Congressional District Health Dashboard Technical Document

Updated March 10, 2023

Table of Contents

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.com) with general feedback or questions not addressed below.

Metric Overview

Domain	Metric Name	Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Clinical Care	Dental Care	Percentage of adults who report visiting a dentist in the past year	PLACES Project, Centers for Disease Control	Not Available	From Tract
	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	Race/Ethnicity	From County
	Preventive Services, 65+	Percentage of adults ≥65 years who are up to date on a core set of clinical preventive services	PLACES Project, Centers for Disease Control	Sex	From Tract
	Routine Checkup, 18+	Percentage of adults who report visiting a doctor for routine checkup in the past year	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Uninsured	Percentage of population ≤64 years without health insurance	American Community Survey	Age, Sex, Race/Ethnicity	From Tract
Health Behavior	Binge Drinking	Percentage of adults who report binge drinking in the past 30 days	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Physical Inactivity	Percentage of adults who report no leisure-time physical activity in the past 30 days	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Smoking	Percentage of adults who report current smoking	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Teen Births	Births to females 15-19 years per 1,000 females in that age group	Natality Data, National Vital Statistics System, National Center for Health Statistics	Race/Ethnicity	From County
Health Outcomes	Breast Cancer Deaths	Deaths due to breast cancer in females per 100,000 female population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Race/Ethnicity	From County
	Cardiovascular Disease Deaths	Deaths due to cardiovascular disease per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Sex, Race/Ethnicity	From County
	Colorectal Cancer Deaths	Deaths due to colorectal cancer per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Sex, Race/Ethnicity	From County

Domain	Metric Name	Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Health Outcomes	COVID Local Risk Index	Index (1-10) developed by the Dashboard, reflecting local social and economic factors and health outcomes for COVID risk	ACS, PLACES, CDC Social Vulnerability Index	Not Available	From Tract
	Diabetes	Percentage of adults who report having diabetes	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Firearm Homicides	Deaths due to firearm homicide per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Sex, Race/Ethnicity	From County
	Firearm Suicides	Deaths due to firearm suicide per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Sex, Race/Ethnicity	From County
	Frequent Mental Distress	Percentage of adults who report ≥14 days of poor mental health in the past 30 days	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Frequent Physical Distress	Percentage of adults who report ≥14 days of poor physical health in the past 30 days	PLACES Project, Centers for Disease Control	Not Available	From Tract
	High Blood Pressure	Percentage of adults who report high blood pressure	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Life Expectancy	Average years of life expectancy at birth	U.S. Small-Area Life Expectancy Estimates Project (USALEEP)	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	Race/Ethnicity	From County
	Obesity	Percentage of adults who report a body mass index (BMI) ≥30 kg/m2	PLACES Project, Centers for Disease Control	Not Available	From Tract
	Opioid Overdose Deaths	Deaths due to opioid overdose per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Sex, Race/Ethnicity	From County
	Premature Deaths (All Causes)	Years of potential life lost before age 75 per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	Sex, Race/Ethnicity	From County
Physical Environment	Air Pollution - Particulate Matter	Average daily concentration ($\mu g/m^3$) of fine particulate matter (PM2.5) per cubic meter of air throughout a year	Community Multiscale Air Quality model, US Environmental Protection Agency	Not Available	From Tract
	Housing with Potential Lead Risk	Percentage of housing stock with potential elevated lead risk	American Community Survey	Not Available	From Tract
	Lead Exposure Risk Index	Index (1-10) reflecting poverty-adjusted risk of housing-based lead exposure	American Community Survey	Not Available	From Tract
	Limited Supermarket Proximity	Percentage of the population living more than ½ mile from the nearest large grocery store	Food Access Research Atlas, Economic Research Service, United States Department of Agriculture	Race/Ethnicity	From Tract

Domain	Metric Name	Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Social and Economic Factors	Broadband Connection	Percentage of households with high speed broadband internet connection (cable, fiber optic, DSL)	American Community Survey	Not Available	From Tract
	Children in Poverty	Percentage of children living in households ≤100% of the federal poverty level	American Community Survey	Race/Ethnicity	From Tract
	High School Completion	Percentage of adults ≥25 years with high school diploma or equivalent, or higher degree	American Community Survey	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	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	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	Not Available	From Tract
	Rent Burden	Percentage of households where ≥30% of income is spent on rent	American Community Survey	Not Available	From Tract
	Unemployment	Percentage of population ≥16 years who are unemployed but seeking work	American Community Survey	Sex, Race/Ethnicity	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

Jacqueline Betro, MPA	Senior Project Coordinator
Samantha Breslin, MPA	Senior Program Coordinator
Alexander Chen, MPH	Data Analyst
Marc N. Gourevitch, MD, MPH	Senior Co-Principal Investigator Muriel and George Singer Professor of Population Health and Chair, Department of Population Health, NYU Langone Health
Neil Kleiman, PhD	City Policy/Partnerships Co-Principal Investigator Professor of Practice & Senior Fellow, Northeastern University Burnes Center for Social Change
Taylor Lampe, MPH	Senior Data Analyst
Yuruo Li, PhD, MPH	Senior Data Analyst
Isabel Nelson, MPH	Senior Data Analyst
Becky Ofrane, MPH	Co-Director, Program and Administration
Caleigh Paster	Program Coordinator
Ben Spoer, PhD, MPH	Co-Director, Metrics and Analytics
Jay Stadelman, MPH	Data Analyst
Lorna E. Thorpe, PhD	Methods Co-Principal Investigator Professor of Epidemiology, NYU School of Medicine Director of the Division of Epidemiology, NYU School of Medicine Vice Chair of Strategy and Planning, Department of Population Health, NYU Langone Health
Anne Vierse, MS	Data Analyst
Noah Zazanis, MS	Data 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 Downloadable Data Codebook, 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 118th Congress

The Dashboard currently provides metric data for the 118th Congress. The 118th Congress is the session of Congress elected November 2022 and inaugurated in January 2023. This election is the first election with districts derived from the 2020 Census. At the time of analysis, the Census Bureau had not released 118th Congress geography information, so the Dashboard gathered geography information state-by-state. See Analytic Decisions for more.

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 Chocktaw 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.

Getting to Congressional District-level Data: Methods and Approaches

Most publicly accessible data are not available at the congressional district level. For this reason, the Congressional District Health Dashboard derives congressional district estimates from source geographies (census tracts and counties). 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]

Creating Block Equivalency Files

In order to derive 118th congressional district estimates, the Dashboard team first acquired information linking 118th congressional districts to 2020 Census blocks. At the time of metric analysis, the Census Bureau had not released 118th Congress geography products. Instead, the Dashboard team obtained block equivalency files (or spatial files, when necessary) directly from each state. A more detailed list of the state files and the location of access can be found in the Appendix.

Once the Census released national block equivalency files for the 118th Congress,[5] the Dashboard team validated our state-based files against the national block equivalency file and found 99.9999% of blocks were matched to the correct congressional district. Please email info@CDhealthdashboard.org for more information.

Formula

First, we combined 118 block equivalency files with 2020 block population counts. Then we summed these block population counts to create population crosswalks which represent the proportion of each source geography's (tract or county) population that overlaps with congressional districts.

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)* 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.

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

We accessed 2020 block population counts from the 2020 Decennial Census P2 Table *(Hispanic or Latino, and Not Hispanic or Latino, by Race)* and used these block counts to derive our proportions for different weights, using these variables:

Population Weight	Variable(s)	Variable Definition(s)
Total	P2_001N	*Total Population*
Hispanic	P2_002N	*Hispanic or Latino*
White	P2_006N	*White, not Hispanic or Latino*
Black	P2_006N	*Black, not Hispanic or Latino*
Asian	P2_008N P2_009N	*Asian, not Hispanic or Latino* *Native Hawaiian or Other Pacific Islander, not Hispanic or Latino*
Other	P2_007N P2_010N P2_011N	*American Indian or Alaska Native, not Hispanic or Latino* *Some Other Race, not Hispanic or Latino* *2 or More Races*

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 portion overlapping with the congressional district. This adjustment is to properly reflect the non-proportional count estimate.

Selecting Population Weights

The Dashboard selects population weights that most closely match the underlying population of the source estimate being aggregated. If a matching variable is unavailable from the 2020 Decennial Census, a total population weight is used. At launch, sex and age block data are not yet available from the 2020 Census.

Selecting Source Geographies

The Dashboard uses census tract or county estimates to derive congressional district estimates, 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 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.

Transforming 2010 Estimates into 2020 Congressional Districts

Some metrics presented on the Dashboard represent pre-2020 data, and are only available in 2010 Census vintages. To aggregate these data into 118th 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.[6] This weight represents the expected proportion of the 2010 block's population and housing units located in each 2020 block.

Combining Multiple Variables

Some metrics require combining multiple variables to calculate the final estimate. If one variable is missing (NA) then we exclude it from the final estimate calculation. Unless otherwise noted in metric-specific sections, the Dashboard calculates the full metric at the source geography, and then aggregates to congressional districts.

Estimate Precision and Margins of Error

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

Censoring/Flagging 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. "Sufficient" criteria were developed by the Dashboard team using the percent of the congressional district's population that is missing source geography data; please email info@CDhealthdashboard.org for more information.

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.

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

For metrics with multiple subgroup/component variables, populations for missing subgroup/component variables may not contribute to censoring or flagging. Censored estimates are removed from the website and downloadable data. Flagged estimates are noted in "Tips and Cautions for using the Data", or in downloadable data.

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

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.

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)).[7] Federal guidelines for reporting data by demographics[7] 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.

State + National Estimates

State and national estimates on the Dashboard represent the unweighted average of congressional district estimates by metric and year for the total population. Estimates for identical or similar metrics that use the state or 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.

District Snapshot: District Facts

Demographic estimates on the District Facts page are sourced from the 2020 Decennial Census (table P2, which represents non-Hispanic or Latino single race categories). This is the same source used by many congressional district offices to report demographics. An exception is age categories. As of Dashboard release date, age breakdowns are not yet released for the 2020 Decennial Census. Instead, age estimates are presented using 2020 ACS data (table DP05) and calculated from tract level via the method described in the section "Getting to Congressional District-level Data: Methods and Approaches". See variable selection below.

Demographic Group	**Variable(s)**
Total population	P2_001N
American Indian and Alaska Native	P2_007N
Asian	P2_008N
Black	P2_006N
Hispanic	P2_002N
Native Hawaiian or Other Pacific Islander	P2_009N
Other race alone	P2_010N
White, non-Hispanic	P2_006N
Two or more races	P2_011N
Age 0-17	DP05_0019E
Age 18-64	DP05_0001E - DP05_0024E - DP05_0019E
Age 65+	DP05_0024E

Website Maps

Congressional district website maps were created by combining Dashboard-created national block equivalency file with 2020 block spatial files. Block shapes were dissolved into congressional districts, then ocean and great lake shorelines were removed using a 5m national cartographic map from the US Census.[8]

Analytic Software

All analyses were performed in R using tidyverse, tidycensus, tigris, and sf packages, among others.[8-12]

SECTION 4: Metric Analyses, by Data Source

American Community Survey

General Notes

The American Community Survey (ACS) is administered by the US Census Bureau[13]. Data are retrieved from the census API using R and the tidycensus package.[14, 10] 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.

Geography-Specific Notes

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.

Broadband Connection

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of households with high speed broadband internet connection (cable, fiber optic, DSL)	American Community Survey	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

Analysis

No additional analysis was conducted by the Dashboard.

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on aggregating source geography estimates to congressional districts.

Children in Poverty

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of children living in households ≤100% of the federal poverty level	American Community Survey	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

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 aggregating source geography estimates to congressional districts.

High School Completion

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults ≥25 years with high school diploma or equivalent, or higher degree	American Community Survey	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. 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 by sex and for total population:

- 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

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

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 aggregating source geography estimates to congressional districts.

Income Inequality

Metric Description	Data Source	Demographic Subgroups	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	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. [15]

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.

Cut points were selected from table B19001 to most closely represent the 20th and 80th household income percentiles[28], as reported by US Census Bureau data Table H-1 (All Races). [16]

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 aggregating source geography estimates to congressional districts.

Housing with Potential Lead Risk

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of housing stock with potential elevated lead risk	American Community Survey	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
- Estimate!!Total:!!Built 2014 or later

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.[17] Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.[18] The Dashboard adapted Vox Media's Python code available on Github[19] for the present analysis, which was conducted by the Dashboard using R v4.1.0 and validated using Python v3.6.[20] Users should note that differences in rounding programming between the two software programs resulted in some minor but appreciable differences in housing risk score. The Washington State Department of Health's analysis uses variables from 2014. [17] In updating the analysis to represent all housing stock built in 2010 or later for years subsequent to 2014 using table B25034, variables were added for housing stock built in 2010-2013 and in 2014 or later.

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.[19] We count the number of housing units in each of five time periods: pre-1938, 1940-59, 1960-79, 1980-99, and 2000 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 Jacobs 2002).[21] This results in an overall percent of housing likely to have some risk of lead exposure.

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

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on aggregating source geography estimates to congressional districts.

Lead Exposure Risk Index

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Index (1-10) reflecting poverty-adjusted risk of housing-based lead exposure	American Community Survey	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
- Estimate!!Total:!!Built 2014 or later

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

The lead analysis was performed as per methodology initially developed by the Washington State Department of Health.[17] Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.[18] The Dashboard adapted Vox Media's Python code available on Github[19] for the present analysis, which was conducted by the Dashboard using R v4.1.0 and

validated using Python v3.6.[20] Users should note that differences in rounding programming between the two software programs resulted in some minor but appreciable differences in housing risk score. The Washington State Department of Health's analysis uses variables from 2014. [17] In updating the analysis to represent all housing stock built in 2010 or later for years subsequent to 2014 using table B25034, variables were added for housing stock built in 2010-2013 and in 2014 or later.

We took the Dashboard Housing with Potential Lead Risk metric (see above) 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 Jacobs 2002[21], 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

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.

Racial/Ethnic Diversity

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Index (0-100) reflecting how evenly distributed the population is across the racial/ethnic groups living in this area	American Community Survey	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). [22]

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

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:

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

Where:

$\boldsymbol{\pi_r}$ refers to a particular racial/ethnic group's proportion of the geography population[22]

As per footnote 5 in Iceland, [22] $\ln\left[\frac{1}{\pi_r}\right]$ is set to 0 when the proportion of a particular group is in a given geography π_r is 0.

Geography-Specific Notes

Geographic aggregation for Racial/Ethnic Diversity is accomplished in a slightly different manner than other metrics due to the nature of the metric calculation. First, DP05 population counts for each racial/ethnic group are aggregated from tract to generate congressional district population counts for each race/ethnic group. 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 entropy/diversity.

Neighborhood Racial/Ethnic Segregation

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Index (0-100) reflecting the geographic clustering of racial/ethnic groups across the area	American Community Survey	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.[22]

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."[22] 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 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:

π_r refers to a particular racial/ethnic group's proportion of the geography's population[22]

As per footnote 5 in Iceland, [22] $\ln\left[\frac{1}{\pi_r}\right]$ is set to 0 when the proportion of a particular group is in a given geography π_r is 0.

Geography-Specific Notes

Geographic aggregation for Neighborhood Racial/Ethnic Segregation is accomplished in a slightly different manner than other metrics due to the nature of the metric calculation. First, DP05 population counts for each racial/ethnic group are aggregated from tract to generate congressional district population counts for each race/ethnicity group. These counts are then used to calculate each racial/ethnic group's proportion of the congressional district (π_r). This value is used in the formula above for congressional district entropy/diversity.

Additionally, 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 summed DP05 counts, to account for the appropriate geographic overlap of each tract in a congressional district.

Rent Burden

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of households where ≥30% of income is spent on rent	American Community Survey	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)

Analysis

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

See above "Geography-Specific Notes" section under "American Community Survey: General Notes" for information on aggregating source geography estimates to congressional districts.

Unemployment

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of population ≥16 years who are unemployed but seeking work	American Community Survey	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 and sex.

The following variables were used to represent Unemployment for total population and by race for White, Hispanic, and Black:

- 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)

Unemployment by race for Asian and Other is represented by the weighted average of the following variables across the racial subcategories 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

Weights are calculated from the relative proportion of each racial subcategory within the summed total population of the full group as per ACS table DP05, using the following variables:

Asian:

- Estimate!!RACE!!Total population!!One race!!Asian
- Estimate!!RACE!!Total population!!One race!!Native Hawaiian and Other Pacific Islander

Other:

- Estimate!!RACE!!Total population!!One race!!American Indian and Alaska Native
- Estimate!!RACE!!Total population!!One race!!Some other race
- Estimate!!RACE!!Total population!!Two or more races

The following variables were used to represent Unemployment by sex. Please note the different age category availability for sex-specific estimates:

- Estimate!!Unemployment rate!!Population 20 to 64 years!!SEX!!Male
- Estimate!!Unemployment rate!!Population 20 to 64 years!!SEX!!Female

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 aggregating source geography estimates to congressional districts.

Uninsured

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of population ≤64 years without health insurance	American Community Survey	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.

Total population

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

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 aggregating source geography estimates to congressional districts.

Community Multiscale Air Quality model, US Environmental Protection Agency (CMAQ, EPA)

General Notes

Data represent modeled estimates produced by CMAQ and do not include estimates for Alaska and Hawaii. An interactive map of locations of active air quality monitors for PM2.5 is available online, through the EPA.[23]

Air Pollution – Particulate Matter

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Average daily concentration of fine particulate matter (PM2.5) per cubic meter of air throughout a year	Community Multiscale Air Quality model, US Environmental Protection Agency	Not Available	From Tract

Data Table(s) + Variable(s)

Tract level data are downloaded from the US Environmental Protection Agency website.[24-27] CONUS PM 2.5 Daily Average files are used in analyses. The estimate variable is labeled as "pm25_daily_average_ug_m3".

Analysis

An annual average of daily concentration for each census tract was calculated, to control for seasonal variation in air pollution. This annual average represents the metric Air Pollution – Particulate Matter. Please refer to the EPA for more information on calculation methods.[28]

Geography-Specific Notes

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

Food Access Research Atlas, Economic Research Service, United States Department of Agriculture (USDA ERS Food Atlas)

General Notes

The Limited Supermarket Proximity metric represents the percent of the population that lives ≥ ½ mile from a supermarket, large grocery store, or supercenter. Data on store locations were obtained by USDA ERS from the 2019 STARS directory of stores authorized to accept SNAP benefits and the 2019 Trade Dimensions TDLinx directory of stores.[29] Stores were defined as a healthy food outlet if they reported at least $2 million in annual sales, were certified to accept SNAP benefits, and sold food in all the following categories: fresh produce, fresh meat and poultry, dairy, dry and packaged goods, and frozen foods. Population data total and by race/ethnicity were obtained by USDA ERS at the block level from US Census.

Race/Ethnicity Definition

Data reported by USDA ERS included estimates by the following race/ethnicity categories: White, Black, Asian, Native Hawaiian/Pacific Islander, American Indian/Alaska Native, Other or two or more races, and Hispanic. Estimates for Asian and Native Hawaiian/Pacific Islander were combined to match the Dashboard Asian/PI designation (see section "Analytic Decisions: Metric Subgroup Race/Ethnicity Categories" for more). Similarly, estimates for other or two or more races and American Indian/Alaska Native were also combined. Note that for this metric Hispanic ethnicity is not mutually exclusive with other race/ethnicity categories.

Limited Supermarket Proximity

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of the population living more than ½ mile from the nearest large grocery store	Food Access Research Atlas, Economic Research Service, United States Department of Agriculture	Race/Ethnicity	From Tract

Data Table(s) + Variable(s)

Tract-level data were downloaded directly from USDA ERS for Limited Supermarket Proximity.[30]

The following variables were used to calculate numerators. See section "Analytic Decisions: Metric Subgroup Race/Ethnicity Categories" for how we combine race/ethnicity subgroups:

- Lapophalf = Population count beyond 1/2 mile from supermarket
- lawhitehalf = White population count beyond 1/2 mile from supermarket
- lablackhalf = Black or African American population count beyond 1/2 mile from supermarket
- laasianhalf = Asian population count beyond 1/2 mile from supermarket
- lanhopihalf = Native Hawaiian or Other Pacific Islander population count beyond 1/2 mile from supermarket
- laaianhalf = American Indian or Alaska Native population count beyond 1/2 mile from supermarket
- laomultirhalf = Other/Multiple race population count beyond 1/2 mile from supermarket
- lahisphalf = Hispanic or Latino ethnicity population count beyond 1/2 mile from supermarket

The following variables were used to calculate denominators. See section "Analytic Decisions: Metric Subgroup Race/Ethnicity Categories" for how we combine race/ethnicity subgroups:

- POP2010 = tract total Population count from 2010 census (number)
- TractWhite = Total count of White population in tract (number)
- TractBlack = Total count of Black or African American population in tract (number)
- TractAsian = Total count of Asian population in tract (number)
- TractNHOPI = Total count of Native Hawaiian and Other Pacific Islander population in tract (number)
- TractAIAN = Total count of American Indian and Alaska Native population in tract (number)
- TractOMultir = Total count of Other/Multiple race population in tract (number)

Analysis

$$\text{Limited Supermarket Proximity} = \frac{\text{Population living} \geq 0.5 \text{ mile from a healthy food store}}{\text{Total population}} \times 100\%$$

Geography-Specific Notes

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

Multiple Data Sources: COVID Local Risk Index

General Notes

The Dashboard team created the COVID Local Risk Index to identify areas susceptible to both higher numbers of COVID cases and more severe COVID cases. The metric was originally developed in response to the COVID-19 pandemic for cities and neighborhoods and is intended to assist public health practitioners in allocating resources to help address the impact of COVID-19.

The index is calculated using data from the U.S. Census American Community Survey (ACS) 2019 5 Year Estimates, the PLACES Project 2019 modeled health outcomes data (2021 release),[31] and informed by the methods and variables used in the Center for Disease Control and Prevention's Social Vulnerability Index (SVI).[32-34] Note: The 2021 PLACES Project release did not include data for New Jersey, due to incomplete data collection. Therefore, New Jersey cities and census tracts use 2018 data.[35,36]

Calculations were completed by Dashboard analytic staff under scientific guidance from Dr. Ben Spoer, Dr. Lorna Thorpe, and Dr. Marc Gourevitch. Methodology was informed by other indices.[37-39] Please note that this index is informed by the best available scientific evidence as of that date; the index's components and weighting may be updated in the future as what is known about COVID changes.

This index was validated against COVID-19 case and death rates for selected cities,[40] but not for all geographies, as no data were available that directly captured these values.

The index represents three conceptual areas (themes):

- **Social vulnerability** The CDC's SVI was selected because it is a validated, peer-reviewed representation of a community's ability to prevent human suffering in the event of a disaster, including disease outbreak.[32] SVI is a well-established and validated index in the scientific literature for emergency preparedness and other health outcomes and has been shown to be associated with COVID outcomes.[34]
- **COVID-related chronic health conditions** were selected because they are known risk factors for COVID and are not included in the original SVI.[41-51] Health conditions with high-quality evidence of increased risk of COVID incidence, morbidity and mortality that were available as modeled estimates from the PLACES Project dataset were included; health outcomes with equivocal evidence were excluded.
- **COVID-related demographics** were selected because of strong evidence that these demographics are at higher risk.[41-51] Some of these demographics already exist in SVI, but were duplicated in this theme to reinforce their prominence within the Dashboard's index.

COVID Local Risk Index

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Index (1-10) developed by the Dashboard, reflecting local social and economic factors and health outcomes for COVID risk	ACS, PLACES, CDC Social Vulnerability Index	Not Available	From Tract

Data Table(s) + Variable(s)

Table: COVID Local Risk Index Data Sources and Conceptual Components

Theme	Data Source	Conceptual Components
Social Vulnerability	American Community Survey (ACS), 2019* 5 Year Estimates (*Variable selection guided by CDC Social Vulnerability Index*)[33,38]	**Group 1: Socioeconomic Status** • Persons below poverty • Civilian (age 16+) unemployed • Per capita income • Persons (aged 25+) with no high school diploma **Group 2: Household Composition & Disability** • Persons aged 65+ • Persons aged 17 and younger • Civilian non-institutionalized population with a disability • Single parent household with children under 18 **Group 3: Race/Ethnicity Status & Language** • All persons except white, non-Hispanic • Persons (age 5+) who speak English "less than well" **Group 4: Housing Type & Transportation** • Housing in structures with 10+ units • Mobile homes • At household level (occupied housing units), more people than rooms • Households with no vehicle available • Persons in institutionalized group quarters
COVID-related Chronic Health Conditions	PLACES Project, 2019* 1 Year Modeled Estimates[31]	• Chronic obstructive pulmonary disease (COPD) among adults aged 18+[43,47,49,50] • Coronary heart disease among adults aged 18+[45,46,48-51] • Diagnosed diabetes among adults aged 18+[43-45,47-49] • Chronic kidney disease among adults aged 18+[41-43,46,47] • Obesity among adults aged 18+[41-44,47,49]
COVID-related Demographics	ACS, 2019* 5 Year Estimates (Table DP05)	• All persons except non-Hispanic white [41,43-46,48,49] • Persons aged 75 to 84[41,42,44-48,50,51] • Persons aged 85+[41,44-48,50,51]

*2018 data are used for New Jersey cities and census tracts[35]

Analysis

The Dashboard adapts the analytic strategy proposed by the CDC's Social Vulnerability Index[32,38] which orders each component's estimates across geographies and assigns the highest percentile rank (100) to the highest value (with the exception of the per capita income component, which assigns the highest percentage rank to the lowest value).[38,52]

The percentile rank of each component's estimate is then multiplied by the component weight. Each component weight sums to the overall theme weight. To maintain fidelity to the equation established by

the CDC's original analysis, SVI components contribute equally to the social vulnerability theme.[32,38] Social vulnerability was *a priori* assigned a theme weight of 30% to shift weight in our March 2021 update towards health conditions and demographic factors. This decision was informed by an in-depth literature review, current as of December 2020, to identify risk factors of severe COVID outcomes. Component weights were developed based on effect sizes found through this literature review and incorporate relative prevalence in the United States. Additional guidance from Dr. Ben Spoer, Dr. Marc Gourevitch, and Dr. Lorna Thorpe informed weighting scheme decisions.

Table: COVID Local Risk Index Theme and Component Weights

Theme	Theme Weight	Component	Component Weight within COVID Local Risk Index
Social Vulnerability	30%	*See list above for complete list (15 components)*	2% per component
COVID-related Chronic Health Conditions	42%	Chronic obstructive pulmonary disease (COPD) among adults aged 18+	4%
		Coronary heart disease among adults aged 18+	5%
		Diagnosed diabetes among adults aged 18+	6%
		Chronic kidney disease among adults aged 18+	9%
		Obesity among adults aged 18+.	18%
COVID-related Demographics	28%	All persons except non-Hispanic white	12%
		Persons aged 75 to 84	11%
		Persons aged 85+	5%

Then, the weighted estimates are summed to establish a "sum of percentiles". This "sum of percentiles" is then categorized into deciles, which is reported as the COVID Local Risk Index.

The formula for the Dashboard's COVID Local Risk Index is:

$$\text{COVID Local Risk Index} = \text{Decile of} \sum_{i=1}^{n} (\text{Percentile rank of component estimate relative to other geographies}) * (\text{Component weight})$$

Where:
n = the number of geographies represented on the Dashboard

Geography-specific notes

The COVID Local Risk index first calculates the "sum of weighted percentiles" at the census tract level, aggregates to the congressional district level, and then categorizes these aggregated scores to generate the decile index. See "SECTION 3: Analytic Decisions" for more details on this method.

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.[53] 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.

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).[54] 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.

Race/Ethnicity Definition

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: Non-Hispanic Black;
- Hispanic: Mexican, Puerto Rican, Cuban, Central or South American and other or unknown Hispanic;
- 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: Non-Hispanic Black;
- Hispanic: Mexican, Puerto Rican, Cuban, Central and South American, and other and unknown Hispanic;
- 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: NHBA
- Hispanic: H
- Asian: NHAA or NHNA
- Other: NHIA or NHTOM
- Total: TOT

Standardized Population Weights

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.

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	Number	Weight	YPLL-75 weight	Standard life expectancy at age of deaths (years)
Total	308745538			
< 5 years	20201362	0.0654	0.0696	72.5
5 to 9 years	20348657	0.0659	0.0701	67.5
10 to 14 years	20677194	0.0670	0.0713	62.5
15 to 19 years	22040343	0.0714	0.0760	57.5
20 to 24 years	21585999	0.0699	0.0744	52.5
25 to 29 years	21101849	0.0683	0.0727	47.5
30 to 34 years	19962099	0.0647	0.0688	42.5
35 to 44 years	41070606	0.1330	0.1415	35
45 to 54 years	45006716	0.1458	0.1551	25
55 to 64 years	36482729	0.1182	0.1257	15
65 to 74 years	21713429	0.0703	0.0748	5
75 to 84 years	13061122	0.0423	0	0
85 years and over	5493433	0.0178	0	0

Geography-Specific Notes

Age-adjusted or crude rates were calculated at the county level for all metrics and then aggregated from the county level to generate congressional district estimates.[55] See "SECTION 3: Analytic Decisions" for more details on this method.

True state values were used for 6 at-large districts including Alaska, Delaware, South Dakota, North Dakota, Vermont and Wyoming.

Breast Cancer Deaths

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Deaths due to breast cancer in females per 100,000 female population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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, 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 aggregating source geography estimates to congressional districts.

Cardiovascular Disease Deaths

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Deaths due to cardiovascular disease per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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[56] 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 aggregating source geography estimates to congressional districts.

Colorectal Cancer Deaths

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Deaths due to colorectal cancer per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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[57] 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 aggregating source geography estimates to congressional districts.

Firearm Homicides

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Deaths due to firearm homicide per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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 aggregating source geography estimates to congressional districts.

Firearm Suicides

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Deaths due to firearm suicide per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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 aggregating source geography estimates to congressional districts.

Opioid Overdose Deaths

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Deaths due to opioid overdose per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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.[58,59]

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 aggregating source geography estimates to congressional districts.

Premature Deaths (All Causes)

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Years of potential life lost before age 75 per 100,000 population	Multiple Cause of Death Data, National Vital Statistics System, National Center for Health Statistics	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.[60]

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)

See above "Geography-Specific Notes" section under "National Vital Statistics System: General Notes" for information on aggregating source geography estimates to congressional districts.

Low Birthweight

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of live births with low birthweight (<2500 grams)	Natality Data, National Vital Statistics System, National Center for Health Statistics	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 aggregating source geography estimates to congressional districts.

Prenatal Care

Metric Description	Data Source	Demographic Subgroups	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	Race/Ethnicity	From County

Data Table(s) + Variable(s)

Prenatal Care estimates represent a slight modification of one component of the Kotelchuck Index.[61] 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. For more information please refer to the natality public use data documentation files.[62-67]

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 aggregating source geography estimates to congressional districts.

Teen Births

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Births to females 15-19 years per 1,000 females in that age group	Natality Data, National Vital Statistics System, National Center for Health Statistics	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 aggregating source geography estimates to congressional districts.

PLACES, Centers for Disease Control and Prevention

General Notes

PLACES apply a multi-level regression with post-stratification (MPR) approach to develop small area estimates (SAE) 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).[68] For more information regarding these metrics, please refer to the PLACES's methodology pages.[69-71]

The Dashboard reports most data as received, with the exception of the preventive service utilization.[36] Some estimates for New Jersey are not available for measures based on BRFSS 2019,[72] as the state did not collect enough BRFSS data to meet minimum requirements for inclusion, and therefore 2018 data are presented.

Geography-Specific Notes

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

Binge Drinking

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report binge drinking in the past 30 days	PLACES Project, Centers for Disease Control	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.[73-78]

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 aggregating source geography estimates to congressional districts.

Dental Care

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report visiting a dentist in the past year	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Diabetes

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report having diabetes	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Frequent Mental Distress

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report ≥14 days of poor mental health in the past 30 days	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Frequent Physical Distress

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report ≥14 days of poor physical health in the past 30 days	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

High Blood Pressure

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report high blood pressure	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Obesity

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report a body mass index (BMI) ≥30 kg/m2	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Physical Inactivity

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report no leisure-time physical activity in the past 30 days	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Preventive Services, 65+

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults ≥65 years who are up to date on a core set of clinical preventive services	PLACES Project, Centers for Disease Control	Sex	From Tract

Data Table(s) + Variable(s)

Preventive Services, 65+ crude prevalence tract level data were downloaded directly from the PLACES website in the GIS friendly format.

Analysis

At the recommendation of the PLACES analytic team[102], overall Preventive Services, 65+ values were calculated as a weighted average of preventive service use by women and preventive service use by men. Per PLACES, we used 2010 Decennial Census Survey for tract level population counts.
The weighted proportion formula is below:

$$\widehat{p_{weighted}} = \frac{\hat{p}_{male\ 65+} * n_{male\ 65+} + \hat{p}_{female\ 65+} * n_{female\ 65+}}{n_{male\ 65+} + n_{female\ 65+}}$$

Where:

$p_{weighted}$ = weighted proportion of overall use of preventive services by men and women 65+

$p_{male\ 65+}$ = reported proportion of overall use of preventive services by men 65+ (from PLACES)

$p_{female\ 65+}$ = reported proportion of overall use of preventive services by women 65+ (from PLACES)

$n_{male\ 65+}$ = population, men 65+ (from 2010 DCS)

$n_{female\ 65+}$ = population, women 65+ (from 2010 DCS)

Geography-Specific Notes

Preventive Services, 65+ crude prevalence for female and male and the calculated weighted data are aggregated from the tract level to generate congressional district estimates. See "SECTION 3: Analytic Decisions" for more details on this method.

Routine Checkup, 18+

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report visiting a doctor for routine checkup in the past year	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

Smoking

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Percentage of adults who report current smoking	PLACES Project, Centers for Disease Control	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 aggregating source geography estimates to congressional districts.

United States Small-Area Life Expectancy Project (USALEEP)

General Notes

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.[79]

Life Expectancy

Metric Description	Data Source	Demographic Subgroups	Geography Aggregation
Average years of life expectancy at birth	U.S. Small-Area Life Expectancy Estimates Project (USALEEP)	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.

Geography-Specific Notes

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.

SECTION 5: Appendix

State-Based 118th Block Equivalency Acquisition

State Name	Date of Access	Method of Access	File Type
Alaska	Not Applicable	assigned to an at large district	NA
Alabama	June 7, 2022	received from Alabama Senate Reapportionment Office	.xlsx
Arkansas	June 13, 2022	received from Arkansas GIS Office	.txt
Arizona	February 11, 2022	downloaded from the Arizona Independent Redistricting Commission website	.txt
California	March 21, 2022	downloaded from the California Commission, We Draw the Lines CA website	.xlsx
Colorado	March 21, 2022	downloaded from the Colorado Independent Redistricting Commission website	.txt
Connecticut	February 25, 2022	downloaded from the Connecticut General Assembly website	.csv
District of Columbia	Not Applicable	assigned to an at large district	NA
Delaware	Not Applicable	assigned to an at large district	NA
Florida	June 10, 2022	downloaded from Florida Office of Economic and Demographic Research website	.txt
Georgia	June 7, 2022	downloaded from Georgia General Assembly website	.xlsx
Hawaii	June 30, 2022	received from Hawaii State Legislature	.csv
Iowa	March 21, 2022	downloaded from Iowa Legislature website	.csv
Idaho	March 21, 2022	downloaded from Idaho Legislature website	.csv
Illinois	March 30, 2022	downloaded from Illinois Redistricting website	shapefile converted to .Rdata
Indiana	March 21, 2022	received from Indiana Secretary of State's Office	shapefile converted to .Rdata
Kansas	June 13, 2022	received from Kansas Legislative Research Department (Redistricting)	.csv
Kentucky	June 7, 2022	downloaded from Kentucky General Assembly website	.xlsx
Louisiana	June 10, 2022	downloaded from Louisiana State Legislature website	.txt
Massachusetts	March 21, 2022	downloaded from Massachusetts State Legislature website	.xlsx
Maryland	June 7, 2022	downloaded from Maryland Department of Planning website	.xlsx
Maine	March 21, 2022	received from Maine State Legislature	shapefile converted to .Rdata
Michigan	June 7, 2022	downloaded from Michigan Independent Citizens Redistricting Commission website	.txt
Minnesota	March 21, 2022	downloaded from Minnesota Legislature website	.csv
Missouri	June 7, 2022	downloaded from Missouri House of Representatives website	.xlsx

State Name	Date of Access	Method of Access	File Type
Mississippi	March 21, 2022	downloaded from Mississippi Automated Resource Information System website	.xlsx
Montana	June 28, 2022	downloaded from Montana Districting and Apportionment Commission	.csv
North Carolina	June 7, 2022	downloaded from North Carolina General Assembly	.csv
North Dakota	Not Applicable	assigned to an at large district	NA
Nebraska	June 28, 2022	downloaded from Nebraska Legislature website	.csv
New Hampshire	June 10, 2022	downloaded from New Hampshire Office of Strategic Initiatives website	.csv
New Jersey	June 29, 2022	downloaded from New Jersey Redistricting Commission website	.txt
New Mexico	June 10, 2022	downloaded from New Mexico Legislature website	.xlsx
Nevada	June 7, 2022	downloaded from Nevada Legislature Website	.xlsx
New York	June 8, 2022	downloaded from New York State Legislative Task Force on Demographic Research and Reapportionment	.xlsx
Ohio	June 10, 2022	downloaded from Ohio Redistricting Commission website	.xlsx
Oklahoma	March 21, 2022	downloaded from Oklahoma State Legislature website	.xlsx
Oregon	March 21, 2022	downloaded from Oregon State Legislature website	.txt
Pennsylvania	June 7, 2022	downloaded from Pennsylvania Redistricting website	.xlsx
Rhode Island	March 2, 2022	downloaded from the Rhode Island Redistricting Project website	shapefile converted to .Rdata
South Carolina	June 7, 2022	downloaded from the South Carolina House of Representatives Redistricting website	.xlsx
South Dakota	Not Applicable	assigned to an at large district	NA
Tennessee	June 7, 2022	downloaded from the Redistricting Data Hub	shapefile converted to .Rdata
Texas	June 7, 2022	downloaded from the Texas Redistricting website	.csv
Utah	June 8, 2022	downloaded from the Utah State Legislature website	.txt
Virginia	March 22, 2022	downloaded from Virginia Redistricting Commission website	.txt
Vermont	Not Applicable	assigned to an at large district	NA
Washington	March 22, 2022	downloaded from the Washington State Redistricting Commission website	.csv
Wisconsin	March 30, 2022	downloaded from the Wisconsin Governor's website	.csv
West Virginia	June 28, 2022	downloaded from West Virginia Legislature website	shapefile converted to .Rdata
Wyoming	Not Applicable	assigned to an at large district	NA

Acknowledgements

The Dashboard acknowledges Allegra Wilson, Sarah Conderino, Miriam Gofine, Dr. Ben Spoer, Taylor Lampe, Peggy Hsieh, Alexander Chen, Dr. Yuruo Li, Dr. Shoshanna Levine, Isabel Nelson, Anne Vierse, Jay Stadelman, Noah Zazanis, 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
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. US Census Bureau. Data from: 118th Congress Block Equivalency Files. 2022.
6. 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*.
7. 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
8. Walker K. tigris: Download and use Census TIGER/Line shapefiles in R. https://github.com/walkerke/tigris
9. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/
10. Walker K, Herman M. tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames. . https://walker-data.com/tidycensus/.
11. 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
12. 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
13. 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
14. US Census Bureau. Available APIs. US Department of Commerce. Accessed January, 2020. https://www.census.gov/data/developers/data-sets.html
15. 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.
16. US Census Bureau. Table H-1. Income Limits for Each Fifth and Top 5 Percent (All Races): Table H-1. Income Limits for Each Fifth and Top 5 Percent of All Households: 1967 to 2018. (Households as of March of the following year. Income in current and 2018 CPI-U-RS adjusted dollars (28)). US Census Bureau,. Updated August 29, 2019. Accessed January 27, 2020. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html
17. 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

18. 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
19. 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
20. Foundation PS. Python Language Reference, version 3.6.
21. Jacobs DE, Clickner RP, Zhou JY, et al. The prevalence of lead-based paint hazards in US housing. *Environmental health perspectives*. 2002;110(10):A599.
22. 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
23. Environmental Protection Agency. Interactive Map of Air Quality Monitors. Updated October 27, 2015. Accessed May 7, 2018, https://www.epa.gov/outdoor-air-quality-data/interactive-map-air-quality-monitors#kmz
24. United States Environmental Protection Agency. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files: Downscaling Output files (files compressed in gzip format): 2013 (CONUS): PM 2.5 Daily Average. United States Environmental Protection Agency,. Updated April 24, 2019. Accessed May 1, 2018. https://www.epa.gov/hesc/rsig-related-downloadable-data-files#output
25. United States Environmental Protection Agency. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files: Downscaling Output files (files compressed in gzip format): 2014 (CONUS): PM 2.5 Daily Average. United States Environmental Protection Agency,. Updated April 24, 2019. Accessed January 28, 2019. https://www.epa.gov/hesc/rsig-related-downloadable-data-files#output
26. United States Environmental Protection Agency. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files: Downscaling Output files (files compressed in gzip format): 2015 (CONUS): PM 2.5 Daily Average. United States Environmental Protection Agency,. Updated April 24, 2019. Accessed May 7, 2019. https://www.epa.gov/hesc/rsig-related-downloadable-data-files#output
27. United States Environmental Protection Agency. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files: Downscaling Output files (files compressed in gzip format): 2016 (CONUS): PM 2.5 Daily Average. United States Environmental Protection Agency,. Updated August 28, 2019. Accessed October 29, 2019. https://www.epa.gov/hesc/rsig-related-downloadable-data-files#output
28. United States Environmental Protection Agency. Community Multiscale Air Quality Monitoring System (CMAQ) Models. https://www.epa.gov/cmaq/cmaq-models-0
29. United States Department of Agriculture Economic Research Service. Food Access Research Atlas Documentation. United States Department of Agriculture,. Updated December 5, 2017. Accessed February 28, 2019. https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/
30. United States Department of Agriculture Economic Research Service. Download the Data: Food Access Research Atlas Data Download 2015. United States Department of Agriculture,. Updated May 18, 2017. Accessed May 1, 2018. https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/
31. Centers for Disease Control and Prevention. About the Project - PLACES: Local Data for Better Health. Accessed February 25, 2021. https://www.cdc.gov/places/about/index.html

32. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. *Journal of homeland security emergency management*. 2011;8(1)
33. Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research Analysis a, Services Program,. Social Vulnerability Index 2018 Database US. Accessed May 25, 2020, https://svi.cdc.gov/data-and-tools-download.html
34. Nayak A, Islam SJ, Mehta A, et al. Impact of Social Vulnerability on COVID-19 Incidence and Outcomes in the United States. 2020;
35. PLACES: Local Data for Better Health. Current Release Notes: 2021. https://www.cdc.gov/places/help/data-notes/index.html
36. SAS Institute Inc. SAS System. Cary, NC 2015.
37. Surgo Foundation. The COVID-19 Community Vulnerability Index (CCVI). Accessed June 4, 2020, https://precisionforcovid.org/ccvi
38. Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research Analysis and Services Program. CDC SVI 2018 Documentation - 1/31/2020. Accessed May 25, 2020, https://svi.cdc.gov/Documents/Data/2018_SVI_Data/SVI2018Documentation.pdf
39. Social Progress Imperative. US Cities Covid-19 Vulnerability Index Methodology. Accessed May 29, 2020, https://socialprogressdotblog.files.wordpress.com/2020/04/methodology-for-us-cities-covid-19-vulnerability-index-2.pdf
40. Spoer BR, McCulley E, Lampe TM, et al. Validation of a neighborhood-level COVID Local Risk Index in 47 large U.S. cities. *Health & place*. 2022;76:102814-102814. doi:10.1016/j.healthplace.2022.102814
41. Rozenfeld Y, Beam J, Maier H, et al. A model of disparities: risk factors associated with COVID-19 infection. *Int J Equity Health*. 2020;doi:10.1186/s12939-020-01242-z
42. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BJM*. 2020;369:m1966. doi:10.1136/bmj.m1966 %J BMJ
43. van Gerwen M, Alsen M, Little C, et al. Risk factors and outcomes of COVID-19 in New York City; a retrospective cohort study. *J Med Virol*. 2020;doi:10.1002/jmv.26337
44. Ebinger JE, Achamallah N, Ji H, et al. Pre-existing traits associated with Covid-19 illness severity. *PLOS ONE*. 2020;15(7):e0236240. doi:10.1371/journal.pone.0236240
45. Azar KMJ, Shen Z, Romanelli RJ, et al. Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California. *Health Aff (Millwood)*. 2020;doi:10.1377/hlthaff.2020.00598
46. Gottlieb M, Sansom S, Frankenberger C, Ward E, Hota B. Clinical Course and Factors Associated with Hospitalization and Critical Illness Among COVID-19 Patients in Chicago, Illinois. *Acad Emerg Med*. 2020;doi:10.1111/acem.14104
47. Kim L, Garg S, O'Halloran A, et al. Risk Factors for Intensive Care Unit Admission and In-hospital Mortality Among Hospitalized Adults Identified through the US Coronavirus Disease 2019 (COVID-19)-Associated Hospitalization Surveillance Network (COVID-NET). *Clinical Infectious Diseases*. 2020;doi:10.1093/cid/ciaa1012
48. Hirsch JS, Ng JH, Ross DW, et al. Acute kidney injury in patients hospitalized with COVID-19. *Kidney International*. 2020/07/01/ 2020;98(1):209-218. doi:https://doi.org/10.1016/j.kint.2020.05.006
49. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020/08/01 2020;584(7821):430-436. doi:10.1038/s41586-020-2521-4
50. Cummings MJ, Baldwin MR, Abrams D, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet*. 2020;doi:10.1016/S0140-6736(20)31189-2
51. Gupta S, Hayek SS, Wang W, et al. Factors Associated With Death in Critically Ill Patients With Coronavirus Disease 2019 in the US. *JAMA Intern Med*. 2020;doi:10.1001/jamainternmed.2020.3596

52. SAS Institute Inc. RANK Procedure. Accessed January 17, 2023. https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/proc/n1hxon9vm350ikn19oeualfap8qy.htm
53. 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
54. US Census Bureau. Population and Housing Unit Estimates. https://www.census.gov/programs-surveys/popest.html
55. Division of Cancer Prevention and Control CfDCaP. Incidence and Death Estimates by Congressional District. Accessed 11/19/2022, https://www.cdc.gov/cancer/uscs/technical_notes/stat_methods/congressional-districts.htm
56. Nolte E, McKee M. *Does health care save lives? Avoidable mortality revisited*. The Nuffield Trust; 2004.
57. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA: a cancer journal for clinicians*. 2017;67(3):177-193.
58. 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
59. 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
60. 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
61. 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.
62. 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
63. 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
64. 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
65. 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
66. 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
67. 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
68. 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.
69. 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
70. 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
71. 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
72. Centers for Disease Control and Prevention. 2019 BRFSS Data. Behavioral Risk Factor Surveillance System. https://www.cdc.gov/brfss/annual_data/annual_2019.html
73. 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
74. 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
75. 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
76. 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
77. 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
78. 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
79. Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski JA. US small-area life expectancy estimates project: methodology and results summary. 2018;
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