Contents

SECTION 1: Overview .................................................................................................................. 6
   Document Mission .................................................................................................................. 6
   Note on State-based Education Data ...................................................................................... 6
   Measure Selection Criteria ..................................................................................................... 7
   City and Tract Selection Criteria ............................................................................................ 7
   Data Analysis .......................................................................................................................... 7
   Updates to Technical Documentation ..................................................................................... 7
   Downloading Dashboard Data ................................................................................................ 9
   Citing Dashboard Data and Technical Document ................................................................. 9
   Feedback or Errors .................................................................................................................. 9
   Measure Overview ................................................................................................................. 10

SECTION 2: Dashboard Analytic Decisions .............................................................................. 12
   Confidence Intervals (CIs) ...................................................................................................... 12
      Dashboard CIs are reported at the 90% level ...................................................................... 12
      Formulas for CI calculation ................................................................................................. 12
      Note on CIs for the Dashboard index values ....................................................................... 12
   Data Censoring ..................................................................................................................... 12
   Data Disclaimer ..................................................................................................................... 12
   Data Rounding ....................................................................................................................... 12
   Federal Information Processing Standards (FIPS) codes .................................................... 12
      Note on Honolulu, HI FIPS code ......................................................................................... 13
      Note on Macon, GA FIPS code ........................................................................................... 13
   Overall (500 Cities) Estimates .............................................................................................. 13
   Race/Ethnicity Categories ...................................................................................................... 13
   Validation ............................................................................................................................... 14
   Introduction to this Section .................................................................................................... 15
   500 Cities Project, Centers for Disease Control and Prevention ......................................... 15
      500 Cities Project: General notes ....................................................................................... 15
      500 Cities Project: Weights ................................................................................................. 15
      500 Cities Project: Categorizing race/ethnicity .................................................................. 15
      500 Cities Project: Metric-specific notes .......................................................................... 15
         Binge drinking .................................................................................................................. 15
         Dental care ..................................................................................................................... 16
         Diabetes ......................................................................................................................... 16
         Frequent physical distress ............................................................................................. 16
         Frequent mental distress ............................................................................................... 16
         High blood pressure ....................................................................................................... 16
Appendix A: Table of US 2010 Standardized Population ......................................................... 48
Appendix B: Summary of Dataset of Origin, Censorship Rules, Estimate Provenance and Date of Download (metrics only) .......................................................... 49
Appendix B: Symbol KEY ........................................................................................................ 50
Appendix C: Glossary of Abbreviations ................................................................................. 51
Appendix D: Detailed Notes on Selection of City and Tract FIPS Codes .............................. 52
  1. Notes on Selection of City and Tract FIPS Codes ............................................................... 52
  2. R Tutorial .............................................................................................................................. 53
Appendix E: Summary of Geographies Reported for Honolulu, HI and Macon, GA (by Metric)........ 55

SECTION 8: Appendices ........................................................................................................... 48
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 State-based Education Technical Document, 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 State-based 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 EDFacts. 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: State-based 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. The Dashboard selected city and tract FIPS codes as census tract boundary shapefiles released by the 500 Cities Project. 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.

Data Analysis

Primary data analysis of values calculated by the Dashboard was performed by Jessica Athens, Sarah Conderino, MPH (Surveillance Data Scientist, Department of Population Health, NYU School of Medicine), Shauna Ford, Miriam Gofine and Susan Kum, PhD (Postdoctoral Fellow, Department of Population Health, NYU School of Medicine). Sarah Conderino, Rania Kanchi, MPH (Data Analyst, Division of Epidemiology, Department of Population Health, NYU School of Medicine), Shauna Ford and Miriam Gofine contributed to data validation.

Updates to Technical Documentation

This technical document will be continuously updated as needed. Please note that the date of last update for this document is provided on the first page and the footer of this document.
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

Jessica Athens, PhD  
Director, Metrics and Analytics, City Health Dashboard  
Assistant Professor, Department of Population Health, New York University School of Medicine

Shoshanna Levine, DrPH, MPH  
Program Director, City Health Dashboard

Shauna Ford, MPH, MS  
Senior Data Analyst, City Health Dashboard

Miriam Gofine, MPH  
Data Analyst, City Health Dashboard

Allegra Wilson, BA  
Research Data Associate, City Health Dashboard
Downloading Dashboard Data

Dashboard data is available for free download at www.cityhealthdashboard.com.

Citing Dashboard Data and Technical Document

Dashboard data:

**NOTE:** as of May 15, 2018, .csv files are not yet available for download

Technical Document:

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.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Metric (Short Name)</th>
<th>Metric (Full Name)</th>
<th>Data Source</th>
<th>Measure Type</th>
<th>Years of Data</th>
<th>Reported at city level</th>
<th>Reported at tract level</th>
<th>Disaggregation by demographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Care</td>
<td>Dental care</td>
<td>Visits to dentist or dental clinic in the previous year among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, Centers for Disease Control and Prevention</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Prenatal care</td>
<td>Births for which prenatal care began in the first trimester (%)</td>
<td>Natality Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)</td>
<td>Percent</td>
<td>2012-2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Preventive services</td>
<td>Adults aged ≥65 years who are up to date on a core set of clinical preventive services (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Primary care physicians</td>
<td>Primary care physicians (per 100,000 population)</td>
<td>American Medical Association Physician Masterfile</td>
<td>Rate</td>
<td>2018</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Uninsured</td>
<td>Current lack of health insurance among adults aged 18-64 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Health Behaviors</td>
<td>Binge drinking</td>
<td>Binge drinking among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Physical inactivity</td>
<td>No leisure-time physical activity in past month among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Smoking</td>
<td>Current smoking among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Teen births</td>
<td>Births to mothers aged 15-19 (per 1,000 females in that age group)</td>
<td>Natality Data, NVSS, NCHS</td>
<td>Rate</td>
<td>2013-2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Breast cancer deaths</td>
<td>Deaths due to breast cancer in females (per 100,000 female population)</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Rate</td>
<td>2013-2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Colorectal cancer deaths</td>
<td>Deaths due to colorectal cancer (per 100,000 population)</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Rate</td>
<td>2013-2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Cardiovascular disease deaths</td>
<td>Deaths due to cardiovascular disease (per 100,000 population)</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Rate</td>
<td>2013-2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Diabetes</td>
<td>Diabetes among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Health Outcomes</td>
<td>Frequent mental distress</td>
<td>Mental health not good for ≥14 days during the past 30 days among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Frequent physical distress</td>
<td>Physical health not good for ≥14 days during the past 30 days among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>High blood pressure</td>
<td>High blood pressure among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Low birthweight</td>
<td>Live births with low birthweight &lt;2500 grams (%)</td>
<td>Natality Data, NVSS, NCHS</td>
<td>Percent</td>
<td>2012-2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Obesity</td>
<td>Adult obesity among adults age ≥18 years (%)</td>
<td>500 Cities Project Data, CDC</td>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Opioid overdose deaths</td>
<td>Deaths due to opioid overdose (per 100,000 population)</td>
<td>Multiple Cause of Death Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)</td>
<td>Rate</td>
<td>2012-2014</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Premature deaths</td>
<td>Years of potential life lost before age 75 (per 100,000 population)</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Rate</td>
<td>2013-2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Physical Environment</td>
<td>Social and Environmental Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Air pollution - particulate matter</strong></td>
<td><strong>Absenteeism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average daily concentration of fine particulate matter (PM2.5) per cubic meter (average)</td>
<td>Public school students who miss ≥15 days of school in an academic year (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Multiscale Air Quality model, US Environmental Protection Agency</td>
<td>Civil Rights Data Collection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>Percent</td>
<td>2013</td>
<td>✓</td>
<td>✓</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing with potential lead risk</td>
<td><strong>Children in poverty</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing stock with potential elevated lead risk (%)</td>
<td>Children living in households ≤100% of the federal poverty line</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Community Survey (ACS)</td>
<td>ACS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent</td>
<td>2016 (5 Year Estimates)</td>
<td>✓</td>
<td>✓</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited access to healthy foods</td>
<td><strong>Housing cost, excessive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population living more than ½ mile from the nearest supermarket, supercenter, or large grocery store (%)</td>
<td>Households where ≥30% of household income is spent on housing costs (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Access Research Atlas, Economic Research Service, United States Department of Agriculture</td>
<td>ACS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lead exposure risk index</td>
<td><strong>Walkability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty-adjusted risk of housing-based lead exposure (index)</td>
<td>Neighborhood amenities accessible by walking as calculated by Walk Score® (index)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACS</td>
<td>Walk Score®</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>2016 (5 Year Estimates)</td>
<td>✓</td>
<td>✓</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park access</td>
<td><strong>Third-grade reading proficiency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population living within a 10 minute walk of green space (%)</td>
<td>Third-graders who score “proficient” or above in reading on standardized tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ParkServe®</td>
<td>See Technical Document Part 2: State-based Education Data (available for download as PDF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent</td>
<td>2015</td>
<td>✓</td>
<td>⨿</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walkability</td>
<td><strong>Income inequality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighborhood racial/ethnic segregation</td>
<td>Households with income at the extremes of the national income distribution (the top 20% or bottom 20%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distribution of the population by race/ethnic group within a census tract relative to the distribution across the city</td>
<td>ACS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>2016 (5 Year Estimates)</td>
<td>✓</td>
<td>✓</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racial/ethnic diversity</td>
<td><strong>Unemployment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distribution of the population by race/ethnic group within a city or census tract (index)</td>
<td>Population aged ≥16 years that is unemployed but seeking work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACS</td>
<td>ACS</td>
<td>2016 (5 Year Estimates)</td>
<td>✓</td>
<td>✓</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social and Environmental Factors</strong></td>
<td><strong>Violent crime</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students who graduate high school within 4 years of entering ninth grade</td>
<td>Violent crime offenses (murder, aggravated assault, robbery, forcible rape) per 100,000 population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>See Technical Document Part 2: State-based Education Data (available for download as PDF)</td>
<td>Uniform Crime Reporting, Federal Bureau of Investigation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income inequality</td>
<td>Percent</td>
<td>2016 (5 Year Estimates)</td>
<td>✓</td>
<td>✓</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighborhood racial/ethnic segregation</td>
<td>Rate</td>
<td>2016</td>
<td>✓</td>
<td>⨿</td>
<td>⨿</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racial/ethnic diversity</td>
<td>Unemployment</td>
<td>(City only)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Third-grade reading proficiency | Violent crime |...
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. 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. 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. The Dashboard selected city and tract FIPS codes as census tract boundary shapefiles released by the 500 Cities Project. As per the CDC 500 Cities Project, 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.

See Appendix E for a summary of the geographic coding used for Macon, GA, per metric.

**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.

**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)). Federal guidelines for reporting data by demographics 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. As of May 15, 2018, the following steps have been completed:

1. Internal data results validation

   All analyses* on the Dashboard were first 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 (beta) site data validation

   Analysts from the City Health Dashboard analysis team web-scraped data on the Dashboard’s beta site in order to compare website data with the .csv datafiles sent directly to the site developers. There were no discrepancies noted.

As of May 15, 2018, further validation steps are ongoing. This Technical Document will be re-released as new validation steps are completed.

*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

500 Cities Project: 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. 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 RStudio v3.3.2).

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). For more information regarding these metrics, please refer to the 500 Cities Project’s methodology pages.

500 Cities Project: 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, we also calculate an average rate for men and women, weighting each group evenly.

500 Cities Project: 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.

500 Cities Project: 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.
Dental care
Percent of respondents aged ≥18 years who report having been to the dentist or dental clinic in the previous year.\textsuperscript{13}

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.\textsuperscript{11}

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.\textsuperscript{11}

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.\textsuperscript{11}

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.\textsuperscript{11}

Obesity
Adult obesity among adults aged ≥18 years.\textsuperscript{12}

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?”\textsuperscript{12}

Preventive services
\textit{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.\textsuperscript{13}

\textit{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.\textsuperscript{13}

Smoking
Respondents aged ≥18 years who report having smoked ≥100 cigarettes in their lifetime and currently smoke every day or some days.\textsuperscript{12}
Uninsured
Respondents aged 18–64 years who report having no current health insurance coverage.13

Data tables
Tract and city-level data were downloaded directly from the 500 Cities Project website.5

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.

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

\[ LCI90 = Est - (1.645 \times SE) \]

\[ UCI90 = Est + (1.645 \times SE) \]

Where:
- SE = approximate standard error
- LCI95 = Reported lower limit for the 95% confidence interval
- UCI95 = Reported upper limit for the 95% confidence interval
- Est = Reported estimate
- LCI90 = Calculated lower limit for the 90% confidence interval
- UCI90 = Calculated upper limit for the 90% confidence interval

American Community Survey (ACS)
ACS: General notes
ACS is administered by the US Census Bureau; data tables are available for download on American FactFinder.14 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. Dashboard analyses using ACS used 5 Year Estimate data tables. All metric analyses using ACS tables used 2016 data. Values derived from ACS that were used as population denominators in metric analysis vary in year (see Section 5 for more details). All analyses of ACS data were performed using SAS v9.4.15

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.

ACS: Weights
Weights were not applied to ACS data as these data do not require weighting.

ACS: 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.14

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.

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

ACS: 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.16

Relevant formulas are presented verbatim here for users’ reference:

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

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

Calculating MOE’s for Derived Proportions (p. A-14, A-15)
MOE_{\text{derived proportion}} = \pm \sqrt{\frac{\text{MOE}_{\text{numerator}}^2}{\hat{X}_{\text{denominator}}} \left( \hat{\rho}^2 \cdot \text{MOE}_{\text{denominator}}^2 \right)}

“where MOE_{\text{numerator}} is the MOE of the numerator; MOE_{\text{denominator}} is the MOE of the denominator; \hat{\rho} = \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 \( \sqrt{\text{MOE}_{\text{numerator}}^2 (\hat{\rho}^2 \cdot \text{MOE}_{\text{denominator}}^2) } \) because \( \text{MOE}_{\text{numerator}}^2 (\hat{\rho}^2 \cdot \text{MOE}_{\text{denominator}}^2) \) 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.
**ACS: Metric-specific notes**

**Children in poverty**

**ACS 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” value presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

**Analysis**
Children in Poverty = \[rac{\text{[Children Age} < \text{ Living in Households below the poverty threshold]}}{\text{[Total number of children age <18 living in households]}} \times 100\%.
\]

The following variables within each data table were summed to calculate the numerator: HD01_VD03, HD01_VD04, HD01_VD05. These variables were summed with HD01_VD11, HD01_VD12, HD01_VD13 to calculate the denominator.

MOE’s for summed estimates were calculated as per published guidance.\(^{16}\) See section “ACS: Calculating MOEs for aggregate count data and derived proportions” for this equation in full.

MOE’s for derived proportions were calculated as per published guidance.\(^{16}\) See section “ACS: Calculating MOEs for aggregate count data and derived proportions” for this equation in full.

**Demographic information**

**ACS data tables**
NOTE: Demographic information is **not** a metric. The graphic information is provided as supplementary information for Dashboard users. This section outlines how these demographic estimates were calculated.

Data table DP05 (2016 5 Year Estimates) was used to provide demographic information about city population values at the city level. Data table S1701 (2016 5 Year Estimates) was used report the percentage of the population with income below <100% of federal poverty level at the city levels.

The demographic information is displayed on “More about [city name]” on each city’s All Metrics View page on the Dashboard.

**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”)

Last updated 5/15/18
- Labelled "Older adults (age 65+)": HC01_VC29 ("Estimate; SEX AND AGE - 65 years and over")
- Labelled "White, non-Hispanic": HC01_VC94 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - White alone")
- Labelled "Black, non-Hispanic": HC01_VC95 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Black or African American alone")
- Labelled "Asian, non-Hispanic": HC01_VC97 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Asian alone")
- Labelled "Other, non-Hispanic": HC01_VC99 ("Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Some other race alone")
- Labelled "Two or more races, any ethnicity": HC01_VC45 ("Estimate; RACE - Total population - Two or more races")
- Labelled "American Indian or Alaska Native (one or more races, any ethnicity)": HC01_VC80 ("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_VC82 ("Estimate; RACE - Race alone or in combination with one or more other races - Total population - Native Hawaiian and Other Pacific Islander")
- Labelled "Hispanic": HC01_VC88 ("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_VC26 ("Estimate; SEX AND AGE - 18 years and over")]

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

**Housing cost, excessive**

*ACS data tables*
Data table DP04 was used to calculate excessive housing cost at both city and tract levels. The "national" value presented on the Dashboard reflects values for the Dashboard’s 500 cities, not the entire United States.

**Analysis**
Excessive housing cost = 100%* [(Selected monthly owner costs (with mortgage): 30.0%-34.9% of monthly income)+ (Selected monthly owner costs (without mortgage): 30.0%-34.9% of monthly income)+Gross Rent as a percentage of household income: 30.0%-34.9% of monthly income) + (Selected monthly owner costs (with mortgage): >=35.0% of monthly income)+ (Selected monthly owner costs (without mortgage): >=35.0% of monthly income)+Gross Rent as a percentage of household income: >=35.0% of monthly income)]/[Total occupied housing units).
In both City and Tract analyses, the following variables in DP04 were summed to calculate the numerator: HC01_VC163, HC01_VC175, HC01_VC203, HC01_VC164, HC01_VC176, and HC01_VC204. The denominator was HC01_VC04.

MOE’s for summed estimates were calculated as per published guidance. See section “ACS: Calculating MOEs for aggregate count data and derived proportions” for this equation in full.

Income inequality

ACS data tables
Data table B19001 was used to calculate income inequality at both city and tract levels. The “national” value 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 et al.

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). As of Dashboard data analysis in March 2018, the most recently available cutpoints for the US 20th and 80th household income percentiles were $24,002 and $121,018, respectively, as per 2016 US Census Bureau data Table H-1 (2016 data, All Races). The following variables in ACS Table B19001 were summed to calculate A(i): HD01_VD14, HD01_VD15, HD01_VD16, HD01_VD17. These variables represent estimates of the number of individuals with income of or greater than $125,000, the closest value to $121,018. The following variables were summed to calculate P(i): HD01_VD02, HD01_VD03, HD01_VD04, HD01_VD05. These variables represent estimates of the number of individuals with $24,999 or less, the closest value to $24,002. In both City and Tract analyses, HD01_VD01 was used to represent T(i).

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

ACS data tables
Data table B25034 was used to calculate housing risk data at both city and tract levels. The “national” value 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.19 Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale.20 The Dashboard adapted Vox Media’s Python code available on Github21 for the present analysis, which was conducted by the Dashboard using SAS v9.416 and validated using Python v3.6.22 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 estimates16; weighting those MOE’s with the same weights used to calculate the numerator; and then calculating an MOE for a derived proportion.16 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 200223 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.21

b. The following variables in B25034 were summed to represent all housing stock built in 2010 or later: “HD01_VD03 Estimate; Total: Built 2010-2013” and “HD01_VD02 Estimate; Total: Built 2014 and later”.

Last updated 5/15/18
Lead exposure risk index

ACS data tables
Data table B25034 was used to calculate housing risk data at both city and tract levels. S1701 was used for calculating poverty risk at both city and tract levels. The “national” value 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. Vox Media worked in conjunction with Washington State Department of Health to apply this methodology on a national scale. The Dashboard adapted Vox Media’s Python code available on Github for the present analysis, which was conducted by the Dashboard using SAS v9.4 and validated using Python v3.6. 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 2002 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 following variables in the B25034 were summed to represent all housing stock built in 2010 or later: “HD01_VD03 Estimate; Total: Built 2010-2013” and “HD01_VD02 Estimate; Total: Built 2014 and later”.

Neighborhood racial/ethnic segregation

ACS data tables
Data table DP05 was used to calculate racial/ethnic segregation at the city level. The “national” value 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.

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: \( H = \sum_{i=1}^{n} \frac{t_i(E_i - E)}{ET} \) 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.” 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.

The following variables were used in the diversity and segregation analyses: HC01_VC88 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Hispanic or Latino (of any race); HC01_VC94 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - White alone); HC01_VC95 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Black or African American alone); HC01_VC96 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - American Indian and Alaska Native alone); HC01_VC97 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Asian alone); HC01_VC98 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Native Hawaiian and Other Pacific Islander alone); HC01_VC99 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Some other race alone); HC01_VC100 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Two or more races).

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

ACS data tables
Data table DP05 was used to calculate racial/ethnic diversity values at the city and tract levels. The “national” value 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. 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 \( \Pi_r \) 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 \( \Pi_{ri} \) refers to a particular racial/ethnic group’s proportion of the population in tract i.”

As per footnote 5 in Iceland, \( \ln \left( \frac{1}{\Pi_r} \right) \) and is set to 0 when the proportion of a particular group is in a given geography (\( \Pi_r \)) 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.

The following variables were used in the diversity and segregation analyses: HC01_VC88 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Hispanic or Latino (of any race); HC01_VC94 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - White alone); HC01_VC95 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Black or African American alone); HC01_VC96 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - American Indian and Alaska Native alone); HC01_VC97 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Hawaiian and Other Pacific Islander alone); HC01_VC99 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Some other race alone); HC01_VC100 (Estimate; HISPANIC OR LATINO AND RACE - Total population - Not Hispanic or Latino - Two or more races).

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 overcounts 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

ACS data tables
Data table S2301 was used to report unemployment rates, disaggregated by race/ethnicity and
sex, at the city level. The “national” value 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
City level: Unemployment rates reported in S2301 are presented as reported in the original
2016 data tables for the following groups: total population (HC04_EST_VC01); black
(HC04_EST_VC16), white (HC04_EST_VC24), Hispanic or Latino origin (of any race)
(HC04_EST_VC23), males (HC04_EST_VC28), females (HC04_EST_VC29).

Estimates and confidence intervals values for “other” and “Asian” are weighted averages of
estimates and confidence intervals for the subgroups that comprise these groups throughout the
Dashboard. The weighted average for other race is comprised of American Indian and Alaska
Native alone (HC04_EST_VC17); Some other race alone (HC04_EST_VC20); and Two or more
races (HC04_EST_VC21). Confidence intervals were calculated using each estimate’s
associated margin of error variable (HC04_MOE_VC17, HC04_MOE_VC20, and HC04_-
MOE_VC21, respectively). Estimate and confidence interval values are weighted by the relative
proportion of each of these groups within the summed total 2016 population of these three
groups within each city as per ACS table DP05 (variables HC01_VC51, HC01_VC69,
HC01_VC70), available for download on American FactFinder.14

The weighted average for Asian is comprised of Asian alone (HC04_EST_VC18)+Native
Hawaiian and Other Pacific islander alone (HC04_EST_VC19)). Confidence intervals were
calculated using each estimate’s associated margin of error variable (HC04_MOE_VC18 and
HC04_MOE_VC19, respectively). Estimate and confidence interval values are weighted by the
relative proportion of each of these groups within the summed total 2016 population of these
two groups within each city as per ACS table DP05 (variables HC01_VC56, HC01_VC64),
available for download on American FactFinder.14

Tract level: Unemployment rate was derived by dividing the estimate for the estimate of
individuals unemployed in the civilian labor force (HD01_VD05) by the estimate for the number
in the civilian labor force (HD01_VD03).
MOE’s for summed estimates were calculated as per published guidance. See section “ACS: Calculating MOEs for aggregate count data and derived proportions” for this equation in full.

MOE’s for derived proportions were calculated as per published guidance. See section “ACS: Calculating MOEs for aggregate count data and derived proportions” for this equation in full.
American Medical Association (AMA) Physician Professional Data

AMA: General notes
The primary care physician metric was calculated using the 2018 Physician Professional Data file (also referred to as Physician Masterfile data) received from the American Medical Association (AMA). The file was current as of April 23, 2018; it is not publicly available at the city level. City-level primary care physician counts were used for city-level analyses. Total population values were calculated using the 2016 American Community Survey’s DP05 data table; refer to section 5 for more information regarding ACS population denominators. All analyses of AMA data were performed using SAS v9.4. All calculated values were rounded to 1 decimal place immediately prior to data export.

AMA: Weights
No weights were applied to AMA data.

AMA: Categorizing race/ethnicity
Not applicable.

AMA: Confidence intervals
CIs for primary care physicians were calculated as follows:

\[
\text{LCI}_{90} = \left[ \frac{100000}{\text{denominator}} \times (\text{numerator} - (1.645 \times \sqrt{\text{numerator}})) \right]
\]
\[
\text{UCI}_{90} = \left[ \frac{100000}{\text{denominator}} \times (\text{numerator} + (1.645 \times \sqrt{\text{numerator}})) \right]
\]

AMA: Metric-specific notes

Primary care physicians

AMA data tables
The 2018 Physician Professional Data was used to calculate total number of primary care physicians. The 2016 American Community Survey’s data table DP05 was used to calculate total population values. See section 5 for more on population denominators.

Analysis
Primary care physicians = [total number of primary care physicians/total population in 2016] \times 100,000.

The following types of physician specialties were summed to calculate the total number of primary care physicians: adolescent medicine, adolescent medicine (family practice), adolescent medicine-peds, family medicine/preventive medicine, family practice/sports medicine, family practice, family practice/geriatric medicine, general practice, general preventive medicine, internal medicine, internal medicine - family practice, internal medicine - geriatrics, internal medicine - pediatrics, internal medicine - preventive medicine, and pediatrics.
Community Multiscale Air Quality model, US Environmental Protection Agency (CMAQ, EPA)

CMAQ, EPA: General notes
The air pollution metric, Air pollution - Average daily concentration of fine particulate matter (PM2.5) per cubic meter, was calculated using 2013 Community Multiscale Air Quality model output for the continental United States.

CMAQ, EPA: Weights
The data available for download from the US Environmental Protection Agency's Community Multiscale Air Quality model provide daily tract-level estimates for the year 2013 for all tracts in the contiguous United States. Population weights were calculated for each tract using the 2016 American Community Survey DP05 total population estimates.

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

Where:
\( w_i \) = tract-specific weight
\( p_i \) = tract total population from ACS Table DP05
\( n \) = total tracts within a city

CMAQ, EPA: Categorizing race/ethnicity
Not applicable.

CMAQ, EPA: Confidence intervals
Not applicable.

CMAQ, EPA: 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 PM2.5 is available online, through the EPA.25

Air pollution - particulate matter

CMAQ, EPA data tables
Tract level data are available for download from the US Environmental Protection Agency.26 As of April 2017, the most current output data files are for the year 2013. Daily estimates of PM 2.5 concentration are reported for census tracts (vintage 2010) within the continental U.S.

The American Community Survey data table DP05 (2016 5-year estimates)14 was used to calculate population weights for each census tract to create population-weighted city averages of PM 2.5 exposure.

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:

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

Civil Rights Data Collection (CRDC)

**CRDC: General notes**
Absenteism data was calculated using the 2013-14 Civil Rights Data Collection survey.\(^{27}\) These data were the most recently available data at the time of analysis in February 2018.

Schools with preschools (SCH_GRADE_PS=“YES”) and Juvenile Justice Facility Schools (JJ=“YES”) were excluded from analysis.

All analyses were performed using SAS v9.4.\(^{15}\)

**CRDC: Weights**
No weights were applied to the CRDC analysis.

**CRDC: Categorizing race/ethnicity**
With the exception of “Other”, racial/ethnic groups were categorized as they were throughout the Dashboard: black, Asian (Asian and Native Hawaiian or Pacific Islander); Hispanic; or White. “Other” represents summed values associated with American Indian or Alaska Native and two or more races only because CRDC data does not provide a “two or more races” category.

**CRDC: Confidence intervals**
CIs are not presented because MOE or SE data were not presented in the underlying dataset, which presents counts.

**CRDC: Metric-specific notes**

Absenteism

**Analysis**
Chronic student absenteeism variables (see variables prefix SCH_ABSENT_) were used in this analysis. Sex-specific values within schools were summed to create totals per school. A spatial join was performed between the school locations (points) and the city boundaries (polygons) to identify schools within the city boundaries. The latitude and longitude of a school, as per the CCD_LATCOD and CCD_LONCOD variables, was used to determine its location within city boundaries defined by the CDC 500 Cities Project.\(^{27}\) See the State-based Data Technical Document for more on geospatial joins.

Schools with preschools (SCH_GRADE_PS=“YES”) and Juvenile Justice Facility Schools (JJ=“YES”) were excluded from analysis.
National Vital Statistics System (NVSS)

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 v9.4.15

The 2014 ACS data table B01001 was used to calculate age-group specific total-population, female, and male values for NVSS data (please refer to section 5 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.28

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.

Multiple Cause of Death Data
The following metrics were calculated using 2012-2014 and 2013-2015 MCDD datasets from NVSS: opioid overdose deaths (2012-2014), breast cancer deaths, cardiovascular disease deaths, colorectal cancer deaths and premature deaths (all causes), respectively. See below for more detail.

Natality Data
The following metrics were calculated using 2012-2014, 2014, and 2013-2015 ND datasets from NVSS: low birthweight, prenatal care and teen births, respectively. See below for more detail.

NVSS: Weights
Multiple Cause of Death Data
Breast cancer, colorectal cancer, cardiovascular disease, and opioid overdose deaths metrics used US 2010 standardized population weights. These weights were calculated via the direct adjustment approach outlined by Klein & Schoenborn29 utilizing the data table “QT-P1, 2010 Decennial Census” downloaded in December 2016 from American Fact Finder.14

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.30
Weights for age-adjusting premature deaths were calculated using the data table “QT-P1, 2010 Decennial Census”.14 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
\[
\text{Weight(age group } i) = 75 - \text{Mid-point age group } i.
\]
Weights and years of life lost are presented in Appendix A.

Natality Data
No weights are applied to ND data.
NVSS: 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

Natality Data

2012 & 2013: 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: Mother’s Hispanic origin recode (mhisp_r) and mother’s bridged race (mbrate) 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
NVSS: 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 Stolley\textsuperscript{31} in a document published by the Utah Department of Health\textsuperscript{32}:

\[
\begin{align*}
\text{LCI90:} &\quad \text{estimate} - (1.645 \times \text{SE(est)}) \\
\text{UCI90:} &\quad \text{estimate} + (1.645 \times \text{SE(est)})
\end{align*}
\]

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

CIs for premature deaths (all causes) were calculated according to the following formula outlined by Vohlonen, Bäckmand, & Korhonen\textsuperscript{33}:

\[
\begin{align*}
\text{LCI90:} &\quad \text{estimate} - (1.645 \times \text{SE(est)}) \\
\text{UCI90:} &\quad \text{estimate} + (1.645 \times \text{SE(est)})
\end{align*}
\]

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

Natality Data

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

\[
\begin{align*}
\text{LCI90} = &\quad \text{estimate} - 1.645 \times \sqrt{\text{estimate} \times ((100-\text{estimate})/\text{numerator})} \\
\text{UCI90} = &\quad \text{estimate} + 1.645 \times \sqrt{\text{estimate} \times ((100-\text{estimate})/\text{numerator})}
\end{align*}
\]

CIs for teen births metric were calculated as follows:

\[
\begin{align*}
\text{LCI90} = &\quad [(1000 / \text{denominator}) \times (\text{numerator} - (1.645 \times \sqrt{\text{numerator}}))] \\
\text{UCI90} = &\quad [(1000 / \text{denominator}) \times (\text{numerator} + (1.645 \times \sqrt{\text{numerator}}))]\]
\]
NVSS: City/County Indicator

Multiple Cause of Death Data

All MCDD calculated values are based on city of residence. However, county of residence data is used to derive city of residence in the event that city of residence is not available in the dataset received by the Dashboard (e.g., due to data censorship by NVSS, etc.). 

The Dashboard created the following indicator, released with the Dashboard’s downloadable datasets (NOTE: as of May 15, 2018, .csv files are not yet available for download), to provide more information on which geography was utilized for all MCDD 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)

Natality Data

All ND calculated values are based on the mother’s city of residence. However, the mother’s county of residence data is used to derive mother’s city of residence in the event that mother’s city of residence is not available in the dataset received by the Dashboard (e.g., due to data censorship by NVSS, etc.).

The Dashboard created the following indicator, released with the Dashboard’s downloadable datasets (NOTE: as of May 15, 2018, .csv files are not yet available for download) to provide more information on which geography was utilized for all MCDD 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)

NVSS: Year(s) of Data Used: Multiplier Indicator

Multiple Cause of Death Data

All MCDD calculated values are based on 2012-2014 or 2013-2015 data. However, fewer years of data is 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 (NOTE: as of May 15, 2018, .csv files are not yet available for download) 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 denominators. Specifically, the 2014 city and county-level population denominators will be multiplied by one, two, or three respectively.
Natality Data

All ND calculated values are based on 2012-2014, 2013-2015, or 2014 data. However, fewer years of data is 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 (NOTE: as of May 15, 2018, .csv files are not yet available for download) 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 denominators. Specifically, the 2014 city and county-level population denominators will be multiplied by one, two, or three respectively.

NVSS: Metric-specific notes

Breast cancer deaths

NVSS data tables
2013, 2014, & 2015 Multiple Cause of Death data files were combined to calculate breast cancer deaths among females at the city level. Please refer to section 5 for more information on population denominators.

Analysis
Breast cancer deaths = \[
\text{sum}\left\{\left(\frac{\text{age-group specific total number of breast cancer deaths amongst females}}{\text{multiplier} \times \text{age-group specific 2014 total female population}}\right) \times \text{US 2010 standardized population age-group specific weight}\right\} \times 100,000
\]

Notes on Analysis
Age-adjusted mortality rates are calculated as per National Association for Public Health Statistics and Information Systems recommendations.\(^{34}\)

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.\(^{35}\)

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

Cardiovascular disease deaths

NVSS data tables
2013, 2014, & 2015 Multiple Cause of Death data files were combined to calculate cardiovascular disease deaths at the city level. Please refer to section 5 for more information on population denominators.
Analysis
Cardiovascular disease deaths = \[ \text{sum}((\text{age-group specific total number of cardiovascular disease deaths})/(\text{multiplier} \times \text{age-group specific 2014 total population}) \times \text{US 2010 standardized population age-group specific weight}) \times 100,000 \]

Notes on Analysis
Age-adjusted mortality rates are calculated as per National Association for Public Health Statistics and Information Systems recommendations.34

The following underlying cause of death ICD-10 codes were summed to calculate the total number of cardiovascular disease deaths:

ICD-10 codes were selected for inclusion based on Nolte & McKee36 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

NVSS data tables
2013, 2014, & 2015 Multiple Cause of Death data files were combined to calculate colorectal cancer deaths at the city level. Please refer to section 5 for more information on population denominators.

Analysis
Colorectal cancer deaths = \[ \text{sum}((\text{age-group specific total number of colorectal cancer deaths})/(\text{multiplier} \times \text{age-group specific 2014 total population}) \times \text{US 2010 standardized population age-group specific weight}) \times 100,000 \]

Notes on Analysis
Age-adjusted mortality rates are calculated as per NAPHSIS recommendations.34

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 al37 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

NVSS data tables
2012, 2013, & 2014 Natality data files were combined to calculate low birthweight at the city level.

Analysis
Low birthweight = (number of live births with birthweight <2500 grams/total number of live births) * 100

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

Opioid overdose deaths

NVSS data tables
2012, 2013, & 2014 Multiple Cause of Death data files were combined to calculate opioid overdose deaths at the city level. Please refer to section 5 for more information on population denominators.

Analysis
Opioid overdose deaths = [sum((age-group specific total number of opioid overdose deaths)/(3 * age-group specific 2013 total population)) * US 2010 standardized population age-group specific weight)] * 100,000

Notes on Analysis
Age-adjusted mortality rates are calculated as per National Association for Public Health Statistics and Information Systems recommendations.34

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.38,39

All deaths with either a missing, unknown, or not stated age are excluded from the analysis.
**Premature deaths (all causes)**

*NVSS data tables*
2013, 2014, & 2015 Multiple Cause of Death data files were combined to calculate premature deaths (all causes) at the city level. Please refer to section 5 for more information on population denominators.

*Analysis*
Premature deaths (all causes)=$\left[\text{sum}\left[\left(\frac{\text{age-group specific total number of deaths}}{\text{multiplier} \times \text{age-group specific 2014 total population}}\right) \times \text{US 2010 standardized population YPLL age-group specific weight} \times \text{age-group specific premature deaths weight - years of life lost}\right]\right] \times 100,000.$

*Notes on Analysis*
Premature deaths (all causes) rates are calculated as per Dranger and Remington’s approach.$^{30}$ 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**

*NVSS data tables*
2014 Natality data file was used to calculate prenatal care at the city level.

*Analysis*
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*
Prenatal care estimates represent a slight modification of one component of the Kotelchuck Index.$^{40}$ All births with missing or unknown prenatal care are excluded from the analysis. Prenatal care data for Connecticut, Rhode Island, and New Jersey are missing because these states use the unrevised birth certificate. 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 2014 natality public use data documentation file.$^{41}$

**Teen births**

*NVSS data tables*
2013, 2014, & 2015 Natality data files were combined to calculate teen births at the city level. Please refer to section 5 for more information on population denominators.

*Analysis*
Teen Births = \([\frac{\text{number of live births to mothers ages 15-19}}{\text{multiplier} \times \text{2014 total female population age 15-19}}]\] \times 1,000
Uniform Crime Reporting, Federal Bureau of Investigation (UCR)

**UCR: 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.\(^{15}\)

**UCR: Weights**
No weights were applied to UCR data.

**UCR: Categorizing race/ethnicity**
Not applicable.

**UCR: Confidence intervals**
CIs for violent crime were calculated as follows:

\[
\text{LCI}_{90} = \left[ \frac{100000}{\text{denominator}} \ast (\text{numerator} - (1.645 \ast \sqrt{\text{numerator}})) \right]
\]
\[
\text{UCI}_{90} = \left[ \frac{100000}{\text{denominator}} \ast (\text{numerator} + (1.645 \ast \sqrt{\text{numerator}})) \right]
\]

**UCR: Metric-specific notes**

**Violent Crime**

**UCR data tables**
The 2016 data table, “Table 6: Offenses Known to Law Enforcement by State by City”\(^{42}\) (https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016/tables/table-6/table-6.xls/view) from Uniform Crime Reporting was used to calculate total number of violent crimes. The 2016 American Community Survey’s data table DP05 was used to calculate total-population values. Please refer to section 5 for more information regarding ACS population denominators.

**Analysis**
Violent crime = [total number of reported violent crimes/total population in 2016]\(\ast\)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.

USDA Food Access Research Atlas: General notes
The limited access to healthy food metric represents the percent of the population that lives ≥ ½ mile from a supermarket, large grocery store, or supercenter. Data on store locations were obtained by USDA ERS from the Trade Dimensions TD Linx directory of stores. Population data total and by race/ethnicity were obtained at the block level from US Census. The Food Access Research Atlas is current as of 2015 (as of the writing of this document in May 2018). The years for the underlying data sets are 2010.

USDA Food Access Research Atlas: Weights
Weights were not applied to USDA ERS data as these data do not require weighting.

USDA Food Access Research Atlas: 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.

USDA Food Access Research Atlas: Metric-specific notes

Limited access to healthy foods
The limited access to healthy food metric represents the percent of the population that lives ≥ ½ 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.

USDA Food Access Research Atlas data tables
Census tract-level data were downloaded directly for this measure.

Analysis
The downloaded data set from USDA ERS reports numerators (population living ≥ ½ mile from a supermarket, large grocery store, or supercenter) and denominators for total population and by race/ethnicity. The measure was calculated as follows.

\[
\frac{\text{Population living ≥ 0.5 mile from a healthy food store}}{\text{Total population}} \times 100
\]

City level estimates were calculated by summing tract numerators and denominators to the city level for total population and by race/ethnicity.
As of May 15, 2018, data are being validated by Park Serve®. Finalized data will be posted on the Dashboard as it becomes available.

City and tract-level data are expected to be released publicly by ParkServe® in Summer 2018.

**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

**ParkServe®: Weights**
Weights were not applied to ParkServe® data as these data do not require weighting.

**ParkServe®: Categorizing race/ethnicity**
City level estimates of park access were also calculated by race/ethnicity. Data reported by ParkServe® included estimates by the following categories: white, black, Asian, 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.

**ParkServe®: Confidence intervals**
CIs are not presented for ParkServe® data.

**ParkServe®: 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.

**ParkServe® data tables**
Preliminary ParkServe® data were shared directly with City Health Dashboard staff under a data use agreement.

**Analysis**
The data from ParkServe® reports numerators (population living ≤ 10-minute walk of a park/green space) for total population and by race/ethnicity. Denominator values were reported for total population, but not for race/ethnicity subgroups. To calculate race/ethnicity park access,
we used population denominators from the American Community Survey table DP05, 2012-2016, for cities (designated as “Places” in American FactFinder). The measure was calculated as follows.

\[
\frac{\text{Population living ≤ 10 minute walk of a park/green space}}{\text{Total population}} \times 100
\]

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

**Walk Score ®**

**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.

**Walk Score ®: Weights**

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

**Walk Score ®: Categorizing race/ethnicity**

Walk Score data are not categorized by race/ethnicity.

**Walk Score ®: Confidence intervals**

CIs are not presented for Walk Score ® data.

**Walk Score: Metric-specific notes**

**Walkability**

**Analysis**

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. Specifically, the html_nodes() and html_table() functions in the rvest package was used to select and extract the data from the tables.

For more detailed information, please refer to the “How Walk Score Works” webpage, available at https://www.redfin.com/how-walk-score-works.
SECTION 5: 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.

   Specifically, for the Violent crime and Primary care physician metrics, the variable labelled “Estimate; SEX AND AGE - Total population” value in ACS table DP05 (2016, 5 Year Estimates) was used to estimate the total population within a city. (Note: This variable is HC01_VC03 in 2016 data, but may change from year to year.) These values were not altered and are used as they are appeared in downloads from American FactFinder. 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. 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- and county-level analyses of total population and sex subgroups, population estimate denominators are generated from 2014 (5 Year Estimates) American Community Survey table B01001.44
   b. For county-level analyses of racial/ethnic subgroups, population estimate denominators are generated from the 2014 National Center for Health Statistics Vintage 2016 Bridged-Race Postcensal Population Estimates data file.44
SECTION 6: Acknowledgements

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 7: References


SECTION 8: 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>
<thead>
<tr>
<th>Age Group</th>
<th>Number</th>
<th>Weight</th>
<th>Weight for YPLL Age-adjustment</th>
<th>Premature Deaths Weight (Years of Life Lost)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>308745538</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5 years</td>
<td>20201362</td>
<td>0.0654</td>
<td>0.0696</td>
<td>72.5</td>
</tr>
<tr>
<td>5 to 9 years</td>
<td>20348657</td>
<td>0.0659</td>
<td>0.0701</td>
<td>67.5</td>
</tr>
<tr>
<td>10 to 14 years</td>
<td>20677194</td>
<td>0.0670</td>
<td>0.0713</td>
<td>62.5</td>
</tr>
<tr>
<td>15 to 19 years</td>
<td>22040343</td>
<td>0.0714</td>
<td>0.0760</td>
<td>57.5</td>
</tr>
<tr>
<td>20 to 24 years</td>
<td>21585999</td>
<td>0.0699</td>
<td>0.0744</td>
<td>52.5</td>
</tr>
<tr>
<td>25 to 29 years</td>
<td>21101849</td>
<td>0.0683</td>
<td>0.0727</td>
<td>47.5</td>
</tr>
<tr>
<td>30 to 34 years</td>
<td>19962099</td>
<td>0.0647</td>
<td>0.0688</td>
<td>42.5</td>
</tr>
<tr>
<td>35 to 44 years</td>
<td>41070606</td>
<td>0.1330</td>
<td>0.1415</td>
<td>35</td>
</tr>
<tr>
<td>45 to 54 years</td>
<td>45006716</td>
<td>0.1458</td>
<td>0.1551</td>
<td>25</td>
</tr>
<tr>
<td>55 to 64 years</td>
<td>36482729</td>
<td>0.1182</td>
<td>0.1257</td>
<td>15</td>
</tr>
<tr>
<td>65 to 74 years</td>
<td>21713429</td>
<td>0.0703</td>
<td>0.0748</td>
<td>5</td>
</tr>
<tr>
<td>75 to 84 years</td>
<td>13061122</td>
<td>0.0423</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>85 years and over</td>
<td>5493433</td>
<td>0.0178</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
### Appendix B: Summary of Dataset of Origin, Censorship Rules, Estimate Provenance and Date of Download (metrics only)

**NOTE:** Symbol key on next page

<table>
<thead>
<tr>
<th>Metric</th>
<th>Dataset of origin</th>
<th>Censorship</th>
<th>Estimate Provenance</th>
<th>Date of Data Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absenteeism</td>
<td>National Center for Education Statistics (NCES) Civil Rights Survey</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>2/2018</td>
</tr>
<tr>
<td>Air pollution - particulate matter</td>
<td>Community Multiscale Air Quality model, US Environmental Protection Agency</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>6/14/17</td>
</tr>
<tr>
<td>Binge drinking</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Breast cancer deaths</td>
<td>MCDD, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>4/25/2018</td>
</tr>
<tr>
<td>Cardiovascular disease deaths</td>
<td>MCDD, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>6/14/17</td>
</tr>
<tr>
<td>Children in poverty</td>
<td>American Community Survey (ACS)**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Colorectal cancer deaths</td>
<td>MCDD, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>4/25/2018</td>
</tr>
<tr>
<td>Dental care</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Diabetes</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>High school graduation</td>
<td>See State-based Education Technical Document (available for download as PDF)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing cost, excessive</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>1/18/2018</td>
</tr>
<tr>
<td>Frequent mental distress</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Frequent physical distress</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Income inequality</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Lead exposure risk index</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Housing with potential lead risk</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Limited access to healthy foods</td>
<td>USDA</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>2/23/18</td>
</tr>
<tr>
<td>Neighborhood racial/ethnic segregation</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>1/18/2018</td>
</tr>
<tr>
<td>Obesity</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Opioid overdose deaths</td>
<td>MCDD, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>4/25/2018</td>
</tr>
<tr>
<td>Park access</td>
<td>ParkServe®</td>
<td>no</td>
<td>Posted as received from ParkServe®</td>
<td>2/20/18</td>
</tr>
<tr>
<td>Physical inactivity</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Premature deaths (all causes)</td>
<td>MCDD, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Prenatal care††</td>
<td>ND, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>5/11/18</td>
</tr>
<tr>
<td>Preventive services</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Primary care physicians</td>
<td>AMA Physician Professional Data</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>4/23/18</td>
</tr>
<tr>
<td>Racial/ethnic diversity</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>1/18/2018</td>
</tr>
<tr>
<td>Smoking</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Teen births</td>
<td>ND, NVSS</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>4/25/2018</td>
</tr>
<tr>
<td>Third-grade reading proficiency</td>
<td>See State-based Education Technical Document (available for download as PDF)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment (City analysis)</td>
<td>ACS**</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>1/18/2018</td>
</tr>
<tr>
<td>Unemployment (Tract analysis)</td>
<td>ACS**</td>
<td>no</td>
<td>Calculated by the Dashboard</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Uninsured</td>
<td>BRFSS-CDC500</td>
<td>no</td>
<td>Posted as received from CDC</td>
<td>12/14/17</td>
</tr>
<tr>
<td>Violent crime</td>
<td>Federal Bureau of Investigation Uniform Crime Reporting</td>
<td>yes</td>
<td>Calculated by the Dashboard</td>
<td>10/20/2017</td>
</tr>
<tr>
<td>Walkability</td>
<td>Walk Score</td>
<td>no</td>
<td>Posted as received from Walk Score®</td>
<td>4/9/2018</td>
</tr>
</tbody>
</table>

Last updated 5/15/18
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 \( \sqrt{MOE^2_{\text{numerator}} - (\hat{p}^2 * MOE^2_{\text{denominator}})} \) because \( MOE^2_{\text{numerator}} - (\hat{p}^2 * MOE^2_{\text{denominator}}) \) 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.5

The 500 Cities: Census Tract Boundaries file (500Cities_Tracts_11082016.zip) was downloaded from the 500 Cities Data Portal.2 The shapefile in the .zip file, 500Cities_Tracts_Clip.shp, was imported into R version 3.4.16 with the rgdal package.7 The field (variable) names were modified; refer to table.

<table>
<thead>
<tr>
<th>Original Variable Names</th>
<th>New Variable Names</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>place2010</td>
<td>PLACE_FIPS</td>
<td>Place FIPS</td>
</tr>
<tr>
<td>tract2010</td>
<td>STATE_COUNTY_TRACT_FIPS</td>
<td>Concatenation of State FIPS, County FIPS, and Tract FIPS</td>
</tr>
<tr>
<td>ST</td>
<td>STATE_FIPS</td>
<td>State FIPS</td>
</tr>
<tr>
<td>PlaceName</td>
<td>PLACE_NAME</td>
<td>Place Name</td>
</tr>
<tr>
<td>plctract10</td>
<td>PLACE_TRACT_FIPS</td>
<td>Concatenation of Place FIPS (hyphen) State FIPS, County FIPS, and Tract FIPS</td>
</tr>
<tr>
<td>PlcTrPop10</td>
<td></td>
<td>Tract Population</td>
</tr>
</tbody>
</table>

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*

```r
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).

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*

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

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)

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*

```r
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)
```

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*

```r
# COUNTY attributes
county_code <- read.table("https://www2.census.gov/geo/docs/reference/codes/files/national_county.txt",
```
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**

```r
# STATE attributes
states_code <- read.table("http://www2.census.gov/geo/docs/reference/state.txt", sep="|", header=T, colClasses=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**

```r
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**

```r
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),]
```

```r
## [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
## <0 rows> (or 0-length row.names)
```

g. **Export data**

Output to .txt file in working directory.

**R Code**

```r
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)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Dataset of origin</th>
<th>Macon, GA: Geography of Reported Value</th>
<th>Honolulu, HI: Geography of Reported Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children in poverty</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Housing cost, excessive</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Income inequality</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Lead exposure risk index</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Housing with potential lead risk</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Neighborhood racial/ethnic segregation</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Racial/ethnic diversity</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Unemployment (City analysis)</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Primary care physicians</td>
<td>AMA Physician Professional Data</td>
<td>Macon (FIPS 13-49000)</td>
<td>Urban Honolulu CDP (FIPS 15-71550)</td>
</tr>
<tr>
<td>Air pollution - particulate matter</td>
<td>American Lung Association</td>
<td>Macon (FIPS 13-49000)</td>
<td>n/a</td>
</tr>
<tr>
<td>Binge drinking</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Dental care</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Frequent mental distress</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Frequent physical distress</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Obesity</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Physical inactivity</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Preventive services</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Smoking</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Uninsured</td>
<td>BRFSS-CDC500</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Violent crime</td>
<td>Federal Bureau of Investigation Uniform Crime Reporting</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Breast cancer deaths</td>
<td>MCDD, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Cardiovascular disease deaths</td>
<td>MCDD, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Colorectal cancer deaths</td>
<td>MCDD, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Opioid overdose deaths</td>
<td>MCDD, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Premature deaths (all causes)</td>
<td>MCDD, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Absenteeism</td>
<td>NCES Civil Rights Survey</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Low birthweight</td>
<td>ND, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Prenatal care</td>
<td>ND, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Teen births</td>
<td>ND, NVSS</td>
<td>Macon (city) and Bibb county - see county_indicator variable</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Park access</td>
<td>Park Serve®</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>High school graduation</td>
<td>See State-based Education Technical Document (available for download as PDF)</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Third-grade reading proficiency</td>
<td>See State-based Education Technical Document (available for download as PDF)</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Limited access to healthy foods</td>
<td>USDA</td>
<td>Macon (FIPS 13-49000)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
<tr>
<td>Walkability</td>
<td>Walk Score®</td>
<td>Macon (FIPS 13-49000)</td>
<td>Urban Honolulu CDP (FIPS 15-71550)</td>
</tr>
<tr>
<td>Demographic information (non-metric)</td>
<td>ACS</td>
<td>Bibb County (FIPS 13-021)</td>
<td>Honolulu County (FIPS 15-003)</td>
</tr>
</tbody>
</table>

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)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Data Source</th>
<th>Hispanic ethnicity is mutually exclusive with racial groups</th>
<th>Definition of Native Hawaiian/ Pacific Islander*</th>
<th>Definition of “Other”</th>
<th>Metric value is available for specific racial/ethnic groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absenteeism</td>
<td>Civil Rights Data Collection</td>
<td>No†</td>
<td>Native Hawaiian or other Pacific Islander†</td>
<td>American Indian or Alaska Native; Two or more races†</td>
<td>Yes</td>
</tr>
<tr>
<td>Breast cancer deaths</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Cardiovascular disease deaths</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Children in poverty</td>
<td>American Community Survey (ACS)</td>
<td>No</td>
<td>Native Hawaiian and Pacific Islander</td>
<td>American Indian or Alaska Native; Two or more races; Some other race</td>
<td>Yes</td>
</tr>
<tr>
<td>Colorectal cancer deaths</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>High school graduation</td>
<td>State-based data</td>
<td></td>
<td></td>
<td>See State-Based Tech Doc</td>
<td>Yes</td>
</tr>
<tr>
<td>Limited access to healthy foods</td>
<td>Food Access Research Atlas, Economic Research Service, United States Department of Agriculture</td>
<td>No</td>
<td>Native Hawaiian or Other Pacific Islander</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Low birthweight</td>
<td>Natality Data, NVSS, NCHS</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Opioid overdose deaths</td>
<td>Multiple Cause of Death Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Premature deaths (all causes)</td>
<td>Multiple Cause of Death Data, NVSS, NCHS</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Prenatal care</td>
<td>Natality Data, National Vital Statistics System (NVSS), National Center for Health Statistics (NCHS)</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Teen births</td>
<td>Natality Data, NVSS, NCHS</td>
<td>Yes</td>
<td>Pacific Islander</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Unemployment</td>
<td>ACS</td>
<td>No</td>
<td>Native Hawaiian and Pacific Islander</td>
<td>American Indian or Alaska Native; Two or more races; Some other race</td>
<td>Yes</td>
</tr>
<tr>
<td>Air pollution - particulate matter</td>
<td>Community Multiscale Air Quality model, US Environmental Protection Agency</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Binge drinking</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Dental care</td>
<td>500 Cities Project Data, Centers for Disease Control and Prevention</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Diabetes</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Frequent mental distress</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Frequent physical distress</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Metric</td>
<td>Data Source</td>
<td>Hispanic ethnicity is mutually exclusive with racial groups</td>
<td>Definition of Asian/Pacific Islander</td>
<td>Definition of “Other”</td>
<td>Metric value is available for specific racial/ethnic groups</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------</td>
<td>-----------------------</td>
<td>------------------------------------------------------------</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Housing cost, excessive</td>
<td>ACS</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Income inequality</td>
<td>ACS</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Lead exposure risk index</td>
<td>ACS</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Housing with potential lead risk</td>
<td>ACS</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Neighborhood racial/ethnic segregation</td>
<td>ACS</td>
<td>Yes – see Technical Document</td>
<td>Native Hawaiian and Pacific Islander</td>
<td>American Indian or Alaska Native; Two or more races; Some other race</td>
<td>No</td>
</tr>
<tr>
<td>Obesity</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Park access</td>
<td>Park Serve®</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Physical inactivity</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Preventive services</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Primary care physicians</td>
<td>American Medical Association Physician Masterfile</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Racial/ethnic diversity</td>
<td>ACS</td>
<td>Yes – see Technical Document</td>
<td>Native Hawaiian and Pacific Islander</td>
<td>American Indian or Alaska Native; Two or more races; Some other race</td>
<td>No</td>
</tr>
<tr>
<td>Smoking</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Third-grade reading proficiency</td>
<td>State-based data</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Uninsured</td>
<td>500 Cities Project Data, CDC</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Violent crime</td>
<td>Uniform Crime Reporting, Federal Bureau of Investigation</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Walkability</td>
<td>Walk Score®</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
</tbody>
</table>

*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