts, then ocean and great lake shorelines were removed using a 5m national cartographic map from the US Census.[11] For 2010 tracts in congressional districts, a spatial intersect and additional cleaning steps were performed between 2010 tract and 119 congressional district boundaries. Please email us at info@CDhealthdashboard.org to learn more.

Customized Areas

This feature allows users to combine census tract estimates to derive customized area data for any of the Dashboard's 30+ neighborhood-level metrics. You can access this feature on the Metric Maps page by clicking the "Customized Areas" tab under "Advanced Tools". From there, you can add tracts to the customized area by clicking on them and the Dashboard will immediately calculate a weighted average for the customized area. Please email info@CDhealthdashboard.org with any further questions.

Population-weighted estimates (i.e. "weighted averages") for customized areas are calculated using the same methods as detailed in the section "Getting to Congressional District level Data: Methods and Approaches."

The formula used is as follows. Population counts from census tracts with missing metric estimates are excluded from the population weight (P) calculation.

$$Est_{\,custom\ area} = \sum_{i=1}^{n} est_{census\ tract\ i} * P_{census\ tract\ i}$$

$$P_{census\ tract\ i} = \frac{census\ tract\ i\ population\ in\ custom\ area}{custom\ area\ population}$$

Where: n represents the number of tracts (*i*) within the custom area

Total population weights are used for all metrics. Variable selection depends on the Census vintage of the metric and data period.

Population Weight	**Decennial Census Vintage**	**File (Table)**	**Variable**

Total	2010	Summary File 1 (P9)	*P009001 – Hispanic or Latino, and Not Hispanic or Latino by Race!!Total*
	2020	Demographic and Housing Characteristics File (P9)	*P9_001N – Hispanic or Latino, and Not Hispanic or Latino by Race!!Total:*

Analytic Software

All analyses were performed in R using tidyverse, tidycensus, tigris, and sf packages, among others.[11-15]

SECTION 4: Metric Analyses, by Data Source

American Community Survey

General Notes

The American Community Survey (ACS) is administered by the US Census Bureau[9]. Data are retrieved from the census API using R and the tidycensus package.[16, 13] Variable labels from the API (e.g., Estimate; SEX AND AGE - Total population), not names (e.g., S2801_C01_017E), are outlined in metric sections.

Race/Ethnicity Definition

Tables ending in the following letters were used to calculate metrics by race/ethnicity

- Asian: Values in tables ending in D (Asian alone) and E (Native Hawaiian and other Pacific Islander alone) were summed
- Black/African American: Tables ending in B (Black or African American alone)
- Hispanic: Tables ending in I (Hispanic or Latino)
- Other: Values in tables ending in C (American Indian and Alaska Native alone), F (Some other race alone), and G (Two or more races) were summed
- White: Tables ending in H (White alone, not Hispanic or Latino)

Users should note that, unless specified otherwise, estimates for some demographic groups derived from ACS data are not mutually exclusive with estimates for Hispanic/Latino ethnicity. Thus, individuals represented in the following racial categories who also identify as Hispanic may also contribute to counts for the Hispanic demographic subgroup: Asian, Black, Native Hawaiian or Pacific Islander, two or more races, or some other race.

Confidence Interval Calculation

90% CIs for relevant ACS data were calculated according to the formula: estimate ± MOE. When confidence intervals extended less than 0 or greater than 100 for % metrics, these were set to 0 or 100, respectively. CIs are not calculated for indices.

When combining multiple ACS variables, approximated MOE's for summed count data and derived proportions/ratios in ACS data were calculated as per the US Census Bureau's publication.[17] The functions moe_sum, moe_prop, and moe_ratio from the tidycensus package were used.[13]

Relevant formulas are presented verbatim here for users' reference:

<u>Calculating MOE's for Summed Count Data[17] (p. A-14)</u>

$$MOE_{\text{aggregated count}} = \pm\sqrt{\sum_c MOE_c^2}$$, "where MOE_c is the MOE of the c^{th} component estimate"

Calculating MOE's for Derived Proportions[17] (p. A-14, A-15)

$$MOE_{derived\ proportion} = \pm \frac{\sqrt{MOE^2_{numerator} - (\hat{p}^2 * MOE^2_{denominator})}}{\hat{X}_{denominator}}$$

"where $MOE_{numerator}$ is the MOE of the numerator; $MOE_{denominator}$ is the MOE of the denominator; $\hat{p} = \frac{\hat{X}_{numerator}}{\hat{X}_{denominator}}$ is the derived proportion; $\hat{X}_{numerator}$ is the estimate used as the numerator of the derived proportion; $\hat{X}_{denominator}$ is the estimate used as the denominator of the derived proportion."

Note: Estimates with particularly large margins of error sometimes resulted in an incalculable value of $\sqrt{MOE^2_{numerator} - (\hat{p}^2 * MOE^2_{denominator})}$ because $MOE^2_{numerator} - (\hat{p}^2 * MOE^2_{denominator})$ resulted in a negative value. In these cases, per the Census' recommendation, the formula for derived ratios was used instead, which provides a conservative estimate of the MOE.

Calculating MOE's for Derived Ratios[17] (p. A-15)

$$MOE_{derived\ ratio} = \pm \frac{\sqrt{MOE^2_{numerator} + (\hat{R}^2 * MOE^2_{denominator})}}{\hat{X}_{denominator}}$$

Geography-Specific Notes

Census Tracts

Census tract estimates and confidence intervals (for non-index metrics) are calculated or provided for each metric using ACS data at the tract-level.

Congressional Districts

Unless otherwise specified in the metric section below, percent or index estimates are aggregated from the tract level to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on this method. Confidence intervals are not calculated.

States

State estimates and confidence intervals (for non-index metrics) are calculated or provided for each metric using ACS data at the state-level.

Broadband Connection

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of households with high speed broadband internet connection (cable, fiber optic, DSL)	American Community Survey	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

The following variable from data table S2801 was used to represent Broadband Connection:

- Estimate!!Percent!!Total households!!TYPE OF INTERNET SUBSCRIPTIONS!!With an Internet subscription:!!Broadband of any type!!Broadband such as cable, fiber optic or DSL

The associated margin of error variable was pulled to calculate confidence intervals.

Analysis

No additional analysis was conducted by the Dashboard.

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on calculation differences between Census tracts, congressional districts, and states.

Children in Poverty

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of children living in households ≤100% of the federal poverty level	American Community Survey	Yes	Race/Ethnicity	From Tract

Data Table(s) + Variable(s)

Data table B17020 and associated race/ethnicity-specific tables were used to calculate Children in Poverty. See above "Race/Ethnicity Definition" section for information on which tables are used for each subgroup.

The following variables in each data table were summed to calculate the numerator:

- Estimate!!Total!!Income in the past 12 months below poverty level!!Under 6 years
- Estimate!!Total!!Income in the past 12 months below poverty level!!6 to 11 years
- Estimate!!Total!!Income in the past 12 months below poverty level!!12 to 17 years

To calculate the denominator, the following variables in each data table were summed with the numerator variables:

- Estimate!!Total!!Income in the past 12 months at or above poverty level!!Under 6 years
- Estimate!!Total!!Income in the past 12 months at or above poverty level!!6 to 11 years
- Estimate!!Total!!Income in the past 12 months at or above poverty level!!12 to 17 years

For non-aggregated geographies, when any of the above variables used for summation were missing, the entire summed estimate was set to missing. Associated margins of error variables are used to calculate confidence intervals associated with these values.

Analysis

$$\text{Children in Poverty} = \frac{\text{Children age < 18 living in households below the poverty threshold}}{\text{Total number of children age < 18 living in households}} \times 100\%$$

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on calculation differences between Census tracts, congressional districts, and states.

High School Completion

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults ≥25 years with high school diploma or equivalent, or higher degree	American Community Survey	Yes	Age, Sex, Race/Ethnicity	From Tract

Data Table(s) + Variable(s)

Data table S1501 was used to calculate High School Completion for the total population and disaggregated by sex and age. Data tables C15002 were used to calculate High School Completion disaggregated by race/ethnicity. See above "American Community Survey: Race/Ethnicity Definition" section for information on which tables are used for each subgroup.

The following variables were used to represent estimates for total population and by sex and age:

- Estimate!!Percent!!Population 25 years and over!!High school graduate or higher
- Estimate!!Percent Male!!Population 25 years and over!!High school graduate or higher
- Estimate!!Percent Female!!Population 25 years and over!!High school graduate or higher
- Percent!!Estimate!!Population 25 to 34 years!!High school graduate or higher
- Percent!!Estimate!!Population 35 to 44 years!!High school graduate or higher
- Percent!!Estimate!!Population 45 to 64 years!!High school graduate or higher
- Percent!!Estimate!!Population 65 years and over!!High school graduate or higher

The following variables were summed to calculate the numerators for disaggregated race/ethnicity estimates:

- Estimate!!Total!!Male!!High school graduate (includes equivalency)
- Estimate!!Total!!Male!!Some college or associate's degree
- Estimate!!Total!!Male!!Bachelor's degree or higher
- Estimate!!Total!!Female!!High school graduate (includes equivalency)
- Estimate!!Total!!Female!!Some college or associate's degree
- Estimate!!Total!!Female!!Bachelor's degree or higher

The following variable was used to represent the denominator for disaggregated race/ethnicity estimates:

- Estimate!!Total

For non-aggregated geographies, when any of the above variables used for summation were missing, the entire summed estimate was set to missing. Associated margins of error variables are used to calculate confidence intervals associated with these values.

Analysis

$$\text{High School Completion} = \frac{\text{Residents aged 25 or older with high school diploma (or equivalent) or higher}}{\text{Total population aged 25 or older}} \times 100$$

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on calculation differences between Census tracts, congressional districts, and states.

Housing with Potential Lead Risk

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of housing stock with potential elevated lead risk	American Community Survey	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Data table B25034 was used to calculate Housing with Potential Lead Risk.

The following variables were used to categorize housing stock by age:

- Estimate!!Total:!!Built 1939 or earlier
- Estimate!!Total:!!Built 1940 to 1949
- Estimate!!Total:!!Built 1950 to 1959
- Estimate!!Total:!!Built 1960 to 1969
- Estimate!!Total:!!Built 1970 to 1979
- Estimate!!Total:!!Built 1980 to 1989
- Estimate!!Total:!!Built 1990 to 1999
- Estimate!!Total:!!Built 2000 to 2009
- Estimate!!Total:!!Built 2010 to 2013 *(data year 2020 only)*
- Estimate!!Total:!!Built 2014 or later *(data year 2020 only)*
- Estimate!!Total:!!Built 2010 to 2019 *(data year 2021-2024)*
- Estimate!!Total:!!Built 2020 or later *(data year 2021-2024)*

The following variable was used to represent total housing stock:

- Estimate!!Total

Analysis

The lead analysis was performed as per methodology initially developed by the Washington State Department of Health.[18] Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.[19] The Dashboard adapted Vox Media's Python code available on Github[20] for the present analysis, which was conducted by the Dashboard using R v4.1.0 and originally validated using Python v3.6.[21] The Washington State Department of Health's analysis uses variables from 2014 and research from 2002 to inform age of housing weights and lead dust hazard guidance. [18] In updating the analysis to represent all housing stock built in 2010 or later for years subsequent to 2014, variables were added for housing stock built using table B25034. See the above *Data Table(s) + Variable(s)* section. Age of housing weights were updated using the 2021 American Healthy Homes Survey II, published by the U.S. Department of Housing and Urban Development.[22] The updated weights table attached below reflect the most recent lead dust hazard guidance, which significantly lowers the acceptable threshold. This update complies with EPA guidance.[23]

Characteristic	% of Housing Units with Significant Lead-Based Paint Hazards
Before 1940	78.0
1940-1959	51.2
1960-1977	17.2
1978-2017	4.7

Housing with Potential Lead Risk is a Dashboard metric sub-analysis based on the Washington State Department of Health/Vox Media analysis intended to report the percentage of housing stock at risk for lead due to the age of the housing. Users can note that this value is the "housing_risk" variable in the original posted Python code.[20] We count the number of housing units in each of four time periods: pre-1940, 1940-59, 1960-79, 1980 or newer. The count of housing units in each time period is weighted by the likelihood of lead exposure due to building age (weights are extrapolated from American Healthy Homes Survey II, 2021. *Please note that available ACS building age variables do not perfectly align with suggested weights. Email* info@CDhealthdashboard.org *for further details.*). This results in an overall percent of housing likely to have some risk of lead exposure.

$$\text{Housing with Potential Lead Risk} = \frac{\text{Weighted sum of housing stock at risk for lead}}{\text{Total housing stock}} \text{ x } 100$$

For non-aggregated geographies, margins of error (MOE) for these estimate values were derived using the following protocol: calculating adjusted MOE's for each housing-age group that had summed estimates; weighting those MOE's with the same weights used to calculate the numerator; and then calculating an MOE for a derived proportion. See section "ACS: Confidence Interval Calculation" for this equation in full.

Geography-Specific Notes

Census Tracts

Census tract estimates and confidence intervals are calculated using ACS data at the tract-level. The Dashboard Team determined that estimates with a 3.25% or greater absolute increase from year to year were unstable and therefore are censored.

Congressional Districts

Estimates are aggregated from the tract level (after censorship is applied) to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on this method. Confidence intervals are not calculated

States

State estimates and confidence intervals are calculated using ACS data at the state-level.

Income Inequality

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Index (-100 to +100) reflecting households with income at the extremes of the national income distribution (the top or bottom 20%)	American Community Survey	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Data table B19001 was used to calculate Income Inequality.

The following variables were summed to calculate the number of households above the 80th percentile:

- Estimate!!Total!!$150,000 to $199,999
- Estimate!!Total!!$200,000 or more

The following variables were summed to calculate the number of households below the 20th percentile:

- Estimate!!Total!!Less than $10,000
- Estimate!!Total!!$10,000 to $14,999
- Estimate!!Total!!$15,000 to $19,999
- Estimate!!Total!!$20,000 to $24,999
- Estimate!!Total!!$25,000 to $29,999

The following variable was used as the total households with known income level:

- Estimate!!Total

Analysis

Income Inequality at the Extremes (ICE) was calculated as per Krieger and colleagues. [24]

The formula for ICE is as follows:

$$\text{ICE} = \frac{\text{Number of households in 80th income percentile - Number of Households in 20th income percentile}}{\text{Total households with known income level in geographic area}} \text{ x } 100$$

Where values of ICE range from -100 to 100.

As of August 2024, the Dashboard limited the presentation of income inequality data to 2022 after discovering temporal variation in this metric related to the metric definition and data availability. While incomes on average increase every year, shifting percentiles up, the income variables used in the calculation are categorical and are not able to be adjusted accordingly each year. Due to this, the Dashboard now only presents the year of data where there is the most alignment between the percentiles and the income variable cut points. The Dashboard concurrently transitioned to using 1- year estimates from ACS Table B19080 to define our percentiles rather than US Census Bureau Table H-1 (All Races).

20th Percentile Cut Point	**80th Percentile Cut Point**
<$29,999	>$150,000

See above “Geography-Specific Notes” section under “American Community Survey: General Notes” for information on calculation differences between Census tracts, congressional districts, and states.

Independent Living Difficulty

Metric Description	**Data Source**	**Tract Estimates**	**Demographic Subgroups**	**CD Geography Aggregation**
Percentage of adults who report difficulty doing errands alone because of a physical, mental, or emotional condition	American Community Survey	Yes	Age	From Tract

Data Table(s) + Variable(s)

Data table S1810 was used to calculate the Independent Living Difficulty metric.

The following variable was used to calculate Independent Living for total population:

- Estimate!!Percent with a disability!!Total civilian noninstitutionalized population!!DISABILITY TYPE BY DETAILED AGE!!With an independent living difficulty

The following variables were used to calculate Independent Living by age:

- Estimate!!Percent with a disability!!Total civilian noninstitutionalized population!!DISABILITY TYPE BY DETAILED AGE!!With an independent living difficulty!!Population 18 to 64 years
- Estimate!!Percent with a disability!!Total civilian noninstitutionalized population!!DISABILITY TYPE BY DETAILED AGE!!With an independent living difficulty!!Population 65 years and over

The associated margin of error variable was pulled to calculate confidence intervals

Analysis

$$\text{Independent Living Difficulty} = \frac{\text{People reporting difficulty living independently}}{\text{Total civilian non-institutional population}} \text{ x } 100\%$$

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on calculation differences between Census tracts, congressional districts, and states.

Lead Exposure Risk Index

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Index (1-10) reflecting poverty-adjusted risk of housing-based lead exposure	American Community Survey	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Data table B25034 was used to calculate housing risk. S1701 was used for calculating poverty risk.

The following variables were used to categorize housing stock by age:

- Estimate!!Total:!!Built 1939 or earlier
- Estimate!!Total:!!Built 1940 to 1949
- Estimate!!Total:!!Built 1950 to 1959
- Estimate!!Total:!!Built 1960 to 1969
- Estimate!!Total:!!Built 1970 to 1979
- Estimate!!Total:!!Built 1980 to 1989
- Estimate!!Total:!!Built 1990 to 1999
- Estimate!!Total:!!Built 2000 to 2009
- Estimate!!Total:!!Built 2010 to 2013 *(data year 2020 only)*
- Estimate!!Total:!!Built 2014 or later *(data year 2020 only)*

- Estimate!!Total:!!Built 2010 to 2019 *(data year 2021-2024)*
- Estimate!!Total:!!Built 2020 or later *(data year 2021-2024)*

The following variable was used to represent total housing stock:

- Estimate!!Total

The following variable was used to represent individuals living in poverty:

- Estimate!!Total!!Population for whom poverty status is determined!!All individuals with income below the following poverty ratios!!125 percent of poverty level

The following variable was used to represent total population for poverty risk calculations:

- Estimate!!Total!!Population for whom poverty status is determined

Analysis

We took the Dashboard Housing with Potential Lead Risk metric (see earlier section for details on weights and lead threshold guidance) and factored in information about the percentage of the population living at or below 125% of the federal poverty level (poverty risk). We z-standardized poverty risk and housing with potential lead risk variables, weighted each by weights extrapolated from American Healthy Homes Survey II, 2021[22], and summed these two components to get a raw lead risk score. We then ranked these scores from 1, or lowest risk, to 10, or highest risk, to create a scale of overall lead exposure risk.

$$\text{Housing risk} = \frac{\text{Weighted sum of housing stock at risk for lead}}{\text{Total housing stock}} \text{ x } 100$$

$$\text{Poverty risk} = \frac{\text{Population below 125\% of poverty level}}{\text{Total population}} \text{ x } 100$$

Raw lead risk score = weighted and z-scored housing risk + weighted and z-scored poverty risk

Lead Exposure Risk Index = decile ranked raw lead risk score

Geography-Specific Notes

Census Tracts

Decile index values are calculated at the tract-level. The Dashboard Team determined that estimates with a +/-2 decile change or greater from year to year were unstable and therefore are censored.

Congressional Districts

The raw lead risk score (pre-decile ranks) is calculated at the tract-level and these weighted estimates are then aggregated from tract to generate congressional district estimates. A decile index ranking is then generated for all congressional districts. See “SECTION 3: Analytic Decisions” for more details on geographic aggregation.

States

State level index values represent the unweighted average of congressional district estimates for that state. This choice was made to allow for more meaningful comparison between congressional districts and their corresponding states.

Neighborhood Racial/Ethnic Segregation

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Index (0-100) reflecting the geographic clustering of racial/ethnic groups across the area	American Community Survey	No	Not Available	From Tract

Data Table(s) + Variable(s)

Data table DP05 and the following variables were used to calculate racial/ethnic segregation using 5 race/ethnicity categories (see above "Race/Ethnicity Definition" section for details about combining categories):

- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Black or African American alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!American Indian and Alaska Native alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Asian alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Native Hawaiian and Other Pacific Islander alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Some other race alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Two or more races
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!White alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Hispanic or Latino (of any race)

Analysis

Segregation was quantified as per Iceland's formula for H, the entropy index.[25]

Iceland defines the entropy index as follows: "The entropy index is the weighted average deviation of each unit's entropy from the metropolitan-wide entropy, expressed as a fraction of the metropolitan area's total entropy."[25] The equation for H provides a raw value between 0-1. The segregation (entropy index) values that are presented on the Dashboard represent H*100 to provide segregation scores that range from 0 to 100.

Neighborhood Racial/Ethnic Segregation on the Dashboard is calculated using the following formula, adapted from the entropy index, where the smaller geography is tracts, and the larger geography is the geography presented on the website (i.e. cities or congressional districts):

$$\text{Neighborhood Racial/Ethnic Segregation} = \sum_{i=i}^{n} \frac{t_i(E\text{-}E_i)}{ET} \times 100$$

Where:

t_i refers to the total population of tract i
T is the larger geography's total population
n is the number of tracts

E is the larger geography's diversity (entropy) score
E_i is tract i's diversity (entropy) score

Iceland defines entropy scores for a given geography (or tract) as follows:

$$E \text{ (entropy/diversity)} = \sum_{r=1}^{r} (\pi_r) \ln[\frac{1}{\pi_r}]$$

Where:

$\boldsymbol{\pi_r}$ refers to a particular racial/ethnic group's proportion of the geography's population[25]

As per footnote 5 in Iceland, [25] $\ln\left[\frac{1}{\pi_r}\right]$ is set to 0 when the proportion of a particular group is in a given geography π_r is 0.

Geography-Specific Notes

Congressional Districts

First, each racial/ethnic group's tract-level proportions ($\boldsymbol{\pi_r}$) are calculated from tract-level DP05 population counts. Then, tract-level DP05 racial/ethnic group population counts aggregated to the congressional district-level and each racial/ethnic group's congressional district proportions ($\boldsymbol{\pi_r}$) are calculated. See "SECTION 3: Analytic Decisions" for more details on geographic aggregation.

These tract and congressional district proportion values ($\boldsymbol{\pi_r}$) are used to calculate census tract congressional district entropy/diversity scores (E_i and E), which are used in the segregation formula above. The total population of each tract (t_i) and total population of each congressional district (T) in the segregation formula are derived from counts from the 2020 Decennial Census P2 Table rather than aggregated DP05 counts, to account for the appropriate geographic overlap of each tract in a congressional district.

States

State level segregation estimates represent the unweighted average of congressional district segregation estimates for that state. This choice was made to allow for more meaningful comparison between congressional district scores and their corresponding state score.

Racial/Ethnic Diversity

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Index (0-100) reflecting how evenly distributed the population is across the racial/ethnic groups living in this area	American Community Survey	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Data table DP05 was used to calculate Racial/Ethnic Diversity values.

The following variables were used to calculate racial/ethnic diversity using 5 race/ethnicity categories (see above "American Community Survey: Race/Ethnicity Definition" for details about combining categories):

- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Black or African American alone

- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!American Indian and Alaska Native alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Asian alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Native Hawaiian and Other Pacific Islander alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Some other race alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!Two or more races
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Not Hispanic or Latino!!White alone
- Estimate!!HISPANIC OR LATINO AND RACE!!Total population!!Hispanic or Latino (of any race)

Analysis

Racial/Ethnic Diversity represents how much of the maximum possible entropy (or diversity) is exhibited in a given area. A lower value (closer to 0) indicates that all residents belong to one racial/ethnic group (low diversity) and a higher value (closer to 100) indicates that all racial/ethnic groups are in equal proportion (high diversity). This metric does not incorporate geographic distributions of racial/ethnic groups. Diversity (or entropy) was quantified using Iceland's formulas for entropy scores (see below). [25]

$$\text{Racial/Ethnic Diversity} = \frac{\text{Entropy score (E)}}{\text{Maximum possible entropy score}} \times 100$$

Where:

Maximum possible entropy score is ln(5), as there are 5 racial/ethnic groups in the calculation
E is the geography's diversity (entropy) score

Iceland defines entropy scores for a given geography (or tract) as follows:

$$\text{E (entropy/diversity)} = \sum_{r=1}^{r} (\pi_r) \ln\left[\frac{1}{\pi_r}\right]$$

Where:

$\boldsymbol{\pi_r}$ refers to a particular racial/ethnic group's proportion of the geography population [25]

As per footnote 5 in Iceland, [25] $\ln\left[\frac{1}{\pi_r}\right]$ is set to 0 when the proportion of a particular group is in a given geography π_r is 0.

Geography-Specific Notes

Census Tracts

Racial/Ethnic Diversity is calculated at the tract-level.

Congressional Districts

DP05 population counts for each racial/ethnic group are aggregated from the tract-level to generate congressional district population counts for each race/ethnic group. See "SECTION 3: Analytic Decisions" for more details on geographic aggregation. These counts are then used to calculate each racial/ethnic group's proportion of the congressional district ($\boldsymbol{\pi_r}$). This value is used in the formula above for congressional district diversity.

States

Racial/Ethnic Diversity is calculated at the state-level.

Rent Burden

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of households where ≥30% of income is spent on rent	American Community Survey	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Data table DP04 was used to calculate Rent Burden.

The following variables were summed to calculate the numerator:

- Estimate!!GROSS RENT AS A PERCENTAGE OF HOUSEHOLD INCOME (GRAPI)!!Occupied units paying rent (excluding units where GRAPI cannot be computed)!!30.0 to 34.9 percent
- Estimate!!GROSS RENT AS A PERCENTAGE OF HOUSEHOLD INCOME (GRAPI)!!Occupied units paying rent (excluding units where GRAPI cannot be computed)!!35.0 percent or more

The following variable was used to represent the denominator:

- Estimate!!GROSS RENT AS A PERCENTAGE OF HOUSEHOLD INCOME (GRAPI)!!Occupied units paying rent (excluding units where GRAPI cannot be computed)

For non-aggregated geographies, when any of the above variables used for summation were missing, the entire summed estimate was set to missing. Associated margins of error variables are used to calculate confidence intervals associated with these values.

$$\text{AnalysisRent Burden} = \frac{\text{Households for which rent} \geq 30\% \text{ of household income}}{\text{Total renter-occupied housing units with reported income}} \times 100\%$$

See above “Geography-Specific Notes” section under “American Community Survey: General Notes” for information on calculation differences between Census tracts, congressional districts, and states.

Unemployment

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of population ≥16 years who are unemployed but seeking work	American Community Survey	Yes	Age, Sex, Race/Ethnicity	From Tract

Data Table(s) + Variable(s)

Data table S2301 was used to report Unemployment total population and disaggregated by race/ethnicity, age, and sex.

The following variables were used to represent Unemployment for total population, age and sex breakdowns, and by race for White, Hispanic, and Black. Please note the different age category availability for sex-specific estimates:

- Estimate!!Unemployment rate!!Population 16 years and over
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Black or African American alone
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!White alone, not Hispanic or Latino
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Hispanic or Latino origin (of any race)
- Estimate!!Unemployment rate!!Population 20 to 64 years!!SEX!!Male
- Estimate!!Unemployment rate!!Population 20 to 64 years!!SEX!!Female
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!16 to 19 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!20 to 24 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!25 to 29 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!30 to 34 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!35 to 44 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!45 to 54 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!55 to 59 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!60 to 64 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!65 to 74 years

Unemployment by race for Asian and Other is represented by the weighted average of the following subgroup variables that comprise the full group.

Asian:

- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Asian alone
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Native Hawaiian and Other Pacific Islander alone

Other:

- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!American Indian and Alaska Native alone
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Some other race alone
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Two or more races

Estimate values are weighted by the relative proportion of each of these groups within the summed total population of these subgroups within each geographic area as per ACS table S2301, using variables with the following labels:

Asian:

- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Asian alone
- Estimate!!Unemployment rate!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Native Hawaiian and Other Pacific Islander alone

Other:

- Estimate!!Total!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!American Indian and Alaska Native alone
- Estimate!!Total!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Some other race alone
- Estimate!!Total!!Population 16 years and over!!RACE AND HISPANIC OR LATINO ORIGIN!!Two or more races

For non-aggregated metrics, when any of the above variables used for summation were missing, the entire summed estimate was set to missing. Where relevant, associated margins of error variables are used to calculate confidence intervals associated with these values.

Analysis

For all estimates except Asian and Other, no additional analysis was conducted by the Dashboard. The formula for combining racial subcategories for Asian and Other is as follows:

$$\text{Unemployment}_{\text{full group}} = \sum_{1=i}^{n} \left(\text{unemployment}_{\text{subcategory i}} * \frac{\text{population}_{\text{subcategory i}}}{\sum_{1=i}^{n} \text{population}_{\text{subcategory i}}} \right)$$

Where:

i = racial/ethnic subcategory contributing to the full racial/ethnic group

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on calculation differences between Census tracts, congressional districts, and states.

Uninsured

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of population ≤64 years without health insurance	American Community Survey	Yes	Age, Sex, Race/Ethnicity	From Tract

Data Table(s) + Variable(s)

Uninsured refers specifically to health insurance status, not lack of any type of insurance.

<u>Total population</u>

Data table S2701 was used to report percent of the civilian noninstitutionalized population without health insurance for ages 0-64; this stratum is referred to as "Total".

To calculate Uninsured, the following variables are summed from table S2701 to calculate the numerator:

- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!Under 19 years
- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!19 to 25 years
- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!26 to 34 years
- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!35 to 44 years
- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!45 to 54 years
- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!55 to 64 years

To calculated Uninsured, the following variables are summed from table S2701 to calculate the denominator:

- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!Under 19 years
- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!19 to 25 years
- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!26 to 34 years
- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!35 to 44 years
- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!45 to 54 years
- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!55 to 64 years

By age category

Data table S2701 was used to report percent of the civilian noninstitutionalized population without health insurance, disaggregated by age.

To calculate Uninsured by age category, the following variables are presented as reported in the S2701 data table:

- Estimate!!Percent Uninsured!!Civilian noninstitutionalized population!!AGE!!Under 19 years
- Estimate!!Percent Uninsured!!Civilian noninstitutionalized population!!AGE!!19 to 25 years
- Estimate!!Percent Uninsured!!Civilian noninstitutionalized population!!AGE!!26 to 34 years
- Estimate!!Percent Uninsured!!Civilian noninstitutionalized population!!AGE!!35 to 44 years

To calculated Uninsured age 45-64, the following variables are summed from table S2701 to calculate the numerator:

- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!45 to 54 years
- Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!55 to 64 years

To calculated Uninsured age 45-64, the following variables are summed from table S2701 to calculate the denominator:

- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!45 to 54 years
- Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!55 to 64 years

By sex

Data table B27001 was used to report uninsured, disaggregated by sex.

To calculated Uninsured by sex, the following variables from table B27001 are summed to calculate the numerator, where *[SEX]* = “Male” or “Female”:

- Estimate!!Total!![*SEX]*!!Under 6 years!!No health insurance coverage
- Estimate!!Total!![*SEX*]!!6 to 18 years!!No health insurance coverage
- Estimate!!Total!![*SEX*]!!19 to 25 years!!No health insurance coverage
- Estimate!!Total!![*SEX*]!!26 to 34 years!!No health insurance coverage
- Estimate!!Total!![*SEX*]!!35 to 44 years!!No health insurance coverage
- Estimate!!Total!![*SEX*]!!45 to 54 years!!No health insurance coverage
- Estimate!!Total!![*SEX*]!!55 to 64 years!!No health insurance coverage

To calculate Uninsured by sex, the following variables from table B27001 are summed to calculate the denominator:

- Estimate!!Total!![*SEX*]!!Under 6 years
- Estimate!!Total!![*SEX*]!!6 to 18 years
- Estimate!!Total!![*SEX*]!!19 to 25 years
- Estimate!!Total!![*SEX*]!!26 to 34 years
- Estimate!!Total!![*SEX*]!!35 to 44 years
- Estimate!!Total!![*SEX*]!!45 to 54 years

By race/ethnicity

Data tables C27001B, C27001C, C27001D, C27001E, C27001F, C27001H, and C27001I were used to calculate uninsured, disaggregated by race/ethnicity. See above "Race/Ethnicity Definition" section for information on which tables are used for each subgroup.

To calculated Uninsured by race/ethnicity, the following variables are summed from the C27001 series to calculate the numerator:

- Estimate!!Total!!Under 19 years!!No health insurance coverage
- Estimate!!Total!!19 to 64 years!!No health insurance coverage

To calculated Uninsured by race/ethnicity, the following variables are summed from the race/ethnicity-specific tables from the C27001 series to calculate the denominator:

- Estimate!!Total!!Under 19 years
- Estimate!!Total!!19 to 64 years

For non-aggregated geographies, when any of the above variables used for summation were missing, the entire summed estimate was set to missing. Where relevant, associated margins of error variables are used to calculate confidence intervals associated with these values.

Analysis

$$\text{Uninsured} = \frac{\text{Persons that have no current health insurance coverage}}{\text{Total population}} \text{ X } 100$$

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on calculation differences between Census tracts, congressional districts, and states.

Youth Not in Work or School

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of youth 16-19 years who are neither working nor in school, also known as "disconnected youth".	American Community Survey	Yes	Sex	From Tract

Data Table(s) + Variable(s)

Data table B14005 was used to report Youth Not in Work or School, disaggregated by sex.

Variables associated with the following labels were summed to calculate the numerator for males:

- Estimate!!Total:!!Male:!!Not enrolled in school:!!High school graduate (includes equivalency):!!Unemployed
- Estimate!!Total:!!Male:!!Not enrolled in school:!!High school graduate (includes equivalency):!!Not in labor force
- Estimate!!Total:!!Male:!!Not enrolled in school:!!Not high school graduate:!!Unemployed
- Estimate!!Total:!!Male:!!Not enrolled in school:!!Not high school graduate:!!Not in labor force

The following variable was used as the denominator for males:

- Estimate!!Total:!!Male:

Variables associated with the following labels were summed to calculate the numerator for females:

- Estimate!!Total:!!Female:!!Not enrolled in school:!!High school graduate (includes equivalency):!!Unemployed
- Estimate!!Total:!!Female:!!Not enrolled in school:!!High school graduate (includes equivalency):!!Not in labor force
- Estimate!!Total:!!Female:!!Not enrolled in school:!!Not high school graduate:!!Unemployed
- Estimate!!Total:!!Female:!!Not enrolled in school:!!Not high school graduate:!!Not in labor force

The following variable was used as the denominator for females:

- Estimate!!Total:!!Female:

Total population numerators and denominators were calculated by summing all the male and female variables.

When any of the above variables used for summation were missing, the entire summed estimate was set to missing. Associated margins of error variables are used to calculate confidence intervals for these values.

Analysis

$$\text{Youth Not in Work or School} = \frac{\text{Youth aged 16-19 years who are neither working nor in school}}{\text{Total youth aged 16-19}} \times 100$$

See above “Geography-Specific Notes” section under “American Community Survey: General Notes” for information on calculation differences between Census tracts, congressional districts, and states.

American Community Survey and Centers for Medicare & Medicaid Services

General Notes

The American Community Survey (ACS) is administered by the US Census Bureau.[9] Data are retrieved from the census API using R and the tidycensus package.[16, 13] Variable labels from the API (e.g., Estimate; SEX AND AGE - Total population), not names (e.g., S2801_C01_017E), are outlined in metric sections.

The Centers for Medicare & Medicaid Services (CMS) requires that all states report monthly data on enrollment within the state Medicaid program. These data are publicly available for download.[26] While we use data marked as "Final Report", CMS data may continue to be updated (e.g., states may continue to adjust their reporting methodologies to better align with CMS's data specifications),[27] potentially leading to estimates changing over time.

Medicaid Enrollment

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of total population enrolled in Medicaid in the past quarter	American Community Survey and Centers for Medicare & Medicaid Services	No	Not Available	From Tract & State

Data Table(s) + Variable(s)

The column "Total Medicaid Enrollment" was used from the monthly "State Medicaid and CHIP Applications, Eligibility Determinations, and Enrollment Data" dataset to access state-level Medicaid enrollment counts. Data were filtered to finalized data (Final Report IS Y).

Using the most recent ACS data available, the following variable from ACS data table S2704 5-year estimates was used to represent total population:

- Estimate!!Total!!Civilian noninstitutionalized population

The following variable from ACS data table S2704 5-year estimates was used to re-allocate state-level estimates to other geographies:

- Estimate!!Public Coverage!!COVERAGE ALONE OR IN COMBINATION!!Medicaid/means-tested public coverage alone or in combination

Analysis

The percentage of individuals enrolled in Medicaid is calculated by dividing enrollment counts (converted to quarterly by calculating three-month averages of monthly enrollment counts) by the total population, represented for this metric by "civilian noninstitutionalized population". Enrollment counts are differentially derived for congressional districts and states (see Geography-Specific Notes below). Estimates for Rhode Island are censored for 2024 Q4 due to the state not reporting data to CMS for December 2024.

$$\text{Medicaid Enrollment } \% = \frac{\text{Medicaid Enrollment Count}}{\text{Civilian Noninstituionalized Population}} \times 100$$

Geography-Specific Notes

Congressional Districts

Congressional district-level Medicaid enrollment data are available directly from the American Community Survey, but are underestimates of the true enrollment numbers due to the sampling design of the survey.[28] For this reason, our estimation of Medicaid enrollment counts at the congressional district-level uses a combination of American Community Survey data and Centers for Medicare & Medicaid Services data.

We first pulled census tract ACS Medicaid enrollment counts and aggregated from the tract level to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on geographic aggregation. Using these values, along with directly pulled ACS state-level enrollment counts, we then calculated the proportion of each state's Medicaid enrollees contained in each congressional district within that state.

$$\text{ACS Proportion}_{CD} = \frac{\text{ACS Medicaid Enrollment}_{CD}}{\text{ACS Medicaid Enrollment}_{State}}$$

We then applied these proportions to the statewide enrollment counts from CMS to get enrollment counts for each congressional district.

$$\text{Medicaid Enrollment}_{CD} = \text{ACS Proportion}_{CD} * \text{CMS Medicaid Enrollment}_{State}$$

Final percent estimates are created by dividing the derived Medicaid enrollment counts by total population ACS estimates that are aggregated from the tract level to generate congressional district counts. See "SECTION 3: Analytic Decisions" for more details on this method.

States

State enrollment counts for the numerator are pulled directly from CMS data. Total population denominators use ACS data at the state-level.

American Community Survey and U.S. Department of Agriculture Food and Nutrition Service

General Notes

The American Community Survey (ACS) is administered by the US Census Bureau.[9] Data are retrieved from the census API using R and the tidycensus package.[16, 13] Variable labels from the API (e.g., Estimate; SEX AND AGE - Total population), not names (e.g., S2801_C01_017E), are outlined in metric sections.

The U.S. Department of Agriculture Food and Nutrition Service (USDA FNS) reports monthly state-level data on Supplemental Nutrition Assistance Program (SNAP) participation. These data are publicly available for download[29]. USDA FNS data may continue to be updated, particularly for the most current numbers which are considered preliminary[30], potentially leading to estimates changing over time.

SNAP Participation

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of households participating in Supplemental Nutrition Assistance Program (SNAP) in the past quarter	American Community Survey and U.S. Department of Agriculture Food and Nutrition Service	No	Not Available	From Tract & State

Data Table(s) + Variable(s)

The Household Participation column from the "National and/or State Level Monthly and/or Annual Data" was used to access state-level SNAP participation counts.

Using the most recent ACS data available, the following variable from ACS data table S2201 5-year estimates was used to represent total households:

- Estimate!!Total!!Households

The following variable from ACS data table S2201 5-year estimates was used to re-allocate state-level estimates to other geographies:

- Estimate!!Households receiving food stamps/SNAP!!Households

Analysis

The percentage of households participating in SNAP is calculated by dividing household participation counts (converted to quarterly by calculating three-month averages of monthly participation counts) by total households. Participation counts are differentially derived for congressional districts and states (see Geography-Specific Notes below). Estimates for Florida are censored for 2024 Q3 and Q4 due to the state not reporting data to USDA FNS for September and October 2024.

$$\text{SNAP Participation } \% = \frac{\text{SNAP Household Participation Count}}{\text{Total Households}} \times 100$$

Geography-Specific Notes

Congressional Districts

Congressional district-level SNAP participation data are available directly from the American Community Survey, but are underestimates of the true participant numbers due to the sampling design of the

survey[31]. For this reason, our estimation of SNAP participation counts at the congressional district-level uses a combination of American Community Survey data and U.S. Department of Agriculture Food and Nutrition Service data.

We first pulled census tract ACS SNAP participation counts (all ACS SNAP data is at the household-level) and aggregated from the tract level to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on geographic aggregation. Using these values, along with directly pulled ACS state-level participation counts, we then calculated the proportion of each state's SNAP participants contained in each congressional district within that state.

$$\text{ACS Proportion}_{CD} = \frac{\text{ACS SNAP Partipcation}_{CD}}{\text{ACS SNAP Participationn}_{State}}$$

We then applied these proportions to the statewide participation counts from USDA FNS to get participation counts for each congressional district.

$$\text{SNAP Participation}_{CD} = \text{ACS Proportion}_{CD} * \text{USDA FNS SNAP Participation}_{State}$$

Final percent estimates are created by dividing the derived SNAP household participation counts by total household ACS estimates that are aggregated from the tract level to generate congressional district counts. See "SECTION 3: Analytic Decisions" for more details on this method.

States

State household participation counts for the numerator are pulled directly from USDA FNS data. Total household denominators use ACS data at the state-level.

Behavioral Risk Factor Surveillance System (BRFSS)

General Notes

The Behavioral Risk Factor Surveillance System (BRFSS) is a national telephone survey created by the Centers for Disease Control and Prevention and administered by state and territorial governments. It uses a complex sampling design to ensure that survey results are representative of each state's population.[32] Estimates for New Jersey in 2019, Florida in 2021, and Pennsylvania and Kentucky in 2023 are not available as the states did not collect enough BRFSS data to meet minimum requirements for inclusion. Estimates for Food Insecurity are available for limited states depending on the year as the BRFSS question used for this metric is part of an optional data module.

BRFSS variables were downloaded directly from the BRFSS website in SAS transport format. Each BRFSS survey year is a separate dataset.[33] Additionally, module data may be present across multiple datasets that represent different survey versions [34].

Using BRFSS data requires applying the following survey design variables: _LLCPWT for weighting, _STSTR for stratification, and _PSU for primary sampling unit (clustering).[32] Module questions require separate weighting variables to account for different survey versions (e.g., _LCPWTV1, _LCPWTV2) [35]. Metric estimates were calculated as proportions using the *survey* package for R,[36] with survey design variables applied.

All metric denominators were calculated from the number of non-missing, non-"don't know/not sure" responses to each respective BRFSS variable.

Confidence Interval Calculation

90% confidence interval of proportions were calculated using Wald formula:

$$p \pm 1.645 * \sqrt{\frac{p * (1 - p)}{n}}$$

Geography-Specific Notes

States

BRFSS data were used to calculate state estimates to accompany the census tract and congressional district estimates calculated using PLACES Project data. The PLACES Project uses BRFSS data to create small area estimates (see "PLACES, Centers for Disease Control and Prevention" section for more).

Binge Drinking

Metric Description	**Data Source**
Percentage of adults who reported binge drinking in the past 30 days	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents grouped as "Yes" (coded 1) in the BRFSS calculated variable X_RFBING5 or X_RFBING6.

Analysis

$$\text{Binge Drinking} = \frac{\text{Weighted sum of adults who report binge drinking in the past 30 days}}{\text{Total adult population}} * 100$$

Dental Care

Metric Description	Data Source
Percentage of adults who reported visiting a dentist in the past year	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents who answered "Within the past year" (coded 1) to LASTDEN4.

Analysis

$$\text{Dental Care} = \frac{\text{Weighted sum of adults who report visiting a dentist in the past year}}{\text{Total adult population}} * 100$$

Diabetes

Metric Description	Data Source
Percentage of adults who reported visiting a dentist in the past year	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents who answered "Yes" (coded 1) to DIABETE4.

Analysis

$$\text{Diabetes} = \frac{\text{Weighted sum of adults who report having diabetes}}{\text{Total adult population}} * 100$$

Food Insecurity

Metric Description	Data Source
Percentage of adults who reported that during the past year they did not have enough food and did not have money to get more	Behavioral Risk Factor Surveillance Survey (Social Determinants and Health Equity optional module), Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents who answered "Always" (coded 1), "Usually" (coded 2), or "Sometimes" (coded 3) to SDHFOOD1.

Analysis

$$\text{Food Insecurity} = \frac{\text{Weighted sum of adults who report they did not have enough food and did not have money to get more in the past year}}{\text{Total adult population}} * 100$$

Frequent Mental Distress

Metric Description	Data Source
Percentage of adults who reported experiencing ≥14 days of poor mental health in the past 30 days	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents grouped as "14+ days when mental health not good" (coded 3) in the BRFSS calculated variable _MENT14D.

Analysis

$$\text{Frequent Mental Distress} = \frac{\text{Weighted sum of adults who report } \geq 14 \text{ days of poor mental health in the past 30 days}}{\text{Total adult population}} * 100$$

Frequent Physical Distress

Metric Description	Data Source
Percentage of adults who reported experiencing ≥14 days of poor physical health in the past 30 days	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents grouped as "14+ days when physical health not good" (coded 3) in the BRFSS calculated variable _PHYS14D.

Analysis

$$\text{Frequent Physical Distress} = \frac{\text{Weighted sum of adults who report } \geq 14 \text{ days of poor physical health in the past 30 days}}{\text{Total adult population}} * 100$$

High Blood Pressure

Metric Description	Data Source
Percentage of adults who reported having high blood pressure	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents who answered "Yes" (coded 1) to BPHIGH4 or BPHIGH6.

Analysis

$$\text{High Blood Pressure} = \frac{\text{Weighted sum of adults who report high blood pressure}}{\text{Total adult population}} * 100$$

Obesity

Metric Description	Data Source
Percentage of adults with a body mass index (BMI) ≥30 kg/m², based on their reported weight and height	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents grouped as "Obese" (coded 1) in the BRFSS calculated variable _BMI5CAT.

Analysis

$$\text{Obesity} = \frac{\text{Weighted sum of adults who report a body mass index (BMI) ≥30 kg/m2}}{\text{Total adult population}} * 100$$

Physical Inactivity

Metric Description	Data Source
Percentage of adults who reported no leisure-time physical activity in the past 30 days	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents who answered "No" (coded 2) to EXERANY2.

Analysis

$$\text{Physical Inactivity} = \frac{\text{Weighted sum of adults who report no leisure-time physical activity in the past 30 days}}{\text{Total adult population}} * 100$$

Routine Checkup, 18+

Metric Description	Data Source
Percentage of adults who reported visiting a doctor for routine checkup in the past year	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents who answered "Within the past year" (coded 1) to CHECKUP1.

Analysis

$$\text{Routine Checkup, 18+=} \frac{\text{Weighted sum of adults who report visiting a doctor for routine checkup in the past year}}{\text{Total adult population}} * 100$$

Smoking

Metric Description	Data Source
Percentage of adults who reported current cigarette smoking	Behavioral Risk Factor Surveillance Survey, Centers for Disease Control

Data Table(s) + Variable(s)

Numerators were calculated from the number of respondents grouped as "Yes" (coded 2) in the BRFSS calculated variable _RFSMOK3.

Analysis

$$\text{Smoking=} \frac{\text{Weighted sum of adults who report current smoking}}{\text{Total adult population}} * 100$$

George Mason University Air Quality Team

General Notes

These data were created by fusing ground observations from the US Environmental Protection Agency (EPA) Air Quality System (AQS) network and computer model prediction from the National Oceanic and Atmospheric Administration (NOAA) National Air Quality Forecast Capability (NAQFC) by the George Mason University air quality team.

Estimates calculated by the George Mason University air quality team will differ from EPA CMAQ RSIG and GMU North America Chemical Reanalysis (NACR), which are commonly used, publicly available data sources for air pollution. While the RSIG presently includes a longer data period, North America Chemical Reanalysis (NACR) uses more up-to-date emission and real-time forecasting data to provide data in a timelier manner (up to yesterday). Both RSIG and NACR provide air pollution data for 12-kilometer square areas, which is larger than many census tracts. EPA CMAQ RSIG further smooths the data to provide census tract-level estimates, while NACR are provided at the 12-kilometer square area level only. As such, adjacent census tracts might share the same ozone pollution value (ppb) or $PM_{2.5}$ pollution value ($\mu g/m^3$).

Geography-Specific Notes

Census Tracts

Census tract estimates are presented as received.

Congressional Districts

Estimates are aggregated from the tract level to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on this method.

States

Estimates are aggregated from the tract level to generate state estimates. See "SECTION 3: Analytic Decisions" for more details on this method.

Air pollution – Ozone

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Average daily maximum concentration (parts per billion) of ground-level ozone throughout a month	George Mason University Air Quality Team	Yes	Not Applicable	From Tract

Data Table(s) + Variable(s)

The data from the George Mason University air quality team provides monthly average of daily maximum tract-level estimates for all tracts in the contiguous United States, except Alaska and Hawaii...

Analysis

The model prediction data are used to fill in gaps between air quality monitors, in particular rural and suburban areas. The model estimates spatial and temporal variations of air pollution based on three major components: emission, meteorology and chemistry. A list of emission sources is provided in table below ("List of emission sources used to estimate ozone"). The meteorology data is provided by the Weather Research and Forecast (WRF) model. The chemistry model is based on the US EPA Community Multiscale Air Quality (CMAQ) model. To merge the model and monitoring data, the optimal interpolation (OI) method is used to generate a fused surface concentration across the Continental

United States at 12km gridding. The gridded data are converted into census tract level by averaging all grid points included in the tract.

List of emission sources used to estimate ozone

Year of Update	Emission Sources Included
2024	Anthropogenic Sources: Agriculture, transportation (vehicle/air/railroad/marine), electricity generation units (EGUs), non-EGU point sources, oil/gas, residential wood combustion Natural Sources: Biogenic; Sea-salt; Wildfires

Air pollution – PM2.5

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Average daily concentration (µg/m³) of fine particulate matter (PM2.5) per cubic meter of air throughout a month	George Mason University Air Quality Team	Yes	Not Applicable	From Tract

Data Table(s) + Variable(s)

The data from the George Mason University air quality team provides monthly average of daily maximum tract-level 8-hour concentration of PM2.5 per cubic meter for all tracts in the contiguous United States, except Alaska and Hawaii.

Analysis

The model prediction data are used to fill in gaps between air quality monitors, in particular rural and suburban areas. The model estimates spatial and temporal variations of air pollution based on three major components: emission, meteorology and chemistry. A list of emission sources is provided in Table below ("List of emission sources used to estimate $PM_{2.5}$"). The meteorology data is provided by the Weather Research and Forecast (WRF) model. The chemistry model is based on the US EPA Community Multiscale Air Quality (CMAQ) model. To merge the model and monitoring data, the optimal interpolation (OI) method is used to generate a fused surface concentration across the Continental United States at 12km gridding. The gridded data are converted into census tract level by averaging all grid points included in the tract.

List of emission sources used to estimate $PM_{2.5}$

Year of Update	Emission Sources Included
2024	Anthropogenic Sources: Agriculture, transportation (vehicle/air/railroad/marine), electricity generation units (EGUs), non-EGU point sources, oil/gas, residential wood combustion Natural Sources: biogenic; sea-salt; windblown dust; biomass burning

2018 data previously presented on the Dashboard were derived from the Environmental Protection Agency, Community Multiscale Air Quality Remote Sensing Information Gateway (CMAQ RSIG). These data can be requested via email at info@CDhealthdashboard.org.

Health Resources & Services Administration

General Notes

Primary Care Health Professional Shortage Area (HPSA) designations are used to identify areas within the United States that are experiencing a shortage of health professionals for primary care. Only one year of data is presented for this metric as HPSA designations may change over time due to updates in the designation process and may not reflect changes in the underlying availability of primary care services.

Designated Primary Care Shortage Area

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of population living in a designed primary care shortage area.	Health Resources & Services Administration	Yes	Not Applicable	From Tract

Data Table(s) + Variable(s)

Primary care health professional shortage areas are retrieved from Health Resources & Services Administration website at December of the data year. This measure only includes geographic HPSAs.

Analysis

The Dashboard utilized the geographic components of primary care Health Professional Shortage Areas (HPSAs), which can include a single county, a county subdivision, or a census tract. These components were assigned to census tracts based on 202 geographies, using population allocations from the 2020 decennial census. The crosswalk from county subdivisions to census tracts was derived from the Geocorr 2022 application provided by the Missouri Census Data Center[37]. For a few census tracts based on 2010 geographies, the Dashboard used the IPUMS National Historical Geographic Information System: Version 19. to convert these tracts to 2020 geographies[7].

Most of the census tract estimates are either 0% or 100%. A 0% estimate indicates that the census tract is not in a primary care HPSA, while a 100% estimate indicates that the census tract is entirely within a primary care HPSA. Estimates between 0% and 100% signify that only part of the census tract is within a primary care HPSA. This partial inclusion typically occurs when the census tract is at the edge of a county subdivision that is part of a HPSA or when the original HPSA-designated census tract is based on 2010 geography. For precise HPSA locations, please refer to https://data.hrsa.gov/tools/shortage-area.

Geography-Specific Notes

Congressional Districts

See Analysis section above.

States

See Analysis section above.

National Center for Education Statistics and U.S. Department of Education

General Notes

Chronic Absenteeism represents the percent of public school students who miss 10% or more school days in an academic year. Public schools for this metric include both charter and district schools. Private schools are not included due to insufficient data availability. This metric should not be used to evaluate school district performance.

Chronic Absenteeism estimates are calculated using school level chronic absenteeism count data published by the U.S. Department of Education initiative Ed Data Express[38] (data group 814), and enrollment data from the National Center for Education Statistics (NCES).[39]

Categorizing race/ethnicity

With the exception of "Other", racial/ethnic groups are categorized as they are throughout the Dashboard: Black, Asian (Asian and Native Hawaiian or Pacific Islander), Hispanic, or White. "Other" represents summed values associated with American Indian or Alaska Native and Two or More Races because the data sources do not provide an "Other" category.

Confidence intervals

CIs are not presented because MOE or SE data were not presented in the underlying dataset, which presents counts.

Chronic Absenteeism

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of public school students who miss 10% or more school days in an academic year	National Center for Education Statistics, U.S. Department of Education	No	Sex, Race/Ethnicity	Not Applicable

Data Table(s) + Variable(s)

School-level chronic absenteeism counts are retrieved from the U.S. Department of Education website, ED Data Express for each school year. School level enrollment counts are accessed through NCES enrollment data for the associated school year. The two datasets are joined using the NCES school id number.

Analysis

The Dashboard assigned schools to congressional districts and states for each school year using the sf package and R along with school data from the National Center for Education Statistics (NCES). ArcGIS Pro 3.0.0 software was used for additional validation. Only schools with full geographic coordinates (latitudes and longitudes) were included.

Schools were assigned to a single congressional district through a spatial join using the school's latitude and longitude data and the Dashboard's spatial boundaries for congressional districts. Due to the unreliability of geocoding, a 200 meter buffer was used in most cases. In a few isolated cases schools were located more than 200 meters from the spatial boundary of the district due to their location on shorelines. In these instances the buffer was increased until these schools were captured, and each case

was checked manually for accurate assignment. Schools were then assigned to corresponding states based on their congressional district.

Schools that meet the following criteria, as detailed in the NCES data, are removed from the crosswalk:

- Private schools
- Pre-K and adult education schools
- Schools classified as special education schools
- Schools that are classified as "closed", "inactive", or otherwise not operating during the given school year
- Schools that are primarily virtual (*note: schools that temporarily taught remotely during the COVID 19 pandemic are not excluded*)
- Schools were additionally removed from the calculation if either the numerator or denominator value was missing or if the chronic absenteeism estimate was greater than 100% and therefore considered unstable.

Chronic absenteeism is calculated by summing the students reported by schools as chronically absent, then dividing by the total number of students enrolled at those schools at time of reporting. See the following formula:

$$\text{Chronic Absenteeism} = \frac{\sum_{i=1}^{n} \text{Students who miss 10\% or more school days in an academic year}}{\sum_{i=1}^{n} \text{Total students enrolled}}$$

Where: n = the number of schools assigned to that geography

Estimates are suppressed if the numerator (number of students chronically absent) is less than 20 students, or if the number of schools with data available for a given congressional district is less than 60% of the total number of schools in that district.

Geography-Specific Notes

Congressional Districts

Some congressional district estimates are suppressed if their associated state estimate was considered unreliable. See States section below.

States

Some state level estimates along with their associated congressional districts are suppressed due to concerns about underlying data reliability. The Dashboard recognizes that our chronic absenteeism estimates will not exactly match the numbers reported by states or the U.S. DOE. This is due to a variety of reasons, most prominently: 1) variations in school or student exclusion criteria; 2) different reported definitions of chronic absenteeism; 3) a lag in data updates in the National source.

As part of the validation process, the Dashboard Team compared state-reported chronic absenteeism rates[40,41] and U.S. DOE-reported rates[42], where available, with our own. Estimates were censored when either difference was large (> 5). This applies to Alaska, Arizona, DC, Idaho, Oregon, Washington, and West Virginia in data period 2022 (SY 2021-2022) and Arizona, Arkansas, Hawaii, and DC in data period 2023 (SY 2022-2023). Please email info@CDhealthdashboard.org if you would like to learn more.

National Vital Statistics System (NVSS)

General Notes

Vital statistics are calculated from data derived from national deaths (Multiple Cause of Death Data (MCDD)) and births (Natality Data (ND)) records. The Dashboard obtained vital statistics micro-data files from the National Center for Health Statistics (NCHS) restricted-use vital statistics data.[43] Metric estimates are calculated by the Dashboard data analytical team.

Users of these data are asked to acknowledge NCHS and the vital statistics jurisdictions as the data source in published reports and studies for which the files were used. NCHS and the vital statistics jurisdictions should also be cited in reports, articles, and news releases in electronic and print media describing the studies or results of the studies. The following is the recommended citation:
National Center for Health Statistics. [*Name of data file(s)*] *([year(s*]), as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.
https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm

Pooled Estimates

Due to low events count in certain demographic subgroups and data suppression policy from NCHS, the Dashboard calculated 1-year estimates for total population, and 3-year pooled estimates for race and sex subgroups. For example, the breast cancer death rate for total population for 2020 was calculated from the 2020 multiple cause of death dataset. However, the breast cancer death rate for Asian for 2020 was calculated from a combined dataset that consisted of 2018, 2019 and 2020 multiple cause of death datasets.

Population Denominators

Population denominators for all NVSS metrics were derived from the Census Bureau's Population Estimates Program (PEP) , and the latest vintage version will be applied.[44] For 3-year pooled estimates for race and sex subgroups, PEP estimates were combined to accurately reflect the population size of the area. For example, if the data were derived from a combined dataset that consisted of 2018, 2019, 2020 multiple cause of death data, the corresponding population denominators were from combined estimates from 2018, 2019, 2020 PEP population estimates.

- Because NCHS transitioned from bridge race to single race group in 2018, the estimates for 2019 race and sex subgroups are 2-year estimates.

Race/Ethnicity Definition

When combining race and ethnicity, Hispanic ethnicity is not mutually exclusive with any race(s) recorded for the decedent. For example, if a person is identified as Hispanic and Asian in their death certificate, this person will be counted in both the Hispanic and Asian death rates. The exception is White, which is defined as White, Non-Hispanic, and this is specified using a combination of the Hispanic origin and Race Recode variables.

Estimates by race for Asian, Black, Hispanic, White and Other for mortality metrics were calculated from Multiple Cause of Death Data (MCDD) Race Recode 40 (position: 489-490) and Hispanic Origin/Race Recode (position: 488). Definitions are as follows:

- White: Non-Hispanic White;
- Black: Black;

- Hispanic: Mexican, Puerto Rican, Cuban, Central or South American and other Hispanic origin;
- Asian: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, other or multiple Asian, Hawaiian, Guamanian, Samoan, and other or multiple islander;
- Other: American Indian or Alaskan Native (AIAN) and more than one race;

Estimates by race for Asian, Black, Hispanic, White and Other for natality metrics were calculated from Natality Data, Mother's Race Recode 6 (MRACE6, position 107) and Mother's Hispanic Origin Recode (MHISP_R, position 115). Definitions are as follows:

- White: Non-Hispanic White;
- Black: Black;
- Hispanic: Mexican, Puerto Rican, Cuban, Central and South American, and other Hispanic origin;
- Asian: Asian, Native Hawaiian and Other Pacific Islander (NHOPI) ;
- Other: American Indian or Alaskan Native (AIAN) and more than one race

The following race/ethnicity definitions were used as the population denominators by race for Asian, Black, Hispanic, White and Other for all NVSS metrics from the Census Bureau's Population Estimates Program (PEP).

- White: NHWA
- Black: BA
- Hispanic: H
- Asian: AA or NA
- Other: IA or TOM
- Total: TOT

Calculation Method

The Dashboard calculated direct age-adjusted death rates for all mortality metrics except for firearm homicides and firearm suicides. Below is the standardized population weight used for age-adjustment.

All NVSS metrics were calculated by first aggregating death or birth counts from the county level to the congressional district level and then applying the direct age-adjustment method.

Variable "YPLL-75 weight" and "standard life expectancy at age of deaths (years)" were used to calculate premature deaths (all cause) and "weight" was used to calculate all other mortality rates.

Table of US 2010 Standardized Population

Age Group	Weight	YPLL-75 weight	Standard life expectancy at age of deaths (years)
Total			
0-19	0.2697	0.287	65
20-34	0.2029	0.2159	47.5
35-44	0.133	0.1415	35
45-54	0.1458	0.1551	25
55-64	0.1182	0.1257	15
65-74	0.0703	0.0748	5
75+	0.0601	0	0

Censoring/Flagging Estimates

Estimates that involve fewer than 10 deaths or births are suppressed due to privacy restrictions imposed by the National Center for Health Statistics. When the population denominators, such as the total number of live births, consist of fewer than 50 individuals, they are also suppressed. Estimates are flagged when death/birth count is less than 50.

Censored estimates are removed from the website and downloadable data. Flagged estimates are noted in "Data Tips" on the website, or in downloadable data. Users should consult the Downloadable Data Dictionary, available at www.congressionaldistricthealthdashboard.org/data-access, for more detail.

Note that these criteria differ from the methods described under "SECTION 3: Analytic Decisions: Censoring/Flagging Aggregated Estimates with Missing Contributing Data"

Confidence Interval Calculation

Multiple Cause of Death Data

We calculated 90% CI using formula below:

$$\text{LCL90} = \text{estimate} - (1.645 \times \text{SE(estimate)})$$

$$\text{UCL90} = \text{estimate} + (1.645 \times \text{SE(estimate)})$$

Standard errors (SE) for age-adjusted rates:

This formula below was applied to breast cancer, colorectal cancer, cardiovascular disease, and opioid overdose deaths metrics to calculate SE as outlined by Lilienfeld and Stolley[45] in a document published by the Utah Department of Health[46]:

$$\text{SE(estimate)} = \sqrt{\left[\sum\left((\text{age-group specific US 2010 standardized population weight})^2 * \frac{\text{age-group specific crude mortality rate}^2}{\text{age-group specific total number of deaths}}\right)\right]}$$

SE for crude rates:

This formula below applied firearm suicides and firearm homicides were calculated according to the following formula outlined by Poisson distributions.

$$\text{SE(estimate)} = \frac{\sqrt{numerator}}{denominator} * 100{,}000$$

Please refer to the metric-specific section below for details on the standard error calculation for Life Expectancy.

Natality Data

CIs for low birthweight and prenatal care metrics were calculated as follows:

$$\text{LCL90} = \text{estimate} - 1.645 * \sqrt{\text{estimate*((100-estimate)/numerator)}}$$

$$\text{UCL90} = \text{estimate} + 1.645 * \sqrt{\text{estimate*((100-estimate)/numerator)}}$$

CIs for teen births metric were calculated as follows:

$$\text{LCL90} = (1000 / \text{denominator}) * (\text{numerator} - (1.645 * \sqrt{\text{numerator}}))$$

$$\text{UCL90} = (1000 / \text{denominator}) * (\text{numerator} + (1.645 * \sqrt{\text{numerator}}))$$

Geography-Specific Notes

Congressional Districts

See "Calculation Method" for more details on aggregation method. Confidence intervals are not calculated.

States

The state estimates and confidence intervals were derived directly from death and birth records, not from aggregation.

Breast Cancer Deaths

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Deaths due to breast cancer per 100,000 females	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Race/Ethnicity	From County

Data Table(s) + Variable(s)

The following underlying cause of death ICD-10 codes were summed to calculate Breast Cancer Deaths (females only): C500, C501, C502, C503, C504, C505, C506, C508, & C509. ICD-10 codes were selected for inclusion as per the 2016 SEER Program Coding and Staging Manual.[81]

Analysis

$$\text{Breast Cancer Deaths} = \left(\sum_{1}^{i} \frac{\text{death}_i}{\text{population}_i} * w_i\right) * 100{,}000$$

Where:

i= total number of age groups (i = 13)
death_i= the number of breast cancer deaths for female population in the i^{th} age group
population_i = the total female population in the i^{th} age group
w_i = US 2010 standardized population weights

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

Cardiovascular Disease Deaths

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Deaths due to cardiovascular disease per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

The following underlying cause of death ICD-10 codes were summed to calculate Cardiovascular Disease Deaths:

I110, I119, I130, I131, I132, I139, I10, I120, I129, I150, I159, I210, I211, I212, I213, I214, I219, I220, I229, I241, I248, I249, I200, I201, I209, I250, I251, I253, I254, I255, I258, I259, I500, I501, I509, I600, I602, I604, I605, I606, I607, I608, I609, I610, I611, I612, I613, I614, I615, I616, I618, I619, I620, I621, I629, I630, I631, I632, I633, I634, I635, I636, I638, I639, I64, I670, I671, I672, I673, I674, I675, I676, I677, I678, I679, I690, I691, I692, I693, I694, I698

ICD-10 codes were selected for inclusion based on Nolte & McKee[47] as well as in consultation with the NYU School of Medicine's Department of Population Health.

Analysis

$$\text{Cardiovascular Disease Deaths} = \left(\sum_{1}^{i} \frac{\text{death}_i}{\text{population}_i} * w_i\right) * 100{,}000$$

Where:

i= total number of age groups (i = 13)
death_i= the number of cardiovascular disease deaths for population in the i^{th} age group
population_i = the total population in the i^{th} age group
w_i = US 2010 standardized population weights

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

Colorectal Cancer Deaths

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Deaths due to colorectal cancer per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

The following underlying cause of death ICD-10 codes were summed to calculate Colorectal Cancer Deaths: C180, C181, C182, C183, C184, C185, C186, C187, C188, C189, C19, & C20. ICD-10 codes were selected for inclusion based on the publication by Siegel, et al[48] and in consultation with the NYU School of Medicine's Division of Gastroenterology.

Analysis

$$\text{Colorectal Cancer Deaths} = \left(\sum_{1}^{i} \frac{\text{death}_i}{\text{population}_i} * w_i\right) * 100{,}000$$

Where:

i= total number of age groups (i = 13)
death_i= the number of colorectal cancer deaths for population in the i^{th} age group
population_i = the total population in the i^{th} age group
w_i = US 2010 standardized population weights

See above “Geography-Specific Notes” section under “National Vital Statistics System: General Notes” for information on calculation differences between congressional districts and states.

Firearm Homicides

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Deaths due to firearm homicide per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

The following underlying cause of death ICD-10 codes were summed to calculate Firearm Homicides: X93, X94 and X95. ICD-10 codes were selected for inclusion in consultation with the NYU School of Medicine with support from Everytown for Gun Safety.

Analysis

$$\text{Firearm Homicides} = \frac{\text{death}}{\text{population}} * 100{,}000$$

Where:

death = the number of firearm related homicides in total population
population = total population

See above “Geography-Specific Notes” section under “National Vital Statistics System: General Notes” for information on calculation differences between congressional districts and states.

Firearm Suicides

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Deaths due to firearm suicide per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

The following underlying cause of death ICD-10 codes were summed to calculate the total number of deaths from intentional self-harm by firearms: X72, X73 and X74. ICD-10 codes were selected for inclusion in consultation with the NYU School of Medicine with support from Everytown for Gun Safety.

Analysis

$$\text{Firearm Suicides} = \frac{\text{death}}{\text{population}} * 100{,}000$$

Where:

death = the number of firearm related homicides in total population
population = total population

See above “Geography-Specific Notes” section under “National Vital Statistics System: General Notes” for information on calculation differences between congressional districts and states.

Opioid Overdose Deaths

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Deaths due to opioid overdose per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

The following underlying cause of death ICD-10 codes were summed to calculate Opioid Overdose Deaths: X40, X41, X42, X43, X44, X60, X61, X62, X63, X64, X85, Y10, Y11, Y12, Y13, & Y14 in combination with T400, T401, T402, T403, T404, & T406 multiple cause of death codes. ICD-10 codes were selected for inclusion as per the CDC's Guide to ICD-9-CM and ICD-10 Codes Related to Poisoning and Pain in addition to the Henry J Kaiser Family Foundation.[49,50]

Due to reporting variability and rapid shifts in opioid use patterns, the reported estimated rates may not accurately reflect current opioids involved deaths.

Analysis

$$\text{Opioid Overdose Deaths} = \left(\sum_{1}^{i} \frac{\text{death}_i}{\text{population}_i} * w_i\right) * 100{,}000$$

Where:

i= total number of age groups (i = 13)
death_i= the number of opioids involved deaths for population in the i^{th} age group
population_i = the total population in the i^{th} age group
w_i = US 2010 standardized population weights

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

Premature Deaths (All Causes)

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Years of potential life lost before age 75 per 100,000 people	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

Premature Deaths (All Causes) rate is defined as years of potential life lost before age 75 (YPLL-75) calculated as per Dranger and Remington's approach.[51]

Analysis

$$\text{Premature Deaths (All Causes)} = \left(\sum_{1}^{i} \frac{\text{death}_i}{\text{population}_i} * w_i * e_i\right) * 100{,}000$$

Where:

i= total number of age groups (i = 13)
death_i= the number of total deaths for population in the i^{th} age group
population_i = the total population in the i^{th} age group
w_i = US 2010 standardized population YPLL-75 age-group specific weight
e_i = standard life expectancy at age of deaths (years)

Standard error for premature deaths (all causes) were calculated according to the following formula outlined by Vohlonen, Bäckmand, & Korhonen:[52]

$$\text{SE(estimate)} = \sqrt{\left[\sum\left(\frac{\text{age-group specific crude mortality rate}^2}{\text{age-group specific total number of deaths}} * (w_1 * w_2)\right)\right]}$$

w_1 = Age-group specific premature deaths weight--years of life lost

w_2 = US 2010 standardized population YPLL age-group specific weight

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

Life Expectancy

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Average life expectancy at birth, in years	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	No	Sex, Race/Ethnicity	From County

Data Table(s) + Variable(s)

Life Expectancy at birth was calculated using Chiang's abridged life table method, with several modifications detailed below[53-58]. Age intervals were 0, 1-4, and 10-year age groups until age 85+. Population denominators for the 0 and 1-4 age intervals were generating by multiplying population counts for the 0-4 age group by 0.2 and 0.8, respectively[59].

Analysis

i = age interval
D_i = number of people dying in age interval
P_i = population in age interval
n_i = size of the age interval in years
a_i = fraction of interval lived by those who died in the age interval *(0.5, except for 0.1 in 0 age interval)*[53]

Analysis Steps	Formulas	Notes
Life Expectancy at Birth		
Age interval-specific death rate (M_i)	$\frac{D_i}{P_i}$	• Set to 0 when $D_i = 0$ and $P_i = 0$ • Set to 1 when $D_i > P_i$ • When $D_{85+} = 0$ or $P_{85+} = 0$, impute state death rate for M_{85+}
Probability of dying in age interval (q_i)	$\frac{n_i * M_i}{1 + (1 - a_i) * n_i * M_i}$	• Set to 1 when $D_i > (P_i / n_i / a_i)$ • $q_{85+} = 1$
Probability of surviving age interval (p_i)	$1 - q_i$	
Number surviving at start of age interval (l_i)	$l_{i-1} * p_{i-1}$	• l_0= 100,000 (Standard hypothetical cohort)
Number dying in age interval (d_i)	$l_i * q_i$	

Analysis Steps	Formulas	Notes
Life Expectancy at Birth		
Number of years lived in age interval (L_i)	$(l_{i+1} * n_i) + (d_i * n_i * a_i)$	• $L_{85+} = \frac{l_{85+}}{M_{85+}}$
Total person-years lived after age interval (T_i)	$\sum_i^{85+} L_i$	
Life expectancy (e_0)	$\frac{T_0}{l_0}$	

Confidence intervals

SE were calculated according to the following formula outlined by Chiang, with an additional adjustment term for the final age interval included as proposed by Silcocks[60,61]:

Analysis Steps	Formulas	Notes
Chiang LE variance		
Variance of probability of dying (var_{qi})	$\frac{q_i^2 * (1 - q_i)}{D_i}$	• Set D_i to 0.693 when $D_i = 0$
Weighted variance ($var1_i$)	$l_i^2 * [(1 - a_i) * n_i + e_i]^2 * var_{qi}$	• $var1_{85+} = 0$
Variance of person-years lived beyond start of interval ($var2_i$)	$\sum_i^{85+} var1_i$	
Variance of life expectancy (var_{ei})	$\frac{var2_i}{l_i^2}$	
Silcocks adjustment term for 85+ age interval		
Variance adjustment (paw)	$\frac{l_{85+}}{l_0}$	
Adjusted variance term (var_{adj})	$\frac{paw^2}{M_{85+}^3 * P_{85+}}$	• Set $P_{85+} = D_{85+}$ when $P_{85+} = 0$
Standard error for life expectancy at birth		
Standard error (SE_0)	$\sqrt{var_{e0} + var_{adj}}$	

We calculated 90% CI using the formula below:

$$LCL90 = estimate - (1.645 \times SE(estimate))$$

$$UCL90 = estimate + (1.645 \times SE(estimate))$$

Notes on Analysis

Life Expectancy estimates were censored in the following instances:

- When the total population was <5000[58]
- When SE was >= 4
- When deaths were greater than population in at least one age interval
- When estimates were over 100

Low Birthweight

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of live births with low birthweight (<2500 grams)	Natality Data, National Vital Statistics System, National Center for Health Statistics	No	Race/Ethnicity	From County

Data Table(s) + Variable(s)

All births with birthweights that are either missing, unknown, or not stated are excluded from the analysis.

Analysis

$$\text{Low Birthweight} = \frac{\text{number of live births with birthweight <2500 grams}}{\text{total number of live births}} * 100$$

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

Prenatal Care

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of births for which prenatal care began in the first trimester	Natality Data, National Vital Statistics System, National Center for Health Statistics	No	Race/Ethnicity	From County

Data Table(s) + Variable(s)

Prenatal Care estimates represent a slight modification of one component of the Kotelchuck Index.[62] All births with missing or unknown prenatal care are excluded from the analysis. Prenatal care data for certain states across years are missing because these states had not implemented 2003 birth certificate revisions. If prenatal care information is missing for 10% or more of live births in a given county, all prenatal care values for that county are censored. See above "SECTION 3: Analytic Decisions" for censorship decisions at congressional district level. For more information please refer to the natality public use data documentation files.[63-68]

Analysis

$$\text{Prenatal Care} = \frac{\text{number of live births with prenatal care beginning between 1 and 3 months}}{\text{total number of live births}} * 100$$

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

Teen Births

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Births per 1,000 females aged 15-19	Natality Data, National Vital Statistics System, National Center for Health Statistics	No	Race/Ethnicity	From County

Data Table(s) + Variable(s)

See analysis below.

Analysis

$$\text{Teen Births} = \frac{\text{number of live births to mothers aged 15-19}}{\text{total female population aged 15-19}} * 100{,}000$$

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on calculation differences between congressional districts and states.

PLACES, Centers for Disease Control and Prevention

General Notes

PLACES apply a multi-level regression with post-stratification (MRP) approach to develop small area estimates (SAE), like census tracts, for key measures captured in the Behavioral Risk Factor Surveillance System (BRFSS). Prior to the PLACES, BRFSS measures were only available at the county, Metropolitan Statistical level or above. For further details on the methodology, see Zhang et al (2014).[69] For more information regarding these metrics, please refer to the PLACES's methodology pages.[70-72]

Estimates for New Jersey in 2019, Florida in 2021, and Kentucky and Pennsylvania in 2023 are not available as the states did not collect enough BRFSS data to meet minimum requirements for inclusion. Estimates for Food Insecurity is subject to additional missingness as the BRFSS module used for this metric is optional. The underlying data source, BRFSS, issued a warning for their 2020 data noting that 2020 data may not be comparable to other years due to difficulties collecting samples during the 2020 pandemic, and an underlying weighting method modification.[73]

Confidence Interval Calculation

Confidence intervals were included with the estimates downloaded from PLACES. However, PLACES reports 95% confidence intervals, rather than the 90% confidence intervals reported by the Dashboard. Upper and lower limits of the 95% confidence intervals were used to calculate an approximate standard error (SE). The SE was then used to calculate 90% confidence intervals. $SE = \frac{UCL95 - LCL95}{1.96 \times 2}$

$$LCL90 = Estimate - (1.645 \times SE)$$

$$UCL90 = Estimate + (1.645 \times SE)$$

Where:

SE = approximate standard error
LCL95 = Reported lower limit for the 95% confidence interval
UCL95 = Reported upper limit for the 95% confidence interval
LCL90 = Calculated lower limit for the 90% confidence interval
UCL90 = Calculated upper limit for the 90% confidence interval

Geography-Specific Notes

Census Tracts

Census tract estimates are provided as received from PLACES. Confidence intervals are provided.

Congressional Districts

PLACES estimates are aggregated from the tract level to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on this method. Confidence intervals are not provided.

States

State estimates are calculated directly from BRFSS. Please see the "Behavioral Risk Factor Surveillance System (BRFSS)" metric analysis section for more details. Confidence intervals are provided.

Binge Drinking

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported binge drinking in the past 30 days	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Binge Drinking crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.[74-79]

Analysis

Binge Drinking is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Dental Care

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported visiting a dentist in the past year	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Dental Care crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Dental Care is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Diabetes

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported having diabetes	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Diabetes crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Diabetes is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Food Insecurity

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported that during the past year they did not have enough food and did not have money to get more	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Food Insecurity crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Food Insecurity is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Frequent Mental Distress

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported experiencing ≥14 days of poor mental health in the past 30 days	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Frequent Mental Distress crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Frequent Mental Distress is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Frequent Physical Distress

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported experiencing ≥14 days of poor physical health in the past 30 days	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Frequent Physical Distress crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Frequent Physical Distress is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

High Blood Pressure

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported having high blood pressure	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

High Blood Pressure crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

High Blood Pressure is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Obesity

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults with a body mass index (BMI) ≥30 kg/m², based on their reported weight and height	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Obesity crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Obesity is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Physical Inactivity

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported no leisure-time physical activity in the past 30 days	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Physical Inactivity crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Physical Inactivity is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Routine Checkup, 18+

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Percentage of adults who reported visiting a doctor for routine checkup in the past year	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Routine Checkup, 18+ crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Routine Checkup, 18+ is reported as received.

See above "Geography-Specific Notes" section under "PLACES, Centers for Disease Control and Prevention: General Notes" for information on calculation differences between Census tracts and congressional districts.

Smoking

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	Geography Aggregation
Percentage of adults who reported current cigarette smoking	PLACES Project, Centers for Disease Control	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Smoking crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

Smoking is reported as received.

See above “Geography-Specific Notes” section under “PLACES, Centers for Disease Control and Prevention: General Notes” for information on calculation differences between Census tracts and congressional districts.

United States Small-Area Life Expectancy Project (USALEEP)

General Notes

As of the 04-06-2026 release, this metric has been archived on the website. It can be accessed through the updated Life Expectancy metric page.

Life expectancy estimates were estimated by the United States Small-Area Life Expectancy Project (USALEEP), a joint effort of The Robert Wood Johnson Foundation, National Association for Public Health Statistics and Information Systems (NAPHSIS) and the National Center for Health Statistics (NCHS) at the Centers for Disease Control (CDC). The methodology used to calculate tract-level data is published.[59]

Life Expectancy - Census Tract-Level, 2015

Metric Description	Data Source	Tract Estimates	Demographic Subgroups	CD Geography Aggregation
Average life expectancy at birth, in years	U.S. Small-Area Life Expectancy Estimates Project (USALEEP)	Yes	Not Available	From Tract

Data Table(s) + Variable(s)

Tract-level data were downloaded from USALEEP; tract-level data and documentation files are available for free download.[80-82]

Analysis

Estimates are calculated by USALEEP and represent the average number of years a person can expect to live from birth.

Confidence Interval Calculation

Standard errors are included in downloadable USALEEP data. Ninety percent confidence intervals for were calculated as per the following formulas:

$$LCL90 = estimate - (1.645 \times SE(estimate))$$

$$UCL90 = estimate + (1.645 \times SE(estimate))$$

Where:

LCL90 = Calculated lower limit for the 90% confidence interval
UCL90 = Calculated upper limit for the 90% confidence interval
SE = approximate standard error

Geography-Specific Notes

Census Tracts

Census tract estimates are presented as received from USALEEP. Confidence intervals are calculated.

Congressional Districts

Life Expectancy estimates are aggregated from the tract level to generate congressional district estimates. See “SECTION 3: Analytic Decisions” for more details on this method. Confidence intervals are not calculated.

States

State estimates are presented as received from USALEEP, but are downloaded from a different location than census tracts.[83] Confidence intervals are calculated.

SECTION 5: Appendix

Acknowledgements

The Dashboard acknowledges Allegra Wilson, Sarah Conderino, Miriam Gofine, Peggy Hsieh, Alexander Chen, Dr. Shoshanna Levine, Anne Vierse, Noah Zazanis, Caleigh Paster, Dr. Jessica Athens, Shauna Ford, Dr. Susan Kum, Leena Abbas, and Farhana Haque for their significant contributions to the Dashboard site and analyses.

The Dashboard acknowledges Rania Kanchi, Alexis Feinberg, and Priscilla Lopez for their significant contributions to the pilot Dashboard site and analyses.

The Dashboard acknowledges the Robert Wood Johnson Foundation and the program staff who support this project.

The Dashboard acknowledges Dr. S V Subramanian (Professor of Population Health and Geography at Harvard University) for his collaboration and expertise on geographic aggregation methods and congressional districts.

For additional expertise surrounding geographic aggregation methods, the Dashboard acknowledges Dr. David Van Riper and Dr. Jonathan Schroeder (IPUMS National Historical Geographic Information System (NHGIS)), Dr. Yan Wang and Dr. Hua Lu (PLACES Project), Glenn Rice (Missouri Census Data Center GeoCorr), and The American Community Survey.

The Dashboard acknowledges the following individuals for their technical assistance: Dr. Ingrid Ellen, Dr. Justin Feldman, Dr. Kevin Kromar, Dr. Jennifer Norton, Susan Kum, Dr. Pasquale Rummo, and the Dashboard Scientific Advisory Committee.

The Dashboard acknowledges Dr. Xingyou Zhang, Senior Mathematical Statistician, U.S. Department of Agriculture, Economic Research Service, for his expertise surrounding the “Air pollution – particulate matter” metric.

The Dashboard acknowledges Dr. Erin Fahle, Assistant Professor, St. John’s University, and Dr. sean f. reardon, Professor, Stanford University, for their expertise surrounding development of the education metrics, particularly third-grade reading scores.

The Dashboard acknowledges Dr. Celia Stall-Meadows (Tribal Research Department, Choctaw Nation), Dr. Carmela Alcántara (Associate Professor, Columbia University School of Social Work), Dr. Stella Yi (Assistant Professor, NYU School of Medicine), Dr. Simona Kwon (Associate Professor, NYU School of Medicine), Jennifer Wong (Senior Program Coordinator, NYU School of Medicine), Center for the Study of Asian American Health, and Dr. Sheri Daniels (Executive Director, Papa Ola Lokahi) for their expertise surrounding demographic maps.

Update History

Update Date	Update Notes

04-28-26	• ACS:
	  ○ Addition of 2024 data
	• SNAP Participation: Release of new metric
	• NVSS:
	  ○ Breast Cancer Deaths: Update to ICD codes
	• Medicaid Enrollment:
	  ○ Change of metric name, previously named Medicaid Enrollment - % of Total population
	• Customized Areas: Addition of new website feature under “Metric Maps”

04-06-2026	• Air pollution – Ozone: Addition of 2024 monthly data
	• Air pollution – Particulate Matter: Addition of 2024 monthly data
	• Designated primary care shortage area: Update to 2025 data
	• Chronic Absenteeism: Addition of 2023 data
	• PLACES Project metrics:
	  ○ Addition of 2023 data for 10 metrics (No updates for Dental Care)
	• NVSS:
	  ○ Addition of 2023 data
	  ○ Update to population denominator to 2024 vintage version
	  ○ Update to aggregation method
	  ○ Update to race and ethnicity definitions
	  ○ Update to age adjustment categories for Opioid Overdose Deaths, Breast Cancer Deaths, Colorectal Cancer Deaths, Premature Deaths (all causes), Cardiovascular Disease Deaths
	• Medicaid Enrollment - % of Total population:
	  ○ Addition of Q2 and Q3 2025 data
	  ○ Updates to previous years of data due to updates from data source
	• Life Expectancy:
	  ○ Replace USALEEP data source with NVSS
	  ○ Release of new Life Expectancy data for 2019-2023, including for race and sex subgroups
	  ○ Archive Life Expectancy from USALEEP as Life Expectancy - Census Tract-Level, 2015
	• District Density Scale:
	  ○ Change name from District Density Index to District Density Scale
	  ○ Update to calculation method

08-19-2025	• Medicaid Enrollment - % of Total population: ○ Addition of Q3 2024 – Q1 2025 data
02-26-2025	• 06-12-2025 revision: corrected an error which had caused 20 tract estimates to be missing in the newest year of data for all PLACES metrics. • Update all metrics to 119th congress. • ACS: ○ Addition of 2023 data

	○ Unemployment: addition of age breakdowns. Update to metric population weighting variable for Asian and Other resulting in small estimate changes for previous years of data. ○ High school completion: addition of age breakdowns. ○ Youth Not in Work or School: Release of new metric (2020-2023), including sex breakdowns ○ Independent Living Difficultly: Release of new metric (2020-2023), including age breakdowns • Air pollution – Ozone: Addition of 2023 monthly data • Air pollution – PM2.5: Addition of 2023 monthly data • Designated primary care shortage area: Release of new metric (2024) • PLACES Project metrics: ○ Addition of 2022 data (uses 2020 geographic boundaries) ○ Update 2021 High Blood Pressure data to use 2020 geographic boundaries ○ Removal of Preventive Services, 65+ metric. ○ Food Insecurity: Release of new metric (2022); only available for 39 states. • NVSS: ○ Addition of 2022 data ○ Update population denominator to 2023 vintage version. • Medicaid Enrollment - % of Total population: ○ Addition of Q2 2024 data ○ Update underlying Medicaid weight to use 2023 ACS data for quarters in 2023 and 2024 • Health Snapshots – District Facts: ○ Update 2023 data • Update aggregation weights for select metrics: ○ PLACES metrics use Adult (18+) weight (previously used Total weight) ○ Broadband Connection and Income Inequality use Total Households weight (previously used Total weight) ○ Lead Exposure Risk Index and Housing with Potential Lead Risk use Housing Units weight (previously used Total weight) ○ Rent Burden uses Renter Households weight (previously used Total weight) ○ Children in Poverty, High School Completion, Unemployment and Uninsured use Hispanic-inclusive weights for Asian, Black, and Other estimates (previously used non-Hispanic weights) ○ Demographic data on Health Snapshots - District Facts use Hispanic-inclusive weights for Asian, Black, and Other estimates (previously used non-Hispanic weights)
08-27-2024	• Medicaid Enrollment - % of Total Population: Addition of Q4 2023 and Q1 2024 data. • Income Inequality: Limit metric to 2022 data year • Addition of State Snapshots, Metric Overview and State Facts pages • District Snapshots – District Facts: ○ Addition of related cities
02-20-2024	• Medicaid Enrollment - % of Total Population: New metric added (Q4 2022 – Q3 2023) • ACS: ○ Addition of 2022 data ○ Calculation update for Lead Exposure Risk Index and Housing with Potential Lead Risk • PLACES Project:

	○ Addition of 2021 data for all metrics except Dental Care and Preventative Services, 65+ ○ Calculation update for 2020 state Preventive Services, 65+ data • NVSS: Addition of 2021 data • Chronic Absenteeism: New metric added (School year ending in 2022) • Air Pollution – Ozone: New monthly metric added (Jan – Dec 2022) • Air Pollution – PM2.5: ○ Update of data source ○ Change in data time period from annual to monthly data (Jan – Dec 2022) • COVID Local Risk Index: Metric removed from website and downloadable data • District Snapshots – District Facts: ○ Updated data source and calculation method ○ New year of data (2022)
06-06-2023	• 08-23-2023 revision: corrected an error in segregation 2021 congressional district and state values that caused a slight inflation of scores. • ACS: Addition of 2021 data • PLACES Project: ○ Addition of 2020 data for all metrics except High Blood Pressure ○ Addition of 2017 data for High Blood Pressure • Air Pollution – Particulate Matter: Addition of 2019 data • Limited Supermarket Proximity: Metric removed from website and downloadable data. • Addition of Census tract estimates (where available) • Addition of state estimates (where available)
03-10-2023	• Limited Supermarket Proximity: Metric name updated from “Limited Access to Healthy Foods”; No change made to underlying construct
01-24-2023	First release of the Congressional District Health Dashboard • 36 metrics • Data for all congressional districts • One year of data released for all metrics

SECTION 6: References

1. Bureau UC. About Congressional Apportionment. https://www.census.gov/topics/public-sector/congressional-apportionment/about.html
2. US Census Bureau. Computing Apportionment. https://www.census.gov/topics/public-sector/congressional-apportionment/about/computing.html
3. Eicher CL, Brewer CA. Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. *Cartography and Geographic Information Science*. 2001/01/01 2001;28(2):125-138. doi:10.1559/152304001782173727
4. Hao Y WE, Jemal A, Pickle LW, Thun MJ. U.S. congressional district death rates. *International Journal of Health Geography*. 2006;
5. Spoer BR, Chen AS, Lampe TM, et al. Validation of a geospatial aggregation method for congressional districts and other US administrative geographies. *SSM Popul Health*. Dec 2023;24:101511. doi:10.1016/j.ssmph.2023.101511
6. 119th Congress Block Equivalency Files (2024).
7. Manson S, Schroeder J, Van Riper D, Kugler T, Ruggles S. Data from: IPUMS National Historic Geographic Information System: Version 17.0 2010 Blocks -> 2020 Blocks Crosswalk. 2022. *Minneapolis, MN*.
8. US Census Bureau. Table & Geography Changes. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes.html
9. Berkley J. Using American Community Survey estimates and margins of error. Presentation. Decennial Statistical Studies Division, US Census Bureau. Updated April 19, 2017. Accessed January 1, 2018. https://www.census.gov/content/dam/Census/programs-surveys/acs/guidance/training-presentations/20170419_MOE.pdf
10. US Department of Education. Final guidance on maintaining, collecting, and reporting racial and ethnic data to the U.S. Department of Education. Updated December 3, 2007. Accessed January 17, 2023. https://www.federalregister.gov/documents/2007/10/19/E7-20613/final-guidance-on-maintaining-collecting-and-reporting-racial-and-ethnic-data-to-the-us-department
11. Walker K. tigris: Download and use Census TIGER/Line shapefiles in R. https://github.com/walkerke/tigris
12. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/
13. Walker K, Herman M. tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames. . https://walker-data.com/tidycensus/.
14. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *Journal of Open Source Software*. 2019;4(43):1686. doi:https://doi.org/10.21105/joss.01686
15. Pebesma E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*. 2018;10(1):439-446. doi:10.32614/RJ-2018-009
16. US Census Bureau. Available APIs. US Department of Commerce. Accessed January, 2020. https://www.census.gov/data/developers/data-sets.html
17. US Census Bureau. *A compass for understanding and using American Community Survey Data: Appendix 1, Understanding and using ACS single-year and multiyear estimates*. Vol. 2018. https://www.census.gov/content/dam/Census/library/publications/2009/acs/ACSResearch.pdf
18. Lofy K, Davies H, Fotinos C, et al. A targeted approach to blood lead screening in children, Washington State: 2015 expert panel recommendations. Washington State Department of Health. Accessed January 9, 2018. https://assets.documentcloud.org/documents/2644455/Expert-Panel-Childhood-Lead-Screening-Guidelines.pdf

19. Frostenson S, Kliff S. Lead Exposure Risk data (readme.md). Vox Media,. Updated April 5, 2016. Accessed January 9, 2017. https://github.com/voxmedia/data-projects/tree/master/vox-lead-exposure-risk
20. Frostenson S, Kliff S. Added data for vox-lead-exposure-risk-map: Lead Exposure Risk data. Vox Media,. Updated April 5, 2016. Accessed January 9, 2017. https://github.com/voxmedia/data-projects/commit/f8ed1caa78af3232f18c0fcee23626a822fc2169
21. Foundation PS. Python Language Reference, version 3.6.
22. U.S. Department of Housing and Urban Development Office of Lead Hazard Control and Healthy Homes. *American Healthy Homes Survey II Lead Findings*. 2021. https://www.hud.gov/sites/dfiles/HH/documents/AHHS_II_Lead_Findings_Report_Final_29oct21.pdf
23. United States Environmental Protection Agency. Hazard Standards and Clearance Levels for Lead in Paint, Dust and Soil (TSCA Sections 402 and 403). Accessed 2/14/2024, 2024. https://www.epa.gov/lead/hazard-standards-and-clearance-levels-lead-paint-dust-and-soil-tsca-sections-402-and-403
24. 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. *American journal of public health*. 2016;106(2):256-263.
25. Iceland J. The multigroup entropy index (also known as Theil's H or the Information Theory Index). US Census Bureau. Accessed January 5, 2018, https://www2.census.gov/programs-surveys/demo/about/housing-patterns/multigroup_entropy.pdf
26. Centers for Medicare & Medicaid Services. Data from: State Medicaid and CHIP Applications, Eligibility Determinations, and Enrollment Data. 2024.
27. Centers for Medicare & Medicaid Services. Methodology: About the Medicaid & CHIP Enrollment Report. Accessed 08/21/2024, 2024. https://www.medicaid.gov/medicaid/national-medicaid-chip-program-information/medicaid-chip-enrollment-data/methodology-about-medicaid-chip-enrollment-report/index.html
28. Boudreaux M, Noon JM, Fried B, Pascale J. Medicaid expansion and the Medicaid undercount in the American Community Survey. *Health Serv Res*. Dec 2019;54(6):1263-1272. doi:10.1111/1475-6773.13213
29. USDA FNS. SNAP Data Tables. Accessed 04/21/2026. https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap
30. USDA FNS. Program Data Overview. Updated 03/24/2026. Accessed 04/21/2026, 2026. https://www.fns.usda.gov/pd/overview
31. Meyer BD, Mittag N, Goerge RM. Errors in Survey Reporting and Imputation and Their Effects on Estimates of Food Stamp Program Participation. *Journal of Human Resources*. 2022;57(5):1605-1644. doi:10.3368/jhr.58.1.0818-9704R2
32. Centers for Disease Control and Prevention. The BRFSS Data User Guide. Updated August 15, 2013. https://www.cdc.gov/brfss/data_documentation/pdf/UserguideJune2013.pdf
33. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System: Annual Survey Data. Updated July 26, 2022. https://www.cdc.gov/brfss/annual_data/annual_data.htm
34. The combined Landline and Cellular Telephone Survey Multiple Questionnaire Version Data including Optional Modules (2023).
35. Complex Sampling Weights and Preparing 2022 BRFSS Module Data for Analysis (2023).
36. Lumley T. survey: Analysis of Complex Survey Samples. Updated May 3, 2023. https://cran.r-project.org/package=survey
37. Missouri Census Data Center. Geocorr 2022: Geographic Correspondence Engine. https://mcdc.missouri.edu/applications/geocorr2022.html

38. Express ED. Data Library. Accessed 9/18/2023. https://eddataexpress.ed.gov/download/data-library?field_year_target_id=All&field_population_value=Chronically+Absent+Students&field_data_topic_target_id=All&field_reporting_level_target_id=28&field_program_target_id=All&field_file_spec_target_id=All&field_data_group_id_target_id=All&combine=
39. Statistics NCfE. ElSi Table Generator. U.S. Department of Education. 2023. https://nces.ed.gov/ccd/elsi/tablegenerator.aspx
40. Dee TS. Higher Chronic Absenteeism Threatens Academic Recovery from the COVID-19 Pandemic. 2023
41. DiMarco B. Tracking State Trends in Chronic Absenteeism. FutureEd. Accessed 3/9/2026, https://www.future-ed.org/tracking-state-trends-in-chronic-absenteeism/
42. U.S. Department of Education. Data Illuminates the Extent of Chronic Absenteeism. Accessed 3/9/2026, https://www.ed.gov/teaching-and-administration/supporting-students/chronic-absenteeism
43. National Center for Health Statistics NVSS. Restricted-use Vital Statistics Data. Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm
44. US Census Bureau. Population and Housing Unit Estimates. https://www.census.gov/programs-surveys/popest.html
45. Lilienfeld D, Stolley P. *Foundations of Epidemiology*. 3rd ed. Oxford University Press; 1994.
46. Office of Public Health Assessment. Confidence intervals in public health. Utah Department of Health, . Accessed January 9, 2018, http://health.utah.gov/opha/IBIShelp/ConfInts.pdf
47. Nolte E, McKee M. *Does health care save lives? Avoidable mortality revisited*. The Nuffield Trust; 2004.
48. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA: a cancer journal for clinicians*. 2017;67(3):177-193.
49. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention. Guide to ICD-9-CM and ICD-10 codes related to poisoning and pain. Accessed January 22, 2018, https://stacks.cdc.gov/view/cdc/59394
50. Henry J Kaiser Family Foundation. Opioid Overdose Death Rates and All Drug Overdose Death Rates per 100,000 Population (Age-Adjusted). Accessed May 5, 2018, https://www.kff.org/other/state-indicator/opioid-overdose-death-rates/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D
51. Dranger E, Remington P. YPLL: A summary measure of premature mortality used in measuring the health of communities Wisconson Public Health & Health Policy Institute https://uwphi.wiscweb.wisc.edu/wp-content/uploads/sites/316/2018/06/issueBriefv05n07.pdf
52. Vohlonen I, Bäckmand H, Korhonen J. Potential years of life lost: The PYLL rate in monitoring the wellbeing of a population. Northen Dimension Partnership in Public Health and Social Well-being. http://www.ndphs.org/documents/2662/Vohlonen%20Ilkka%20PYLL_article_2007.pdf
53. Chiang CL. *The Life Table and Its Applications*. Krieger Pub Co; 1983.
54. 74 United States Life Tables, 2023 (National Vital Statistics Reports) (2025).
55. Roubal AM, Pollock EA, Gennuso KP, Blomme CK, Givens ML. Comparative Methodologic and Practical Considerations for Life Expectancy as a Public Health Mortality Measure. *Public Health Rep*. Mar-Apr 2022;137(2):255-262. doi:10.1177/0033354921999407
56. County Health Rankings and Rodmaps. Mortality and Life Expectancy Calculator. Updated 11/15/2019. Accessed 07/15/2025, 2025. https://www.countyhealthrankings.org/resources/mortality-and-life-expectancy-calculator
57. Public Health England. PHE Life Expectancy Calculator. Accessed 07/15/2025, 2025. https://fingertips.phe.org.uk/documents/phe%20life%20expectancy%20calculator.xlsm

58. Eayres D, Williams E. Evaluation of methodologies for small area life expectancy estimation. *Journal of Epidemiology & Community Health*. 2004;58(3):243-249.
59. Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski JA. US small-area life expectancy estimates project: methodology and results summary. 2018;
60. Silcocks P, Jenner D, Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. *Journal of Epidemiology & Community Health*. 2001;55(1):38-43.
61. Lo E, Vatnik D, Benedetti A, Bourbeau R. Variance models of the last age interval and their impact on life expectancy at subnational scales. *Demographic Research*. 2016;35:399-454.
62. Kotelchuck M. An evaluation of the Kessner Adequacy of Prenatal Care Index and a proposed Adequacy of Prenatal Care Utilization Index. *Am J Public Health*. Sep 1994;84(9):1414-20.
63. Centers for Disease Control and Prevention. User guide to the 2015 natality public use file. Accessed December 7, 2018, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2015.pdf
64. Centers for Disease Control and Prevention. User guide to the 2014 natality public use file. Accessed January 22, 2018, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2014.pdf
65. Centers for Disease Control and Prevention. User guide to the 2012 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2012.pdf
66. Centers for Disease Control and Prevention. User guide to the 2013 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2013.pdf
67. Centers for Disease Control and Prevention. User guide to the 2016 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2016.pdf
68. Centers for Disease Control and Prevention. User guide to the 2017 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2017.pdf
69. Zhang X, Holt JB, Lu H, et al. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *American journal of epidemiology*. 2014;179(8):1025-1033.
70. PLACES: Local Data for Better Health. Health outcomes. Centers for Disease Control and Prevention. Updated December 8, 2020. Accessed December 8, 2020. https://www.cdc.gov/places/measure-definitions/health-outcomes/index.html
71. PLACES: Local Data for Better Health. Unhealthy behaviors. Centers for Disease Control and Prevention. Updated December 8, 2020. Accessed December 8, 2020. https://www.cdc.gov/places/measure-definitions/unhealthy-behaviors/index.html
72. PLACES: Local Data for Better Health. Prevention. Centers for Disease Control and Prevention. Updated December 8, 2020. Accessed December 8, 2020. https://www.cdc.gov/places/measure-definitions/prevention/index.html

73. PLACES: Local Data for Better Health. Current Release Notes: 2021. https://www.cdc.gov/places/help/data-notes/index.html
74. 500 Cities: Local Data for Better Health. Local Data for Better Health, 2017 Release. Updated December 4, 2017. Accessed September 27, 2017. https://chronicdata.cdc.gov/500-Cities-Places/500-Cities-Local-Data-for-Better-Health-2017-relea/vurf-k5wr
75. 500 Cities: Local Data for Better Health. Local Data for Better Health, 2018 Release. Updated January 4, 2021. Accessed December 4, 2018. https://chronicdata.cdc.gov/500-Cities-Places/500-Cities-Local-Data-for-Better-Health-2018-relea/rja3-32tc
76. 500 Cities: Local Data for Better Health. Local Data for Better Health, 2016 Release. Updated January 4, 2021. Accessed May 6, 2019. https://chronicdata.cdc.gov/500-Cities-Places/500-Cities-Local-Data-for-Better-Health-2016-relea/9z78-nsfp
77. PLACES: Local Data for Better Health. Local Data for Better Health, 2020 Tract Release. Accessed December 8, 2020, https://chronicdata.cdc.gov/500-Cities-Places/PLACES-Census-Tract-Data-GIS-Friendly-Format-2020-/yjkw-uj5s
78. PLACES: Local Data for Better Health. Local Data for Better Health, 2020 Place Release. Accessed December 8, 2020, https://chronicdata.cdc.gov/500-Cities-Places/PLACES-Place-Data-GIS-Friendly-Format-2020-release/vgc8-iyc4
79. PLACES: Local Data for Better Health. Local Data for Better Health, 2020 County Release. Accessed December 8, 2020, https://chronicdata.cdc.gov/500-Cities-Places/PLACES-County-Data-GIS-Friendly-Format-2020-releas/i46a-9kgh
80. National Center for Health Statistics. U.S. Small-Area Life Expectancy Estimates Project - USALEEP: Life Expectancy Estimates File for United States, 2010-2015. National Center for Health Statistics,. Updated March 6, 2020. Accessed April 8, 2020. https://www.cdc.gov/nchs/nvss/usaleep/usaleep.html
81. National Center for Health Statistics. U.S. Small-Area Life Expectancy Estimates Project - USALEEP: Life Expectancy Estimates File for Maine, 2010-2015. National Center for Health Statistics,. Updated March 6, 2020. Accessed April 8, 2020. https://www.cdc.gov/nchs/nvss/usaleep/usaleep.html
82. National Center for Health Statistics. U.S. Small-Area Life Expectancy Estimates Project - USALEEP: Life Expectancy Estimates File for Wisconsin, 2010-2015. National Center for Health Statistics,. Updated March 6, 2020. Accessed April 8, 2020. https://www.cdc.gov/nchs/nvss/usaleep/usaleep.html
83. U.S. Department of Health & Human Services. U.S. Life Expectancy at Birth by State and Census Tract - 2010-2015. Centers for Disease Control and Prevention,. Updated April 21, 2022. https://catalog.data.gov/dataset/u-s-life-expectancy-at-birth-by-state-and-census-tract-2010-2015