City Health Dashboard Technical Document

Updated July 17, 2024

Contents

SECTION 1: Overview

The City Health Dashboard (the Dashboard) is a one-stop resource allowing users to view and compare data from multiple sources on health and the factors that shape health to guide local solutions. Through a vigorous selection process, the City Health Dashboard selected 40+ metrics spanning 5 domains — clinical care, health behaviors, health outcomes, physical environment and social and economic factors — to quantify health, health determinants, and equity at the city and, where available, census tract level.

Metrics are derived from both private and publicly available data sources, with some data sources contributing several metrics and others contributing only a single metric.

Document Mission

This document is written for an audience interested in the technical attributes of the Dashboard. It provides details 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@cityhealthdashboard.com) with general feedback or questions not addressed below.

Measure Selection Criteria

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

- Rigorous methods underlying the original data collection
- Feasible data acquisition by the Dashboard analytic team
- Evidence of importance and validity in academic literature
- Metrics that are amenable to city-level 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 100% of cities included in CDC's 500 Cities, PLACES projects
    - Aligned with city preferences based on input from the Dashboard pilot cities and City Advisory Board

City and Tract Selection Criteria

The Dashboard reports data for a select sample of U.S. Census Bureau geographies (collectively referred to as "cities") and their associated Census tracts. These geographies include Places (Incorporated and Designated), County Subdivisions (Minor Civil Divisions), and Hawaiian Counties.[1-4] See Appendix B ("Detailed Notes on Selection of Cities and Tracts") for more detail on city and tract selection criteria and usage of FIPS codes. Contact info@cityhealthdashboard.com with any questions.

Updates to Technical Documentation

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

Please see Appendix F: Updates Summary for an outline of changes made to each version of this document.

Multi-year Data: Appropriate Usage for Evaluating Trends Over Time

The Dashboard displays multiple years of data for many of its metrics. Before evaluating trends over time, users should be aware of the caveats associated with multi-year data from specific data sources.

Please refer to https://www.cityhealthdashboard.com/multi-year-data for caveats associated with specific metrics and data sources.

City Health Dashboard Team

Jacqueline Betro, MPA	Research Project Manager
Samantha Breslin, MPA	Senior Program Coordinator
Dana Duong, MPH	Research Analyst
Marc N. Gourevitch, MD, MPH	Primary Investigator
Yoomin Hwang	Project Coordinator
Neil Kleiman, PhD	Co-Primary Investigator (City Policy/Partnerships)
Taylor Lampe, MPH	Senior Data Analyst
Yuruo Li, PhD, MPH	Senior Data Analyst
Shirley Liang, PharmD, MSPH	Data Analyst
Isabel Nelson, MPH	Senior Data Analyst
Eileen Shea, MPH	Data Analyst
Ben Spoer, PhD, MPH	Program Director
Jay Stadelman, MPH	Data Analyst
Lorna E. Thorpe, PhD	Co-Primary Investigator (Methods)
Anne Vierse, MS	Data Analyst
Hannah Wade, MPH	Senior Population Health Policy Analyst

Downloading Dashboard Data

Users should note that much of the data outlined in this document is available for free download at www.cityhealthdashboard.com/data-downloads.

Users should consult the Downloadable Data Codebook, available at www.cityhealthdashboard.com/data-downloads, for more detail.

Please contact the Dashboard at info@cityhealthdashboard.com with any questions or concerns.

Citing Dashboard Data and Technical Document

City Health 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. City Health Dashboard. https://www.cityhealthdashboard.com/. Accessed [INSERT DATE OF ACCESS].

To cite our Technical Document, we suggest the following citation:

> City Health Dashboard Team. *City Health Dashboard Technical Document*. New York: City Health Dashboard; [YEAR]. Available at www.cityhealthdashboard.com/technical-documentation. Accessed [INSERT DATE OF ACCESS].

Feedback or Errors

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

Measure Overview

The Dashboard presents measures in one of three different formats: percentage, rate, or index. The type of measure is determined by the data that are analyzed to derive each estimate. All measures are calculated at the city level; measures are also calculated by demographic subgroups or at the tract level if the underlying data allow for such disaggregation. For more information about these measures, including data sources, years of data, and measure calculation, visit the website Metric page or refer to the rest of the City Health Dashboard Technical Document.

Domain	Metric Name	Metric Description	Data Source	Tract Estimates	Demographic Subgroups	Most Current Data Period
Clinical Care	Dental Care	Percentage of adults who report visiting a dentist in the past year	PLACES Project, Centers for Disease Control	Yes	Not Available	2020, 1 year modeled estimate
	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	No	Race/Ethnicity	2020, 3 year estimate
	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	Yes	Sex	2020, 1 year modeled estimate
	Routine Checkup, 18+	Percentage of adults who report visiting a doctor for routine checkup in the past year	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	Uninsured	Percentage of population ≤64 years without health insurance	American Community Survey, U.S. Census Bureau	Yes	Age, Sex, Race/Ethnicity	2022, 5 year estimate
Health Behavior	Binge Drinking	Percentage of adults who report binge drinking in the past 30 days	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	Physical Inactivity	Percentage of adults who report no leisure-time physical activity in the past 30 days	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	Smoking	Percentage of adults who report current smoking	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	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	No	Race/Ethnicity	2020, 3 year estimate
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	No	Race/Ethnicity	2021, 3 year estimate
	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	No	Sex, Race/Ethnicity	2021, 3 year estimate
	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	No	Sex, Race/Ethnicity	2021, 3 year estimate
	Diabetes	Percentage of adults who report having diabetes	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	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	No	Sex, Race/Ethnicity	2021, 4 year estimate

Domain	Metric Name	Metric Description	Data Source	Tract Estimates	Demographic Subgroups	Most Current Data Period
	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	No	Sex, Race/Ethnicity	2021, 4 year estimate
	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	Yes	Not Available	2021, 1 year modeled estimate
	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	Yes	Not Available	2021, 1 year modeled estimate
	High Blood Pressure	Percentage of adults who report high blood pressure	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	Independent Living Difficulty	Percentage of adults who report difficulty doing errands alone because of a physical, mental, or emotional condition.	American Community Survey, U.S. Census Bureau	Yes	Age	2022, 5 year estimate
	Life Expectancy	Average years of life expectancy at birth	U.S. Small-area Life Expectancy Estimates Project (USALEEP)	Yes	Not Available	2015, 6 year modeled estimate
	Low Birthweight	Percentage of live births with low birthweight (<2500 grams)	Natality Data, National Vital Statistics System, National Center for Health Statistics	No	Race/Ethnicity	2020, 3 year estimate
	Obesity	Percentage of adults who report a body mass index (BMI) ≥30 kg/m2	PLACES Project, Centers for Disease Control	Yes	Not Available	2021, 1 year modeled estimate
	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	No	Sex, Race/Ethnicity	2021, 3 year estimate
	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	No	Sex, Race/Ethnicity	2021, 3 year estimate
Physical Environment	Air Pollution - Ozone	Average daily maximum concentration (parts per billion) of ground-level ozone throughout a month	George Mason University	Yes	Not Available	12/2023, 1 month average estimate
	Air Pollution - Particulate Matter	Average daily concentration (µg/m³) of fine particulate matter (PM2.5) per cubic meter of air throughout a month	George Mason University	Yes	Not Available	12/2023, 1 month average estimate
	Housing with Potential Lead Risk	Percentage of housing stock with potential elevated lead risk	American Community Survey, U.S. Census Bureau	Yes	Not Available	2022, 5 year estimate
	Lead Exposure Risk Index	Index (1-10) reflecting poverty-adjusted risk of housing-based lead exposure	American Community Survey, U.S. Census Bureau	Yes	Not Available	2022, 5 year estimate
	Park Access	Percentage of population living within a 10 minute walk of green space	ParkServe®	Yes	Race/Ethnicity	2022 *
	Walkability	Index (0-100) reflecting amenities accessible by walking as calculated by Walk Score	Walk Score®	Yes	Not Available	2022 *

Domain	Metric Name	Metric Description	Data Source	Tract Estimates	Demographic Subgroups	Most Current Data Period
Social and Economic Factors	Broadband Connection	Percentage of households with high speed broadband internet connection (cable, fiber optic, DSL)	American Community Survey, U.S. Census Bureau	Yes	Not Available	2022, 5 year estimate
	Children in Poverty	Percentage of children living in households ≤100% of the federal poverty level	American Community Survey, U.S. Census Bureau	Yes	Race/Ethnicity	2022, 5 year estimate
	Chronic Absenteeism	Percentage of public school students who miss 10% or more school days in an academic year. Note: this metric is at the city (not school district) level.	National Center for Education Statistics, U.S. Department of Education	No	Sex, Race/Ethnicity	School year ending in 2022
	Credit Insecurity Index	Index (0-100) reflecting community-level limited credit access	New York Fed Consumer Credit Panel/Equifax	Yes	Not Available	2020 *
	High School Completion	Percentage of adults ≥25 years with high school diploma or equivalent, or higher degree	American Community Survey, U.S. Census Bureau	Yes	Age, Sex, Race/Ethnicity	2022, 5 year estimate
	Income Inequality	Index (-100 to +100) reflecting households with income at the extremes of the national income distribution (the top or bottom 20%)	American Community Survey, U.S. Census Bureau	Yes	Not Available	2022, 5 year estimate
	Neighborhood Racial/Ethnic Segregation	Index (0-100) reflecting the geographic clustering of racial/ethnic groups across the area	American Community Survey, U.S. Census Bureau	No	Not Available	2022, 5 year estimate
	Racial/Ethnic Diversity	Index (0-100) reflecting how evenly distributed the population is across the racial/ethnic groups living in this area	American Community Survey, U.S. Census Bureau	Yes	Not Available	2022, 5 year estimate
	Rent Burden	Percentage of households where ≥30% of income is spent on rent	American Community Survey, U.S. Census Bureau	Yes	Not Available	2022, 5 year estimate
	Third-Grade Reading Scores	Average reading test scores (in grade levels) of third graders in public schools. Note: this metric is at the city (not school district) level.	Stanford Education Data Archive	No	Sex, Race/Ethnicity	School year ending in 2019
	Unemployment - Annual, Neighborhood-Level	Percentage of population ≥16 years who are unemployed but seeking work	American Community Survey, U.S. Census Bureau	Yes	Age, Sex, Race/Ethnicity	2022, 5 year estimate
	Unemployment - Current, City-Level	Percentage of civilian labor force who are unemployed, by month	Local Area Unemployment Statistics, U.S. Bureau of Labor Statistics	No	Not Available	02/2024, 1 month modeled estimate
	Voter Turnout	Percentage of voting-eligible population who voted in the past general election.	L2 and American Community Survey	Yes	Age, Sex	2020 General Election*

** Only one year of data is available for this metric*

SECTION 2: Dashboard Analytic Decisions

Confidence Intervals (CIs)

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

Dashboard CIs are reported at the 90% level

Ninety-five percent CIs are most commonly reported in the scientific literature. However, the Dashboard reports 90% CIs for a number of reasons. Most notably, the Census Bureau recommends calculation of 90% CIs when using American Community Survey data.[5] The Dashboard opted to construct 90% CIs from standard errors where necessary to ensure consistency between measures.

Formulas for CI calculation

There are a number of formulas for deriving CIs; selection depends on properties of the underlying data. See Section 3 below for specifics on the formula used.

Confidence intervals for percentages were manually restricted to minimum 0 and maximum 100 when raw values exceeded these bounds.

Note on CIs for the Dashboard index values

As a rule, CIs were not calculated for the Dashboard's index values because indices reflect a weighted composite of measures that are then scaled, making CI calculation relatively complicated.

Data Censoring and Flagging

The Dashboard censors and flags for users any city-level estimates that are missing sufficient contributing data or have insufficient sample size to generate reliable estimates.

Censoring

The Dashboard censors city-level estimates that meet these criteria:

- For city-level estimates that are aggregated from census tracts (see Appendix E for more about this method), estimates are censored when > 30% of the city's population are missing contributing data
- For city-level estimates with available numerator/denominator, estimates are censored when num<10 and denom<50
- For city-level education-based metrics and city-level race/ethnicity estimates for NVSS mortality metrics, estimates are censored when the number of eventscontributing to the estimate is <20.

There are some additional metric-specific censorship criteria; please see metric sections below for more information.

Censored estimates are noted on the website and in downloadable data. See the Downloadable Data Codebook, available at www.cityhealthdashboard.com/data-downloads, for more information. Please email info@cityhealthdashboard.com with any questions.

Flagging

The Dashboard flags city-level estimates that meet these criteria:

- For city-level estimates that are aggregated from census tracts (see Appendix E for more about this method), estimates are flagged when > 15% and < 30% of the city's population are missing contributing data
- For city-level estimates with available numerator/denominator, estimates are flagged when numerator<50

Flagged estimates are noted on the website and in downloadable data. See the Downloadable Data Codebook, available at www.cityhealthdashboard.com/data-downloads, for more information. Please email info@cityhealthdashboard.com with any questions.

Data Disclaimer

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

Data Rounding

All calculated values were rounded to one decimal place immediately prior to data export.[6,7,8,9]

Use of County-Level Data on the Dashboard

County-level data are used in some instances where city-level data are unavailable/censored. Given that county geographies and populations do not always align with cities, this is only done in certain cases:

- For coterminous city/county where the spatial extent between city and county are the same
- For consolidated cities where city and county share governments and the spatial extents are similar, but the county may include some unincorporated areas.
- Hawaiian counties are presented across the website. See Appendix B for more information.

When this occurs, it is noted on the website and in downloadable data. Please see the Downloadable Data Codebook, available at www.cityhealthdashboard.com/data-downloads, for more details.

Dashboard City Average Estimates

Dashboard city average estimates on the Dashboard averages data from the cities represented on the Dashboard, by metric, demographic group, and year. The estimates are not intended to reflect estimates for the United States nationally, unless otherwise noted.

Averages are calculated after censoring criteria are applied.

Population Percentages

Text describing population breakdowns by racial/ethnic demographic group (and by sex, for the preventive services, 65+ metric only) accompanies metric values on the Demographic Detail page. The values are calculated using American Community Survey (ACS) data. These values are not available for download; please email info@cityhealthdashboard.com for more information on their calculation.

Race/Ethnicity Categories

Where possible, the Dashboard disaggregates metrics by the following demographic groups: Asian (Asian or Native Hawaiian or Pacific Islander (NHOPI)); black/African American; Hispanic/Latino; white (not Hispanic or Latino); and other (some other race, 2 or more races, or American Indian/Alaska Native (AIAN)).[10] Federal guidelines for reporting data by demographics[10] mandate separate categories for AIAN and NHOPI. However, the geographic areas reported on the Dashboard generally lack large enough

populations for reporting stable estimates for these groups. The Dashboard consequently combines NHOPI with Asian and AIAN with "other race" and two or more races, as data availability allows. See data source-specific sections for information about when Hispanic ethnicity is mutually exclusive of the other racial groups and definitions of NHOPI and other.

To ensure these population groups are represented on the Dashboard, the City Overview for each city includes a granular breakdown of each city's racial/ethnic composition by city and census tract to enable a more nuanced understanding of each area. See "Demographics Overview" and "Demographics by Census Tracts" on the City Overview page for more.

Validation

The Dashboard implements a multi-step data validation process to ensure the accuracy of metric value calculation and data uploaded to the website display. Please email info@cityhealthdashboard.com if you have questions about this process.

ZIP Codes

Estimates at the ZIP code level are not provided on the Dashboard.

However, City Health Dashboard's Metric Detail pages indicate a ZIP code associated with each tract estimate. While tracts may cross the boundaries of more than one ZIP code, the Dashboard identifies only the ZIP code with the most overlap with the tract. This ZIP code is identified using a ZIP-TRACT Crosswalk from the USPS and HUD. Current versions used on the Dashboard are 4th Quarter 2022 for tracts using 2010 Census boundaries, and 1st Quarter 2024 for tracts using 2020 Census boundaries.[11]

City and Tract Demographics Overview

Users interested in the demographic make-up of their city and census tracts can explore demographic tables and maps on the "Demographics Overview" and "Demographics by Census Tract" tabs of the City Overview web page (a sample link for New York, NY is provided here). Estimates come from the American Community Survey (ACS) 5-year estimates. A list of ACS table names used for each demographic group is included below. Please contact info@cityhealthdashboard.com for specific variable names. Users can find more information about these demographic estimates in our blog

Demographic Groups	ACS Table Name	ACS Table Description
Age (all groups)	DP05	ACS Demographic and Housing Estimates
American Indian and Alaska Native	DP05	ACS Demographic and Housing Estimates
Asian (all groups)	B02018	Asian Alone or in Any Combination by Selected Groups
Black or African American	DP05	ACS Demographic and Housing Estimates
Foreign-born (all groups)	B06004 B, D, H, & I	Place of Birth
Foreign-born (total only)	B05012	Nativity in the United States
Hispanic or Latino (all groups)	B03001	Hispanic or Latino Origin by Specific Origin
Native Hawaiian or Other Pacific Islander (all groups)	B02019	Native Hawaiian and Other Pacific Islander Alone or in Any Combination by Selected Groups
White, non-Hispanic	DP05	ACS Demographic and Housing Estimates
Total population	DP05	ACS Demographic and Housing Estimates

"Estimate should be interpreted with caution"

Demographic estimates are labeled with a caution symbol when they meet one of the following two criteria:

1. The demographic group's count estimate is 0.
2. The relative standard error (RSE) of the demographic group's count estimate is > 30%, where:

$$RSE = \left(\frac{SE}{Estimate}\right) * 100$$

Interested users are welcome to email info@cityhealthdashboard.com with further questions.

Analytic Software

All analyses were performed in R using tidyverse, tidycensus, tigris, and sf packages, among others.[12-16]

SECTION 3: Data Sources and Metric Analyses

Introduction to this Section

This section is organized by data source, with notes on elements specific to individual metrics.

American Community Survey (ACS)

General notes

ACS is administered by the US Census Bureau. All data was obtained from the Census API[17] using R and the tidycensus package.[13] Place and county subdivision data were used for city-level analyses; Tract data were used for tract-level analyses; County data were used for county-level analyses (*Honolulu only*). All data from ACS are 5-year estimates.

Variable labels (e.g., Estimate; SEX AND AGE - Total population), not names (e.g., S2801_C01_017E), are outlined in this section. Variable labelling conventions used by the US Census Bureau's API are recorded below, except where otherwise noted.

Multi-year data

Data from 2013 - 2022 (5 Year Estimates) are used on the Dashboard.

Variable name changes in annual data releases are assessed by the Dashboard's analytic staff as per the US Census Bureau's technical documentation regarding table and geography changes.[18-23] Metric-specific sections note where analogous labels change over time. For parsimony and clarity, only the most recently available labels are listed in this document. Please contact info@cityhealthdashboard.com with any questions about specific variable names or labels used in prior year analyses.

Weights

Weights were not applied to ACS data because data do not require additional weighting.

Categorizing race/ethnicity

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 (i.e., certain values from data table DP05, see Racial/ethnic diversity and Neighborhood racial/ethnic segregation sections below), estimates for Asian, black/African American, and other demographic groups derived from ACS data are not mutually exclusive with estimates for Hispanic/Latino ethnicity. Values presented for white are always for "White, non-Hispanic", as per the data available for download from ACS. 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. These categorizations reflect those defined by ACS.

Refer to Section 2 "Race/ethnicity categories" (above) for more detail on Dashboard definitions.

Confidence intervals

CIs for all ACS data were calculated according to the formula estimate ± MOE. Approximated MOE's for aggregated count data and derived proportions in ACS data were calculated as per the US Census Bureau's publication.[24] Relevant formulas are presented verbatim here for users' reference:

Calculating MOE's for Aggregated Count Data[24] (p. A-14)

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

Note: When calculating the MOE for a summed estimate which includes multiple zero estimates, per the Census' recommendation, only the MOE for one of the zero estimates was used in the calculation. This prevents the MOE from being artificially inflated.[25]

Calculating MOE's for Derived Proportions[24] (p. A-14, A-15)

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

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

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

Calculating MOE's for Derived Ratios[24] (p. A-15)

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

When confidence intervals extended less than 0 or greater than 100 for proportion metrics, these were set to 0 or 100, respectively.

Data censoring and flagging

The Dashboard censors and flags certain estimates that meet criteria for instability. Refer to Section 2 "Data Censoring and Flagging" (above) for detailed criteria. Only estimates that are calculated using numerator and denominator are censored or flagged; indexes and metrics that are received as percentages directly from ACS are not censored or flagged by the Dashboard. Certain metrics have additional censoring or flagging criteria as noted below in metric-specific notes.

Metric-specific notes

Broadband connection

Data tables

Data table S2801 was used to calculate percentage of households with connections to high speed broadband internet (including cable, fiber optic, and DSL connections). The metric is available for years 2017 forward.

Analysis

$$\text{Broadband connection} = \frac{\text{[Households with connections to high speed broadband internet]}}{\text{Total households}} \text{ x } 100$$

Broadband connection is presented as reported using the variable labelled as:

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

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

Children in poverty

Data tables

Data table B17020 and associated race/ethnicity-specific tables were used to calculate percentage of children in poverty at city and tract levels.

Analysis

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

Variables with the following labels within 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

Numerator variables were summed with variables with the following labels within each data table to calculate the denominator:

- 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

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

High school completion

Data tables

High school completion represents the percent of the population over age 25 that has completed at least a high school degree or equivalent.

Data table S1501 was used to calculate high school completion, for “total population” and disaggregated by sex, for all years. Variable selection changed between years (see below).

Data tables C15002B, C15002C, C15002D, C15002E, C15002F, C15002G, C15002H, and C15002I were used to calculate high school completion, disaggregated by race/ethnicity, for all years.

Analysis

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

Total population; sex: 2017 Forward Analyses (Table S1501)

High school completion is presented as reported using variables labelled as:

- 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

Total population; sex: 2015-2016 Analyses (Table S1501)

Variables with the following labels were summed to calculate the numerator (replace “Total” with “Male” or “Female” for sex-specific estimates):

- Total!!Estimate!!Population 25 years and over!!High school graduate (includes equivalency)
- Total!!Estimate!!Population 25 years and over!!Some college, no degree
- Total!!Estimate!!Population 25 years and over!!Associate's degree
- Total!!Estimate!!Population 25 years and over!!Bachelor's degree
- Total!!Estimate!!Population 25 years and over!!Graduate or professional degree

This variable was used for the denominator (replace “Total” with “Male” or “Female” for sex-specific estimates):

- Total!!Estimate!!Population 25 years and over

Total population; sex: 2013-2014 Analyses (Table S1501)

High school completion was calculated by summing the variables with the following labels (replace “Total” with “Male” or “Female” for sex-specific estimates):

- Total!!Estimate!!Population 25 years and over!!High school graduate (includes equivalency)
- Total!!Estimate!!Population 25 years and over!!Some college, no degree
- Total!!Estimate!!Population 25 years and over!!Associate's degree
- Total!!Estimate!!Population 25 years and over!!Bachelor's degree
- Total!!Estimate!!Population 25 years and over!!Graduate or professional degree

Race/ethnicity Analyses (Tables C15002x)

Variables with the following labels were summed to calculate the numerator. See above “Categorizing race/ethnicity” section for information on which tables are used for each subgroup.

- 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

This variable was used for the denominator:

- Estimate!!Total

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

Age: 2015 Forward Analyses (Table S1501)

High school completion is presented as reported using variables labelled as:

- Percent!!Estimate!!Population 25 to 34 years!!High school graduate or higher

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

Age: 2013-2014 Analyses (Table S1501)

High school completion is presented as reported using variables labelled as:

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

Notes on analysis

As of October 2020, the Dashboard revised its high school graduation metric, by changing the data source from state-based education data to the American Community Survey. While the old metric looked at public school students who graduated within four years of entering ninth grade in high schools geographically located within the city (whether or not students lived in the city), this new metric looks at high school completion rates among all city (or census tract) residents aged 25+ years old, regardless of where they went to high school.

Income inequality

Data tables

Data table B19001 was used to calculate income inequality at both city and tract levels. Data table B19080 was used to define the 20th and 80th household income percentiles for each year.

Analysis

Income Inequality at the Extremes (ICE) was calculated as per Krieger and colleagues.[26]

The formula for ICE is as follows:

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

Where values of ICE range from -100 to 100.

Data are presented for years where the cut points available from table B19001 most closely represent the 20th and 80th household income percentiles[26], defined using 1-year estimates from table B19080.

Year	20th Percentile Cut Point	80th Percentile Cut Point
2022	< $29,999	> $150,000
2018	< $24,999	> $125,000

Variables with the following labels in ACS Table B19001 were summed to calculate the number of households above the 80th percentile for each year:

- Estimate!!Total!!$125,000 to $149,999
- Estimate!!Total!!$150,000 to $199,999
- Estimate!!Total!!$200,000 or more

Variables with the following labels were summed to calculate the number of households below the 20th percentile for each year:

- 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

In both City and Tract analyses, the variable with the following label was used to represent total households with known income level:

- Estimate!!Total

Notes on analysis

Confidence intervals were not calculated because ICE is an index.

As of July 2024, the Dashboard limited the presentation of income inequality data to 2018 and 2022 after discovering temporal variation in this metric related to the metric definition and data availability. While incomes on average increase every year, shifting percentiles up, the income variables used in the calculation are categorical and are not able to be adjusted accordingly each year. Due to this, the Dashboard now only presents the years of data where there is the most alignment between the percentiles and the income variable cut points. The Dashboard concurrently transitioned to using ACS Table B19080 to define our percentiles rather than US Census Bureau Table H-1 (All Races).

Independent Living Difficulty

Data tables

Data table S1810 was used to calculate the independent living difficulty metric at both city and tract levels.

Analysis

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

For 2015 forward, independent living difficulty is presented as reported using the variable labelled as:

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

For 2015 forward, independent living difficulty by age is presented as reported using the variables labelled as:

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

Notes on analysis

Associated margins of error variables are used to calculate confidence intervals associated with these values.

Housing with potential lead risk

Data tables

Data table B25034 was used to calculate housing risk data at both city and tract levels.

Analysis

The lead analysis was performed as per methodology initially developed by the Washington State Department of Health.[27] Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.[28] The Dashboard adapted Vox Media's Python code available on Github[29] for the present analysis, which was conducted by the Dashboard using R v4.1.0 and originally validated using Python v3.6.[30] The Washington State Department of Health's analysis uses variables from 2014 and research from 2002 to inform age of housing weights and lead dust hazard guidance. [27] In updating the analysis to represent all housing stock built in 2010 or later for years subsequent to 2014, variables were added for housing stock built using table B25034. Age of housing weights were updated using the 2021 American Healthy Homes Survey II, published by the U.S. Department of Housing and Urban Development.[31] The updated weights reflect the most recent lead dust hazard guidance, which significantly lowers the acceptable threshold. This update complies with EPA guidance.[32]

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.[29] We count the number of housing units in each of four time periods: pre-1940, 1940-59, 1960-79, 1980 or newer. The count of housing units in each time period is weighted by the likelihood of lead exposure due to building age (weights are extrapolated from American Healthy Homes Survey II, 2021. *Please note that available ACS building age variables do not perfectly align with suggested weights. Email info@cityhealthdashboard.com for further details.*). This results in an overall percent of housing likely to have some risk of lead exposure.

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

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

- Estimate!!Total:!!Built 1939 or earlier
- Estimate!!Total:!!Built 1940 to 1949
- Estimate!!Total:!!Built 1950 to 1959
- Estimate!!Total:!!Built 1960 to 1969
- Estimate!!Total:!!Built 1970 to 1979
- Estimate!!Total:!!Built 1980 to 1989
- Estimate!!Total:!!Built 1990 to 1999
- Estimate!!Total:!!Built 2000 to 2009
- Estimate!!Total:!!Built 2010 to 2013 *(data year 2020 only)*
- Estimate!!Total:!!Built 2014 or later *(data year 2020 only)*
- Estimate!!Total:!!Built 2010 to 2019 *(data year 2021 and 2022)*
- Estimate!!Total:!!Built 2020 or later *(data year 2021 and 2022)*

The following variable was used to represent total housing stock:

- Estimate!!Total

Margins of error (MOE) for these estimate values were derived using the following protocol: calculating adjusted MOE's for each housing-age group that had summed estimates[24]; weighting those MOE's with the same weights used to calculate the numerator; and then calculating an MOE for a derived proportion.[24] See section "*ACS: Calculating MOEs for aggregate count data and derived proportions*" for this equation in full.

Notes on analysis

The Dashboard Team determined that estimates with a 3.25% or greater absolute increase from year to year were unstable and therefore are censored.

Lead exposure risk index

Data tables

Data table B25034 was used to calculate housing risk at both city and tract levels. S1701 was used for calculating poverty risk at both city and tract levels.

Analysis

We took the Dashboard Housing with Potential Lead Risk metric (see earlier section for details on weights and lead threshold guidance) and factored in information about the percentage of the population living at or below 125% of the federal poverty level (poverty risk). We z-standardized poverty risk and housing with potential lead risk variables, weighted each by weights extrapolated from American Healthy Homes Survey II, 2021[31], 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

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

- Estimate!!Total:!!Built 1939 or earlier
- Estimate!!Total:!!Built 1940 to 1949
- Estimate!!Total:!!Built 1950 to 1959
- Estimate!!Total:!!Built 1960 to 1969
- Estimate!!Total:!!Built 1970 to 1979
- Estimate!!Total:!!Built 1980 to 1989
- Estimate!!Total:!!Built 1990 to 1999
- Estimate!!Total:!!Built 2000 to 2009
- Estimate!!Total:!!Built 2010 to 2013 *(data year 2020 only)*
- Estimate!!Total:!!Built 2014 or later *(data year 2020 only)*
- Estimate!!Total:!!Built 2010 to 2019 *(data year 2021 and 2022)*
- Estimate!!Total:!!Built 2020 or later *(data year 2021 and 2022)*

The following variable was used to represent total housing stock:

- Confidence intervals were not calculated because lead exposure risk is a ranked index. 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

Notes on analysis

a. Confidence intervals were not calculated because lead exposure risk is a ranked index.
b. The Dashboard Team determined that estimates with a change of 2 deciles or more year to year were unstable and therefore are censored.
c. The decile ranking ranks risk of lead exposure risk relative to the other cities included on the Dashboard, not all US cities. Thus, a city/tract level estimate might change when other cities/tracts are added to the calculation pool.

Neighborhood racial/ethnic segregation

Data tables

Data table DP05 was used to calculate racial/ethnic segregation at the city level.

Analysis

Segregation was quantified as per Iceland's formula for H, the entropy index.[33]

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." [33] 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.

Segregation on the Dashboard is calculated using the following formula, adapted from the entropy index:

$$\text{Neighborhood racial/ethnic segregation} = \sum_{i=i}^{n} \frac{t_i(E-E_i)}{ET} \times 100$$

Where:

t_i refers to the total population of tract i
T is the metropolitan area population
n is the number of tracts
E is the metropolitan area diversity (entropy) score
E_i is the tract i's diversity (entropy) score

Iceland defines entropy scores for cities and tracts as follows: "A metropolitan area's entropy score is calculated as:

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

Where:

$\boldsymbol{\pi_r}$ refers to a particular racial/ethnic group's proportion of the whole metropolitan area population…

A unit within the metropolitan area, such as a census tract, would analogously have its entropy score, or diversity, defined as:

$$E_i \text{ (tract entropy/diversity)} = \sum_{r=1}^{r} (\pi_{ri}) \ln\left[\frac{1}{\pi_{ri}}\right]$$

Where:

$\boldsymbol{\pi_{ri}}$ refers to a particular racial/ethnic group's proportion of the population in tract i."[33]

As per footnote 5 in Iceland,[33] $\ln\left[\frac{1}{\pi_r}\right]$ is set to 0 when the proportion of a particular group is in a given geography (π_r) is 0. This is done for calculations of both E and E_i.

Variables with the following labels were used in the diversity and segregation analyses:

- 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)

Notes on analysis

The estimates of persons in each racial/ethnic group within a city's tracts were summed to calculate the total population within each city. This calculated total population is not reported on the Dashboard. Users should note that this value sometimes equals the city's actual total population estimate reported in DP05. However, the summed total of tract total populations sometimes over-counts the total population of a city. This is because Census tract boundaries are not perfectly nested within Census place (city) boundaries. The Dashboard used this method for the purposes of calculating denominators for Diversity and Segregation (E, E(i) and H) analyses because the entropy index analyses demand that proportions of racial/ethnic groups sum to a total of 1. Thus, for the purposes of our calculation, the "total population" of a geographic area was necessarily the sum of the total population of each mutually exclusive racial/ethnic group within the area. Further, the entropy index analysis examines the relationship between populations at the city and tract level; analysis thus required use of all the tracts associated with a given city.

The Dashboard does not release segregation (H) scores for cities with a single census tract because the entropy index is not valid in cities in which residents may only live in a single census tract. Due to the nature of how this metric is calculated and changes in census tract boundaries over time, estimates are not always comparable before and after 2020.

Confidence intervals were not calculated because the entropy scores are components of an index.

Racial/ethnic diversity

Data tables

Data table DP05 was used to calculate racial/ethnic diversity values at the city and tract levels.

Analysis

Racial/ethnic diversity represents how much of the maximum possible entropy (or diversity) is exhibited in a given tract or city. 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 measure does not incorporate geographic distributions of racial/ethnic groups. Diversity (or entropy) was quantified using Iceland's formulas for E and E_i entropy scores (see below).[33] In our analysis, E and E_i represent city and tract racial/ethnic diversity scores (or entropy), respectively.

$$\text{Racial/ethnic diversity} = \frac{\text{City or tract entropy score (E or } E_i)}{\text{Maximum possible entropy score}} \times 100$$

Where:

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

Iceland defines entropy scores for cities and tracts as follows: “A metropolitan area’s entropy score is calculated as:

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

Where:

$\boldsymbol{\pi_r}$ refers to a particular racial/ethnic group’s proportion of the whole metropolitan area population…

A unit within the metropolitan area, such as a census tract, would analogously have its entropy score, or diversity, defined as:

$$E_i \text{ (tract entropy/diversity)} = \sum_{r=1}^{r} (\pi_{ri}) \ln[\frac{1}{\pi_{ri}}]$$

Where:

$\boldsymbol{\pi_{ri}}$ refers to a particular racial/ethnic group’s proportion of the population in tract i.”[33]

As per footnote 5 in Iceland,[33] $\ln\left[\frac{1}{\pi_r}\right]$ and is set to 0 when the proportion of a particular group is in a given geography (π_r) is 0. This is done for calculations of both E and E_i.

Variables with the following labels were used in the diversity and segregation analyses:

- 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)

Notes on analysis

The estimates of persons in each racial/ethnic group within a city’s tracts were summed to calculate the total population within each city. This calculated total population is not reported on the Dashboard. Users should note that while this value sometimes equals the city's actual total population estimate reported in DP05, the summed total of tract total populations sometimes over counts the total population of a city. This is likely because Census tract boundaries are not perfectly nested within Census place (city) boundaries. The Dashboard used this method for the purposes of calculating denominators for diversity and segregation (E, E(i) and H) analyses because the entropy index analyses demands that proportions of racial/ethnic groups sum to a total of 1. Thus, for the purposes of our calculation, the "total population" of a geographic area was necessarily the sum of the total population of each mutually exclusive racial/ethnic group within the area. Further, the entropy index analysis examines the relationship between populations at the city and tract level; analysis thus required use of all the tracts associated with a given city.

Confidence intervals were not calculated because the entropy scores are components of an index. See the “Confidence intervals” section in Section 2 above for more details.

Rent burden

Data tables

Data table DP04 was used to calculate rent burden at both city and tract levels.

Analysis

$$\text{Rent burden} = \frac{\text{Gross Rent as a percentage of household income: Occupied units paying rent - 30.0\%-34.9\% of monthly income} + \text{Gross Rent as a percentage of household income: Occupied units paying rent - ≥35.0\% of monthly income}}{\text{Total renter-occupied units with reported income}} \text{ x } 100\%$$

In both City and Tract analyses, the variables in DP04 with the following labels 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

In both City and Tract analyses, the variable in DP04 with the following labels were summed to calculate the denominator:

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

- NOTE: Variable labels for 2013 and 2014 tables are slightly different from labels for 2015 Forward. Only the most recently available labels are provided here. Please email info@cityhealthdashboard.com with any questions or for more detailed information about variable naming over time.

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

Unemployment – annual, neighborhood level

Data tables

Data table S2301 was used to report annual unemployment rates, disaggregated by race/ethnicity and sex, at the city and tract levels.

Analysis

Total population; sex breakdown, age breakdown; Black; White; Hispanic: 2013 Forward City Analyses (Table S2301)

Annual unemployment rates reported in S2301 are presented as reported using variables labelled as:

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

- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!35 to 44 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!45 to 54 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!55 to 59 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!60 to 64 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!65 to 74 years
- Estimate!!Unemployment rate!!Population 16 years and over!!AGE!!75 years and over

Other, Asian: 2013 Forward City Analyses (Table S2301)

Estimates and confidence intervals values for “other race” and “Asian” are weighted averages of estimates and confidence intervals for the subgroups that comprise these groups throughout the Dashboard.

The value for “other race” is a weighted average of the variables associated with the following labels in S2301:

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

Estimate and confidence interval values are weighted by the relative proportion of each of these groups within the summed total population of these three groups within each city as per ACS table S2301, using the variables with the following labels:

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

The value for “Asian” is a weighted average of the variables associated with the following labels in S2301:

- 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

Estimate and confidence interval values are weighted by the relative proportion of each of these groups within the summed total population of these two groups within each city as per ACS table S2301, using the variables with the following labels:

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

When any of the above variables used for summation were missing, the entire summed estimate was set to missing.

Associated margins of error variables were used to calculate confidence intervals associated with these values for all demographic subgroups and total population.

Total population: 2013 Forward Tract Analyses (Table S2301)

Annual unemployment rate was reported using the variable with the following label in table S2301:

- Estimate!!Unemployment rate!!Population 16 years and over

Uninsured

Data tables

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

Data table S2701 was used to report percent of the civilian noninstitutionalized population without health insurance for ages 0-64 at the city and tract levels; this stratum is referred to as “total population”.

Data table S2701 was used to report percent of the civilian noninstitutionalized population without health insurance, disaggregated by age, at the city level for 2015 Forward. Age strata change over time in accordance with data availability. Of note, the Census Bureau changed age categories as of the 2017 data release to better align with the current health insurance landscape.[34]

Summary of available age strata for uninsured, by year			
	2013, 2014	**2015, 2016**	**2017 Forward**
Age breakdown (*city only)*	*unavailable*	0-17	0-18
		18-24	19-25
		25-34	26-34
		35-44	35-44
		45-64	45-64

Data table B27001 was used to report uninsured, disaggregated by sex, at the city level.

Data tables C27001B, C27001C, C27001D, C27001E, C27001F, C27001H, and C27001I were used to calculate uninsured, disaggregated by race/ethnicity, at the city level.

Analysis

The percentage of uninsured people for all subgroup population disaggregations (total population, sex, and age) are calculated using the following formula:

$$\text{Uninsured}_{\text{subgroup}} = \frac{\text{Uninsured: Estimate}_{\text{subgroup}}}{\text{Total: Estimate}_{\text{subgroup}}} \text{ x } 100\%$$

Total population, Age: 2013 Forward City Analyses (Table S2701)

NOTE: Detailed age breakdowns are not available in 2013 and 2014. Additionally, variable labeling conventions and age strata changed slightly between years, which is noted below.

Variables associated with the following labels are summed to calculate the numerator:

- 2013, 2014:
  - Number Uninsured!!Estimate!!AGE!!Under 18 years
  - Number Uninsured!!Estimate!!AGE!!18 to 64 years
- 2015, 2016:
  - Uninsured!!Estimate!!AGE!!Under 18 years
  - Uninsured!!Estimate!!AGE!!18 to 64 years
- 2017 Forward:
  - 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

Variables associated with the following labels are summed to calculate the denominator:

- 2013, 2014, 2015, 2016:
  - Total!!Estimate!!AGE!!Under 18 years
  - Total!!Estimate!!AGE!!18 to 64 years
- 2017 Forward:
  - 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

Percent uninsured disaggregated by age are presented as reported in the S2701 data table using the variables labelled as:

- 2015, 2016:
  - Percent Uninsured!!Estimate!!AGE!!Under 18 years
  - Percent Uninsured!!Estimate!!AGE!!18 to 64 years!!18 to 24 years
  - Percent Uninsured!!Estimate!!AGE!!18 to 64 years!!25 to 34 years
  - Percent Uninsured!!Estimate!!AGE!!18 to 64 years!!35 to 44 years
- 2017 Forward:
  - 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

Variables associated with the following labels are summed to calculate the numerator:

- 2015, 2016:
  - Uninsured!!Estimate!!AGE!!18 to 64 years!!45 to 54 years
  - Uninsured!!Estimate!!AGE!!18 to 64 years!!55 to 64 years
- 2017 Forward:
  - Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!45 to 54 years
  - Estimate!!Uninsured!!Civilian noninstitutionalized population!!AGE!!55 to 64 years

Variables associated with the following labels are summed to calculate the denominator:

- 2015, 2016:
  - Total!!Estimate!!AGE!!18 to 64 years!!45 to 54 years
  - Total!!Estimate!!AGE!!18 to 64 years!!55 to 64 years
- 2017 Forward:
  - Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!45 to 54 years
  - Estimate!!Total!!Civilian noninstitutionalized population!!AGE!!55 to 64 years

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

Sex: 2013 Forward City Analyses (Table B27001)

Table B27001 is used to calculate percentage of uninsured male and female populations at the city level only

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

- Estimate!!Total!!Male!!Under 6 years!!No health insurance coverage
- Estimate!!Total!!Male!!6 to 18 years!!No health insurance coverage
- Estimate!!Total!!Male!!19 to 25 years!!No health insurance coverage
- Estimate!!Total!!Male!!26 to 34 years!!No health insurance coverage
- Estimate!!Total!!Male!!35 to 44 years!!No health insurance coverage

- Estimate!!Total!!Male!!45 to 54 years!!No health insurance coverage
- Estimate!!Total!!Male!!55 to 64 years!!No health insurance coverage

Variables associated with the following labels are summed to calculate the denominator for males:

- Estimate!!Total!!Male!!Under 6 years
- Estimate!!Total!!Male!!6 to 18 years
- Estimate!!Total!!Male!!19 to 25 years
- Estimate!!Total!!Male!!26 to 34 years
- Estimate!!Total!!Male!!35 to 44 years
- Estimate!!Total!!Male!!45 to 54 years
- Estimate!!Total!!Male!!55 to 64 years

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

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

Variables associated with the following labels are summed to calculate the denominator for females:

- Estimate!!Total!!Female!!Under 6 years
- Estimate!!Total!!Female!!6 to 18 years
- Estimate!!Total!!Female!!19 to 25 years
- Estimate!!Total!!Female!!26 to 34 years
- Estimate!!Total!!Female!!35 to 44 years
- Estimate!!Total!!Female!!45 to 54 years
- Estimate!!Total!!Female!!55 to 64 years

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

Race/ethnicity: 2013 Forward City Analyses
(Tables C27001B, C27001C, C27001D, C27001E, C27001F, C27001G, C27001H, C27001I)

Race/ethnicity-specific tables from the C27001 series are used to calculate percent of uninsured populations, by race/ethnicity and at the city level only.

Variables associated with the following labels are summed to calculate the numerator, per racial/ethnic group:

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

Variables associated with the following labels are summed to calculate the denominator, per racial/ethnic group:

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

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

Total population: 2013 Forward Tract Analyses (Table S2701)

Variables associated with the following labels are summed to calculate the numerator:

- 2013, 2014:
  - Number Uninsured!!Estimate!!AGE!!Under 18 years
  - Number Uninsured!!Estimate!!AGE!!18 to 64 years
- 2015, 2016:
  - Uninsured!!Estimate!!AGE!!Under 18 years
  - Uninsured!!Estimate!!AGE!!18 to 64 years
- 2017 Forward:
  - 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

Variables associated with the following labels are summed to calculate the denominator:

- 2013, 2014, 2015, 2016:
  - Total!!Estimate!!AGE!!Under 18 years
  - Total!!Estimate!!AGE!!18 to 64 years
- 2017 Forward:
  - 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

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

George Mason University Air Quality Team

General note

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

Multi-year data

Monthly Data from January 2018 - are presented on the Dashboard.

Weights

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

The Dashboard uses a block population-weighted aggregation method to calculate city-level estimates. Please see *Appendix E: Block Population-Weighted Aggregation for Selected Metrics* for more information about how this method.

Categorizing race/ethnicity

Not applicable.

Confidence intervals

Not applicable.

Metric-specific notes

Air pollution – ozone

Data represent modeled estimates and do not include estimates for Alaska and Hawaii.

EPA CMAQ RSIG and GMU North America Chemical Reanalysis (NACR) are commonly used, publicly available data sources for ozone pollution. While the RSIG presently includes a longer data period, North America Chemical Reanalysis (NACR) uses more up-to-date emission and real-time forecasting data to provide ozone data in a timelier manner (up to yesterday). Both RSIG and NACR provide ozone pollution (parts per billion (ppb)) data for 12 kilometer square areas, which is larger than many census tracts. EPA CMAQ RSIG further smooths the data to provide census tract-level estimates, while NACR are provided at the 12-kilometer square area level only. As such, adjacent census tracts might share the same ozone pollution value (ppb).

Data tables

Tract level data were received directly from George Mason University. Monthly average of daily maximum 8-hour concentration of ground-level ozone in parts per billion are reported at the census tract level.

Analysis

The model prediction data are used to fill in gaps between air quality monitors, in particular in rural and suburban areas. The model estimates spatial and temporal variations of air pollution based on three major components: emission, meteorology and chemistry. A list of emission sources is provided in Table 1. The meteorology data are provided by the Weather Research and Forecast (WRF) model. The chemistry model is based on the US EPA Community Multiscale Air Quality (CMAQ) model. To merge the

model and monitoring data, the optimal interpolation (OI) method is used to generate a fused surface concentration across the Continental United States at 12km gridding. The gridded data are converted into census tract level by averaging all grid points included in the tract.

Table 2. List of emission sources used to estimate ozone

Year of Update	Emission Sources Included
2018	Anthropogenic Sources: Agriculture, transportation (vehicle/air/railroad/marine), electricity generation units (EGUs), non-EGU point sources, oil/gas, residential wood combustion Natural Sources: biogenic; sea-salt;
2019	Anthropogenic Sources: Agriculture, transportation (vehicle/air/railroad/marine), electricity generation units (EGUs), non-EGU point sources, oil/gas, residential wood combustion Natural Sources: biogenic; sea-salt;
2020-forward	Anthropogenic Sources: Agriculture, transportation (vehicle/air/railroad/marine), electricity generation units (EGUs), non-EGU point sources, oil/gas, residential wood combustion Natural Sources: Biogenic; Sea-salt; Wildfires

To calculate city-level estimates, block-level population weights were applied to the census tract annual averages and the weighted block values were summed to the city level. Please see *Appendix E: Block Population-Weighted Aggregation for Selected Metrics* for more information about how block-level weights are calculated and applied.

Air pollution – PM2.5

Data represent modeled estimates and do not include estimates for Alaska and Hawaii.

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

Data tables

Tract level data were received directly from George Mason University. Monthly average of daily maximum 8-hour concentration of PM2.5 per cubic meter are reported at the census tract level.

Analysis

The model prediction data are used to fill in gaps between air quality monitors, in particular in rural and suburban areas. The model estimates spatial and temporal variations of air pollution based on three major components: emission, meteorology and chemistry. A list of emission sources is provided in Table 2. The meteorology data are provided by the Weather Research and Forecast (WRF) model. The chemistry model is based on the US EPA Community Multiscale Air Quality (CMAQ) model. To merge the model and monitoring data, the optimal interpolation (OI) method is used to generate a fused surface concentration across the Continental United States at 12km gridding. The gridded data are converted into census tract level by averaging all grid points included in the tract.

Table 2. List of emission sources used to estimate $PM_{2.5}$

Year of Update	Emission Sources Included
2022	Anthropogenic Sources: Agriculture, transportation (vehicle/air/railroad/marine), electricity generation units (EGUs), non-EGU point sources, oil/gas, residential wood combustion Natural Sources: biogenic; sea-salt; windblown dust; biomass burning

2013-2018 data previously presented on the Dashboard were derived from the Environmental Protection Agency, Community Multiscale Air Quality Remote Sensing Information Gateway (CMAQ RSIG). These data can be requested via email at info@cityhealthdashboard.com

L2

General notes

Voter turnout represents the percentage of the voting eligible population who voted in a general election.

The voter turnout data were derived from the L2, accessed through NYU library. For additional information please email info@cityhealthdashboard.com.

Weights

No weights were applied to the analysis.

Confidence intervals

CIs are not presented because MOE or SE data were not presented in the underlying dataset.

Metric-specific notes

Voter turnout

Data tables

The counts for voter turnout were retrieved from the National Voter File, provided by L2, for the 2020 General Election.

The voting-eligible population data were sourced from the Citizen Voting Age Population (CVAP) by Race and Ethnicity - A Special Tabulation from the 2022 ACS 5-Year Estimates (derived from 2018-2022 data). We chose to use the 2018-2022 5-year estimates for the 2020 general election because we believe these estimates better represent the underlying population trend in 2020 than the 2016-2020 5-year estimates.

For 109 cities not provided by American Community Survey special tabulation directly, the Dashboard uses a block population-weighted aggregation method to calculate city-level estimates. An aggregation indicator was applied for city estimates calculated using this aggregation method. Please see Appendix E: Block Population-Weighted Aggregation for Selected Metrics for more information about how this method.

We provided voter turnout estimates by sex and age groups: Age 18-29, Age 30-44, Age 45-64, Age 65+.

Analysis

To obtain tract and city-level estimates, the Dashboard used the longitude and latitude of voters' residential addresses, geocoded by L2 and NYU Langone Health Geocoding Center, and performed a spatial join with the city and tract boundary to determine voter turnout for the 2020 general election.

$$\text{Voter turnout} = \frac{\text{\# of voters voted in a general election}}{\text{Voting-eligible population of the election year}} * 100\%$$

Variables from ACS CVAP with the following labels were used as our denominators:

- Total population: Estimate!!Total!!Citizens 18 years and over
- Age 18-29: Estimate!!Total!!Citizens 18 years and over!!AGE!!18 to 29 years
- Age 30-44: Estimate!!Total!!Citizens 18 years and over!!AGE!!30 to 44 years
- Age 45-64: Estimate!!Total!!Citizens 18 years and over!!AGE!!45 to 64 years
- Age 65+: Estimate!!Total!!Citizens 18 years and over!!AGE!!65 years and over
- Male: Estimate!!Total!!Citizens 18 years and over!!SEX!!Male

- Female: Estimate!!Total!!Citizens 18 years and over!!SEX!!Female

Notes on analysis

As part of the validation process, the Dashboard Team compared the reported county-level voter turnout count with official county numbers. It's important to note that our voter turnout estimates may not exactly match the numbers reported by the state. This discrepancy can be attributed to several factors, including:

1) The imprecision of the geocoding process.
2) Time lags between election day and when states release their voter file.
3) The imprecision of the ACS population denominator estimate.

Some estimates have been suppressed due to concerns regarding the reliability of the underlying data. For more information, please email info@cityhealthdashboard.com.

As not all states release the birth year of voters in their voter file, we have only provided age categories for the following 25 states:

Alabama	Louisiana	New Mexico
Arizona	Maine	Pennsylvania
Colorado	Michigan	South Carolina
Delaware	Minnesota	Tennessee
Georgia	Missouri	Texas
Illinois	Montana	Vermont
Iowa	Nebraska	Virginia
Kansas	Nevada	
Kentucky	New Jersey	

Local Area Unemployment Statistics, U.S. Bureau of Labor Statistics

General notes

Local Area Unemployment Statistics (LAUS) data are published by the U.S. Bureau of Labor Statistics. The LAUS program provides monthly and annual unemployment and labor force data for many geographies, by place of residence. Estimates are derived from the Current Population Survey (CPS)[35], Current Unemployment Statistics (CES), and State Unemployment Insurance (UI) and disaggregated using data from Decennial Census (DCS) and annual population estimates from the Population Estimates Program[36] by applying a handbook approach and a blended base method.[37]

Multi-year data

LAUS unemployment rates are released monthly. Monthly estimates are, on average, released by LAUS three months after the month has passed. The Dashboard provides estimates from January 2018 through present, updated monthly.

NOTE: January 2018-December 2020 data were updated in June 2021 to reflect revised estimates released by the Bureau of Labor Statistics. Estimates may have changed due to state-level model specifications and data correction.[38] The Dashboard recommends using the most recently updated data published.

Weights

The Dashboard reports LAUS data as downloaded. No weights were applied.

Categorizing race/ethnicity

LAUS data are not categorized by race/ethnicity.

Confidence intervals

CIs are not presented for LAUS data.

Metric-specific notes

Unemployment – current, city-level

Data tables

City-level data are downloaded monthly directly from the Bureau of Labor Statistics[39].

Analysis

City-level estimates represent the proportion of the civilian non-institutionalized population that is unemployed and actively seeking work:

$$\frac{[\text{Civilian population aged} \geq 16 \text{ that is unemployed and actively seeking work}]}{\text{Civilian population aged} \geq 16} \times 100$$

LAUS provides Series Identifiers[40], which are 20-digit unique identifiers that contain information about the geography and analysis performed. The Dashboard made the following selections to obtain monthly unemployment data for cities (see Appendix B for more information on city selection):

For Incorporated Places:

- Prefix = “LA”
- Seasonal Adjustment Code = “S” (“Seasonally Adjusted”)

- Area Code = “CT” + FIPS code[41] + “000000”
- Measure Code = “03” (“Unemployment rate”)

For County Subdivisions:

- Prefix = “LA”
- Seasonal Adjustment Code = “S” (“Seasonally Adjusted”
- Area Code = “CS” + State FIPS code (2 digit) + County Subdivision FIPS code (5 digit)+ “000000”
- Measure Code = “03” (“Unemployment rate”)

Data are unavailable for cities smaller than 25,000.

Manual FIPS code matching was required for four consolidated cities because LAUS only releases (consolidated) estimates. Consolidated cities share a government with their county, but retain unincorporated areas with distinct governments. (balance) estimates do not include these unincorporated areas, while (consolidated) estimates do. Therefore, data for these cities are equivalent to county estimates and include populations that live within unincorporated areas[42].

Dashboard Geography Reported	***LAUS Geography Reported***
Athens-Clarke County unified government (balance)	Athens-Clarke County (consolidated) city
Augusta-Richmond County consolidated government (balance)	Augusta-Richmond County (consolidated) city
Louisville/Jefferson County metro government (balance)	Louisville-Jefferson County (consolidated) city
Nashville-Davidson metropolitan government (balance)	Nashville-Davidson (consolidated) city

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

General notes

Chronic Absenteeism represents the percent of public school elementary and secondary students who miss 10% or more school days in an academic year. NOTE: this metric is calculated at the city-level, not the school district-level, and thus should not be used to evaluate school district performance.

Chronic Absenteeism estimates are calculated using school level chronic absenteeism count data published by the U.S. Department of Education initiatives Ed*Facts*[43] and Ed Data Express[44] (data group 814), and enrollment data from the National Center for Education Statistics (NCES).[45] This metric has been modified twice during its tenure on the Dashboard: once in November 2021 and once in December 2023. Both modifications included source updates and refinements to school exclusion criteria, therefore currently published data should not be compared with previously published data. Current estimates consider consultations with the Stanford Education Data Archive along with expanded data availability to provide multiple years of data. If you would more information on these changes, email info@cityhealthdashboard.com.

Weights

No weights were applied to the analysis.

Categorizing race/ethnicity

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

Confidence intervals

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

Metric-specific notes

Chronic Absenteeism

Data tables

School-level chronic absenteeism counts are retrieved from the U.S. Department of Education website, ED Data Express[44], for school years (SY) 2019-2020 onward. SY 2017-2018 chronic absenteeism counts are retrieved via the U.S. Department of Education initiative, Ed*Facts*[43]. SY 2018-2019 absenteeism counts are not available through either data initiative as of the last Dashboard update date. School level enrollment counts are accessed through NCES enrollment data for the associated school year[45]. The two datasets are joined using the NCES school id number.

Analysis

To obtain city-level estimates, the Dashboard creates a school to city crosswalk for each school year, assigning schools to cities with a 200 meter buffer. Note that schools that fall within a city’s 200 meter buffer range may later be assigned to an adjacent city as the Dashboard adds more cities, which may slightly change city-level estimates. See Appendix D for more details.

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

- Private schools
- Pre-K and adult education schools

- Schools classified as special education schools
- Schools that are classified as “closed”, “inactive”, or otherwise not operating during the given school year
- Schools that are primarily virtual (*note: schools that temporarily taught remotely during the COVID 19 pandemic are not excluded*)
    - For SY 2020-2021 data, the Dashboard team found that all schools in Danville, VA, plus 14 additional schools in other cities changed their virtual status for only this school year. The Dashboard team made the decision to classify these schools as non-virtual for the SY 2020-2021 year based on the surrounding school years of data so that they might be included in the calculation.
- Schools were additionally removed from the calculation if either the numerator or denominator value was missing or if the chronic absenteeism estimate was greater than 100% and therefore considered unstable.

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

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

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

City-level estimates are suppressed if the numerator (number of students chronically absent) is less than 20 students. City-level estimates are also censored if the state-level estimate was deemed unreliable through external validation. This applies to cities in Alaska, Arizona, DC, Idaho, Oregon, Washington, and West Virginia in data year 2022 (SY 2021-2022).

National Vital Statistics System (NVSS)

General notes

Unless otherwise specified (see section "City/County Indicator" below), deaths are assigned to the reported city of residence of the deceased; births are assigned to the city of residence reported by the mother. All NVSS data were analyzed using R.

Restricted use NVSS data are available through the National Association for Public Health Statistics and Information Systems (NAPHSIS). Data request forms are available online.[46]

Users should be cautious when comparing values from different states because of variation in classifying cause of death across locations. This is particularly true for deaths related to opioid overdose.

The downloadable data tables shared on the City Health Dashboard website were not released as a micro-level downloadable datasets from NCHS/RDC, rather .csv aggregated data tables whose analyses were conducted per NCHS disclosure requirements in a secure environment and released as approved output. The findings and conclusions on this website are those of the author(s) and do not represent the views of the Research Data Center, the National Center for Health Statistics, or the Centers for Disease Control and Prevention. NCHS does not recommend further analysis of these tables because linking them to individually identifiable data from other NCHS or non-NCHS datasets could potentially cause disclosure. If you believe a disclosure has occurred please contact info@cityhealthdashboard.com and RDCA@cdc.gov.

NOTE: To provide vital statistics estimates for a subset of New Jersey cities (Burlington, Chatham, Clayton, Cranbury, Egg Harbor City, Fanwood, Garwood, Highlands, Lawnside, Long Beach, Mountainside, Penns Grove, Salem, South Bound Brook, Wildwood Crest), the Dashboard uses data from New Jersey State Health Assessment Data (NJSHAD),[47] a public health data source managed by the New Jersey Department of Health. Estimates from these data sources are valid for comparison with NCHS/RDC data presented on the Dashboard. For further information, please see Appendix C. Please contact info@cityhealthdashboard.com with any questions.

Multi-year data

NOTE: A single year is used to refer to NVSS data throughout the Dashboard. However, multiple years of data are used in the Dashboard's analyses of NVSS data. Nomenclature is determined based on the most recent year of data used in an analysis.

Multiple Cause of Death Data

MCDD data for 2012 through 2021 for the following metrics are presented on the Dashboard as 3 year pooled estimates (For example, 2012 analyses use data from 2010, 2011 and 2012): opioid overdose deaths, breast cancer deaths, cardiovascular disease deaths, colorectal cancer deaths, premature deaths (all causes).

MCDD data for 2014 through 2020 for the following metrics are presented on the Dashboard as 5 year pooled estimates (For example, 2014 analyses use data from 2010-2014): firearm homicides and firearm suicides. Due to bridge-single race transition (see *Categorizing race/ethnicity* section below), in 2021 we presented 4 year pooled estimates for these two metrics.

However, fewer years of data within each set are used in the event that city and/or county-level estimates are not available for all specified years in the dataset received by the Dashboard.

The Dashboard provides more information on how many years of data were utilized for all MCDD reported values under "Source" on our website and "period_type" in downloadable data. They specify the how many years of data were used in calculation and also serve as a multiplier for population. Population

denominators from the endpoint of each wave of data are used as the multiplier. However, data availability for each year varies by city. For example, "Calculated by the Dashboard Team using data from 2021, 1-year estimate" indicates that only 1 year of data were used between 2018-2021 for breast cancer deaths. Please email info@cityhealthdashboard.com for more details.

Natality Data

ND data for 2012 through 2020 for the following metrics are presented on the Dashboard as 3 year pooled estimates (For example, 2012 analyses use data from 2010, 2011 and 2012): prenatal care, low birthweight, and teen births, respectively. 2015 prenatal care estimates are calculated using data from 2014 and 2015 only; see section "Prenatal care" below for more detail.

However, fewer years of data within each set are used in the event that city and/or county-level estimates are not available for all specified years in the dataset received by the Dashboard.

The Dashboard provides more information on how many years of data were utilized for all MCDD reported values under "Source" on our website and "period_type" in downloadable data. They specify the how many years of data were used in calculation and also serve as a multiplier for population. Population denominators from the endpoint of each wave of data are used as the multiplier. However, data availability for each year varies by city. For example, "Calculated by the Dashboard Team using data from 2021, 1-year estimate" indicates that only 1 year of data were used between 2018-2021 for prenatal care. Please email info@cityhealthdashboard.com for more details.

Population denominators

For city-level analyses, population estimate denominators are generated from American Community Survey table B01001 (5 Year Estimates) for Place.[48]

- Data from the endpoint year of a pooled estimate is used in each analysis. For example, B01001 (2016, Year Estimate) is used for 2016, 3 Year Estimate NVSS estimates; B01001 (2015, Year Estimate) is used for 2015, 3 Year Estimate NVSS estimates, etc.

For county-level, population estimate denominators are generated from the National Center for Health Statistics.

- Data from the midpoint year of a pooled estimate is used in each analysis. For example, 2013-2015 estimates use vintage 2016 Bridged-Race Postcensal Population Estimates data file for 2014.[49] 2014-2016 estimates use vintage 2016 Bridged-Race Postcensal Population Estimates data file for 2015.[50]

As noted above, population denominators from the endpoint of each wave of data are used as the multiplier.

Weights

Multiple Cause of Death Data

Breast cancer, colorectal cancer, cardiovascular disease, and opioid overdose deaths metrics use US 2010 standardized population weights. These weights were calculated via the direct adjustment approach outlined by Klein & Schoenborn[51] utilizing the data table "QT-P1, 2010 Decennial Census" downloaded in December 2016 from American Fact Finder.[48]

Premature deaths (all causes) used premature death weights and years of life lost derived from the US 2010 standardized population weights using Dranger and Remington's approach.[52] Weights for age-adjusting premature deaths were calculated using the data table "QT-P1, 2010 Decennial Census".[48] The weights are similar to those used for other mortality metrics, but were adjusted to include only the population aged 74 and younger. Weights for years of potential life lost (i.e., the number of years of life

"lost" for each death within an age group) were calculated as the mid-point of the age group subtracted from the reference age using the following formula Weight (age group *i*) = 75 – Mid-point age group *i*. Weights and years of life lost are presented in Appendix A.

No weights are applied to firearm suicides and firearm homicides.

Natality Data

No weights are applied to ND data.

Categorizing race/ethnicity

Multiple Cause of Death Data

2012-2019 Data:

The National Center for Health Statistics releases bridged-race estimates in their deaths records data from 2012 to 2020. These estimates result from bridging the 31 race categories as specified in the 1997 Office of Management and Budget (OMB) standards for the collection of data on race and ethnicity, to the four race categories specified in the 1977 OMB standards. In short, they used a statistical model to reassign the people who identified themselves as "Two or more races" back to the "main" race categories (White, Black, Asian/Pacific Islander and American Indian).

"Hispanic origin/race recode" and "race recode 5" variables were used to categorize race/ethnicity for all mortality metrics. "Other" race/ethnicity category is not provided. Refer to Section 2 "Race/ethnicity categories" (above) for more detail on Dashboard definitions.

- Hispanic/Latino: "Hispanic origin/race recode" code of either Mexican, Puerto Rican, Cuban, Central or South American, or other or unknown Hispanic, as well as a "race recode 5" code of either: white, black, American Indian, or Asian/Pacific Islander.
- White: "Hispanic origin/race recode code" of non-Hispanic white
- Black: "Hispanic origin/race recode code" of non-Hispanic black
- Asian: "Hispanic origin/race recode code" of non-Hispanic other races, as well as a race recode 5 code of Asian/Pacific Islander

2020 – Data

Starting from 2018, The National Center for Health Statistics releases single-race estimates in their deaths records data according to the 31 race categories as specified in the 1997 Office of Management and Budget (OMB) standards for the collection of data on race and ethnicity. Since we release 3 year pooled estimates data, we provide single-race estimates starting from 2020, 3 year estimate.

When combining race and ethnicity, Hispanic ethnicity takes precedence over whatever race(s) are recorded for the decedent. For example, if a person is identified as Hispanic and Asian in their death certificate, this person will be classified as Hispanic, not Asian.

Hispanic Origin:

- Non-Hispanic: Non – Hispanic, Unknown (assigned to Asian, Black, Other, or White groups);
- Hispanic: Spaniard, Mexican, Puerto Rican, Cuban, Dominican, and South American, Central American, South American, Latin American, Other Hispanic;

Race Recode 40:

- White: White;
- Black: Black;

- Asian: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Other or Multiple Asian, Hawaiian, Guamanian, Samoan, Other or Multiple Pacific Islander;
- Other: American Indian or Alaskan Native (AIAN), Black and White, Black and AIAN, Black and Asian, Black and Native Hawaiian or Other Pacific Islander (NHOPI), AIAN and White, AIAN and Asian, AIAN and NHOPI, Asian and White, Asian and NHOPI, NHOPI and White, Black, AIAN and White, Black, AIAN and Asian, Black, AIAN and NHOPI, Black, Asian and White, Black, Asian and NHOPI, Black, NHOPI and White, AIAN, Asian and White, AIAN, NHOPI and White, AIAN, Asian and NHOPI, Asian, NHOPI and White, Black, AIAN, Asian and White, Black, AIAN, Asian and NHOPI, Black, AIAN, NHOPI and White, Black, Asian, NHOPI and White, AIAN, Asian, NHOPI and White, Black, AIAN, Asian, NHOPI and White, Unknown and Other Race;

Natality Data

2012, 2013 Data:
Mother's race/Hispanic origin (mracehisp) and mother's race recode (mracerec) variables were used to categorize race/ethnicity for all natality metrics. "Other" race/ethnicity category is not provided. Refer to Section 2 "Race/ethnicity categories" (above) for more detail on Dashboard definitions.

- Hispanic/Latino: Mother's race/Hispanic origin code of either Mexican, Puerto Rican,
- Cuban, Central or South American, or other or unknown Hispanic, as well as a mother's race recode code of either: white, black, American Indian/Alaskan Native, or Asian/Pacific Islander.
- White: Mother's race/Hispanic origin code of non-Hispanic white
- Black: Mother's race/Hispanic origin code of non-Hispanic black
- Asian: Mother's race/Hispanic origin code of non-Hispanic other races, as well as a mother's race recode code of Asian/Pacific Islander

2014, 2015, 2016, 2017, 2018, 2019 Data:
Mother's Hispanic origin recode (mhisp_r) and mother's bridged race (mbrace) variables were used to categorize race/ethnicity for all natality metrics:

- Hispanic/Latino: Mother's Hispanic origin recode code of either Mexican, Puerto Rican, Cuban, Central or South American, or other or unknown Hispanic, as well as a mother's bridged race code of either: white, black, American Indian/Alaskan Native, or Asian/Pacific Islander.
- White: Mother's Hispanic origin recode code non-Hispanic, as well as a mother's bridged race code of white
- Black: Mother's Hispanic origin recode code of non-Hispanic, as well as a mother's bridged race code of black
- Asian: Mother's Hispanic origin recode code of non-Hispanic, as well as a mother's bridged race code of Asian/Pacific Islander

2020 Data:

- Due to the changes in NVSS birth files, racial/ethnic breakdowns for 2020 natality metrics are temporarily not available.

Confidence intervals

Multiple Cause of Death Data

We calculated 90% CI using formula below:

$$LCL90 = estimate - (1.645 \times SE(estimate))$$
$$UCL90 = estimate + (1.645 \times SE(estimate))$$

Standard errors (SE) for breast cancer, colorectal cancer, cardiovascular disease, and opioid overdose deaths metrics were calculated according to following formula outlined by Lilienfeld and Stolley[53] in a document published by the Utah Department of Health[54]:

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

SE for firearm suicides and firearm homicides were calculated according to the following formula outlined by Poisson distributions.

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

SE for premature deaths (all causes) were calculated according to the following formula outlined by Vohlonen, Bäckmand, & Korhonen:[55]

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

w_1 = Age-group specific premature deaths weight--years of life lost
w_2 = US 2010 standardized population YPLL age-group specific weight

Natality Data

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

$$\text{LCL90} = \text{estimate} - 1.645 * \sqrt{\text{estimate*((100-estimate)/numerator)}}$$
$$\text{UCL90} = \text{estimate} + 1.645 * \sqrt{\text{estimate*((100-estimate)/numerator)}}$$

CIs for teen births metric were calculated as follows:

$$\text{LCL90} = (1000 / \text{denominator}) * (\text{numerator} - (1.645 * \sqrt{\text{numerator}}))$$
$$\text{UCL90} = (1000 / \text{denominator}) * (\text{numerator} + (1.645 * \sqrt{\text{numerator}}))$$

City/County indicator

Multiple Cause of Death Data

Please note that some years of estimates for Honolulu, HI, Macon, GA, and select consolidated cities (Athens, GA; Augusta, GA; Indianapolis, IN; Louisville, KY; Nashville, TN) use data from county of residence. Consolidated cities share a government with their county, but retain unincorporated areas with distinct governments. County estimates include populations that live within unincorporated areas; city estimates do not.[42]

The Dashboard indicates where county values were utilized for MCDD reported values using an alert in the site's "Data tips" box on the Dashboard in downloadable data, available at www.cityhealthdashboard.com/data-access.

Race/ethnicity-specific estimates: Beginning of 07/17/2024, the Dashboard calculates race/ethnicity estimates based on city of residence. Due to data quality concerns, we applied a new censorship

criterion for city-level race/ethnicity data. See section “Data censoring and flagging” above or email info@cityhealthdashboard.com for more detail.

Natality Data

Total population, sex-specific estimates: The Dashboard calculates total population and sex estimates based on city of residence.

Please note that some years of total population and sex estimates for Honolulu, HI, Macon, GA, and select consolidated cities (Athens, GA; Augusta, GA; Indianapolis, IN; Louisville, KY; Nashville, TN) use county of residence. Consolidated cities share a government with their county, but retain unincorporated areas with distinct governments. County estimates include populations that live within unincorporated areas; city estimates do not.[42]

The Dashboard indicates where county values were utilized for ND reported values using an alert in the site’s “Data tips” box on the Dashboard and in downloadable data, available at www.cityhealthdashboard.com/data-access.

Race/ethnicity-specific estimates: The Dashboard calculates race/ethnicity estimates based on county of residence due to data quality issues with city-level race/ethnicity data.
See section “Use of County-Level Data on the Dashboard” above or email info@cityhealthdashboard.com for more detail.

Metric-specific notes

Breast cancer deaths

Data tables

The following underlying cause of death ICD-10 codes were summed to calculate the total number of 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.[56]

All deaths with either a missing, unknown, or not stated age are excluded from the analysis.

Analysis

$$\text{Breast cancer deaths} = \sum\left(\frac{\text{age-group specific total number of breast cancer deaths among females}}{\text{multiplier} * (\text{midpoint-year age-group specific total female population})} * \text{US 2010 standardized population age-group specific weight}\right) * 100{,}000$$

Notes on Analysis

Age-adjusted mortality rates are calculated as per National Association for Public Health Statistics and Information Systems recommendations.[57]

Cardiovascular disease deaths

Data tables

The following underlying cause of death ICD-10 codes were summed to calculate the total number of 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[58] as well as in consultation with the NYU School of Medicine's Department of Population Health.

All deaths with either a missing, unknown, or not stated age are excluded from the analysis.

Analysis

$$\text{Cardiovascular disease deaths} = \sum\left(\frac{\text{age-group specific total number of cardiovascular disease deaths}}{\text{multiplier} * (\text{midpoint-year age-group specific total population})} * \text{US 2010 standardized population age-group specific weight}\right) * 100{,}000$$

Notes on Analysis

Age-adjusted mortality rates are calculated as per National Association for Public Health Statistics and Information Systems recommendations.[57]

Colorectal cancer deaths

Data tables

The following underlying cause of death ICD-10 codes were summed to calculate the total number of 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[59] and in consultation with the NYU School of Medicine's Division of Gastroenterology.

All deaths with either a missing, unknown, or not stated age are excluded from the analysis.

Analysis

$$\text{Colorectal cancer deaths} = \sum\left(\frac{\text{age-group specific total number of colorectal cancer deaths}}{\text{multiplier} * (\text{midpoint-year age-group specific total population})} * \text{US 2010 standardized population age-group specific weight}\right) * 100{,}000$$

Notes on Analysis

Age-adjusted mortality rates are calculated as per NAPHSIS recommendations.[57]

Firearm Homicides

Data tables

The following underlying cause of death ICD-10 codes were summed to calculate the total number of deaths from assault by firearms: 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.

All deaths with either a missing, or unknown are excluded from the analysis.

Analysis

$$\text{Firearm-related Homicides} = \frac{\text{Firearm related deaths by homicide}}{\text{Total population}} \times 100{,}000$$

Firearm Suicides

Data tables

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.

All deaths with either a missing, or unknown are excluded from the analysis.

Analysis

$$\text{Firearm-related Suicides} = \frac{\text{Firearm related deaths by suicide}}{\text{Total population}} \times 100{,}000$$

Low birthweight

Data tables

Low birthweight is defined as birthweight < 2500 grams (dbwt).

Analysis

$$\text{Low birthweight} = \frac{\text{Number of live births with birthweight <2500 grams}}{\text{Total number of live births}} \times 100$$

Notes on Analysis

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

Opioid overdose deaths

Data tables

The following underlying cause of death ICD-10 codes were summed to calculate the total number of 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.[60,61]

All deaths with either a missing, unknown, or not stated age are excluded from the analysis.

Analysis

$$\text{Opioid overdose deaths} = \sum\left(\frac{\text{age-group specific total number of opioid overdose deaths}}{\text{multiplier} * (\text{midpoint-year age-group specific total population})} * \text{US 2010 standardized population age-group specific weight}\right) * 100{,}000$$

Notes on Analysis

Age-adjusted mortality rates are calculated as per National Association for Public Health Statistics and Information Systems recommendations.[57]

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

Premature deaths (all causes)

Data tables

Premature deaths (all causes) rates are calculated as per Dranger and Remington's approach.[52] Refer to NVSS: Weights above and Appendix A for more detail.

All deaths with either a missing, unknown, or not stated age are excluded from the analysis.

Analysis

$$\text{Premature deaths (all causes)} = \sum \left(\left(\frac{\text{age-group specific total number of deaths}}{\text{multiplier} * \text{(midpoint-year age-group specific total population)}} \right) * \text{(US 2010 standardized population YPLL age-group specific weight)} * \text{(age-group specific premature death weight - years of life lost)} \right) * 100{,}000$$

Prenatal care

Data tables

Prenatal care is defined as a live birth with prenatal care beginning within the first trimester.

Analysis

$$\text{Prenatal Care} = \frac{\text{Number of live births with prenatal care beginning between the first and third month}}{\text{Total number of live births}} \times 100$$

Notes on analysis

Due to changes in maternal gestational age reporting before 2014 that introduce unknown bias into the estimate across years, we do not release prenatal care estimates for 2012, 2013, and 2014. We only provided 2- year estimates for data year 2015.

Prenatal care estimates represent a slight modification of one component of the Kotelchuck Index.[62] All births with missing or unknown prenatal care are excluded from the analysis. Prenatal care data for certain states across years are missing because these states had not implemented 2003 birth certificate revisions. If prenatal care information is missing for 10% or more of a given city, all prenatal care values for that city are censored. For more information please refer to the natality public use data documentation files.[63-68]

Teen births

Teen birth is defined as a live birth to a moth ages 15-19.

Data tables

Analysis

$$\text{Teen Births} = \frac{\text{Number of live births to mothers ages 15-19}}{\text{multiplier} * \text{(midpoint-year total female population age 15-19)}} \times 1{,}000$$

New York Fed/Equifax

General notes

The Credit Insecurity Index is a measure of limited access to credit among people living in a community. The index is calculated by The New York Fed Consumer Credit Panel and Equifax and incorporates both the adult population with no formal credit history and the adult population with credit-limiting outcomes. It was calculated for each tract or city using a sample of individuals. Estimates for cities or tracts with less than 50 people sampled were censored by the data source.

Multi-year data

Multi-year data for this metric are unavailable.

Weights

Weights were not applied to tract or city level data received from The New York Fed Consumer Credit Panel and Equifax.

Categorizing race/ethnicity

The Credit Insecurity Index does not include sub-group estimates by race/ethnicity.

Confidence intervals

Confidence intervals for this metric are unavailable.

Metric-specific notes

Credit insecurity index

The Credit Insecurity Index is a community-level indicator of the population with limited credit access. The index is comprised of a weighted sum of the populations:

- With no formal credit history;
- Without a revolving credit product;
- Who have completely- or over-utilized their credit limits; and
- With low credit scores (ex: an Equifax Credit Risk score of 580 or less)

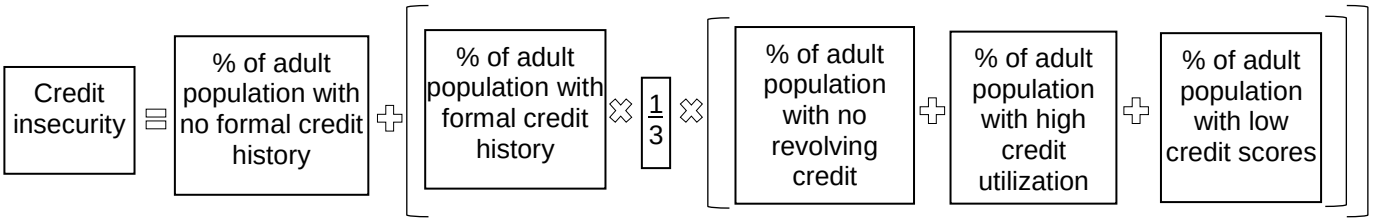
Higher Credit Insecurity Index values indicate greater credit insecurity, and lower credit access.

Data tables

Full-country tract data and city-level, for 766 selected cities only, were downloaded from the Federal Reserve Bank of New York.

Analysis

Credit insecurity is calculated by the following formula at both the tract and city level for data downloaded from the Federal Reserve Bank of New York directly:



For cities not provided by New York Federal Reserve Bank directly, the Dashboard uses a block population-weighted aggregation method to calculate city-level estimates. An aggregation indicator was applied for city estimates calculated using aggregation method. Please see Appendix E: Block Population-Weighted Aggregation for Selected Metrics for more information about how this method.

ParkServe ®

General notes

Park access represents the percent of the population living within a 10 minute walk of green space. ParkServe ® obtained 2022 GIS data on parks through outreach to cities, towns and communities with direct request for parks data. If GIS data was not provided, park or green space locations were collected from a series of resources, including municipal websites, county or state GIS data, and satellite imagery.

Properties included in ParkServe ® analyses:

- Publicly-owned local, state, and national parks
- School parks with a joint-use agreement with the local government. Considering the scale of the project, only the joint-use agreements collected through ParkScore ® were used.
- Privately-owned parks that are managed for full public use

Multi-year data

Multi-year data for this metric are unavailable.

Weights

Weights were not applied to city level ParkServe ® data received from ParkServe.

Categorizing race/ethnicity

City level estimates of park access were also calculated by race/ethnicity. Data reported by ParkServe ® included estimates by the following categories: White, Black, Asian, Pacific Islander, Native American, other race, two or more races, and Hispanic. Estimates for Asian and Pacific Islander were aggregated to match our Asian/PI designation. Similarly, estimates for other race, two or more races, and Native were also aggregated. Refer to Section 2 “Race/Ethnicity Categories” (above) for more detail on Dashboard definitions.

Confidence intervals

CIs for Park Access were calculated as follows:

$$LCL90 = estimate - (1.645 \times SE(estimate))$$
$$UCL90 = estimate + (1.645 \times SE(estimate))$$

Metric-specific notes

Park access

This metric represents the percent of the population that lives within a 10-minute walk of a park or publicly accessible green space.

Data tables

The Dashboard reports ParkServe ® calculated city level estimates as received directly from ParkServe ®; data are not available for download from the ParkServe ® website.

City Health Dashboard analytic staff calculates tract level estimates using non-publicly available 2022 ParkServe ® GIS data representing the 10 minute walk radius around each park and 2022 Esri block group demographic estimates.

Analysis

City level data reports numerators (population living ≤10-minute walk of a park/green space) and denominators for total population and racial/ethnic subgroups. The formula for city-level values is:

$$\frac{\text{Population living} \leq \text{10 minute walk of a park/green space}}{\text{Total population}} \times 100\%$$

Tract level park access values were calculated using block group population estimates. Block groups are "building blocks" for census tracts. The formula for tract-level values is:

$$\frac{\sum(\text{(\% of block group covered by 10 minute walk radii)(Total block group population)})}{\sum(\text{Total block group population})} \times 100\%$$

Manual FIPS recoding was required for seven Pennsylvania cities which were identified as county subdivisions in the ParkServe ® data. Comparison of the boundaries showed that geographies identified by the county subdivision FIPS and place FIPS are equivalent, except in the case of Bethlehem, Pennsylvania where the place is equivalent to two separate county subdivisions combined. Additionally, due to minor rounding inconsistencies in the data, the population with park access in two instances for the city of Hoboken exceeded the total population by one individual and were therefore rounded down.

PLACES Project (formerly 500 Cities Project), Centers for Disease Control and Prevention

General notes

Measures of health status, health behaviors, and clinical care were estimated by the Centers for Disease Control and Prevention's 500 Cities Project[69], renamed the PLACES Project in 2020 to reflect its expansion to more geographies[70]. The PLACES Project uses identical methods to the 500 Cities Project, except that PLACES Project only releases estimates for full census tracts, whereas 500 Cities Project releases estimates for census tract portions within city boundaries (see "Geographic Coverage" section here)[70].

The PLACES and 500 Cities Projects 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 500 Cities Project, BRFSS measures were only available at the county, Metropolitan Statistical level or above. For further details on the methodology, see Zhang et al (2014).[71] For more information regarding these metrics, please refer to the PLACES Project's methodology pages.[72-74]

The Dashboard reports most data as received, with the exception of the preventive service utilization values and CI values. Estimates for New Jersey in 2019 and Florida in 2021 are not available as the states did not collect enough BRFSS data to meet minimum requirements for inclusion.[75] The underlying data source, BRFSS, issued a warning for their 2020 data noting that 2020 data may not be comparable to other years due to difficulties collecting samples during the 2020 pandemic, and an underlying weighting method modification.[75]

The following table outlines the years of PLACES Project and 500 Cities Project data available on the Dashboard, per metric. Some metrics listed are unavailable for certain years because BRFSS asks some questions every other year. For more information, consult the PLACES Project website.[76]

Metric	2013	2014	2015	2016	2017	2018	2019	2020	2021
Binge drinking		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
Dental care		✓		✓		✓ (PLACES)		✓ (PLACES)	
Diabetes		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
Frequent mental distress		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
Frequent physical distress		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
High blood pressure	✓		✓		✓ (PLACES)		✓ (PLACES)		✓ (PLACES)
Obesity		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
Physical inactivity		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
Preventive services, 65+		✓		✓		✓ (PLACES)		✓ (PLACES)	
Routine checkup, 18+		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)
Smoking		✓	✓	✓	✓	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)	✓ (PLACES)

Weights

The Dashboard reports PLACES Project and 500 Cities Project data as received, so in general, no weights are applied in the calculation of the estimates by the Dashboard analysts. (Please refer to the previous citations to learn more about how post-stratification weights are applied in the modeling process.) The one exception is the measure of preventive service utilization, 65+, which is reported separately for men and women in the 500 Cities data. Though the Dashboard reports the rates by sex, a weighted average rate is also calculated to get a total population rate.

The Dashboard uses a block population-weighted aggregation method to calculate city-level estimates for cities represented by County Subdivisions (cities represented by Incorporated Places are available directly from PLACES). Please see *Appendix E: Block Population-Weighted Aggregation for Selected Metrics* for more information about how this method, and *Appendix B* for more information about city selection.

Categorizing race/ethnicity

Estimates from the PLACES Project and 500 Cities Project do not include sub-group estimates by race/ethnicity. Race/ethnicity, age, and income are included as covariates in the MPR approach used to calculate modeled estimates.

Importantly, only crude (not age-adjusted) measures are available at the census tract level. The PLACES Project and 500 Cities Project report both crude and age-adjusted estimates at the city level. For consistency and comparability between tract and city estimates, the Dashboard reports crude estimates for both tracts and cities.

Confidence intervals

Confidence intervals were included with the estimates downloaded from the 500 Cities Project. However, the 500 Cities Project reports 95% confidence intervals, rather than the 90% confidence intervals reported by the Dashboard. Upper and lower limits of the 95% confidence intervals were used to calculate an approximate standard error (SE). The SE was then used to calculate 90% confidence intervals. See Preventive services, 65+ below for metric-specific confidence interval calculations.

$$SE = \frac{UCL95 - LCL95}{1.96 \times 2}$$

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

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

Where:

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

Data tables

Tract and city-level data were downloaded directly from the PLACES Project and 500 Cities Project website.[7,76-80]

Metric-specific notes

Definitions are taken verbatim from the PLACES Project and 500 Cities Project:

Binge drinking

Definition: Adults aged ≥18 years who report having five or more drinks (men) or four or more drinks (women) on an occasion in the past 30 days.[73]

Dental care

Definition: Percent of respondents aged ≥18 years who report having been to the dentist or dental clinic in the previous year.[74]

Diabetes

Definition: Respondents aged ≥18 years who report ever been told by a doctor, nurse, or other health professional that they have diabetes other than diabetes during pregnancy.[72]

Frequent physical distress

Definition: Respondents aged ≥18 years who report 14 or more days during the past 30 days during which their physical health was not good.[72]

Frequent mental distress

Definition: Respondents aged ≥18 years who report 14 or more days during the past 30 days during which their mental health was not good.[72]

High blood pressure

Definition: Respondents aged ≥18 years who report ever having been told by a doctor, nurse, or other health professional that they have high blood pressure. Women who were told high blood pressure only during pregnancy and those who were told they had borderline hypertension were not included.[72]

Obesity

Definition: Adult obesity among adults aged ≥18 years.[73]

Physical inactivity

Definition: Respondents aged ≥18 years who answered "no" to the following question: "During the past month, other than your regular job, did you participate in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise?"[73]

Preventive services, 65+

Please note that PLACES modified their inclusion criteria for being up to date on core preventative services in 2020.

Definition, Women: Number of women aged ≥65 years reporting having received all of the following: an influenza vaccination in the past year; a pneumococcal vaccination (PPV) ever; either a fecal occult blood test (FOBT) within the past year, a sigmoidoscopy within the past 5 years and a FOBT within the past 3 years, or a colonoscopy within the previous 10 years; and a mammogram in the past 2 years.[74]

Definition, Men: Number of men aged ≥65 years reporting having received all of the following: an influenza vaccination in the past year; a PPV ever; and either a fecal occult blood test (FOBT) within the past year, a sigmoidoscopy within the past 5 years and a FOBT within the past 3 years, or a colonoscopy within the past 10 years.[74]

Analysis

At the recommendation of the PLACES analytic team[81], 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 (DCS) population counts for place and tract weights, and the Census' Population Estimates Program (PEP) population estimates for county[70,81].

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+
$p_{female\ 65+}$ = reported proportion of overall use of preventive services by women 65+
$n_{male\ 65+}$ = population, men 65+
$n_{female\ 65+}$ = population, women 65+

Note that $n_{male\ 65+}$ and $n_{female\ 65+}$ represent the 65+ population in an entire tract, while $p_{male\ 65+}$ and $p_{female\ 65+}$ represent only the 65+ population living within city limits for 500 Cities data. This discrepancy is only relevant for years 2014 and 2016 where 500 Cities is the data source. After 2018 (PLACES data), $p_{male\ 65+}$ and $p_{female\ 65+}$ are both full tract proportions..

To calculate our pooled MOE, we performed a series of steps:

1. For male and female, convert *upper* MOE to standard error (SE)

$$SE=\frac{MOE_{upper}-\hat{p}}{1.96}$$

2. For male and female, transform standard error into variance (var)

$$var=\left(SE*\sqrt{n}\right)^2$$

3. Pool the variances into a pooled standard deviation

$$SD_{pooled}=\sqrt{\frac{(n_{male\ 65+}-1)*var_{male}+(n_{female\ 65+}-1)*var_{female}}{n_{male\ 65+}+n_{female\ 65+}-2}}$$

4. Transform pooled standard deviation into standard error

$$SE_{pooled}=SD_{pooled}*\sqrt{\frac{1}{n_{male\ 65+}}+\frac{1}{n_{female\ 65+}}}$$

5. Compute pooled MOE at the 90% confidence level.

$$MOE_{pooled,\ 90\%}=SE_{pooled}*1.645$$

Where:
n = population (by sex)
SD_{pooled} = pooled standard deviation
SE_{pooled} = pooled standard error
MOE_{pooled} = pooled margin of error

Note that, for Preventive services, 65+ *only, the* MOE remains the same on both sides compared to other PLACES-derived metrics. PLACES uses other methods to generate their confidence limits[81], where the Dashboard uses population parameters to calculate MOE and confidence limits.

Routine checkup, 18+

Definition: Respondents aged ≥18 years who report having been to a doctor for a routine checkup (eg., a general physical exam, not an exam for a specific injury, illness, condition) in the previous year).[74]

Smoking

Definition: Respondents aged ≥18 years who report having smoked ≥100 cigarettes in their lifetime and currently smoke every day or some days.[73]

Stanford Education Data Archive (SEDA)

General notes

SEDA is part of the Educational Opportunity Project at Standford University, an initiative aimed at harnessing data to help scholars, policymakers, educators, and parents learn how to improve educational opportunity for all children. SEDA 5.0 achievement data are derived from the ED*Facts* data system, provided by the United States Department of Education. States report to ED*Facts* the number of students in each school, subgroup, grade and year that perform at that state's defined achievement level. SEDA obtains these data and applies a multi-step HETOP (Heteroskedastic Ordered Probit) and a pooled HETOP model to scale and aggregate average test scores from the individual school reported data to the city-level. For further details on the methodology, please see SEDA Technical Documentation[82].

The Dashboard provides SEDA a school-to-city crosswalk. SEDA uses this crosswalk to calculate city-level third grade reading estimates. The provided education estimates are city-level estimates, not school district-level estimates. Census tract-level estimates are not available. Formore detailed information about how schools were assigned to cities, please refer to Appendix D.

Multi-year data

Multi-year data are available from academic years 2010-2011 through 2018-2019.

Weights

No weights are applied by Dashboard analysts.

Categorizing race/ethnicity

SEDA provides, where available, disaggregated data by race/ethnicity and by sex for third-grade reading scores. Hispanic category is mutually exclusive with other race categories. The 'Other' category includes American Indian or Alaska Native, Two or more races, and Some other race. Asian includes Asian, Native Hawaiian and Pacific Islander. Refer to Section 2 "Race/ethnicity categories" (above) for more detail.

Confidence intervals

Confidence intervals are calculated and provided by SEDA, and follow the calculation:

$$LCL90 = estimate - (1.645 \times SE(estimate))$$
$$UCL90 = estimate + (1.645 \times SE(estimate))$$

Metric-specific notes

Third-grade reading scores

This measure represents the average reading test scores (in grade levels) of third graders in public schools in the city.

Data tables

SEDA reports reading test scores scaled around a national average grade level of 0. Dashboard analysts added 3 to each estimate calculated by SEDA to scale city estimates around a national average of 3 (for 3rd grade).

Analysis

Please see SEDA Technical Documentation[82] for detailed information on the derivation of the measures.

Additional notes

City-level data may be presented as missing for a variety of reasons. Data for some years may not have been reported by a state to the Department of Education, inducing missing values. Other data may lack a sufficient number of observations to generate a stable estimate – city estimates were censored if they represented less than 20 students or if there were fewer than 100 assessments completed by a group in that year. SEDA includes state and year level reasons for missing data in their Technical Documentation, Table 4b.

SEDA recommends users to exercise caution when comparing estimates over time, as total number of students may fluctuate based on district and state reporting to ED*Facts*.

Users may see a substantial decrease in city-level estimates in 2014. That year, 15 states received double testing waivers, which means students may have taken an alternate assessment in lieu of the reported state assessment. As a consequence, cities in these states may not have data for 2014 because data were not collected nor reported.

California
Connecticut
Idaho
Illinois
Iowa
Maryland
Massachusetts
Mississippi
Montana
Nevada
Oregon
South Dakota
Vermont
Washington
Wyoming

Data from Texas is not available across all years due to different tests.

Users should note that with the release of SEDA 5.0, some city level estimates were updated to reflect more rigorous methods. Most apparently, Texas city-level estimates for all years that were reported in SEDA 4.1 are now suppressed, along with estimates in East Orange, NJ for 2016 and Gresham, OR for 2018.

As the value for a city does not reflect a city's reading proficiency but rather a city's students' reading scores in context of a national modeled score, the Dashboard does not release numerators or denominators for this measure.

United States Small-Area Life Expectancy Project (USALEEP)

General notes

Tract-level 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.[83]

Multi-year data

Multi-year data for this metric are unavailable.

Weights

Weights are used to derived city-level estimates from tract data; see “Analysis” section below for more details.

Categorizing race/ethnicity

Not applicable.

Confidence intervals

Tract-level standard errors are included in downloadable USALEEP data. Ninety percent confidence intervals for tract-level data were calculated as per the following formulas:

$$LCL90 = estimate - (1.645 \times SE(estimate))$$
$$UCL90 = estimate + (1.645 \times SE(estimate))$$

Where:

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

Confidence intervals are not provided for city estimates.

Metric-specific notes

Life expectancy

Data tables

The Dashboard downloaded tract-level from USALEEP.[84-86]

Analysis

Tract-level data are reported as received.

To calculate city-level estimates, block-level population weights were applied to the census tract annual averages and the weighted block values were summed to the city level. Please see *Appendix E: Block Population-Weighted Aggregation for Selected Metrics* for more information about how block-level weights are calculated and applied. This method was approved per email correspondence with USALEEP analytic staff at the National Center for Health Statistics.

Walk Score ®

General notes

Walk Score measures the walkability of any address[87] by incorporating walking routes to nearby amenities from multiple categories.[88] This metric is a value from 0 to 100, with 100 being more walkable.

Multi-year data

Multi-year data for this metric are unavailable.

Weights

The Dashboard reports Walk Score data as received. No weights were applied.

Categorizing race/ethnicity

Walk Score data are not categorized by race/ethnicity.

Confidence intervals

CIs are not presented for Walk Score ® data.

Metric-specific notes

Walkability

The Dashboard reports 2022 aggregated city scores calculated from tract level data received from Redfin Real Estate.[89]

Analysis

Tract level data are presented as received.

To calculate city-level estimates, block-level population weights were applied to the Redfin Real Estate tract scores and the weighted block values were summed to the city level. Please see *Appendix E: Block Population-Weighted Aggregation for Selected Metrics* for more information about how block-level weights are calculated and applied.

For more detailed information, please refer to the "How Walk Score Works" webpage, available at https://www.redfin.com/how-walk-score-works, and the Walk Score Methodology page, available at https://www.walkscore.com/methodology.shtml.

SECTION 4: 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, Eileen Shea, Anne Vierse, Jay Stadelman, Noah Zazanis, Shirley Liang, 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 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.

The Dashboard acknowledges Dr. David Van Riper and Dr. Jonathan Schroeder from Integrated Public Use Microdata Series (IPUMS), Dr. Yan Wang and Dr. Hua Lu from PLACES Project, Glenn Rice (Programmer / Web Manager) from MCDC GeoCorr, and ACS, for their expertise surrounding geographic aggregation methods.

SECTION 5: Appendices

Appendix A: Table of US 2010 Standardized Population

Refer to NVSS: Weights (MCDD) and Premature deaths (all causes): Notes on analysis above for detail on how these weights were calculated.

Table of US 2010 Standardized Population

Age Group	Number	Weight	Weight for YPLL Age-adjustment	Premature Deaths Weight (Years of Life Lost)
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

Appendix B: Detailed Notes on Selection of Cities and Tracts

City Selection

The Dashboard defines 'cities' using multiple U.S. Census Bureau geographies that represent locally recognized communities or legal entities with active governments:

1) Incorporated and Designated Places[1,3]
    a. NOTE: Florida designated places are currently excluded.
2) County Subdivisions, specifically Minor Civil Divisions.[2]
    a. NOTE: Only select states (CT, MA, ME, MI, MN, NH, NJ, NY, PA, RI, VT, WI) contain Minor Civil Divisions with active governments.
3) Counties, for Hawaii only
    a. The Dashboard is committed to providing data for all U.S. states. Hawaii is the only state without Incorporated Places, and HI's Designated Places do not meet our population cut-off (see below). Therefore, we have elected to provide data for HI counties, which serve similar functions as city governments in Hawaii.

Incorporated and Designated Places, Minor Civil Divisions, and Hawaiian Counties across the country are included on the Dashboard if their population in the 2020 US Census is greater than 50,000. Select smaller cities with population <50,000 are also included:

1) Cities already included from the 2018 Dashboard launch based on the CDC's 500 Cities Project[69]
2) New Jersey cities added with financial support provided by New Jersey Health Initiatives[4]
3) Cities added through the Dashboard's 2022, 2023, and 2024 "Put Us on the Map" Challenges[90,91]
4) Cities for which the federal government's 1930's Home Owners' Loan Corporation (HOLC) program created redlining maps[92]

All city selection is performed in R with support of the tidycensus, tigris and sf packages.[13,14,16] Please email info@cityhealthdashboard.com for more information about this process.

Tract Selection

Census tracts are included on the Dashboard if they overlap with a Dashboard city boundary. Overlap is determined using Census blocks (the lowest-level US Census geography).[93] If at least one block within the census tract overlaps with a city boundary, then the tract is assigned to that city. Therefore, tracts may be assigned to multiple cities. Website maps represent the portion of the tract within the city boundary, not the entire census tract, whereas tract estimates on the website represent the entire tract.

Block relationship files for places and county subdivisions are pulled from the US Census Bureau and the Missouri Census Data Center.[94,95] All tract selection is performed in R with support of the tigris package.[14] Please email info@cityhealthdashboard.com for more information about this process [96]

FIPS Codes

The Dashboard uniquely identifies geographies using the US Census Bureau FIPS Codes system.[97] Details for the number of digits in FIPS codes for Dashboard cities and tracts are below, and also detailed in the Downloadable Data Codebook, available at www.cityhealthdashboard.com/data-downloads.

A unique 11-digit tract FIPS is created by combining the 2-digit state FIPS, 3-digit county FIPS, and 6-digit tract FIPS. Place FIPS combines the 2-digit state FIPS and 5-digit place FIPS. And the county subdivision FIPS combines the 2-digit state FIPS, 3-digit county FIPS, and 5-digit county subdivision FIPS.

Note on Hawaiian FIPS code

County FIPS codes (2-digit state FIPS, 3-digit county FIPS) and county data are used to represent the Hawaiian counties presented on the Dashboard.

Census Vintages

The Dashboard presents cities and census tracts from the 2010 and 2020 Census, depending on the underlying Census vintage of metric data sources. Any tract-level FIPS code updates that occurred between Decennial Census are not reflected on the Dashboard, and data sources that have incorporated these changes are reverted to the Decennial Census FIPS codes.[96] City-level FIPS code updates are incorporated on the website. Please email info@cityhealthdashboard.com if you would like more details.

Note on Macon, GA

As of 2014, the city of Macon, GA and its county (Bibb County) share a consolidated government.[8,9] The US Census Bureau began to issue data for Macon, GA using FIPS code 13-49008 (Macon-Bibb County, Georgia) in 2014; FIPS 13-49000 (Macon city, Georgia) was used through 2013.

For all years of data available on the Dashboard, the Dashboard uses the geographic boundaries and FIPS codes associated with the consolidated city (FIPS 13-49008), which is equivalent to the county (FIPS 13-021). This decision was made to keep the geographic area and population represented in the data as consistent as possible across time, despite the changing governmental structure. For some years for certain metrics, only values for Macon city, GA (FIPS 13-49000) are available. When this occurs, it is noted on the website and in downloadable data. Please see the Downloadable Data Codebook, available at www.cityhealthdashboard.com/data-downloads, for more details, or email info@cityhealthdashboard.com.

Appendix C: Summary of State-Based Vital Statistics Data Sources

New Jersey State Health Assessment Data (NJSHAD)

General notes

To provide vital statistics estimates for a subset of New Jersey cities (Burlington, Clayton, Egg Harbor City, Glassboro, Hammonton, Lawnside, Millville, Penns Grove, Pleasantville, Salem), the Dashboard uses data from New Jersey State Health Assessment Data (NJSHAD),[47] a public health data source managed by the New Jersey Department of Health.

Data were accessed and analyzed using procedures identical to those in NCHS/RDC analyses of NVSS data, as defined above in Section 3.

As per private correspondence with NJSHAD staff, analyses by Dashboard analytic staff, and consultation with an NYU School of Medicine faculty expert in vital statistics, estimates from these data sources are valid for comparison with NCHS/RDC NVSS data presented on the Dashboard. To establish comparability between NJSHAD and RDC/NCHS data, Dashboard staff compared estimates for New Jersey cities already on the Dashboard generated from NJSHAD data with estimates generated at RDC/NCHS. In consultation with a NYU School of Medicine faculty expert in vital statistics, Dashboard staff concluded that estimates from both datasets were identical or of inconsequential difference. Please contact info@cityhealthdashboard.com for more information on the analyses that establish comparability between NJSHAD and NCHS/RDC data or any other questions or concerns.

Weights

Note that NJSHAD provides counts of deaths and births by 5-year age groups. A subset of age strata used in Dashboard age adjusted analyses use 10-year age groups (see Section 7, Appendix A). When necessary, age groups were summed to mirror the age groups defined above in Section 7, Appendix A.

Metric-specific notes

Opioid overdose deaths

Not calculated by Dashboard staff due to insufficient data availability on NJSHAD.

Prenatal care

Not calculated by Dashboard staff due to insufficient data availability on NJSHAD.

Appendix D: School-to-City Crosswalk Creation

In order to calculate city-level chronic absenteeism and third-grade reading scores metrics, the Dashboard uses crosswalks that identify which schools fall within 200 meters of each Dashboard city.

School to Dashboard city crosswalks were created for each school year using ArcGIS Pro 3.0.0 software and school data from the National Center for Education Statistics (NCES). Only schools with full geographic coordinates (latitudes and longitudes) were included.

School latitude and longitude data was uploaded to ArcGIS and converted to XY Data. Dashboard city shape files were uploaded to ArcGIS and a Near Analysis was completed between the XY Data and city shape layers using a location and geodesic method with a search radius of 200 meters. Only schools that were matched to a city boundary were kept in the final sample of schools.

Each school was matched to only one city, with the exception of schools in cities with overlapping boundaries. These are: 1) Hempstead Village, NY and Hempstead Town, NY; 2) Levittown, PA and Bristol Township, PA.

When cities are adjacent to one another (not overlapping), schools are assigned to the city they are located within and the 200 meter boundary does not apply. This may result in schools getting "reassigned" to a different city as new cities are added to the Dashboard.

Appendix E: Block Population-Weighted Aggregation from Census Tract to City for Selected Metrics

General notes

Census tract boundaries do not perfectly align with city boundaries. When city-level estimates are unavailable and where appropriate, the Dashboard uses a block population-weighted aggregation method to estimate city estimates from tract[98-100]. This method is conceptually similar to a dasymetric approach, which accounts for population density distribution across the census tract and assumes that the underlying population and metric estimate values are equally distributed across the entire block.

Weights

Population weights are created using information at the block-level. Block relationship files for places and county subdivisions are pulled from the US Census Bureau and the Missouri Census Data Center.[94,95] Block populations come from the Decennial Census.

Different weights are applied depending on the underlying population of the metric being aggregated. Variables used for different weights are detailed below.

Weight	Decennial Census	File (Table)	Variable(s)
Total population	2010	Supplemental File – 1 (P-9)	*P009001 – Hispanic or Latino, and Not Hispanic or Latino by Race!!Total*
Age 18+	2010	Supplemental File – 1 (P-10)	*P010001 – Race for the Population 18 Years and Over!!Total*
Age 65+	2010	Supplemental File – 1 (P-12)	• *P012020 – Sex by Age!!Total!!Male!!65 and 66 years* • *P012021 – Sex by Age!!Total!!Male!!67 to 69 years* • *P012022 – Sex by Age!!Total!!Male!!70 to 74 years* • *P012023 – Sex by Age!!Total!!Male!!75 to 79 years* • *P012024 – Sex by Age!!Total!!Male!!80 to 84 years* • *P012025 – Sex by Age!!Total!!Male!!85 years and over* • *P012044 – Sex by Age!!Total!!Female!!65 and 66 years* • *P012045 – Sex by Age!!Total!!Female!!67 to 69 years* • *P012046 – Sex by Age!!Total!!Female!!70 to 74 years* • *P012047 – Sex by Age!!Total!!Female!!75 to 79 years* • *P012048 – Sex by Age!!Total!!Female!!80 to 84 years* • *P012049 – Sex by Age!!Total!!Female!!85 years and over*
Total population	2020	Census Demographic and Housing Characteristics File (P9)	*P9_001N – Hispanic or Latino, and Not Hispanic or Latino by Race!!Total:*

Analysis

Using the block relationship files to assign census tracts to cities, we then created a population weight (*P)* by dividing the overlapping population count by the full city population count. Population counts from source geographies with missing estimates were dropped from the calculation. We multiplied this population weight by the census tract estimate, then summed all weighted estimates to calculate the final derived city estimate (see Equation).

$$Est_{derived,\ city} = \sum_{i=1}^{n} est_{census\ tract\ i} * P_{(census\ tract\ i\ population\ in\ city\ |\ city\ population)}$$

Where: n represents the number of tracts (*i*) within the city boundary,

Metrics

This aggregation method is currently applied to the following metrics and geographies.

Metric	Weight Used	Places	County Subdivisions	HI Counties
Air pollution – PM 2.5	Total population	✓	✓	✓
Air pollution – Ozone	Total population	✓	✓	✓
Binge drinking	Age 18+		✓	
Credit Insecurity Index	Total population	✓ *(select cities)*	✓	✓
Dental care	Age 18+		✓	
Frequent mental distress	Age 18+		✓	
Frequent physical distress	Age 18+		✓	
High blood pressure	Age 18+		✓	
Life expectancy	Total population	✓	✓	✓
Obesity	Age 18+		✓	
Physical inactivity	Age 18+		✓	
Preventive services, 65+	Age 65+		✓	
Routine checkup, 18+	Age 18+		✓	
Smoking	Age 18+		✓	
Voter Turnout	Total population		✓ *(ACS only)*	
Walkability	Total population	✓	✓	✓

Appendix F: Updates Summary

Technical Document Version	Date Posted Online	Update Notes

	7/27/2024	• Addition of 130 new cities, including Hawaiian counties, designated places, 2024 "Put Us On the Map" cities, and historically redlined cities
		  ○ Slight modification to tract inclusion for a few existing cities due to updated methodology
		• Updated city names for 48 townships in MI, NJ, and PA
		• ACS metrics: Add 2022 (5 year estimate) data
		  ○ High school completion: Addition of age breakdowns
		  ○ Housing with potential lead risk & Lead exposure risk index: Apply city-level censorship criteria to tracts; Update housing weights
		  ○ Income inequality: Limit metric to 2018 and 2022 data years
		  ○ Independent living difficultly: Release of new metric (2015-2022), including age breakdowns
		  ○ Racial/ethnic diversity & Neighborhood racial/ethnic segregation: Marginal change in estimates for select cities impacted by changes in tract inclusion
		  ○ Unemployment: Addition of age breakdowns; Update to population weighting variable for Asian and Other resulting in small estimate changes for previous years of data
		• Air pollution – Ozone: Add 2023 monthly data
		• Air pollution – PM2.5: Add 2023 monthly data
		• Chronic absenteeism
		  ○ Small estimate changes for select due to addition of new cities
		  ○ Select cities newly censored in 2022 due to unreliable state data
		• Demographics overview: Update to 2022 data; Addition of new breakdowns
		• NVSS/NJSHAD (Mortality metrics only)
		  ○ Add 2021 data
		  ○ Update race/ethnicity estimates to represent cities instead of counties, including new censorship criteria
		  ○ Data newly available for more cities
		• PLACES Project metrics: Update population weight for select aggregated city-level estimates
		• Third-grade reading scores:
		  ○ Add 2019 data
		  ○ Data newly available for more cities
		  ○ Select cities newly missing due to updated data source methodologies
		• Voter turnout: Release of new metric (2020), including sex and age breakdowns

	12/5/2023	• PLACES Project
		  ○ Add 2021 data for 9 metrics
		  ○ Preventive services, 65+: Update data source for Honolulu calculation weights
		• Chronic Absenteeism
		  ○ Update metric name from "Absenteeism" to "Chronic Absenteeism"
		  ○ Add 2020 – 2022 school years data using updated data sources
		  ○ Replace 2018 school year estimates using updated data sources
		  ○ Expand data availability for cities below ~66,000 population, including "Put Us On the Map" cities
		• Unemployment – current, city-level: Add data through July 2023

	7/26/2023
	(going forward, versioning will represent date of data release)
		• ACS metrics: Add 2021 (5 year estimate) data
		  ○ Uninsured: removal of age breakdowns in 2013, 2014
		  ○ Minor change to confidence interval calculation methods
		• Air pollution – Ozone: Add 2022 monthly data
		• Air pollution – PM2.5: New data source (monthly data for 2022); Removal of existing multi-year metric
		• COVID Local Risk Index: Removal of metric
		• Credit insecurity index: New method for generating some city-level estimates
		• Life expectancy: New method for generating city-level estimates
		• Limited access to healthy foods: Removal of metric
		• NVSS/NJSHAD: Correction of 2014 and 2017 estimates for breast cancer deaths, colorectal cancer deaths, cardiovascular disease deaths, premature deaths (all cause), and 2012, 2015, 2018 estimates for low birthweight and teen births. These

		corrections were applied to 10 New Jersey cities using New Jersey State Health Assessment. County estimates for these metrics for these cities were unintentionally posted on the website between 7/27/2022 – 08-02-2023. This error was also in downloadable data version 15.0, 15.1 and 16.0. • NVSS Breast Cancer Deaths 2013: We discovered our breast cancer deaths for 2013 was a complete duplication of 2012 data from 1/13/2020 (version 8.0) to 7/21/22 (version 15.0). This error has been fixed in version 15.0 release. • Park Access: Update to 2022 data • PLACES Project: Add 2020 data for 10 metrics • Unemployment – current, city-level: Add data through December 2022 • Walkability: Update to 2022 data; New method for generating city-level estimates • Addition of new city estimate censorship criteria • Addition of 67 smaller cities from the 2023 "Put Us On the Map" Challenge ○ Addition of prior years of data for 2022 "Put Us On the Map" cities for most metrics, where available • Modification of FIPS code for Macon, GA when using 2010 geographies • City overview, demographic estimates: Update to 2021 data
16.0	12/14/22	• Housing cost, excessive: replace housing cost, excessive with new metric Rent burden (ACS) • Air pollution -ozone: Add 2021 monthly data • Unemployment – current, city-level: Add data through Aug 2022 • Violent Crime: Removal of metric
15.1	9/6/22	• Firearm homicides: Release of new metric; 2014 - 2020 data (5 Year Estimates) • Firearm suicides: Release of new metric; 2014 - 2020 data (5 Year Estimates) • Unemployment – current, city-level: Add data through May 2022

15.0	7/21/22	• ACS metrics: Add 2020 (5 year estimate) data
		• NVSS metrics: Add 2018, 2019, and 2020 (3 year estimate) data
		• Addition of 111 new cities > 50,000 population
		  ○ 66 represented by County Subdivisions (Minor Civil Divisions)
		  ○ 45 represented by Incorporated Places
		• Addition of 29 smaller cities from the "Put Us On the Map" Challenge[90]

14.0	3/29/22	• PLACES Project: Add 2021 release (2019 data) for 9 metrics • Air pollution – particulate matter: Add 2018 data; city calculation method update • Air pollution – ozone: Release of new metric; monthly data from 1/2018-12/2020 • Credit insecurity index: Release of new metric; 2020 data
13.0	11/1/21	• Routine Checkup: Release of new metric; 2014, 2015, 2016, 2017, 2018; 500 Cities/PLACES Project (1 year estimate) data • NVSS metrics: Change data_year_type format to "2019, 3 Year Estimate" (from "2017-2019") • Limited Access to Healthy Foods update: add 2019 data • Third-Grade Reading scores: Change data source from state-based education metrics to SEDA; add data for 2011-2018 • Absenteeism update: update to 2018 data, change data source to CRDC and ED*Facts*
12.0	6/1/21	• ACS metrics: Add 2019 (5 year estimate) data • Air pollution: Add 2017 CMAQ data • Violent crime: Add 2019 UCR data • Broadband connection: Release of new metric; 2017, 2018, 2019 ACS (5 year estimate) data • Unemployment – annual, neighborhood-level: Slight revision of estimates to improve precision; see page 41 for more details.
11.0	3/1/21	• COVID Local Risk Index: Replace with updated estimates (new components and weighting); expand to all 766 Dashboard cities • PLACES Project (formerly 500 Cities Project): 2020 release; expand to all 766 cities for 2018 estimates and High blood pressure (2017 estimates)

		• Preventive services, 65+: Name changed from "Preventive services"; Revision of 2014, 2016 estimates and confidence levels to improve precision
10.3	2/1/21	• Income inequality: 20th and 80th percentile cut points updated for 2013, 2014 to more closely approximate underlying distribution
10.2	1/7/21	• Unemployment – annual, neighborhood-level: Metric added back to website; still available for download
10.1	10/29/20	• High school completion: Revision of metric analysis and data source; addition of tract-level data; addition of multi-year data; revision of metric name (formerly "High school graduation") • Unemployment – current, city-level: Release of new metric • Unemployment – annual, neighborhood-level: Revision of metric name (previously "Unemployment"); temporarily removed from website; still available for download • Technical Document Part 2, Education Data no longer released: Addition of Absenteeism, High school completion and Third-grade reading proficiency metrics to Technical Document Part 1
10.0	6/4/20	• COVID Local Risk Index: Release of metric for cities represented in 500 Cities Project
9.0	5/1/20	• Addition of 256 cities with populations >50,000 and not already included in 500 Cities[101] *(financial support provided by the Robert Wood Johnson Foundation)* • ACS metrics: Updated to 2018 (5 year estimate) data; 2018 ACS data for Macon, GA represented by FIPS 1349008 (Macon-Bibb County, GA) instead of 1349000 (Macon, GA) or 13021 (Bibb County, GA) • Life expectancy: Addition of data for Maine and Wisconsin as per update to USALEEP • NVSS metrics: Addition of data for some years and metrics for consolidated cities (Athens, GA; Augusta, GA; Indianapolis, IN; Louisville, KY; Nashville, TN) • Walkability: Updated to 2019; Methodology updated from tract centroid value to entire tract population weight
8.0	1/23/20	• Addition of subset of New Jersey cities (Burlington, Clayton, Egg Harbor City, Glassboro, Hammonton, Lawnside, Millville, Penns Grove, Pleasantville, Salem) *(financial support provided by New Jersey Health Initiatives)* • Air pollution: Added 2016 CMAQ data • Park access: Updated to 2018 (city); addition of tract data • Violent crime: Added 2018 UCR data • 500 Cities Project metrics: Added 2019 release • NVSS metrics: Use of NJSHAD data source for NJ cities listed above
7.0	8/29/19	• NVSS metrics: Added 2010-2012, 2011-2013, 2012-2014, 2015-2017 data ○ Prenatal care: 2010-2012, 2011-2013, 2012-2014 not released • NVSS metrics: Revisions to 2013-2015 and 2014-2016 data for improved accuracy
6.0	7/15/19	• Correction of Table of Contents and section label enumeration; no changes to Technical Document context or version number
6.0	6/5/19	• Information on Absenteeism moved to Technical Document Part 2: Education Data • Multi-year data added to site and associated metadata added to Technical Documentation ○ Added multi-year data, by metric: 500 Cities Project metrics: 2016 and 2017 releases; Air pollution: 2015; ACS metrics: 2013, 2014, 2015, 2016 (5 year estimate) data; CMAQ: 2013 data; NVSS metrics: 2013-2015 data; UCR metric: 2016 data
5.0	2/14/19	• ACS metrics: Updated to 2017 (5 year estimate) data ○ Uninsured: Note revision of age strata • Air pollution: Updated to 2014 CMAQ data; city population denominator changed from ACS DP05 2016 (5 year estimate) to ACS DP05 2014 (5 year estimate)

4.1	12/20/18	• Note regarding temporary removal of primary care physicians data from Dashboard inserted in this document
4.0	12/12/18	• 500 Cities metrics: Updated to 2018 release • NVSS metrics: Updated to 2014-2016 data
3.1	October 2018	• Absenteeism: Data source updated to 2015-16 from 2013-14 (10/29/18) • Life expectancy: Metric posted 10/2/18; minor revision to city values posted 10/29/18 • Uninsured: Data source changed to ACS from BRFSS-CDC 500 (10/29/18); revised strata • Violent crime: Updated to 2017 UCR data (10/29/18) • Revised Technical Documentation (v3.0), downloadable data (v3.0) and codebook (v3.0) (10/29/18) • Minor typos corrected in Technical Documentation (v3.1) (10/31/18)

SECTION 6: References

1. Ratcliffe M. Understanding "Place" in Census Bureau Data Products. Presentation. Geography Division, US Census Bureau. Accessed January 3, 2020. https://www.census.gov/content/dam/Census/data/developers/understandingplace.pdf
2. US Census Bureau. Chapter 8 - County Subdivisions. *Geographic Areas Reference Manual*. 1994:chap 8.
3. US Census Bureau. Chapter 9 - Places. *Geographic Areas Reference Manual*. 1994:chap 9.
4. City Health Dashboard. New Data, New Cities: City Health Dashboard Updates - Introducing 10 New Cities. Updated January 27, 2020. Accessed April 24, 2020, https://www.cityhealthdashboard.com/story/1152
5. 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
6. 500 Cities: Local Data for Better Health. 500 Cities: Local Data for Better Health, Census Tract Boundaries (Last Update Date: December 8, 2020). Updated December 8, 2020. Accessed December 4, 2017. https://chronicdata.cdc.gov/500-Cities-Places/500-Cities-Census-Tract-Boundaries/x7zy-2xmx
7. 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
8. Stucka M. Macon-Bibb County consolidation wins with strong majorities. August 5, 2015. http://www.macon.com/news/politics-government/election/article30109740.html
9. Macon-Bibb County. Macon-Bibb County. Accessed April 10, 2018. http://www.maconbibb.us/
10. US Department of Education. Final guidance on maintaining, collecting, and reporting racial and ethnic data to the U.S. Department of Education. Updated December 3, 2007. Accessed January 17, 2023. https://www.federalregister.gov/documents/2007/10/19/E7-20613/final-guidance-on-maintaining-collecting-and-reporting-racial-and-ethnic-data-to-the-us-department
11. Office of Policy Development and Research - US Department of Housing and Urban Development. HUD USPS ZIP Code Crosswalk Files (ZIP-TRACT, 2nd Quarter 2019). US Department of Housing and Urban Development. Accessed December 19, 2019. https://www.huduser.gov/portal/datasets/usps_crosswalk.html
12. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/
13. Walker K, Herman M. tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames. . https://walker-data.com/tidycensus/.
14. Walker K. tigris: Download and use Census TIGER/Line shapefiles in R. https://github.com/walkerke/tigris
15. 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
16. 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
17. US Census Bureau. Available APIs. US Department of Commerce. Accessed January, 2020. https://www.census.gov/data/developers/data-sets.html
18. US Census Bureau. Technical Documentation: Table & Geography Changes: 2013-2017 ACS 5-year Estimates. US Census Bureau. Updated November 28, 2018. Accessed March 22, 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/5-year.html
19. US Census Bureau. Technical Documentation: Table & Geography Changes: 2012-2016 ACS 5-year Estimates. United States Census Bureau. Updated December 12, 2017. Accessed March 22, 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2016/5-year.html
20. US Census Bureau. Technical Documentation: Table & Geography Changes: 2010-2015 ACS 5-year Estimates. United States Census Bureau. Updated December 8, 2016. Accessed March 22, 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2015/5-year.html
21. US Census Bureau. Technical Documentation: Table & Geography Changes: 2009-2014 ACS 5-year Estimates. United States Census Bureau. Updated November 24, 2015. Accessed March 22, 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2014/5-year.html
22. US Census Bureau. Technical Documentation: Table & Geography Changes: 2008-2013 ACS 5-year Estimates. United States Census Bureau. Updated May 18, 2015. Accessed March 22, 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2013/5-year.html

23. US Census Bureau. Technical Documentation: Table & Geography Changes: 2014-2018 ACS 5-year Estimates. US Census Bureau. Updated December 13, 2019. Accessed March 3, 2020. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2018/5-year.html
24. US Census Bureau. *A compass for understanding and using American Community Survey Data: Appendix 1, Understanding and using ACS single-year and multiyear estimates*. Vol. 2018. https://www.census.gov/content/dam/Census/library/publications/2009/acs/ACSResearch.pdf
25. US Census Bureau. Using ACS Estimates and Margins of Error. https://www.census.gov/content/dam/Census/programs-surveys/acs/guidance/training-presentations/2016_MOE_Slides_01.pdf
26. 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.
27. 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
28. 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
29. 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
30. Foundation PS. Python Language Reference, version 3.6.
31. U.S. Department of Housing and Urban Development Office of Lead Hazard Control and Healthy Homes. *American Healthy Homes Survey II Lead Findings*. 2021. https://www.hud.gov/sites/dfiles/HH/documents/AHHS_II_Lead_Findings_Report_Final_29oct21.pdf
32. United States Environmental Protection Agency. Hazard Standards and Clearance Levels for Lead in Paint, Dust and Soil (TSCA Sections 402 and 403). Accessed 2/14/2024, 2024. https://www.epa.gov/lead/hazard-standards-and-clearance-levels-lead-paint-dust-and-soil-tsca-sections-402-and-403
33. 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
34. US Census Bureau. Technical Documentation: Table & Geography Changes: 2013-2017 ACS 5-year Estimates: Modified Tables: Health Insurance. United States Census Bureau. Updated November 28, 2018. Accessed March 22, 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/5-year.html
35. US Census Bureau. CPS Methodology. United States Census Bureau. 2020. https://www.census.gov/programs-surveys/cps/technical-documentation/methodology.html
36. US Census Bureau. Methodology Updates for the Vintgae 2022 Population Estimates. Updated December 5, 2022. https://www2.census.gov/about/training-workshops/2022/2022-12-05-vintage-2022-transcript.pdf
37. U.S. Bureau of Labor Statistics. Overview: Local Area Unempoyment Statistics. Accessed 9/30/20, https://www.bls.gov/lau/lauov.htm
38. U.S. Bureau of Labor Statistics. Local Area Unemployment Statistics: Estimation Methodology. Accessed May 20, 2021, https://www.bls.gov/lau/laumthd.htm
39. U.S. Bureau of Labor Statistics. Download City LAUS Time Series Data. Accessed 9/30/20, https://download.bls.gov/pub/time.series/la/la.data.65.City
40. U.S. Bureau of Labor Statistics. Series ID Formats: Local Area Unemployment Statistics. Accessed 9/30/20, https://www.bls.gov/help/hlpforma.htm#LA
41. U.S. Bureau of Labor Statistics. Download LAUS Time Series Area Codes. Accessed 9/30/20, https://download.bls.gov/pub/time.series/la/la.area
42. US Census Bureau. Population and Housing Unit Estimates: Terms and Definitions. US Department of Commerce. Updated December 5, 2016. Accessed April 17, 2020. https://www.census.gov/programs-surveys/popest/guidance-geographies/terms-and-definitions.html
43. Collection ECRD. Downloadable Data Files. U.S. Department of Education. https://ocrdata.ed.gov/resources/downloaddatafile
44. Express ED. Data Library. Accessed 9/18/2023. https://eddataexpress.ed.gov/download/data-library?field_year_target_id=All&field_population_value=Chronically+Absent+Students&field_data_topic_target_id=All&field_reporting_level_target_id=28&field_program_target_id=All&field_file_spec_target_id=All&field_data_group_id_target_id=All&combine=
45. Statistics NCfE. ElSi Table Generator. U.S. Department of Education. 2023. https://nces.ed.gov/ccd/elsi/tablegenerator.aspx
46. National Association for Public Health Statistics and Information Systems. NAPHSIS Research Requests. Accessed January 30, 2018. https://www.naphsis.org/research-requests

47. State of New Jersey Department of Health. New Jersey State Health Assessment Data (NJSHAD). State of New Jersey Department of Health,. Updated Febuary 19, 2019. Accessed December 6-16, 2019. https://www-doh.state.nj.us/doh-shad/home/Welcome.html
48. American FactFinder. US Census Bureau. https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml
49. CDC/National Center for Health Statistics. Bridged-Race Population Estimates - Data Files and Documentation: Vintage 2016 Bridged-Race Postcensal Population Estimates for 2014 (pcen_v2016_y14.sas7bdat). Updated June 26, 2017. Accessed June 14, 2017. https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm#vintage2016
50. CDC/National Center for Health Statistics. Bridged-Race Population Estimates - Data Files and Documentation: Vintage 2016 Bridged-Race Postcensal Population Estimates for 2015 (pcen_v2016_y15.sas7bdat). Updated June 26, 2017. Accessed September 3, 2018. https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm#vintage2016
51. Klein R, Schoenbon C. Age adjustment using the 2000 projected US population. Centers for Disease Control and Prevention/National Center for Health Statistics https://www.cdc.gov/nchs/data/statnt/statnt20.pdf
52. 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
53. Lilienfeld D, Stolley P. *Foundations of Epidemiology*. 3rd ed. Oxford University Press; 1994.
54. Office of Public Health Assessment. Confidence intervals in public health. Utah Department of Health, . Accessed January 9, 2018, http://health.utah.gov/opha/IBIShelp/ConfInts.pdf
55. Vohlonen I, Bäckmand H, Korhonen J. Potential years of life lost: The PYLL rate in monitoring the wellbeing of a population. Northen Dimension Partnership in Public Health and Social Well-being. http://www.ndphs.org/documents/2662/Vohlonen%20Ilkka%20PYLL_article_2007.pdf
56. National Cancer Institute SEER Program. SEER Program Coding Guidelines: Breast C500-C509. Accessed January 22, 2018, https://seer.cancer.gov/manuals/2016/AppendixC/Coding_Guidelines_Breast_2016.pdf
57. National Association for Public Health Statistics and Information Systems. Age-adjusted death rate. Accessed January 9, 2018. https://naphsis-web.sharepoint.com/about/Documents/Mortality_AgeAdj_Final_Lois.pdf
58. Nolte E, McKee M. *Does health care save lives? Avoidable mortality revisited*. The Nuffield Trust; 2004.
59. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA: a cancer journal for clinicians*. 2017;67(3):177-193.
60. 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
61. 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
62. Kotelchuck M. An evaluation of the Kessner Adequacy of Prenatal Care Index and a proposed Adequacy of Prenatal Care Utilization Index. *Am J Public Health*. Sep 1994;84(9):1414-20.
63. Centers for Disease Control and Prevention. User guide to the 2015 natality public use file. Accessed December 7, 2018, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2015.pdf
64. Centers for Disease Control and Prevention. User guide to the 2014 natality public use file. Accessed January 22, 2018, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2014.pdf
65. Centers for Disease Control and Prevention. User guide to the 2012 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2012.pdf
66. Centers for Disease Control and Prevention. User guide to the 2013 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2013.pdf
67. Centers for Disease Control and Prevention. User guide to the 2016 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2016.pdf
68. Centers for Disease Control and Prevention. User guide to the 2017 natality public use file. Accessed August 22, 2019, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2017.pdf
69. Centers for Disease Control and Prevention. About the Project - 500 Cities: Local Data for Better Health. Accessed April 10, 2018. https://www.cdc.gov/500cities/about.htm
70. 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

71. 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.
72. 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
73. 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
74. 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
75. PLACES: Local Data for Better Health. Current Release Notes: 2021. https://www.cdc.gov/places/help/data-notes/index.html
76. 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
77. 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
78. 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
79. 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
80. 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
81. PLACES: Local Data for Better Health, Wang Y, Lu H. Personal Email Correspondence: PLACES Release Questions. In: City Health Dashboard Analysts, editor. 2020.
82. Fahle EM, Chavez B, Kalogrides D, Shear BR, Reardon SF, Ho AD. Stanford Education Data Archive Technical Documentation Version 4.1 June 2021. 2021;
83. Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski JA. US small-area life expectancy estimates project: methodology and results summary. 2018;
84. 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
85. 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
86. 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
87. Walk Score®. How Walk Score Works. Walk Score. Accessed December 24, 2019. https://www.walkscore.com/how-it-works/
88. Walk Score®. Walk Score Methodology. Walk Score. Accessed December 24, 2019. https://www.walkscore.com/methodology.shtml
89. Redfin Real Estate. Redfin. https://www.redfin.com/
90. City Health Dashboard. The City Health Dashboard Selects 29 Cities for the Inagural Put Us on the Map Challenge. Updated April 13, 2022. Accessed June 24, 2022,
91. City Health Dashboard. Putting Small Cities on the Map to Advance Health Equity: Checking in on last year's cities and announcing 2023 cities! Updated April 13, 2023. July 6, 2023. https://www.cityhealthdashboard.com/blog-media/put-us-on-the-map-2023-city-announcement
92. Nelson RK, LaDale Winling, et al. Mapping Inequality: Redlining in New Deal America. Digital Scholarship Lab. https://dsl.richmond.edu/panorama/redlining
93. US Census Bureau. Standard Hierarchy of Census Geographic Entities. https://www2.census.gov/geo/pdfs/reference/geodiagram.pdf
94. US Census Bureau. Data from: Block Assignment Files: . 2010, 2020.
95. Missouri Census Data Center. Geocorr 2018: Geographic Correspondence Engine. https://mcdc.missouri.edu/applications/geocorr2018.html

96. US Census Bureau. Table & Geography Changes. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes.html
97. US Census Bureau. Understanding Geographic Identifiers (GEOIDs). https://www.census.gov/programs-surveys/geography/guidance/geo-identifiers.html
98. Hao Y, Ward EM, Jemal A, Pickle LW, Thun MJ. U.S. congressional district cancer death rates. *International journal of health geographics*. 2006;5:28-28. doi:10.1186/1476-072X-5-28
99. 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
100. Rolheiser LA, Cordes J, Subramanian SV. Opioid Prescribing Rates by Congressional Districts, United States, 2016. *American journal of public health*. 2018;108(9):1214-1219. doi:10.2105/AJPH.2018.304532
101. City Health Dashboard. Small and Midsized Cities. Accessed April 24, 2020, www.cityhealthdashboard.com/citytypes