



City Health Dashboard Technical Document Part 1

Version 7.0: Updated August 29, 2019

Contents

SECTION 1: Overview	7
Document Mission	7
Note on Education Data.....	7
Measure Selection Criteria	8
City and Tract Selection Criteria	8
Updates to Technical Documentation.....	8
Multi-year Data: Appropriate Usage for Evaluating Trends Over Time	8
Downloading Dashboard Data.....	10
Citing Dashboard Data and Technical Document	10
Feedback or Errors	10
Measure Overview	11
SECTION 2: Dashboard Analytic Decisions	13
Confidence Intervals (CIs)	13
Dashboard CIs are reported at the 90% level	13
Formulas for CI calculation.....	13
Note on CIs for the Dashboard index values	13
Data Censoring	13
Data Disclaimer.....	13
Data Rounding	13
Federal Information Processing Standards (FIPS) codes	13
Note on Honolulu, HI FIPS code	14
Note on Macon, GA FIPS code	14
Use of County-Level Data on the Dashboard	14
Overall (500 Cities) Estimates	14
Population Percentages.....	14
Race/Ethnicity Categories	15
Validation	15
SECTION 3: Data Sources and Metric Analyses	16
Introduction to this Section	16
500 Cities Project, Centers for Disease Control and Prevention.....	16
General notes	16
Multi-year data.....	16
Weights.....	17
Categorizing race/ethnicity	17
Metric-specific notes.....	17
Binge drinking	17

Dental care	17
Diabetes	17
Frequent physical distress	17
Frequent mental distress.....	17
High blood pressure	17
Obesity	18
Physical inactivity	18
Preventive services	18
Smoking	18
American Community Survey (ACS)	19
General notes	19
Multi-year data	19
Weights.....	19
Categorizing race/ethnicity	19
Confidence intervals	20
Calculating MOEs for aggregate count data and derived proportions	20
Metric-specific notes.....	21
Children in poverty	21
Demographic information	22
Housing cost, excessive.....	23
Income inequality	24
Housing with potential lead risk	25
Lead exposure risk index	25
Neighborhood racial/ethnic segregation	26
Racial/ethnic diversity	27
Unemployment	28
Uninsured.....	30
American Medical Association (AMA) Physician Professional Data	35
Community Multiscale Air Quality model, US Environmental Protection Agency (CMAQ, EPA)	36
General notes	36
Multi-year data.....	36
Weights.....	36
Categorizing race/ethnicity	36
Confidence intervals	36
Metric-specific notes.....	36
Air pollution - particulate matter	36
National Vital Statistics System (NVSS)	38
General notes	38
Multi-year data.....	38

Multiple Cause of Death Data	38
Natality Data.....	38
Weights.....	39
Multiple Cause of Death Data	39
Natality Data.....	39
Categorizing race/ethnicity	39
Multiple Cause of Death Data	39
Natality Data.....	39
Confidence intervals	40
Multiple Cause of Death Data	40
Natality Data.....	40
City/County indicator	40
Multiple Cause of Death Data	40
Natality Data.....	41
Year(s) of data used: Multiplier indicator.....	41
Multiple Cause of Death Data	41
Natality Data.....	41
Metric-specific notes.....	42
Breast cancer deaths	42
Cardiovascular disease deaths	42
Colorectal cancer deaths	43
Low birthweight	44
Opioid overdose deaths	44
Premature deaths (all causes)	45
Prenatal care	45
Teen births	46
United States Small-Area Life Expectancy Project (USALEEP)	47
General notes	47
Multi-year data.....	47
Weights.....	47
Categorizing race/ethnicity	47
Confidence intervals	47
Metric-specific notes.....	47
Life expectancy	47
Uniform Crime Reporting, Federal Bureau of Investigation.....	49
General notes	49
Multi-year data.....	49
Weights.....	49
Categorizing race/ethnicity	49

Confidence intervals	49
Metric-specific notes	49
Violent Crime	49
Food Access Research Atlas, Economic Research Service, United States Department of Agriculture (USDA Food Atlas)	51
General notes	51
Multi-year data	51
Weights	51
Categorizing race/ethnicity	51
Metric-specific notes	51
Limited access to healthy foods	51
ParkServe ®	52
General notes	52
Multi-year data	52
Weights	52
Categorizing race/ethnicity	52
Confidence intervals	52
Metric-specific notes	52
Park access	52
Walk Score ®	53
General notes	53
Multi-year data	53
Weights	53
Categorizing race/ethnicity	53
Confidence intervals	53
Metric-specific notes	53
Walkability	53
SECTION 4: Population Estimates	54
SECTION 5: Acknowledgements	55
SECTION 6: References	56
SECTION 7: Appendices	59
Appendix A: Table of US 2010 Standardized Population	59
Appendix B: Symbol KEY	61
Appendix C: Glossary of Abbreviations	62
Appendix D: Detailed Notes on Selection of City and Tract FIPS Codes	63
1. Notes on Selection of City and Tract FIPS Codes	63
2. R Tutorial	63
Appendix E: Summary of Geographies Reported for Honolulu, HI and Macon, GA (by Metric)	66

Appendix F: Summary of Definition of Hispanic, NHOPI and Other Race for Metrics with Demographic-Specific Values (by Metric)	67
Appendix G: Updates Summary	69

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 36 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 level 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 (with the exception of the high school graduation and third grade reading proficiency metrics) and explains the rationale for analytic decisions. It should be used in conjunction with the Technical Document Part 2: Education Data, which outlines technical attributes of the state-based education data sources used on the Dashboard.

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

Note on Education Data

By applying the measure selection criteria, the Dashboard chose to use state-based education data sources for high school graduation and third-grade reading proficiency over federally reported data sources through the U.S. Department of Education ED Facts. State-based education data sources are updated more regularly and provide data at a more granular level than federally reported data. Thoroughly outlining the attributes of state-based data sources demanded a separate manual. It is available for download on the Dashboard website, titled Technical Document Part 2: Education Data.

Please note that absenteeism, outlined below, is an education metric; however, it comes from a national data source rather than a state-based education data source.

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 CHDB 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 project
 - 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 the 500 most populous cities in the nation as selected by the CDC's 500 Cities Project.¹ The Dashboard selected city and tract FIPS codes as census tract boundary shapefiles released by the 500 Cities Project.² See the "Federal Information Processing Standards (FIPS) codes" section and Appendix D ("Detailed Notes on Selection of City and Tract FIPS Codes using R") below for more detail.

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 G: 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

Marc Gourevitch, MD, MPH	Senior Co-Principal Investigator, City Health Dashboard Muriel and George Singer Professor of Population Health and Chair, Department of Population Health, NYU Langone Health
Lorna Thorpe, PhD	Methods Co-Principal Investigator, City Health Dashboard Professor of Epidemiology, NYU School of Medicine Director of the Division of Epidemiology, NYU School of Medicine Vice Chair of Strategy and Planning, Department of Population Health, NYU Langone Health
Neil Kleiman, PhD	City Policy/Partnerships Co-Principal Investigator, City Health Dashboard Clinical Assistant Professor of Public Service, NYU Wagner Graduate School of Public Service
Shoshanna Levine, DrPH, MPH	Program Director, City Health Dashboard
Becky Ofrane, MPH	Manager, Engagements and Partnerships, City Health Dashboard
Ben Spoer, PhD, MPH	Manager, Data and Analytics, City Health Dashboard
Samantha Breslin, MPA	Project Coordinator, City Health Dashboard
Miriam Gofine, MPH	Senior Data Analyst, City Health Dashboard
Taylor Lampe, MPH	Data Analyst, City Health Dashboard
Allegra Wilson, BA	Research Analyst, City Health Dashboard

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 us at info@cityhealthdashboard.com with any questions or concerns.

Citing Dashboard Data and Technical Document

Dashboard data:

City Health Dashboard. *City Health Dashboard Data*. New York: City Health Dashboard; 2019. Available for download at www.cityhealthdashboard.com.

Technical Document:

Gofine M, Ford S, Wilson A, Lampe T, Kum S, Levine S, Athens J, Spoer B. *City Health Dashboard Technical Document*. New York: City Health Dashboard; 2019. Available at www.cityhealthdashboard.com/technical-documentation

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

On the next page, 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 subgroup or at the tract level if the underlying data allow for such disaggregation.

Domain	Metric (Short Name)	Metric (Long Name)	Data Source	City level	Tract level	Demo subgroups	Multi-year available
Clinical Care	Dental care	Visits to dentist or dental clinic in the previous year among adults aged ≥18 years (%)	500 Cities Project Data, Centers for Disease Control and Prevention	✓	✓	✗	✓
	Prenatal care	Births for which prenatal care began in the first trimester (%)	Nativity Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)	✓	✗	✓	✓
	Preventive services	Adults aged ≥65 years who are up to date on a core set of clinical preventive services (%)	500 Cities Project Data, CDC	✓	✓	✓	✓
	Primary care physicians	Primary care physicians (per 100,000 population)	American Medical Association Physician Masterfile	✓	✗	✗	✗
	Uninsured	Current lack of health insurance among people aged 0–64 years (%)	American Community Survey (ACS)	✓	✓	✓	✓
Health Behaviors	Binge drinking	Binge drinking among adults aged ≥ 18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Physical inactivity	No leisure-time physical activity in past month among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Smoking	Current smoking among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Teen births	Births to mothers aged 15-19 (per 1,000 females in that age group)	Nativity Data, NVSS, NCHS	✓	✗	✓	✓
Health Outcomes	Breast cancer deaths	Deaths due to breast cancer in females (per 100,000 female population)	Multiple Cause of Death Data, NVSS, NCHS	✓	✗	✓	✓
	Colorectal cancer deaths	Deaths due to colorectal cancer (per 100,000 population)	Multiple Cause of Death Data, NVSS, NCHS	✓	✗	✓	✓
	Cardiovascular disease deaths	Deaths due to cardiovascular disease (per 100,000 population)	Multiple Cause of Death Data, NVSS, NCHS	✓	✗	✓	✓
	Diabetes	Diabetes among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Frequent mental distress	Mental health not good for ≥14 days during the past 30 days among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Frequent physical distress	Physical health not good for ≥14 days during the past 30 days among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	High blood pressure	High blood pressure among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Life expectancy	Life expectancy at birth (average)	U.S. Small-area Life Expectancy Estimates Project Data (USALEEP), NCHS	✓	✓	✗	✗
	Low birthweight	Live births with low birthweight <2500 grams (%)	Nativity Data, NVSS, NCHS	✓	✗	✓	✓
	Obesity	Adult obesity among adults aged ≥18 years (%)	500 Cities Project Data, CDC	✓	✓	✗	✓
	Opioid overdose deaths	Deaths due to opioid overdose (per 100,000 population)	Multiple Cause of Death Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)	✓	✗	✗	✓
	Premature deaths (all causes)	Years of potential life lost before age 75 (per 100,000 population)	Multiple Cause of Death Data, NVSS, NCHS	✓	✗	✓	✓

Physical Environment	Air pollution - particulate matter	Average daily concentration of fine particulate matter (PM2.5) per cubic meter (average)	Community Multiscale Air Quality model, US Environmental Protection Agency	✓	✓	✗	✓
	Housing with potential lead risk	Housing stock with potential elevated lead risk (%)	ACS	✓	✓	✗	✓
	Limited access to healthy foods	Population living more than ½ mile from the nearest supermarket, supercenter, or large grocery store (%)	Food Access Research Atlas, Economic Research Service, United States Department of Agriculture	✓	✓	✓	✗
	Lead exposure risk index	Poverty-adjusted risk of housing-based lead exposure (index)	ACS	✓	✓	✗	✓
	Park access	Population living within a 10 minute walk of green space (%)	ParkServe®	✓	✗	✓	✓
	Walkability	Neighborhood amenities accessible by walking as calculated by Walk Score® (index)	Walk Score®	✓	✓	✗	✗
Social and Environmental Factors	Absenteeism	Public school students who miss ≥15 days of school in an academic year (%)	See Technical Document Part 2: Education Data (available for download as PDF)				
	Children in poverty	Children living in households ≤100% of the federal poverty line (%)	ACS	✓	✓	✓	✓
	Housing cost, excessive	Households where ≥30% of household income is spent on housing costs (%)	ACS	✓	✓	✗	✓
	High school graduation	Students who graduate high school within 4 years of entering ninth grade (%)	See Technical Document Part 2: Education Data (available for download as PDF)				
	Income inequality	Households with income at the extremes of the national income distribution (the top 20% or bottom 20%)	ACS	✓	✓	✗	✓
	Neighborhood racial/ethnic segregation	Distribution of the population by race/ethnic group within a census tract relative to the distribution across the city (index)	ACS	✓	✗	✗	✓
	Racial/ethnic diversity	Distribution of the population by race/ethnic group within a city or census tract (index)	ACS	✓	✓	✗	✓
	Third-grade reading proficiency	Third-graders who score "proficient" or above in reading on standardized tests (%)	See Technical Document Part 2: Education Data (available for download as PDF)				
	Unemployment	Population aged ≥16 years that is unemployed but seeking work (%)	ACS	✓	✓	✓	✓
	Violent crime	Violent crime offenses (murder, aggravated assault, robbery, forcible rape) per 100,000 population	Uniform Crime Reporting, Federal Bureau of Investigation	✓	✗	✗	✓

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 wider 90% CIs for a number of reasons. First, the Census Bureau recommends calculation of 90% CIs when using American Community Survey data.³ The Dashboard uses a consistent degree of confidence to ensure clarity for its users.

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

See Appendix B for a summary of where and how censoring was applied.

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.

Federal Information Processing Standards (FIPS) codes

The Federal Information Processing Series (FIPS), formerly Federal Information Processing Standards, are codes for geographic entities maintained and issued by the Census Bureau. When concatenated as State-County, State-Place, or State-County-Tract, FIPS codes function as unique identifiers for geographic entities. The Census Bureau assigns codes to geographic entities such as tracts, which are not covered by FIPS.⁴ Note: Census Bureau codes for tracts are referred to as Tract FIPS within the Dashboard. For more detailed information, refer to Appendix Section D.

Note on Honolulu, HI FIPS code

The Dashboard reports data for the 500 most populous cities in the nation as selected by the CDC's 500 Cities Project.¹ The Dashboard selected city and tract FIPS codes as census tract boundary shapefiles released by the 500 Cities Project.² As per the CDC 500 Cities Project,⁵ the Dashboard uses the FIPS code for the county of Honolulu, Hawaii (15-003) to represent the geographic area associated with the city of Honolulu (Urban Honolulu CDP, FIPS code 15-71550). Dashboard metric values for the city of Honolulu, HI are calculated using values for Honolulu County (FIPS 15-003) where county-level data are available; otherwise, metric values for Honolulu city (FIPS code 15-71550) are presented. See Appendix E for a summary of the geographic coding used for Honolulu, HI, per metric.

Note on Macon, GA FIPS code

As of 2013, American Community Survey data do not publish data for the city of Macon, GA (FIPS code 13-49000). Metrics calculated using American Community Survey data present data for Bibb County (FIPS code 13-021), which shares a consolidated government with Macon, for the city of Macon, GA.^{6,7} See Appendix E for a summary of the geographic coding used for Macon, GA, per metric.

Use of County-Level Data on the Dashboard

County-level data is used where city-level data is unavailable/censored (see NVSS: City/County indicator section below) and as outlined in section "Federal Information Processing Standards (FIPS) codes" above. The Dashboard provides the county_indicator variable in the downloadable datasets to indicate which geography was utilized for all reported values:

- 0 = estimate is calculated from city-specific values
- 1 = estimate is calculated from an average of component counties' values (i.e. city falls under one or more counties)
- 2 = estimate is calculated from its single corresponding county values (i.e. city falls under one specific county)

The Dashboard indicates when county data is displayed on a page (i.e., where county_indicator = 1 or county_indicator = 2) under the "Tips and Cautions for Using the Data" sub-header.

Overall (500 Cities) Estimates

"National" estimates on the Dashboard averages data from the 500 cities represented on the Dashboard by metric. The estimates are not intended to reflect estimates for the United States nationally.

National estimates are calculated after censoring criteria defined below (see Appendix Table B) are applied.

Population Percentages

Text describing population breakdowns by racial/ethnic demographic group (and by sex, for the preventive services metric only) accompanies metric values on the Demographic Detail page. 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)).⁸ Federal guidelines for reporting data by demographics⁸ 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. To ensure these population groups are represented on the Dashboard, the demographic overview for each city includes a granular breakdown of each city’s racial/ethnic composition to enable a more nuanced understanding of each area (scroll down to “More about...” on the All Metrics View page on the Dashboard). See Appendix F for a metric- and data source-specific summary of where Hispanic ethnicity is mutually exclusive of the other racial groups and definitions of NHOPi and other.

Validation

The Dashboard implemented a multi-step data validation process to ensure the accuracy of (1) metric value calculation and (2) data uploaded to the website display.

1. Internal data results validation (*ongoing*)

All analyses* on the Dashboard were initially calculated by a primary analyst from the City Health Dashboard analysis team. All analyses* were then independently replicated by a secondary analyst within the group. Results were directly compared and if applicable, discrepancies were iteratively investigated, addressed, and internally documented until the two separate analyses generated identical values.

2. The Dashboard development site data validation (*May 2019*)

A quality assurance audit by the site’s web developers at Forum One ensured that values calculated by Dashboard staff in SAS correctly appear on the site.

*Please refer to Appendix B for a table listing metric values that were posted as received from the data source.

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.

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.⁹ The Dashboard reports most 500 Cities Project data as received, with the exception of the preventive service utilization values and CI values (see below; these analyses were performed using SAS v9.4).¹⁰

The 500 Cities Project applies 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 available at the county or Metropolitan Statistical level or above. For further details on the methodology used by the 500 Cities Project, see Zhang et al (2014).¹¹ For more information regarding these metrics, please refer to the 500 Cities Project's methodology pages.¹²⁻¹⁴

Multi-year data

The Dashboard reports values from the following 500 Cities: Local Data for Better Health releases:

- 2016 release: The Dashboard labels these data as 2014, 1 Year Modeled Estimates. The 2016 release uses the following data sources: BRFSS data (2014, 2013); Census Bureau 2010 census population data, and ACS 2009-2013, 2010-2014 estimates.¹⁵
- 2017 release: The Dashboard labels these data as 2015, 1 Year Modeled Estimates. The 2017 release uses the following data sources: BRFSS data (2015, 2014); Census Bureau 2010 census population data; and ACS 2011-2015, 2010-2014 estimates.²
- 2018 release: The Dashboard labels these data as 2016, 1 Year Modeled Estimates. The 2018 release uses the following data sources: Behavioral Risk Factor Surveillance System (BRFSS) data (2016, 2015); Census Bureau 2010 census population data; and American Community Survey (ACS) 2012-2016, 2011-2015 estimates.⁹

The following table outlines the years of 500 Cities Project data available on the Dashboard, per metric. Multi-year data for some metrics listed below are unavailable because BRFSS asks some questions every other year. For more information, consult the 500 Cities Project website.⁹

Metric	2013, 1 Year Modeled Estimates	2014, 1 Year Modeled Estimates	2015, 1 Year Modeled Estimates	2016, 1 Year Modeled Estimates
Binge drinking	x	✓	✓	✓
Dental care	x	✓	x	✓
Diabetes	x	✓	✓	✓
Frequent mental distress	x	✓	✓	✓
Frequent physical distress	x	✓	✓	✓
High blood pressure	✓	x	✓	x
Obesity	x	✓	✓	✓
Physical inactivity	x	✓	✓	✓
Preventive services	x	✓	x	✓
Smoking	x	✓	✓	✓

Weights

The Dashboard reports 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, which is reported separately for men and women in the 500 Cities data. Though the Dashboard reports the rates by sex, an average rate for men and women is also calculated, weighting each group evenly.

Categorizing race/ethnicity

Estimates from the 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 500 Cities Project does 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.

Metric-specific notes

The following definitions are taken verbatim from the 500 Cities Project:

Binge drinking

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.¹³

Dental care

Percent of respondents aged ≥ 18 years who report having been to the dentist or dental clinic in the previous year.¹⁴

Diabetes

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.¹²

Frequent physical distress

Respondents aged ≥ 18 years who report 14 or more days during the past 30 days during which their physical health was not good.¹²

Frequent mental distress

Respondents aged ≥ 18 years who report 14 or more days during the past 30 days during which their mental health was not good.¹²

High blood pressure

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.¹²

Obesity

Adult obesity among adults aged ≥18 years.¹³

Physical inactivity

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?”¹³

Preventive services

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.¹⁴

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.¹⁴

Smoking

Respondents aged ≥18 years who report having smoked ≥100 cigarettes in their lifetime and currently smoke every day or some days.¹³

Data tables

Tract and city-level data were downloaded directly from the 500 Cities Project website.^{5,9,15}

Analysis

No analysis by the Dashboard’s staff was required for 500 Cities Project data, with the exception of a) deriving 90% CIs from reported 95% CIs and b) calculating overall preventive service use by older adults aged 65+. Overall preventive services values were calculated as an average of preventive service use by women and preventive service use by men.

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.

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

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

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

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

American Community Survey (ACS)

General notes

ACS is administered by the US Census Bureau; data tables are available for download on American FactFinder.¹⁶ County (050) tables were used for county-level analyses; Place (160) tables were used for city-level analyses; Tract (140) tables were used for tract-level analyses. Values derived from ACS that were used as population denominators in metric analysis vary in year (see Section 4 for more details). All analyses of ACS data were performed using SAS v9.4 unless specified otherwise.¹⁰

All values for Honolulu, HI generated using ACS data represent values associated with the county of Honolulu, HI. All values for Macon, GA generated using ACS data represent values associated with Bibb County, GA. See section “Federal Information Processing Standards (FIPS) codes” (above) and Appendix E for a summary of the geographic coding used for each metric for more detail.

Multi-year data

Data from 2013 (5 Year Estimates), 2014 (5 Year Estimates), (2015 5 Year Estimates), (2016 5 Year Estimates), and (2017 5 Year Estimates) are used on the Dashboard.

Variable labels (e.g., Estimate; SEX AND AGE - Total population), not names (e.g., HC01_VC03), are outlined in this section. 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.¹⁷⁻²¹

Metric-specific sections note where analogous labels change over time as the Census adjusts variable definitions. For parsimony and clarity, only the most recently available labels are listed in this document. Please email info@cityhealthdashboard.com with any questions about specific variable names or labels used in multi-year analyses.

Weights

Weights were not applied to ACS data because data do not require 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), 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), 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, Neighborhood racial/ethnic segregation, and Demographic Information 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 in the data tables available for download on American Fact Finder.¹⁶

Refer to Section 2 “Race/ethnicity categories” (above) for more detail. See Appendix F for a metric- and data source-specific summary of where Hispanic ethnicity is mutually exclusive of the other racial groups and definitions of NHOPI and Other.

Confidence intervals

CIs for all ACS data were calculated according to the formula estimate \pm MOE. See section “Calculating MOEs for Aggregate Count Data and Derived Proportions” below for more on how MOE’s were calculated for summed estimates and derived proportions.

Calculating MOEs for aggregate count data and derived proportions

Approximated MOE’s for aggregate count data and derived proportions in ACS data were calculated as per the US Census Bureau’s publication.²²

Relevant formulas are presented verbatim here for users’ reference:

Calculating MOE’s for Aggregated Count Data²² (p. A-14)

$$\text{MOE}_{\text{aggregated count}} = \pm \sqrt{\sum_c \text{MOE}_c^2}, \text{ “where MOE}_c \text{ is the of the } c^{\text{th}} \text{ component estimate”}$$

Calculating MOE’s for Derived Proportions²² (p. A-14, A-15)

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

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

$\sqrt{\text{MOE}_{\text{numerator}}^2 - (\hat{p}^2 * \text{MOE}_{\text{denominator}}^2)}$ because $\text{MOE}_{\text{numerator}}^2 - (\hat{p}^2 * \text{MOE}_{\text{denominator}}^2)$ resulted in a negative value. In these cases, confidence intervals could not be calculated and associated estimates were censored on the Dashboard. No other censoring of ACS data was performed.

Metric-specific notes

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. The “national” average presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

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; Income in the past 12 months below poverty level: - Under 6 years
- Estimate; Income in the past 12 months below poverty level: - 6 to 11 years
- Estimate; 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; Income in the past 12 months at or above poverty level: - Under 6 years
- Estimate; Income in the past 12 months at or above poverty level: - 6 to 11 years
- Estimate; Income in the past 12 months at or above poverty level: - 12 to 17 years

Demographic information

Data tables

NOTE: The demographic information is displayed on “More about [city name]” on each city’s All Metrics View page on the Dashboard; this demographic information is **not** a metric. The graphic information is provided as supplementary information for Dashboard users. This section outlines how these demographic estimates are calculated.

Longitudinal demographic information data are not presented on the Dashboard. Data table DP05 (2017 5 Year Estimates) was used to provide demographic information about city population values at the city level. Data table S1701 (2017 5 Year Estimates) was used report the percentage of the population with income below <100% of federal poverty level at the city levels.

Analysis

With the exception of “Children (age 0-17)” and “Adults (age 18-64)” (see below), demographic values on the Dashboard do not have analysis applied to them, other than conversion of estimate values to percentages using HC01_VC03 (Total population) as the denominator.

Table DP05

- Labelled "Total population": HC01_VC03 ("Estimate; SEX AND AGE - Total population")
- Labelled "Male": HC01_VC04 ("Estimate; SEX AND AGE - Total population - Male")
- Labelled "Female": HC01_VC05 ("Estimate; SEX AND AGE - Total population - Female")
- Labelled "Older adults (age 65+)": HC01_VC32 ("Estimate; SEX AND AGE - 65 years and over")
- Labelled "White, non-Hispanic": HC01_VC99 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - White alone")
- Labelled "Black, non-Hispanic": HC01_VC100 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Black or African American alone")
- Labelled "Asian, non-Hispanic": HC01_VC102 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Asian alone")
- Labelled "Other, non-Hispanic": HC01_VC104 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Some other race alone")
- Labelled "Two or more races, any ethnicity": HC01_VC50 ("Estimate; RACE - Total population - Two or more races")
- Labelled "American Indian or Alaska Native (one or more races, any ethnicity)": HC01_VC85 ("Estimate; RACE - Race alone or in combination with one or more other races - Total population - American Indian and Alaska Native")
- Labelled "Native Hawaiian and other Pacific Islanders (one or more races, any ethnicity)": HC01_VC87 ("Estimate; RACE - Race alone or in combination with one or more other races - Total population - Native Hawaiian and Other Pacific Islander")
- Labelled "Hispanic": HC01_VC93 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Hispanic or Latino (of any race)")

Table S1701

- Labelled "<100% of federal poverty level": HC03_EST_VC01 ("Percent below poverty level; Estimate; Population for whom poverty status is determined")

Calculated by the Dashboard

- Labelled “Children (age 0-17)” = [HC01_VC03 (“Estimate; SEX AND AGE - Total population”)] – [HC01_VC29 (“Estimate; SEX AND AGE - 18 years and over”)]
- Labelled “Adults (age 18-64)” = [HC01_VC03 (“Estimate; SEX AND AGE - Total population”)] – [HC01_VC32 (“Estimate; SEX AND AGE - 65 years and over”)] – [(calculated total aged 0-17)]

Housing cost, excessive

Data tables

Data table DP04 was used to calculate excessive housing cost at both city and tract levels. The “national” average presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

Analysis

$$\text{Excessive housing cost} = \frac{\begin{array}{l} \text{Selected monthly owner costs (with mortgage): 30.0\%-34.9\% of monthly income} \\ + \\ \text{Selected monthly owner costs (without mortgage): 30.0\%-34.9\% of monthly income} \\ + \\ \text{Gross Rent as a percentage of household income: 30.0\%-34.9\% of monthly income} \\ + \\ \text{Selected monthly owner costs (with mortgage): } \geq 35.0\% \text{ of monthly income} \\ + \\ \text{Selected monthly owner costs (without mortgage): } \geq 35.0\% \text{ of monthly income} \\ + \\ \text{Gross Rent as a percentage of household income: } \geq 35.0\% \text{ of monthly income} \end{array}}{\text{Total occupied housing units}} \times 100\%$$

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

- Estimate; SELECTED MONTHLY OWNER COSTS AS A PERCENTAGE OF HOUSEHOLD INCOME (SMOCAPI) - Housing units with a mortgage (excluding units where SMOCAPI cannot be computed) - 30.0 to 34.9 percent
- Estimate; SELECTED MONTHLY OWNER COSTS AS A PERCENTAGE OF HOUSEHOLD INCOME (SMOCAPI) - Housing unit without a mortgage (excluding units where SMOCAPI 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) - 30.0 to 34.9 percent
- Estimate; SELECTED MONTHLY OWNER COSTS AS A PERCENTAGE OF HOUSEHOLD INCOME (SMOCAPI) - Housing units with a mortgage (excluding units where SMOCAPI cannot be computed) - 35.0 percent or more
- Estimate; SELECTED MONTHLY OWNER COSTS AS A PERCENTAGE OF HOUSEHOLD INCOME (SMOCAPI) - Housing unit without a mortgage (excluding units where SMOCAPI cannot be computed) - 35.0 percent or more
- 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; HOUSING OCCUPANCY - Total housing units - Occupied housing units

NOTE: Variable labels for 2013 and 2014 tables are slightly different from labels for 2015-2017 variables. 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.

Income inequality

Data tables

Data table B19001 was used to calculate income inequality at both city and tract levels. The “national” average presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

Analysis

Income Inequality at the Extremes (ICE) was calculated as per Krieger and colleagues.²³

The formula for ICE is as follows: $ICE(i) = (A(i) - P(i)) / T(i)$, where A(i) is equal to number of persons in 80th income percentile; P(i) is equal to number of persons in 20th percentile and T(i) is equal to total population with known income level in the geographic area. This formula produces values within the range -1 to 1. The Dashboard multiplied ICE values by 100 to provide values that range between -100 and 100.

Cutpoints were selected to represent the 20th and 80th percentiles, as per Krieger et al: “The ICE for income set as the extremes the ACS household income categories that most closely approximated cutpoints for the US 20th and 80th household income percentiles...which for this time period were less than \$25,000 and greater than or equal to \$100,000” (p. 258).²³ As of Dashboard analysis of 2017 ACS data in January 2019, the most recently available cutpoints for the US 20th and 80th household income percentiles were \$24,638 and \$126,855, respectively, as per 2017 US Census Bureau data Table H-1 (2017 data, All Races).²⁴

Variables with the following labels in ACS Table B19001 were summed to calculate A(i). These variables represent estimates of the number of households with incomes equal to or greater than \$125,000, the closest value to \$126,855:

- 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 P(i). These variables represent estimates of the number of households with income of or greater than of \$24,999 or less, the closest value to \$24,638:

- 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

In both City and Tract analyses, the variable with the following label was used to represent T(i):

- Estimate; Total

Notes on analysis

Confidence intervals were not calculated because ICE is an index. See the “Confidence intervals” in Section 2 above for further detail.

Housing with potential lead risk

Data tables

Data table B25034 was used to calculate housing risk data at both city and tract levels. The “national” average presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

Analysis

The lead analysis was performed as per methodology initially developed by the Washington State Department of Health.²⁵ Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.²⁶ The Dashboard adapted Vox Media’s Python code available on Github²⁷ for the present analysis, which was conducted by the Dashboard using SAS v9.4¹⁰ and validated using Python v3.6.²⁸ Users should note that differences in rounding programming between the two softwares resulted in some minor but appreciable differences in housing risk score.

Dashboard’s lead in housing metric reports the risk-adjusted percentage of housing stock at risk for lead and associated confidence intervals. Users can note that this value is the “housing_risk” variable in Washington State Department of Health/Vox Media’s posted Python code. 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 estimates²²; weighting those MOE’s with the same weights used to calculate the numerator; and then calculating an MOE for a derived proportion.²² See section “ACS: Calculating MOEs for aggregate count data and derived proportions” for this equation in full.

Notes on analysis

- a. Washington State Department of Health/Vox Media’s analysis incorporates data on poverty, age of housing, and weights extrapolated from Jacobs 2002²⁹ to generate a decile ranking of lead risk in a given geography; see “Lead exposure risk, overall” metric below. The “Housing with potential lead risk” metric is a Dashboard sub-analysis intended to illustrate the lead-related quality of housing stock for the site’s users. The “housing with potential lead risk metric that is presented on the Dashboard uses the “housing_risk” variable in the code available on Github.²⁷
- b. The Washington State Department of Health’s analysis uses variables from 2014.²⁵ In updating the analysis to represent all housing stock built in 2010 or later for years subsequent to 2014 using table B25034, variables with the following labels were summed: “Estimate; Total: Built 2010-2013” and “Estimate; Total: Built 2014 and later”.

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. The “national” average presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States. The decile ranking ranks risk of lead exposure risk relative to the other cities included on the Dashboard, not all US cities.

Analysis

The lead analysis was performed as per methodology initially developed by the Washington State Department of Health.²⁵ Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.²⁶ The Dashboard adapted Vox Media’s Python code available on Github²⁷ for the present analysis, which was conducted by the Dashboard using SAS v9.4¹⁰ and validated using Python v3.6.²⁸ Users should note that differences in rounding programming between the two softwares resulted in minor but appreciable differences in overall lead exposure risk score and, consequently, the decile ranking of these values.

The analysis uses data on poverty and age of housing and weights extrapolated from Jacobs 2002²⁹ to generate a decile index ranking of lead risk in a given geography; 1 represents “low risk” and 10 represents “high risk”. The decile ranking ranks risk of overall lead exposure risk relative to the other cities included on the Dashboard, not all US cities.

Confidence intervals were not calculated because lead exposure risk is a ranked index. See the “Confidence intervals” section in Section 2 above for more details.

Notes on analysis

The Washington State Department of Health’s analysis uses variables from 2014.²⁵ In updating the analysis to represent all housing stock built in 2010 or later for years subsequent to 2014 using table B25034, variables with the following labels were summed: “Estimate; Total: Built 2010-2013” and “Estimate; Total: Built 2014 and later”.

Neighborhood racial/ethnic segregation

Data tables

Data table DP05 was used to calculate racial/ethnic segregation at the city level. The “national” average presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

Analysis

Segregation was quantified as per Iceland’s formula for H, the entropy index.³⁰

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: $H = \sum_{i=1}^n \frac{t_i(E-E_i)}{ET}$ where t_i refers to the total population of tract i , T is the metropolitan area population, n is the number of tracts, and E_i and E represent tract i ’s diversity (entropy) and metropolitan area diversity respectively.”³⁰ The equation for H above 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.

See the section on “Racial/ethnic diversity” below for more on E and E_i entropy scores. Note that these values are referred to by the Dashboard as city and tract diversity scores, respectively.

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.

Confidence intervals were not calculated because the entropy scores are components of an index. See the "Confidence intervals" above for more details.

Racial/ethnic diversity

Data tables

Data table DP05 was used to calculate racial/ethnic diversity values at the city and tract levels. The "national" average presented on the Dashboard reflects values for the Dashboard's 500 cities, not the entire United States.

Analysis

Diversity was quantified per Iceland's formulas for E and E_i entropy scores.³⁰ In our analysis, E (a metropolitan area's entropy score) and E_i represent city and tract racial/ethnic diversity scores, respectively.

Iceland defines entropy scores for cities and tracts as follows: "A metropolitan area's entropy score is calculated as $E = \sum_{r=1}^R (\pi_r) \ln \left[\frac{1}{\pi_r} \right]$, where π_r refers to a particular racial/ethnic group's proportion of the whole metropolitan area population. All logarithmic calculations use the natural log... A unit within the metropolitan area, such as a census tract, would analogously have its entropy score, or diversity, defined as $E_i = \sum_{r=1}^R (\pi_{ri}) \ln \left[\frac{1}{\pi_{ri}} \right]$ where π_{ri} refers to a particular racial/ethnic group's proportion of the population in tract i."³⁰

As per footnote 5 in Iceland,³⁰ $\ln \left[\frac{1}{\pi_r} \right]$ 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.

The diversity (entropy scores) values that are presented on the Dashboard represent ((the calculated entropy score value)/(maximum possible entropy score, i.e., $\ln(5)$)*100%. $\ln(5)$ is the maximum entropy score because 5 racial/ethnic groups are used in the entropy score calculation.

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.

Unemployment

Data tables

Data table S2301 was used to report unemployment rates, disaggregated by race/ethnicity and sex, at the city level. The "national" average presented on the Dashboard reflects values for the Dashboard's 500 cities, not the entire United States.

S2301 is unavailable at the tract level; data table B23025 was used to report unemployment rates at the tract level. Values in B23025 are not disaggregated by sex and race/ethnicity.

Analysis

Total population; sex; black; white; Hispanic: 2013-2017 City Analyses (Table S2301)

Unemployment rates reported in S2301 are presented as reported using variables labelled as:

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

Other, Asian: 2013-2017 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:

- Unemployment rate; Estimate; RACE AND HISPANIC OR LATINO ORIGIN - American Indian and Alaska Native alone
- Unemployment rate; Estimate; RACE AND HISPANIC OR LATINO ORIGIN - Some other race alone
- Unemployment rate; Estimate; 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 DP05, using the variables with the following labels:

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

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

- Unemployment rate; Estimate; RACE AND HISPANIC OR LATINO ORIGIN - Asian alone
- Unemployment rate; Estimate; 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 DP05, using the variables with the following labels:

- Estimate; RACE - One race – Asian
- Estimate; RACE - One race - Native Hawaiian and Other Pacific Islander

Confidence intervals were calculated using each estimate's associated margin of error variable.

NOTE: Variable labels for 2013 and 2014 S2301 tables are slightly different from labels for variables in 2015, 2016 and 2017 tables. 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.

Total population: 2013-2017 Tract Analyses (Table B23025)

Unemployment rate was derived by dividing the estimate for individuals unemployed in the civilian labor force by the total number in the civilian labor force using the variables associated with the following labels in table B23025:

- Estimate; In labor force: - Civilian labor force: - Unemployed
- Estimate; In labor force: - Civilian labor force:

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 level; this stratum is referred to as “total population”. Table B23025 was used for tract-level analysis.

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, 2016 and 2017; S2702 was used for 2013 and 2014. 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.³¹

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.

The “national” values presented on the Dashboard reflect values for the Dashboard’s 500 cities, not the entire United States.

Analysis

Summary of data table provenance and strata for uninsured (total population and age), 2013-2017						
		Age Strata*				
		Table S2702		Table S2701		
		2013	2014	2015	2016	2017
City only	Children	0-17	0-17	0-17	0-17	0-18
	Adult	18-24	18-24	18-24	18-24	19-25
		25-34	25-34	25-34	25-34	26-34
		35-44	35-44	35-44	35-44	35-44
		45-64 unavailable	45-64 unavailable	45-64	45-64	45-64
		Table S2701				
City and tract	Total population	0-64	0-64	0-64	0-64	0-64

Total population, Age: 2013, 2014 City Analyses (Table S2702)

NOTE: Table S2702 is used instead of S2701 for 2013 and 2014 analyses because of data availability. The percentage of uninsured people (any race/ethnicity, any sex, age 45-64) is not calculated for 2013 and 2014 because the data necessary for this analysis are unavailable.

Percent uninsured are presented as reported in the S2702 data table using the variables labelled as:

- Uninsured Population; Estimate; AGE - Under 18 years
- Uninsured Population; Estimate; AGE - 18 to 64 years - 18 to 24 years
- Uninsured Population; Estimate; AGE - 18 to 64 years - 25 to 34 years
- Uninsured Population; Estimate; AGE - 18 to 64 years - 35 to 44 years

The percentage of uninsured people (any race/ethnicity, any sex, age 0-64) at the city level is calculated using the following formula:

$$\text{Uninsured}_{\text{total population, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{0 \text{ to } 64}}{\text{Total: Estimate}_{0 \text{ to } 64}} \times 100\%$$

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

- Number Uninsured; Estimate; AGE - Under 18 years
- Number Uninsured; Estimate; AGE - 18 to 64 years

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

- Total; Estimate; AGE - Under 18 years
- Total; Estimate; AGE - 18 to 64 years

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

Total population, Age: 2015, 2016 City Analyses (Table S2701)

The percentage of total population uninsured people (any race/ethnicity, any sex, age 0-64) at the city level is calculated using the following formula:

$$\text{Uninsured}_{\text{total population, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{0 \text{ to } 64}}{\text{Total: Estimate}_{0 \text{ to } 64}} \times 100\%$$

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

- Uninsured; Estimate; AGE - Under 18 years
- Uninsured; Estimate; AGE - 18 to 64 years

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

- Total; Estimate; AGE - Under 18 years
- Total; Estimate; AGE - 18 to 64 years

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

- 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

The percentage of uninsured people (any race/ethnicity, any sex, age 45-64) is calculated by the Dashboard using the following formula:

$$\text{Uninsured}_{\text{total population, 45 to 64}} = \frac{\text{Uninsured: Estimate}_{45 \text{ to } 64}}{\text{Total: Estimate}_{45 \text{ to } 64}} \times 100\%$$

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

- Uninsured; Estimate; AGE - 18 to 64 years - 45 to 54 years
- Uninsured; Estimate; AGE - 18 to 64 years - 55 to 64 years

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

- Total; Estimate; AGE - 18 to 64 years - 45 to 54 years
- Total; Estimate; AGE - 18 to 64 years - 55 to 64 years

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

Total population, Age: 2017 City Analyses (Table S2701)

The percentage of total population uninsured people (any race/ethnicity, any sex, age 0-64) at the city level is calculated using the following formula:

$$\text{Uninsured}_{\text{total population, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{0 \text{ to } 64}}{\text{Total: Estimate}_{0 \text{ to } 64}} \times 100\%$$

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

- Uninsured; Estimate; Under 19 years
- Uninsured; Estimate; AGE - 19 to 25 years
- Uninsured; Estimate; AGE - 26 to 34 years
- Uninsured; Estimate; AGE - 35 to 44 years
- Uninsured; Estimate; AGE - 45 to 54 years
- Uninsured; Estimate; AGE - 55 to 64 years

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

- Total; Estimate; Under 19 years
- Total; Estimate; AGE - 19 to 25 years
- Total; Estimate; AGE - 26 to 34 years
- Total; Estimate; AGE - 35 to 44 years
- Total; Estimate; AGE - 45 to 54 years
- Total; Estimate; AGE - 55 to 64 years

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

- Percent Uninsured; Estimate; Under 19 years
- Percent Uninsured; Estimate; AGE - 19 to 25 years
- Percent Uninsured; Estimate; AGE - 26 to 34 years
- Percent Uninsured; Estimate; AGE - 35 to 44 years

The percentage of uninsured people (any race/ethnicity, any sex, age 45-64) at both the city and tract level is calculated by the Dashboard using the following formula:

$$\text{Uninsured}_{\text{total population, 45 to 64}} = \frac{\text{Uninsured: Estimate}_{45 \text{ to } 64}}{\text{Total: Estimate}_{45 \text{ to } 64}} \times 100\%$$

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

- Uninsured; Estimate; AGE - 45 to 54 years
- Uninsured; Estimate; AGE - 55 to 64 years

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

- Total; Estimate; AGE - 45 to 54 years
- Total; Estimate; AGE - 55 to 64 years

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

Sex: 2013-2017 City Analyses (Table B27001)

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

$$\text{Uninsured}_{\text{male, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{\text{male, 0 to 64}}}{\text{Total: Estimate}_{\text{male, 0 to 64}}} \times 100\%$$

$$\text{Uninsured}_{\text{female, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{\text{female, 0 to 64}}}{\text{Total: Estimate}_{\text{female, 0 to 64}}} \times 100\%$$

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

- Estimate; Male: - Under 6 years: - No health insurance coverage
- Estimate; Male: - 6 to 18 years: - No health insurance coverage
- Estimate; Male: - 19 to 25 years: - No health insurance coverage
- Estimate; Male: - 26 to 34 years: - No health insurance coverage
- Estimate; Male: - 35 to 44 years: - No health insurance coverage
- Estimate; Male: - 45 to 54 years: - No health insurance coverage
- Estimate; Male: - 55 to 64 years: - No health insurance coverage

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

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

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

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

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

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

NOTE: Variable labels for 2017 B27001 tables are slightly different from labels for variables in 2013, 2014, 2015 and 2016 tables. 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.

Race/ethnicity: 2013-2017 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, using the following formula:

$$\text{Uninsured}_{\text{racial/ethnic group, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{\text{racial/ethnic group, 0 to 64}}}{\text{Total: Estimate}_{\text{racial/ethnic group, 0 to 64}}} \times 100\%$$

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

- Estimate; Under 19 years: - No health insurance coverage
- Estimate; 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; Under 19 years
- Estimate; 19 to 64 years

NOTE: Variable labels for 2017 C27001B-I tables are slightly different from labels for variables in 2013, 2014, 2015 and 2016 tables. 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.

Total population: 2013, 2014, 2015, 2016 Tract Analyses (Table S2701)

The percentage of uninsured people (any race/ethnicity, any sex, age 0-64) at the tract level is calculated using the following formula:

$$\text{Uninsured}_{\text{total population, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{0 \text{ to } 64}}{\text{Total: Estimate}_{0 \text{ to } 64}} \times 100\%$$

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

- Uninsured; Estimate; AGE - Under 18 years
- Uninsured; Estimate; AGE - 18 to 64 years

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

- Total; Estimate; AGE - Under 18 years
- Total; Estimate; AGE - 18 to 64 years

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

Total population: 2017 Tract Analyses (Table S2701)

The percentage of uninsured people (any race/ethnicity, any sex, age 0-64) at the city level is calculated using the following formula:

$$\text{Uninsured}_{\text{total population, 0 to 64}} = \frac{\text{Uninsured: Estimate}_{0 \text{ to } 64}}{\text{Total: Estimate}_{0 \text{ to } 64}} \times 100\%$$

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

- Uninsured; Estimate; Under 19 years
- Uninsured; Estimate; AGE - 19 to 25 years
- Uninsured; Estimate; AGE - 26 to 34 years
- Uninsured; Estimate; AGE - 35 to 44 years
- Uninsured; Estimate; AGE - 45 to 54 years
- Uninsured; Estimate; AGE - 55 to 64 years

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

- Total; Estimate; Under 19 years
- Total; Estimate; AGE - 19 to 25 years
- Total; Estimate; AGE - 26 to 34 years
- Total; Estimate; AGE - 35 to 44 years
- Total; Estimate; AGE - 45 to 54 years
- Total; Estimate; AGE - 55 to 64 years

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

American Medical Association (AMA) Physician Professional Data

NOTE: As of December 20, 2018, primary care physicians data are temporarily removed from the Dashboard. The Dashboard will post updates as they are available.

Data for this metric are not released in downloadable data versions 4.0 or later. Users should exercise caution in interpreting primary care physician estimates released in previous downloadable data versions.

Please email info@cityhealthdashboard.com with any questions.

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

General notes

The air pollution metric, Air pollution - Average daily concentration of fine particulate matter (PM_{2.5}) per cubic meter, was calculated using Community Multiscale Air Quality model output for the continental United States. CONUS PM 2.5 Daily Average files are used in analyses. Data are downloaded from the US Environmental Protection Agency website.^{32,33}

Multi-year data

Data from 2013, 2014 and 2015 are presented on the Dashboard.

Please note that the estimate variable 2013 and 2014 datasets is labelled as “pm25_daily_average_ug_m3_” but the estimate variable in the 2015 dataset is labelled as “Prediction”. Dashboard staff confirmed the equivalency of these variables in private correspondence with EPA staff in May 2019; please contact info@cityhealthdashboard.com with any questions.

Weights

The data available for download from the US Environmental Protection Agency’s Community Multiscale Air Quality model provide daily tract-level estimates for all tracts in the contiguous United States for the given year. Population weights were calculated for each tract using the American Community Survey DP05 5 Year estimates (variable labelled “Estimate; SEX AND AGE - Total population”) for the associated year (e.g., 2015 DP05 is used for 2015 CMAQ data; 2014 DP05 is used for 2014 CMAQ data, etc.)

$$w_i = \frac{p_i}{\sum_{i=1}^n p_i}$$

Where:

w_i = tract-specific weight

p_i = tract total population from ACS Table DP05

n = total tracts within a city

Categorizing race/ethnicity

Not applicable.

Confidence intervals

Not applicable.

Metric-specific notes

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

An interactive map of locations of active air quality monitors for PM_{2.5} is available online, through the EPA.³⁴

Air pollution - particulate matter

Data tables

Tract level data are available for download from the US Environmental Protection Agency. Daily estimates of PM 2.5 concentration are reported for census tracts (vintage 2010) within the continental U.S.

For information about weighting of these values, see the Weights section above.

Analysis

An annual average for each census tract reported in the CMAQ data was calculated.

To calculate city-level estimates, the population weights were applied to the census tract annual averages and the weighted tract values were summed to the city level using the following formula:

$$PM2.5_j = \sum_{i=1}^n w_i \times PM2.5_i$$

National Vital Statistics System (NVSS)

General notes

Unless otherwise specified, 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 SAS v9.4.¹⁰

American Community Survey data were used to calculate age-group specific total-population, female, and male values for NVSS data (please refer to section 4 for more information on population denominators).

Restricted use NVSS data are available through the National Association for Public Health Statistics and Information Systems (NAPHSIS). Data request forms are available online.³⁵

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.

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. For example, 2012 analyses use data from 2010, 2011 and 2012; 2013 analyses use data from 2011, 2012, and 2013; 2014 analyses use data from 2012, 2013, and 2014, etc.

Multiple Cause of Death Data

MCDD data for 2012, 2013, 2014, 2015, 2016, and 2017 for the following metrics are presented on the Dashboard: opioid overdose deaths, breast cancer deaths, cardiovascular disease deaths, colorectal cancer deaths and premature deaths (all causes).

Nativity Data

ND data for 2012, 2013, 2014, 2015, 2016, and 2017 for the following metrics are presented on the Dashboard: prenatal care, low birthweight, and teen births, respectively. See below for more detail.

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 & Schoenborn³⁶ utilizing the data table “QT-P1, 2010 Decennial Census” downloaded in December 2016 from American Fact Finder.¹⁶

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.³⁷ Weights for age-adjusting premature deaths were calculated using the data table “QT-P1, 2010 Decennial Census”.¹⁶ 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 $\text{Weight}(\text{age group } i) = 75 - \text{Mid-point age group } i$. Weights and years of life lost are presented in Appendix A.

Nativity Data

No weights are applied to ND data.

Categorizing race/ethnicity

Multiple Cause of Death Data

“Hispanic origin/race recode” and “race recode 5” variables were used to categorize race/ethnicity for all mortality metrics:

- 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

Nativity 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:

- 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 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

Confidence intervals

Multiple Cause of Death Data

CIs for breast cancer, colorectal cancer, cardiovascular disease, and opioid overdose deaths metrics were calculated according to following formula outlined by Lilienfeld and Stolley³⁸ in a document published by the Utah Department of Health³⁹:

$$\begin{aligned} \text{LCL90} &= \text{estimate} - (1.645 \times \text{SE}(\text{estimate})) \\ \text{UCL90} &= \text{estimate} + (1.645 \times \text{SE}(\text{estimate})) \end{aligned}$$

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

CIs for premature deaths (all causes) were calculated according to the following formula outlined by Vohlonen, Bäckmand, & Korhonen:⁴⁰

$$\begin{aligned} \text{LCL90} &= \text{estimate} - (1.645 \times \text{SE}(\text{est})) \\ \text{UCL90} &= \text{estimate} + (1.645 \times \text{SE}(\text{est})) \end{aligned}$$

$$\text{SE}(\text{est}) = \text{SQRT} [\text{sum}((((\text{age-group specific crude mortality rate}^2)/(\text{age-group specific total number of deaths})) * \text{age-group specific premature deaths weight--years of life lost}) * \text{US 2010 standardized population YPLL age-group specific weight})]$$

Nativity Data

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

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

CIs for teen births metric were calculated as follows:

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

City/County indicator

Multiple Cause of Death Data

Total population, sex-specific estimates: The Dashboard calculates total population and sex estimates based on city of residence. Please note that total population and sex estimates for Honolulu, HI and Macon, GA use data from county of residence; see section "Federal Information Processing Standards (FIPS) codes" above for more or email info@cityhealthdashboard.com.

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.

Nativity Data

Total population, sex-specific estimates: The Dashboard calculates total population and sex estimates based on city of residence. Please note that total population and sex estimates for Honolulu, HI and Macon, GA use county of residence; see section “Federal Information Processing Standards (FIPS) codes” above for more detail or email info@cityhealthdashboard.com.

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.

Year(s) of data used: Multiplier indicator

Multiple Cause of Death Data

All MCDD calculated values are based on data from three-year periods. For example, the estimate for 2015 on the Dashboard is calculated from 2013, 2014 and 2015 data. 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 created the following indicator to provide more information on how many years of data were utilized for all MCDD reported values:

- 1 = estimate is calculated from one year of data
- 2 = estimate is calculated from two years of data
- 3 = estimate is calculated from three years of data

This indicator also serves as a multiplier for population. Population denominators from the midpoint of each wave of data are used as the multiplier (please refer to section 4 for more information on population denominators). The multiplier is included as multiplier_indicator in the Dashboard’s downloadable data.

Nativity Data

All ND calculated values are based on data from three-year periods (with the exception of 2015 prenatal care estimates, which are calculated using data from 2014 and 2015 only; see section “Prenatal care” below for more detail). For example, the estimate for 2015 on the Dashboard is calculated from 2013, 2014 and 2015. 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 created the following indicator, released with the Dashboard’s downloadable datasets as multiplier_indicator, to provide more information on how many years of data were utilized for all ND reported values:

- 1 = estimate is calculated from one year of data
- 2 = estimate is calculated from two years of data
- 3 = estimate is calculated from three years of data

This indicator also serves as a multiplier for population. Population denominators from the midpoint of each wave of data are used as the multiplier (please refer to section 4 for more information on population denominators). The multiplier is included as multiplier_indicator in the Dashboard’s downloadable data.

Metric-specific notes

Breast cancer deaths

Data tables

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Multiple Cause of Death data files were combined across years to calculate breast cancer deaths among females at the city level. Please refer to section 4 for more information on population denominators.

Analysis

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.⁴¹

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.⁴²

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

Cardiovascular disease deaths

Data tables

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Multiple Cause of Death data files were combined to calculate cardiovascular disease deaths at the city level. Please refer to section 4 for more information on population denominators.

Analysis

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.⁴¹

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 & McKee⁴³ 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.

Colorectal cancer deaths

Data tables

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Multiple Cause of Death data files were combined to calculate colorectal cancer deaths at the city level. Please refer to section 4 for more information on population denominators.

Analysis

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.⁴¹

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 al⁴⁴ 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.

Low birthweight

Data tables

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Nativity data files were combined to calculate low birthweight at the city level.

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

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Multiple Cause of Death data files were combined to calculate opioid overdose deaths at the city level. Please refer to section 4 for more information on population denominators.

Analysis

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.⁴¹

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.^{45,46}

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

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

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Multiple Cause of Death data files were combined to calculate premature deaths (all causes) at the city level. Please refer to section 4 for more information on population denominators.

Analysis

Premature deaths (all causes) =

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

Notes on Analysis

Premature deaths (all causes) rates are calculated as per Dranger and Remington's approach.³⁷ 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.

Prenatal care

Data tables

Year	Data Files Used
2015	2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Nativity data files were combined to calculate prenatal care at the city level.

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.

Prenatal care estimates represent a slight modification of one component of the Kotelchuck Index.⁴⁷ 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.⁴⁸⁻⁵³

Teen births

Data tables

Year	Data Files Used
2012	2010, 2011, 2012
2013	2011, 2012, 2013
2014	2012, 2013, 2014
2015	2013, 2014, 2015
2016	2014, 2015, 2016
2017	2015, 2016, 2017

Natality data files were combined to calculate teen births at the city level. Please refer to section 4 for more information on population denominators.

Analysis

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

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.⁵⁴

Multi-year data

Multi-year data for this metric are unavailable.

Weights

Tract values are presented on the Dashboard as downloaded (see below). The methodology used to calculate tract-level data is published.⁵⁴ City values represent a population weighted average of tract values associated with respective cities. 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:

$$\begin{aligned} \text{LCL90} &= \text{estimate} - (1.645 \times \text{SE}(\text{estimate})) \\ \text{UCL90} &= \text{estimate} + (1.645 \times \text{SE}(\text{estimate})) \end{aligned}$$

Where:

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

City-level CI's were calculated using the formulas for 90% confidence intervals listed above. The Satterthwaite approximation was used to calculate a pooled standard error of the tracts associated with each city.⁵⁵ The formula for the Satterthwaite approximation is:

$$se_s = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where s is the tract standard error provided in USALEEP downloadable data⁵⁶ (variable `se_e_0__`) and n is the six-year population of the tract (see “Analysis” section below).

Metric-specific notes

Life expectancy

Data tables

The Dashboard reports tract-level data as received from USALEEP; tract-level data and documentation files are available for free download.⁵⁶

City-level life expectancy estimates are calculated by Dashboard analytic staff using SAS¹⁰ as outlined below.

Analysis

This section outlines the calculation of city-level weighted averages of tract life expectancy estimates. As per email correspondence with USALEEP analytic staff at the National Center for Health Statistics, weights are calculated as followed:

$$\text{Tract weight} = \frac{\text{Six-year tract population}}{\text{Six-year city population}}$$

Where:

$$\text{six year tract population} = (\text{tract population}_{2010 \text{ Census}}) + (5 \times \text{tract population}_{2015 \text{ American Community Survey, 5 Year Estimates}})$$

$$\text{six-year city population} = \sum \text{six year tract populations in city}$$

Variable HD01_S001 in table “2010 SF1 100% Data, DP-1 Profile of General Population and Housing Characteristics: 2010” is used to define tract population in 2010 Census data. Variable HC01_VC03 in table DP05 is used to define tract population in 2015 ACS (5 Year Estimates) data. Both tables are available for download on American Fact Finder.¹⁶

Tracts with missing life expectancy, Decennial Census, or ACS values were excluded from analysis.

City life expectancy estimates that are calculated where data for 20% or more tracts are missing are indicated with a note in the Tips and Cautions for Using the Data box on the Dashboard.

Uniform Crime Reporting, Federal Bureau of Investigation

General notes

Uniform Crime Reporting data are published by the Federal Bureau of Investigation. All analyses of UCR data were performed using SAS v9.4.¹⁰

Multi-year data

Data from 2016 and 2017 are presented on the Dashboard.

Weights

No weights were applied to UCR data.

Categorizing race/ethnicity

Not applicable.

Confidence intervals

CIs for violent crime were calculated as follows:

$$\text{LCL } 90 = \frac{100,000}{\text{denominator}} \times (\text{numerator} - (1.645 \times \sqrt{\text{numerator}}))$$

$$\text{UCL } 90 = \frac{100,000}{\text{denominator}} \times (\text{numerator} + (1.645 \times \sqrt{\text{numerator}}))$$

Metric-specific notes

City-level specific values are not available for all cities within the Uniform Crime Reporting dataset. In the event that a city-level specific value is not available, the Dashboard reports the most approximate geography reported by the Uniform Crime Reporting dataset. The following table lists corresponding geographies:

<i>Dashboard City Name</i>	<i>UCR Geography Reported</i>
Athens, Georgia	Athens-Clarke County, Georgia
Camden, New Jersey	Camden County Police Department, New Jersey
Charlotte, North Carolina	Charlotte-Mecklenburg, North Carolina
Las Vegas, Nevada	Las Vegas Metropolitan Police Department, Nevada
Louisville, Kentucky	Louisville Metro, Kentucky
Nashville, Tennessee	Nashville Metropolitan, Tennessee
Savannah, Georgia	Savannah-Chatham Metropolitan, Georgia

Violent Crime

Data tables

The 2017 data table, "Table 8: Offenses Known to Law Enforcement by State by City, 2017"⁵⁷ was used to calculate total number of violent crimes for 2017. The 2016 American Community Survey's data table DP05 was used to calculate total-population values. Please refer to section 4 for more information regarding ACS population denominators.

The 2016 data table, “Table 6: Offenses Known to Law Enforcement by State by City ⁵⁸” from Uniform Crime Reporting was used to calculate total number of violent crimes for 2016. The 2016 American Community Survey’s data table DP05 was used to calculate total-population values for both 2016 and 2017 analyses. Please refer to section 4 for more information regarding ACS population denominators.

Analysis

The measure was calculated as follows:

$$\frac{\text{Total number of reported violent crimes}}{\text{Total population}} \times 100,000$$

The following types of violent crimes were summed to calculate the total number of violent crimes: murder and non-negligent manslaughter, rape (revised definition), rape (legacy definition), robbery, and aggravated assault.

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

General notes

The limited access to healthy food metric represents the percent of the population that lives $\geq \frac{1}{2}$ mile from a supermarket, large grocery store, or supercenter. . Data on store locations were obtained by USDA ERS from the 2015 STARS directory of stores authorized to accept SNAP benefits and the 2015 Trade Dimensions TDLinx directory of stores.⁵⁹ Population data total and by race/ethnicity were obtained at the block level from US Census. Data are available for download online.⁶⁰

Multi-year data

Multi-year data for this metric are unavailable.

Weights

Weights were not applied to USDA ERS data as these data do not require weighting.

Categorizing race/ethnicity

City level estimates of limited access to healthy food were also calculated by race/ethnicity. Data reported by USDA ERS included estimates by the following categories: white, black, Asian, Native Hawaiian/Pacific Islander, American Indian/Alaska Native, other or two or more Races, and Hispanic. Estimates for Asian and Native Hawaiian/Pacific Islander were aggregated to match our Asian/PI designation. Similarly, estimates for other or two or more Races and American Indian/Alaska Native were also aggregated.

Metric-specific notes

Limited access to healthy foods

The limited access to healthy food metric represents the percent of the population that lives $\geq \frac{1}{2}$ mile from a supermarket, large grocery store, or supercenter. Stores were defined as a healthy food outlet if they reported at least \$2 million in annual sales, were certified to accept SNAP benefits, and sold food in all the following categories: fresh produce, fresh meat and poultry, dairy, dry and packaged goods, and frozen foods.

Data tables

Census tract-level data were downloaded directly for this measure.

Analysis

The downloaded data set from USDA ERS reports numerators (population living $\geq \frac{1}{2}$ mile from a supermarket, large grocery store, or supercenter) and denominators for total population and by race/ethnicity. The measure was calculated as follows:

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

City level estimates were calculated by summing tract numerators and denominators to the city level for total population and by race/ethnicity.

ParkServe ®

General notes

Park access represents the percent of the population living within a 10 minute walk of green space. GIS data on parks was obtained by the ParkServe® team through outreach to cities, towns and communities with a request for their 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 ®:

- 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 ParkServe ® data as these data do not require weighting.

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, Native Hawaiian/Pacific Islander, American Indian/Alaska Native, other or two or more races, and Hispanic. Estimates for Asian and Native Hawaiian/Pacific Islander were aggregated to match our Asian/PI designation. Similarly, estimates for other or two or more races and American Indian/Alaska Native were also aggregated.

Confidence intervals

CIs are not presented for ParkServe ® data.

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

ParkServe ® GIS data is publicly available. The Dashboard reports ParkServe ® calculated data as received.

Analysis

The data from ParkServe ® reports numerators (population living ≤ 10-minute walk of a park/green space) and denominators for total population and by race/ethnicity. The measure was calculated as follows:

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

Walk Score®

General notes

Walk Score® present a web page, organized by state, that contain a table with walk score, transit score, and bike score as well as zip code and population for select cities in the state. R version 3.4.1 was used to extract the walk score from the tables on state web pages and export the extracted data as .csv files.

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

Analysis

The Dashboard reports Walk Score® data as received. For more detailed information, please refer to the "How Walk Score Works" webpage, available at <https://www.redfin.com/how-walk-score-works>.

SECTION 4: Population Estimates

Population denominators are **not** a metric and are not displayed on the Dashboard. The values are used as denominators in Dashboard analyses that result in metric values presented on the website. This section outlines the analytics that generate the population denominator analyses and documents the sources of the values.

Population denominators are used in two instances:

1. Where the total population of a geographic area is required for an analysis.

The variable labelled “Estimate; SEX AND AGE - Total population” value in ACS table DP05 was used to estimate the total population within a city. (See metric-specific notes for the year of DP05 that is selected for the analysis.) These values were not altered and are used as presented in downloads from American FactFinder.¹⁶ These values from DP05 were also used as the denominator for the “Demographic information” percentages (see section “Demographic information” above for more detail).

As noted elsewhere in this document, values for Honolulu, HI generated using ACS data represent values associated with the county of Honolulu, HI. All values for Macon, GA generated using ACS data represent values associated with Bibb County, GA. All values for Macon, GA generated using NVSS data from 2015 and later represent values associated with Bibb County, GA. See section “Federal Information Processing Standards (FIPS) codes” (above) and Appendix E (below) for a summary of the geographic coding used for each metric for more detail.

2. Where sex- and/or age-specific population estimates are required for adjustment in data from National Vital Statistics System (see section “National Vital Statistics System (NVSS)” above):

- a. For city-level analyses of total population and sex subgroups, population estimate denominators are generated from American Community Survey table B01001 (5 Year Estimates) for Place.¹⁶

Data from the midpoint year of a three-year wave is used in each analysis. For example, B01001 (2015, Year Estimate) is used for 2014-2016 NVSS estimates; B01001 (2014, Year Estimate) is used for 2013-2015 NVSS estimates, etc.

- b. For county-level analyses of racial/ethnic subgroups, population estimate denominators are generated from the National Center for Health Statistics.

Data from the midpoint year of a three-year wave is used in each analysis. For example, 2013-2015 estimates use vintage 2016 Bridged-Race Postcensal Population Estimates data file for 2014.⁶¹ 2014-2016 estimates use vintage 2016 Bridged-Race Postcensal Population Estimates data file for 2015.⁶²

SECTION 5: Acknowledgements

The Dashboard acknowledges Sarah Conderino, Miriam Gofine, Ben Spoer, Taylor Lampe, Allegra Wilson, Dr. Jessica Athens, Shauna Ford and Dr. Susan Kum 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. 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.

SECTION 6: References

1. Centers for Disease Control and Prevention. About the Project - 500 Cities: Local Data for Better Health. 2018; <https://www.cdc.gov/500cities/about.htm>. Accessed April 10, 2018.
2. 500 Cities: Local Data for Better Health. 500 Cities: Local Data for Better Health, Census Tract Boundaries (Last Update Date: December 4, 2017). 2017; <https://chronicdata.cdc.gov/500-Cities/500-Cities-Census-Tract-Boundaries/x7zy-2xmx>. Accessed December 4, 2017.
3. Berkley J. Using American Community Survey estimates and margins of error. [Presentation]. 2017; <https://www.census.gov/mso/www/training/pdf/acs-estimates-moe4-19-17.pdf>. Accessed January 1, 2018.
4. Geographic Terms and Concepts – Codes for Geographic Entities U.S. Census Bureau; https://www.census.gov/geo/reference/gtc/gtc_codes.html#fips. Accessed January 24, 2018.
5. 500 Cities: Local Data for Better Health. Local Data for Better Health, 2017 Release. 2017; <https://chronicdata.cdc.gov/500-Cities/500-Cities-Local-Data-for-Better-Health-2017-relea/6vp6-wxug>. Accessed September 27, 2017.
6. Stucka M. Macon-Bibb County consolidation wins with strong majorities. August 5, 2015, 2012.
7. Macon-Bibb County. Macon-Bibb County. 2018; <http://www.maconbibb.us/>. Accessed April 10, 2018.
8. US Department of Education. Policy questions on the Department of Education's 2007 guidance on collecting, maintaining and reporting data by race or ethnicity. 2008; <https://www2.ed.gov/policy/rschstat/guid/raceethnicity/questions.html>. Accessed January 9, 2018.
9. 500 Cities: Local Data for Better Health. Local Data for Better Health, 2018 Release. 2018; <https://chronicdata.cdc.gov/500-Cities/500-Cities-Local-Data-for-Better-Health-2018-relea/6vp6-wxug>. Accessed December 4, 2018.
10. SAS Institute Inc. SAS System. In. Vol 9.4. Cary, NC 2015.
11. 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.
12. 500 Cities: Local Data for Better Health. Health outcomes. 2016; <https://www.cdc.gov/500cities/definitions/health-outcomes.htm>. Accessed January 9, 2018.
13. 500 Cities: Local Data for Better Health. Unhealthy behaviors. 2016; <https://www.cdc.gov/500cities/definitions/unhealthy-behaviors.htm>. Accessed January 9, 2018.
14. 500 Cities: Local Data for Better Health. Prevention. 2016; <https://www.cdc.gov/500cities/definitions/prevention.htm>. Accessed January 9, 2018.
15. 500 Cities: Local Data for Better Health. Local Data for Better Health, 2016 Release. 2018; <https://chronicdata.cdc.gov/500-Cities/500-Cities-Local-Data-for-Better-Health-2016-relea/9z78-nfsp>. Accessed May 6, 2019.
16. American FactFinder. US Census Bureau. 2017; <https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml>.
17. Technical Documentation: Table & Geography Changes: 2013-2017 ACS 5-year Estimates. 2018; <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/5-year.html>. Accessed March 22, 2019.
18. Technical Documentation: Table & Geography Changes: 2012-2016 ACS 5-year Estimates. 2017; <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2016/5-year.html>. Accessed March 22, 2019.
19. Technical Documentation: Table & Geography Changes: 2010-2015 ACS 5-year Estimates. 2016; <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2015/5-year.html>. Accessed March 22, 2019.
20. Technical Documentation: Table & Geography Changes: 2009-2014 ACS 5-year Estimates. 2015; <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2014/5-year.html>. Accessed March 22, 2019.
21. Technical Documentation: Table & Geography Changes: 2008-2013 ACS 5-year Estimates. 2015; <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2013/5-year.html>. Accessed March 22, 2019.
22. US Census Bureau. *A compass for understanding and using American Community Survey Data: Appendix 1, Understanding and using ACS single-year and multiyear estimates*. US Census Bureau,.
23. 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.
24. US Census Bureau. Table H-1. Income Limits for Each Fifth and Top 5 Percent (All Races): Table H-1. Income Limits for Each Fifth and Top 5 Percent of All Households: 1967 to 2017. (Households as of March of the following year. Income in current and 2017 CPI-U-RS adjusted dollars (28)). 2019;

- <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html>. Accessed January 28, 2019.
25. Lofy K, Davies H, Fotinos C, et al. A targeted approach to blood lead screening in children, Washington State: 2015 expert panel recommendations. 2015; <https://assets.documentcloud.org/documents/2644455/Expert-Panel-Childhood-Lead-Screening-Guidelines.pdf>. Accessed January 9, 2018.
 26. Frostenson S, Kliff S. Lead Exposure Risk data (readme.md). 2016; <https://github.com/voxmedia/data-projects/tree/master/vox-lead-exposure-risk>. Accessed January 9, 2017.
 27. Frostenson S, Kliff S. Added data for vox-lead-exposure-risk-map: Lead Exposure Risk data. 2016; <https://github.com/voxmedia/data-projects/commit/f8ed1caa78af3232f18c0fcee23626a822fc2169>. Accessed January 9, 2017.
 28. Foundation PS. Python Language Reference, version 3.6. In.
 29. Jacobs DE, Clickner RP, Zhou JY, et al. The prevalence of lead-based paint hazards in US housing. *Environmental health perspectives*. 2002;110(10):A599.
 30. Iceland J. The multigroup entropy index (also known as Theil's H or the Information Theory Index). 2004; https://www.census.gov/hhes/www/housing/resseg/multigroup_entropy.pdf. Accessed January 5, 2018.
 31. Technical Documentation: Table & Geography Changes: 2013-2017 ACS 5-year Estimates: Modified Tables: Health Insurance. 2018; <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/5-year.html>. Accessed March 22, 2019.
 32. United States Environmental Protection Agency. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files: Downscaling Output files (files compressed in gzip format): 2013 (CONUS): PM 2.5 Daily Average. 2019; <https://www.epa.gov/hesc/rsig-related-downloadable-data-files#output>. Accessed May 1, 2018.
 33. United States Environmental Protection Agency. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files: Downscaling Output files (files compressed in gzip format): 2014 (CONUS): PM 2.5 Daily Average. 2019; <https://www.epa.gov/hesc/rsig-related-downloadable-data-files#output>. Accessed January 28, 2019.
 34. Environmental Protection Agency. Interactive Map of Air Quality Monitors. 2017; <https://www.epa.gov/outdoor-air-quality-data/interactive-map-air-quality-monitors#kmz>. Accessed May 7, 2018.
 35. National Association for Public Health Statistics and Information Systems. NAPHSIS Research Requests. 2018; <https://www.naphsis.org/research-requests>. Accessed January 30, 2018.
 36. Klein R, Schoenbon C. Age adjustment using the 2000 projected US population. 2007; <https://www.cdc.gov/nchs/data/statnt/statnt20.pdf>.
 37. Dranger E, Remington P. YPLL: A summary measure of premature mortality used in measuring the health of communities 2004; <https://uwphi.pophealth.wisc.edu/publications/issue-briefs/issueBriefv05n07.pdf>.
 38. Lilienfeld D, Stolley P. *Foundations of Epidemiology*. 3rd ed. New York: Oxford University Press; 1994.
 39. Office of Public Health Assessment. Confidence intervals in public health. <http://health.utah.gov/opha/IBIShelp/ConfIntns.pdf>. Accessed January 9, 2018.
 40. Vohlonen I, Bäckmand H, Korhonen J. Potential years of life lost: The PYLL rate in monitoring the wellbeing of a population. 2007; http://www.ndphs.org/documents/2662/Vohlonen%20Ilkka%20PYLL_article_2007.pdf.
 41. National Association for Public Health Statistics and Information Systems. Age-adjusted death rate. https://naphsis-web.sharepoint.com/about/Documents/Mortality_AgeAdj_Final_Lois.pdf. Accessed January 9, 2018.
 42. National Cancer Institute SEER Program. SEER Program Coding Guidelines: Breast C500-C509. 2016; https://seer.cancer.gov/manuals/2016/AppendixC/Coding_Guidelines_Breast_2016.pdf. Accessed January 22, 2018.
 43. Nolte E, McKee M. *Does health care save lives? Avoidable mortality revisited*. The Nuffield Trust; 2004.
 44. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA: a cancer journal for clinicians*. 2017;67(3):177-193.
 45. 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. 2013; https://www.cdc.gov/drugoverdose/pdf/pdo_guide_to_icd-9-cm_and_icd-10_codes-a.pdf. Accessed January 22, 2018.
 46. Henry J Kaiser Family Foundation. Opioid Overdose Death Rates and All Drug Overdose Death Rates per 100,000 Population (Age-Adjusted). 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>. Accessed May 5, 2018.
 47. 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*. 1994;84(9):1414-1420.

48. Centers for Disease Control and Prevention. User guide to the 2015 natality public use file. 2015; ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2015.pdf. Accessed December 7, 2018.
49. Centers for Disease Control and Prevention. User guide to the 2014 natality public use file. 2014; ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2014.pdf. Accessed January 22, 2018.
50. Centers for Disease Control and Prevention. User guide to the 2012 natality public use file. 2012; ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2012.pdf. Accessed August 22, 2019.
51. Centers for Disease Control and Prevention. User guide to the 2013 natality public use file. 2013; ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2013.pdf. Accessed August 22, 2019.
52. Centers for Disease Control and Prevention. User guide to the 2016 natality public use file. 2016; ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2016.pdf. Accessed August 22, 2019.
53. Centers for Disease Control and Prevention. User guide to the 2017 natality public use file. 2017; ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2017.pdf. Accessed August 22, 2019.
54. Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski JA. US small-area life expectancy estimates project: methodology and results summary. 2018.
55. Dyer L. Pooled Standard Error versus the Satterthwaite Approximation. 2018; <https://wolfweb.unr.edu/~ldyer/classes/396/PSE.pdf>. Accessed October 18, 2018.
56. National Center for Health Statistics. U.S. Small-Area Life Expectancy Estimates Project - USALEEP: Life Expectancy Estimates File for United States, 2010-2015. 2018; <https://www.cdc.gov/nchs/nvss/usaleep/usaleep.html>. Accessed September 26, 2018.
57. Federal Bureau of Investigations. Offenses known to law enforcement by state by city, 2017. 2017; <https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/tables/table-8/table-8.xls/view>. Accessed October 26, 2018.
58. Federal Bureau of Investigations. Offenses known to law enforcement by state by city, 2016. 2016; <https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016/tables/table-6/table-6.xls/view>.
59. United States Department of Agriculture Economic Research Service. Food Access Research Atlas Documentation. 2017; <https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/>. Accessed February 28, 2019.
60. United States Department of Agriculture Economic Research Service. Download the Data: Food Access Research Atlas Data Download 2015. 2017; <https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/>. Accessed May 1, 2018.
61. 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). 2017; https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm#vintage2016. Accessed June 14, 2017.
62. 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). 2017; https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm#vintage2016. Accessed September 3, 2018.
63. US Census Bureau. Geographic Terms and Concepts – Census Tract. 2018; https://www.census.gov/geo/reference/gtc/gtc_ct.html. Accessed January 24, 2018.
64. Bivand R, Keitt T, Rowlingson B. rgdal: Bindings for the Geospatial Data Abstraction Library (R package version 1.2-8). 2017; <https://CRAN.R-project.org/package=rgdal>.
65. US Department of Education. Civil Rights Data Collection (CRDC) for the 2013-14 School Year. Access 2013-14 CRDC Data <https://www2.ed.gov/about/offices/list/ocr/docs/crdc-2013-14.html>. Accessed October 19, 2017.

SECTION 7: 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: Summary of Dataset of Origin, Censorship Rules, Estimate Provenance and Date of Download (metrics only)

NOTE: Symbol key on next page

Metric <i>Summary applies to city and tract analyses except where otherwise noted</i>	Dataset of origin	Censorship		Estimate Provenance	Date of Data Access*
		Internal Dashboard Guidelines	NVSS Guideline (as per DUA) [‡]	<i>Estimates posted on the Dashboard as received from source or calculated by the Dashboard's staff</i>	
		Censor where denominator <50	Censor where numerator <10		
Absenteeism	See Technical Document Part 2: Education Data (available for download as PDF)				
Air pollution - particulate matter	CMAQ, US EPA	no	no	Calculated by the Dashboard	2013: 5/1/18 2014: 1/28/19 2015: 5/7/19
Binge drinking	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 2018 releases: 12/4/18
Breast cancer deaths	MCDD, NVSS	yes	yes	Calculated by the Dashboard	6/2019
Cardiovascular disease deaths	MCDD, NVSS	yes	yes	Calculated by the Dashboard	
Children in poverty	ACS ^{**}	no	no	Calculated by the Dashboard	2013, 14, 15: 11/20/2017 2016: 1/18/18 2017: 12/12/19
Colorectal cancer deaths	MCDD, NVSS	yes	yes	Calculated by the Dashboard	6/2019
Dental care	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 18 releases: 12/4/18
Diabetes	BRFSS-CDC500	no	no	Posted as received from CDC	
High blood pressure	BRFSS-CDC500	no	no	Posted as received from CDC	
High school graduation	See Technical Document Part 2: Education Data (available for download as PDF)				
Housing cost, excessive	ACS ^{**}	no	no	Calculated by the Dashboard	2013, 14, 15 : 10/30/17 2016: 1/24/18 2017: 12/12/19
Frequent mental distress	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 2018 releases: 12/4/18
Frequent physical distress	BRFSS-CDC500	no	no	Posted as received from CDC	
Income inequality	ACS ^{**}	no	no	Calculated by the Dashboard	2013, 14, 15 : 12.7.17 2016: 1.25.18 2017: 12.12.18
Lead exposure risk index	ACS ^{**}	no	no	Calculated by the Dashboard	2013, 14, 15 : 11.13.17 2016: 1.25.18 2017: 12.12.18
Housing with potential lead risk	ACS ^{**}	no	no	Calculated by the Dashboard	
Life expectancy	USALEEP	no	no	Tract: Posted as received City: Calculated by the Dashboard	9/12/18
Limited access to healthy foods	USDA	yes	no	Calculated by the Dashboard	2/23/18
Neighborhood racial/ethnic segregation	ACS ^{**}	no	no	Calculated by the Dashboard	2013, 14, 15 : 2.11.19 2016: 1/18/18 2017: 12/12/18
Low birthweight	ND, NVSS	yes	yes	Calculated by the Dashboard	6/2019
Obesity	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 2018 releases: 12/4/18
Opioid overdose deaths	MCDD, NVSS	yes	yes	Calculated by the Dashboard	6/2019
Park access	ParkServe	no	no	Posted as received from ParkServe®	2/20/18
Physical inactivity	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 18 releases: 12/4/18
Premature deaths (all causes)	MCDD, NVSS	yes	yes	Calculated by the Dashboard	6/2019
Prenatal care ^{†††}	ND, NVSS	yes	yes	Calculated by the Dashboard	
Preventive services	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 18 releases: 12/4/18
Primary care physicians	AMA Physician Professional Data	yes	no	Calculated by the Dashboard	4/23/18
Racial/ethnic diversity	ACS ^{**}	no	no	Calculated by the Dashboard	2013, 14, 15 : 2.11.19 2016: 1/18/18 2017: 12/12/18
Smoking	BRFSS-CDC500	no	no	Posted as received from CDC	2016 release: 4.18.19; 2017, 18 releases: 12/4/18

Teen births	ND, NVSS	yes	yes	Calculated by the Dashboard	6/2019
Third-grade reading proficiency	See Technical Document Part 2: Education Data (available for download as PDF)				
Unemployment (City analysis)	ACS**	no	no	Posted as received***	2013, 14, 15: 11/16/17 2016: 1.26.18 2017: 12/12/19
Unemployment (Tract analysis)	ACS**	no	no	Calculated by the Dashboard	2013, 14, 15: 11/20/17 2016: 1/24/18 2017: 12/12/19
Uninsured	ACS**	no	no	Calculated by the Dashboard	2013, 14, 15: 2/19; 2016: 8/18 (B27001), 1/18 (S2701, C27001x); 2017: 12/18
Violent crime	FBI Uniform Crime Reporting	yes	no	Calculated by the Dashboard	2016, 17: 10/5/18;
Walkability	Walk Score	no	no	Posted as received from Walk Score®	4/9/18

Appendix B: Symbol KEY

* Datasets are sometimes updated after the initial release if the administrator identifies an error. This column records the date of dataset download to indicate to users which version of the underlying dataset informs our analyses.

‡ NVSS Data Use Agreement censorship guidelines require censorship of values with numerator less than 10

§ Data are posted as received from CDC; no analysis applied by the Dashboard

*** With exception of weighted averages for Asian and Other (see section on Unemployment above)

††† If prenatal care information is missing for 10% or more of a given city, all prenatal care values for that city are censored.

** As noted in “American Community Survey: ACS: Calculating MOEs for aggregate count data and derived proportions” above, 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 $MOE_{numerator}^2 - (\hat{p}^2 * MOE_{denominator}^2)$ resulted in a negative value. In these cases, confidence intervals could not be calculated and associated estimates were censored on the Dashboard. No other censoring of ACS data was performed.

Appendix C: Glossary of Abbreviations

ACS American Community Survey
AMA American Medical Association
BRFSS Behavioral Risk Factor Surveillance System
BRFSS-CDC500 Behavioral Risk Factor Surveillance System (BRFSS) data, reported by 500 Cities Project, Centers for Disease Control
CDC-500 500 Cities Project, Centers for Disease Control
CDC Centers for Disease Control
CI Confidence Interval
CIs Confidence intervals
CRDC Civil Rights Data Collection
FBI Federal Bureau of Investigation
FIPS Federal Information Processing Standards
ICD-10 International Statistical Classification of Diseases and Related Health Problems, 10th Revision
MCDD Multiple Cause of Death Data, National Vital Statistics System
NCES National Center for Education Statistics
NCHS National Center for Health Statistics
ND Natality Data, National Vital Statistics System
NVSS National Vital Statistics System
RSE Relative Standard Error
SE Standard Error
UCR Uniform Crime Reporting

Appendix D: Detailed Notes on Selection of City and Tract FIPS Codes

1. Notes on Selection of City and Tract FIPS Codes

Census tracts are subdivisions of county or equivalent entity with population size between 1200 and 8000 people. Census tract boundaries are defined to be maintained over time to facilitate statistical comparisons from census to census. Census tracts may be split because of population growth or merged because population decline. Census tracts are labelled with an integer basic tract number of maximum four digits and an optional two-digit suffix. Hence, census tract codes consist of six digits, with leading zeroes for basic tract number and trailing zeroes for suffix.⁶³

The 500 Cities: Census Tract Boundaries file (500Cities_Tracts_11082016.zip) was downloaded from the 500 Cities Data Portal.⁵ The shapefile in the .zip file, 500Cities_Tracts_Clip.shp, was imported into R version 3.4.16 with the rgdal package.⁶⁴ The field (variable) names were modified; refer to table below:

Original Variable Names	New Variable Names	Description
place2010	PLACE_FIPS	Place FIPS
tract2010	STATE_COUNTY_TRACT_FIPS	Concatenation of State FIPS, County FIPS, and Tract FIPS
ST	STATE_FIPS	State FIPS
PlaceName	PLACE_NAME	Place Name
plctract10	PLACE_TRACT_FIPS	Concatenation of Place FIPS (hyphen) State FIPS, County FIPS, and Tract FIPS
PlcTrPop10		Tract Population

STATE_COUNTY_TRACT_FIPS (formerly tract 2010) is a character string comprised of state FIPS, county FIPS, and tract FIPS; substrings were extracted from this string to create two new variables, COUNTY_FIPS and TRACT_FIPS.

- COUNTY_FIPS is a substring of STATE_COUNTY_TRACT_FIPS, start position 3 to end position 5.
- TRACT_FIPS is a substring of STATE_COUNTY_TRACT_FIPS, start position 6 to end position 11.

County names, COUNTY_NAME, and state abbreviations, STATE, were appended by joining the “2010 FIPS Codes for Counties and County Equivalent Entities” text file by state FIPS and county FIPS.⁸ State names, STATE_NAME, were appended by joining the “National FIPS and GNIS Codes File” by state FIPS.⁹

The final file used in the Dashboard analyses includes the following variables: PLACE_TRACT_FIPS, STATE_COUNTY_TRACT_FIPS, STATE_FIPS, STATE, STATE_NAME, COUNTY_FIPS, COUNTY_NAME, TRACT_FIPS, PLACE_FIPS, PLACE_NAME, PlcTrPop10.

Notable changes to geography: As of 2012, there was a change in the numbering of census tracts in Pima County, Tucson, Arizona.¹⁰ The CDC 500 Cities data reflects these changes.

2. R Tutorial

a. Setup

Select working directory, i.e. the directory where input and output files are saved. Load in dplyr and rgdal.

R Code

```
setwd(choose.dir())

library(dplyr)
library(rgdal)
```

b. Census tract boundaries shapefile from 500 Cities Open Data Portal

Download, save, and extract census tract boundaries shapefile in working directory from CDC 500 Cities: Census Tract Boundaries (version: updated December 4, 2017).⁶⁵

Downloaded file: 500Cities_Tracts_11082016.zip

Extracted file: 500Cities_Tracts_Clip(.cpf,.dbf,.prj,.sbn,.sbx,.shp,.xml,.shx)

Read in shapefile as save as R object named cdc_tracts.

R Code

```
cdc_tracts <- rgdal::readOGR("/500Cities_Tracts_Clip.shp", "500Cities_Tracts_Clip", stringsAsFactors = F)

## OGR data source with driver: ESRI Shapefile
## Source: "/500Cities_Tracts_Clip.shp", layer: "500Cities_Tracts_Clip"
## with 29305 features
## It has 6 fields
## Integer64 fields read as strings: PlcTrPop10
```

Save attributes from shapefile in R object named cdc_data. Rename the variables: "PLACE_FIPS", "STATE_COUNTY_TRACT_FIPS", "STATE_FIPS", "PLACE_NAME", "PLACE_TRACT_FIPS", and "PlcTrPop10". Create new variables, COUNTY_FIPS and TRACT_FIPS, from "STATE_COUNTY_TRACT_FIPS".

R Code

```
cdc_data <- cdc_tracts@data
names(cdc_data) <- c("PLACE_FIPS", "STATE_COUNTY_TRACT_FIPS", "STATE_FIPS",
                    "PLACE_NAME", "PLACE_TRACT_FIPS", "PlcTrPop10")

cdc_data$COUNTY_FIPS <- substr(cdc_data$STATE_COUNTY_TRACT_FIPS, 3, 5)
cdc_data$TRACT_FIPS <- substr(cdc_data$STATE_COUNTY_TRACT_FIPS, 6, 11)
```

Note Census 2010 FIPS code length for each geography: STATE (max length = 2), COUNTY (max length = 3), TRACT (max length = 6), PUMA (max length = 5)

c. 2010 FIPS Codes for Counties and County Equivalent Entities

Read text file of all counties in the United States, and save in R object named county_code. Assign variable names in county_code with names() function.

R Code

```
# COUNTY attributes
county_code <- read.table("https://www2.census.gov/geo/docs/reference/codes/files/national_county.txt",
                        sep=",", stringsAsFactors=FALSE, quote="",
                        colClasses=c("character", "character", "character",
                                     "character", "character"))
names(county_code) <- c("STATE", "STATE_FIPS", "COUNTY_FIPS", "COUNTY_NAME", "CLASS_FIPS")
county_attr <- county_code[,1:4]
```


d. National FIPS and GNIS Codes File

Read text file of all states in the United States, and save in R object named `states_code`. Create `data.frame` of `STATE_FIPS` and `STATE_NAME` and save as R object named `states_code`.

R Code

```
# STATE attributes
states_code <- read.table("http://www2.census.gov/geo/docs/reference/state.txt",
  sep="|", header=T,
  colClass=c("character", "character", "character", "character"))
states_attr <- data.frame(STATE_FIPS=states_code$STATE,
  STATE_NAME=states_code$STATE_NAME,
  stringsAsFactors = F)
```

e. Join selected variables from `cdc_geo`, `census2010_geo`, `county_code`, and `states_code`

Use `left_join` function from `dplyr` package to perform data joins

R Code

```
tracts_county <- dplyr::left_join(cdc_data, county_attr, by=c("STATE_FIPS", "COUNTY_FIPS"))
tracts_county_state <- dplyr::left_join(tracts_county, states_attr, by="STATE_FIPS")
```

f. Format and create final data file

R Code

```
final <- with(tracts_county_state,
  data.frame(PLACE_TRACT_FIPS, STATE_COUNTY_TRACT_FIPS,
    STATE_FIPS, STATE, STATE_NAME,
    COUNTY_FIPS, COUNTY_NAME, TRACT_FIPS,
    PLACE_FIPS, PLACE_NAME, PlcTrPop10),
  stringsAsFactors = F)
# Check for NA
(findNA <- final[!complete.cases(final),])

## [1] PLACE_TRACT_FIPS    STATE_COUNTY_TRACT_FIPS
## [3] STATE_FIPS          STATE
## [5] STATE_NAME          COUNTY_FIPS
## [7] COUNTY_NAME         TRACT_FIPS
## [9] PLACE_FIPS          PLACE_NAME
## [11] PlcTrPop10
## <0 rows> (or 0-length row.names)
```

g. Export data

Output to `.txt` file in working directory.

R Code

```
write.table(final, file="/500Cities_FIPS_Names.txt", quote=T, sep=",")
```

Appendix E: Summary of Geographies Reported for Honolulu, HI and Macon, GA (by Metric)

Metric	Dataset of origin	Macon, GA: Geography of Reported Value	Honolulu, HI: Geography of Reported Value
Children in poverty	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Housing cost, excessive	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Income inequality	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Lead exposure risk index	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Housing with potential lead risk	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Neighborhood racial/ethnic segregation	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Racial/ethnic diversity	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Unemployment (City analysis)	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)
Primary care physicians	AMA Physician Professional Data	Macon (FIPS 13-49000)	Urban Honolulu CDP (FIPS 15-71550)
Air pollution - particulate matter	CMAQ, EPA	Macon (FIPS 13-49000)	n/a
Binge drinking	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Dental care	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Diabetes	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
High blood pressure	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Frequent mental distress	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Frequent physical distress	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Obesity	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Physical inactivity	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Preventive services	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Smoking	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Uninsured	BRFSS-CDC500	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Violent crime	FBI UCR	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Breast cancer deaths	MCDD, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Cardiovascular disease deaths	MCDD, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Colorectal cancer deaths	MCDD, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Opioid overdose deaths	MCDD, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Premature deaths (all causes)	MCDD, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Absenteeism	See Technical Document Part 2: Education Data (available for download as PDF)		
Low birthweight	ND, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Prenatal care	ND, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Teen births	ND, NVSS	Macon (city) and Bibb county - see county_indicator variable	Honolulu County (FIPS 15-003)
Park access	ParkServe®	Macon (FIPS 13-49008)	Urban Honolulu CDP (FIPS 15-71550)
High school graduation	See Technical Document Part 2: Education Data (available for download as PDF)		
Third-grade reading proficiency	See Technical Document Part 2: Education Data (available for download as PDF)		
Life expectancy	USALEEP, NCHS	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Limited access to healthy foods	USDA	Macon (FIPS 13-49000)	Honolulu County (FIPS 15-003)
Walkability	Walk Score®	Macon (FIPS 13-49000)	Urban Honolulu CDP (FIPS 15-71550)
Demographic information (non-metric) - All Metrics View > More about...	ACS	Bibb County (FIPS 13-021)	Honolulu County (FIPS 15-003)

Note: 13 is the state FIPS for Georgia; 15 is the state FIPS for Hawaii

Appendix F: Summary of Definition of Hispanic, NHOPI and Other Race for Metrics with Demographic-Specific Values (by Metric)

Metric	Data Source	Hispanic ethnicity is mutually exclusive with racial groups	Definition of Native Hawaiian/ Pacific Islander*	Definition of "Other"	Metric value is available for specific racial/ethnic groups
Absenteeism	Civil Rights Data Collection	See Technical Document Part 2: Education Data			Yes
Breast cancer deaths	Multiple Cause of Death Data, NVSS, NCHS	Yes	Pacific Islander	n/a	Yes
Cardiovascular disease deaths	Multiple Cause of Death Data, NVSS, NCHS	Yes	Pacific Islander	n/a	Yes
Children in poverty	American Community Survey (ACS)	No	Native Hawaiian and Pacific Islander	American Indian or Alaska Native; Two or more races; Some other race	Yes
Colorectal cancer deaths	Multiple Cause of Death Data, NVSS, NCHS	Yes	Pacific Islander	n/a	Yes
High school graduation	data	See Technical Document Part 2: Education Data			Yes
Limited access to healthy foods	Food Access Research Atlas, Economic Research Service, United States Department of Agriculture	No	Native Hawaiian or Other Pacific Islander		Yes
Low birthweight	Nativity Data, NVSS, NCHS	Yes	Pacific Islander	n/a	Yes
Opioid overdose deaths	Multiple Cause of Death Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)	n/a	n/a	n/a	No
Premature deaths (all causes)	Multiple Cause of Death Data, NVSS, NCHS	Yes	Pacific Islander	n/a	Yes
Prenatal care	Nativity Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)	Yes	Pacific Islander	n/a	Yes
Teen births	Nativity Data, NVSS, NCHS	Yes	Pacific Islander	n/a	Yes
Unemployment	ACS	No	Native Hawaiian and Pacific Islander	American Indian or Alaska Native; Two or more races; Some other race	Yes
Air pollution - particulate matter	Community Multiscale Air Quality model, US Environmental Protection Agency	n/a	n/a	n/a	No
Binge drinking	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Dental care	500 Cities Project Data, Centers for Disease Control and Prevention	n/a	n/a	n/a	No
Diabetes	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Frequent mental distress	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Frequent physical distress	500 Cities Project Data, CDC	n/a	n/a	n/a	No

Metric	Data Source	Hispanic ethnicity is mutually exclusive with racial groups	Definition of Asian/Pacific Islander	Definition of “Other”	Metric value is available for specific racial/ethnic groups
High blood pressure	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Housing cost, excessive	ACS	n/a	n/a	n/a	No
Income inequality	ACS	n/a	n/a	n/a	No
Lead exposure risk index	ACS	n/a	n/a	n/a	No
Housing with potential lead risk	ACS	n/a	n/a	n/a	No
Life expectancy	USALEEP	n/a	n/a	n/a	No
Neighborhood racial/ethnic segregation	ACS	Yes – see Technical Document	Native Hawaiian and Pacific Islander	American Indian or Alaska Native; Two or more races; Some other race	No
Obesity	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Park access	ParkServe®	No	Native Hawaiian and Pacific Islander	American Indian or Alaska Native; Two or more races; Some other race	Yes
Physical inactivity	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Preventive services	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Primary care physicians	American Medical Association Physician Masterfile	n/a	n/a	n/a	No
Racial/ethnic diversity	ACS	Yes – see Technical Document	Native Hawaiian and Pacific Islander	American Indian or Alaska Native; Two or more races; Some other race	No
Smoking	500 Cities Project Data, CDC	n/a	n/a	n/a	No
Third-grade reading proficiency	data	n/a	n/a	n/a	No
Uninsured	ACS	n/a	n/a	n/a	No
Violent crime	Uniform Crime Reporting, Federal Bureau of Investigation	n/a	n/a	n/a	No
Walkability	Walk Score®	n/a	n/a	n/a	No

*Native Hawaiian/Pacific Islander is combined with Asian on the Demographic Detail display on the website

†Civil Rights Data Collection. Survey Forms: 2013-14 CRDC School Form (downloadable MS Word Document) - see table headers on page 10. Available at: <https://www2.ed.gov/about/offices/list/ocr/data.html>. Accessed April 26, 2018

Appendix G: Updates Summary

Technical Document Version	Date Technical Document Posted Online	Update Notes
v7.0	8/29/19	<ul style="list-style-type: none"> NVSS metrics: Added 2010-2012, 2011-2013, 2012-2014, 2015-2017 data <ul style="list-style-type: none"> Prenatal care: 2010-2012, 2011-2013, 2012-2014 <u>not</u> released NVSS metrics: Revisions to 2013-2015 and 2014-2016 data for improved accuracy
v6.0	7/15/19	<ul style="list-style-type: none"> Correction of Table of Contents and section label enumeration; no changes to Technical Document context or version number
v6.0	6/5/19	<ul style="list-style-type: none"> Information on Absenteeism moved to Technical Document Part 2: Education Data Multi-year data added to site and associated metadata added to Technical Documentation <ul style="list-style-type: none"> <u>Added multi-year data, by metric</u>: 500 Cities 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 FBI metric: 2015 data
v5.0	2/14/19	<ul style="list-style-type: none"> ACS metrics: Updated to 2017 (5 year estimate) data <ul style="list-style-type: none"> 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)
v4.1	12/20/18	<ul style="list-style-type: none"> Note regarding temporary removal of primary care physicians data from Dashboard inserted in this document
v4.0	12/12/18	<ul style="list-style-type: none"> 500 Cities metrics: Updated to 2018 release NVSS metrics: Updated to 2014-2016 data
v3.1	October 2018	<ul style="list-style-type: none"> 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 FBI 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)