TECHNICAL DOCUMENTATION v1

ABOUT THE DATA & METRIC SELECTION 4
DATA SOURCES 4
GEOGRAPHIC REPRESENTATION 4
YEARS OF DATA AVAILABLE 6
MISSING DATA, DATA CALCULATIONS, METRIC SUB-GROUPS 6
Missing Data 6
Data Calculations 6
Metric sub-groups 6
Sex 6
Race and Ethnicity 6
EXPLANATION OF METRICS 7
Access to Health Services 7
Health Insurance 7
Oral Health 7
Births 7
Chronic Health Conditions 7
Physical Activity 7
Respiratory Disease 8
Cancer 8
Cardiovascular Disease 8
Diabetes and Obesity 9
Disability 9
Dietary Quality 9
Infectious Diseases 10
Respiratory Infection 10
Sexually Transmitted Disease 10
Life Expectancy and Deaths 11
Deaths 11
Life Expectancy at Birth 11
Maternal and Child Health 12
Births 12

Deaths 12
Birth Control 12
Mental Health and Substance Use 12
Mental Health 12
Substance Use 13
Physical and Built Environment 14
Air Quality 14
Food Access 14
Park Access 14
Active Transportation 15
Transportation 15
Poisoning 16
Lead Poisoning 16
Demographics 16
Population 16
Race/Ethnicity 17
Language and Nativity 17
Social and economic factors 17
Housing 17
Education 18
Income 18
Violence and Injury 20
Crime Incidents 20
Deaths 21
School Violence 21
GLOSSARY OF TERMS 22
APPENDICES 23
Appendix 1. List of data sources 23
Appendix 2. Details on data sources. 23
America Community Survey (ACS) 23
Youth Risk Behavior Surveillance System (YRBS) 23
Centers for disease control and prevention (CDC)-Places 24
National Vital Statistics System (NVSS) 24

Mortality Rates.....26
Birth Outcome.....28
Appendix 3. References.....29

ABOUT THE DATA & METRIC SELECTION

As of October 2022, the Big Cities Health Inventory (BCHI) data platform has over 100 metrics. The metrics encompass broad categories of public health importance: Access to Health Services, Chronic Health Conditions, Demographics, Infectious Diseases, Life Expectancy and Deaths, Maternal and Child Health, Mental Health and Substance Use, Physical and Built Environment, Poisoning, Social and Economic Factors, Violence and Injury. These categories were chosen because of their relationship to leading causes of morbidity and mortality in the United States and their role in creating healthier, safer communities.

Platform metrics were selected if they met 'substantive' and 'coverage & standardization' criteria.

Metrics were selected if they met at least 1 of the following 'substantive criteria': 1. public health relevance, in particular their alignment with CDC's Healthy People goals; 2. ability to serve as benchmarks for new policy initiatives; or 3. can be used to highlight demographic and socioeconomic disparities in health and healthy environments. Note that about one-half of the metrics were in BCHC's prior version of the platform (before this new/enhanced platform was developed).

Metrics were selected if they met all of the following 'coverage & standardization criteria': 1. there is a uniform source of the data, thereby ensuring comparability of the metric across jurisdictions; 2. data are available for at least 50% of the BCHC cities; and 3. data are available for recent years.

Additionally, we prioritized publicly available data (to facilitate ease of updating data and to enable data sharing) and city-level data (as opposed to county-level).

DATA SOURCES

Most of the data in the Big Cities Health Inventory come from publicly available data sources for which city data were available. A few of the sources required data use agreements to release data at the city level.

Currently, data for 2010 to 2020 are included. The platform will continue to be updated with additional data as it becomes available. Details about the data sources are given in **Table 2**.

GEOGRAPHIC REPRESENTATION

Effort has been made to provide city level data for the metrics. Some metrics are at the county level due to unavailability of city level data. This is footnoted in the data displays, where applicable. The county proxy used for each city were based on population weighted place to county allocation factor published by Missouri Census Data Center of the University of Missouri[1]. **Table 1** shows all the county proxies used.

Table 1: County proxies used for members when city level data were unavailable

City	County Proxy
Austin, TX	Travis County, TX
Baltimore, MD	City has same boundaries as Baltimore City County, MD
Boston, MA	Suffolk County, MA
Charlotte, NC	Mecklenburg County, NC
Chicago, IL	Cook County, IL
Cleveland, OH	Cuyahoga County, OH
Columbus, OH	Franklin County, OH
Dallas, TX	Dallas County, TX
Denver, CO	City has same boundaries as Denver County, CO
Detroit, MI	Wayne County, MI
Fort Worth, TX	Tarrant County, TX
Houston, TX	Harris County, TX
Indianapolis, IN	Marion County, IN
Kansas City, MO	Jackson County, MO
Las Vegas, NV	Clark County, NV
Long Beach, CA	Los Angeles County, CA
Los Angeles, CA	Los Angeles County, CA
Miami, FL	Miami-Dade County, FL
Minneapolis, MN	Hennepin County, MN
New York City, NY	City has same boundaries as combination of 5 counties: Kings County, Queens County, New York County, Bronx County, Richmond County, NY
Oakland, CA	Alameda County, CA
Philadelphia, PA	City has same boundaries as Philadelphia County, PA
Phoenix, AZ	Maricopa County, AZ
Portland, OR	Multnomah County, OR
San Antonio, TX	Bexar County, TX
San Diego, CA	San Diego County, CA
San Francisco, CA	City has same boundaries as San Francisco County, CA
San Jose, CA	Santa Clara County, CA
Seattle, WA	King County, WA
Washington, DC	The city is synonymous with District of Columbia County, DC

YEARS OF DATA AVAILABLE

The years for which data were available for the metrics from each source is given in Appendix 1. Where missing, data were imputed from latest available year to facilitate visualization and comparison across different metrics.

MISSING DATA, DATA CALCULATIONS, METRIC SUB-GROUPS

MISSING DATA

All metrics are not available for all cities. Further, for cities that have data for a metric, those data may not exist for each year. Not all metrics have data for subgroups because the data were not reported or in cases where denominators or incidences were too small, certain rates for subpopulations are not displayed. This is notated either as an empty column in the graph, a label on the graph (data unavailable, "N/A"), or the pull-down menu category is grey and not able to be selected.

DATA CALCULATIONS

As is customary, communicable disease indicators are reported using crude rates. Mortality rates are age-adjusted to compare relative mortality risks among cities, different demographic groups, and over time. Standardization was done using the 2000 U.S. standard million population. All mortality rates are presented per 100,000 people exception is mortality from police violence. This is reported as crude rate per 1,000,000 people. Metric names and subtitles reflect differences regarding age-adjustment and/or crude rates.

METRIC SUB-GROUPS

Sex

Where available, data for sex categories are reported for male and female.

Race and Ethnicity

Categorization of race and ethnicity were limited by the data source. For most metrics, the default options were Hispanic, White non-Hispanic, Black non-Hispanic, and Asian non-Hispanic. For the metrics which were sourced from the census, race is classified as Hispanic, White non-Hispanic, Black Hispanic and non-Hispanic, Asian Hispanic and non-Hispanic. For the metrics which were sourced from the National Vital Statistics System (NVSS), race is classified as Hispanic, White non-Hispanic, Black Hispanic and non-Hispanic, Asian/Pacific Islander Hispanic and non-Hispanic.

EXPLANATION OF METRICS

Each of the metrics and associated methodology are further explained below.

ACCESS TO HEALTH SERVICES

Health Insurance

Uninsured, All Ages and **Uninsured, Child** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated for all civilian non-institutionalized population from table B27010. Uninsured, All Ages is population of all ages who don't have any private or public insurance and the denominator is total population. Uninsured, Child, is the population less than 18 who don't have any private or public insurance and the denominator is population who are less than 18.

Oral Health

Dental Care was reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. The numerator is the respondents aged 18+ who visited dentist or dental clinic in the previous year and the denominator were the total respondents. For more details, please see Appendix 2.

Births

Prenatal Care was calculated from National Vital Statistics System (NVSS) Natality files. For details, please refer to Appendix 2. The numerator is the number of live births where prenatal care began between the first and third month of pregnancy and denominator is the total number of live births.

CHRONIC HEALTH CONDITIONS

Physical Activity

Adult Physical Inactivity was reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. The BRFSS question used is "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?". The numerator is respondents who answered 'no' to the question and denominator is all respondents. For more details, please see Appendix 2.

Teen Physical Inactivity, Teen Physical Activity Levels, Teen Physical Education, Teen Computer Time, and Teen TV Time were reported as received from the YRBS. Teen Physical Inactivity is the percent of respondents who were not physically active for at least 60 minutes even 1 day in the 7 days before the survey. Teen Physical Activity Levels is the percent of respondents who did not participate in at least 60 minutes of physical activity per day for at least 5 days in the 7 days before the survey i.e., did not meet the physical activity recommendations set by the CDC. Teen Physical Education is the percent of high school students who did not attend PE class 1 or more days in in an average week when they were in school. Teen Computer Time is the percent of respondents who played video or computer games or used a computer three or more hours per day on an average school day. Teen TV Time is the percent of high school students who watched television for 3 or more hours per day on an average school day. The denominator for all metrics is all respondents. For more details, please see Appendix 2.

Respiratory Disease

Teen Asthma was reported as received from the YRBS and is the percent of respondents who were ever told by a doctor or nurse that they had asthma. For more details, please see Appendix 2.

Cancer

All Cancer Deaths, Lung Cancer Deaths and **Breast Cancer Deaths** were calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes used are given in **Table 6**.

Cardiovascular Disease

Cardiovascular Disease Deaths and **Heart Disease Deaths** were calculated from the NVSS Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

High Blood Pressure prevalence was reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. The numerator is 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 they have high blood pressure only during pregnancy and those who were told they had borderline hypertension were not included. The denominator is all respondents. For more details, please see Appendix 2.

Blood Pressure Medication was reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates

(CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. The numerator is respondents aged ≥18 years who reported taking medicine for high blood pressure. The denominator is respondents aged ≥18 years who report having been told by a doctor, nurse, or other health professional of having high blood pressure other than during pregnancy (excluding those who refused to answer, had a missing answer, or answered "don't know/not sure"). For more details, please see Appendix 2.

Diabetes and Obesity

Diabetes Deaths were calculated from the NVSS Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Diabetes prevalence was reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. The percentage of diagnosed diabetes were number of respondents who reported ever been told by a doctor, nurse, or other health professional that they have diabetes. Denominator is all respondents aged ≥18 years who report or do not report ever been told by a doctor, nurse, or other health professional that they have diabetes. Diabetes during pregnancy was excluded. Crude prevalence is reported.

Adult Obesity was reported as reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. It is the percentage of the population 18 years or over who had body mass index (BMI) of 30 or above. The BMI was derived from self-reported height and weight. Respondents who reported being pregnant, extremes of height (<3 ft or ≥8 ft), extremes of weight (<50 lbs or ≥650 lbs) or if height or weight data are missing were removed from the analysis.

Teen Obesity was reported as received from the YRBS. Obesity for children is defined as BMI at or above the 95th percentile of children of the same age or sex. The denominator is all respondents. For more details, please see Appendix 2.

Disability

People with Disabilities was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B18101. The numerator is population of all ages who reported having any long-term disability and the denominator is all civilian noninstitutionalized population.

Dietary Quality

Teen Soda and **Teen Breakfast** data are reported as received from the YRBS. Teen Soda is the percent who responded yes to the YRBS question "Whether they drank a can, bottle, or glass of soda or pop such as Coke, Pepsi, or Sprite, not counting diet soda or diet pop one or more times per day, during the 7 days before the survey". Teen Breakfast is percent who reported eating no breakfast during the 7 days before the survey. The denominator is all respondents. For more details, please see Appendix 2.

INFECTIOUS DISEASES

Respiratory Infection

Pneumonia or Influenza Deaths was calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

New Tuberculosis Cases was reported from the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention Atlas Plus of the CDC**.** Tuberculosis Cases which met the clinical case classification or laboratory confirmed are counted. Crude rates per 100,000 population were calculated for adults aged 18+.

Flu Vaccine, Adult was reported as received from the Centers for Medicare and Medicaid Services (CMMS), of the U.S. Department of Health and Human Services. The denominator represents all Medicare beneficiaries enrolled in the fee-for-service (FFS) program. The numerator represents the subset of beneficiaries who received a flu vaccine in the year specified; CMMS derived this from administrative claims for flu vaccination.

COVID-19 Deaths was calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Sexually Transmitted Disease

HIV/AIDS Prevalence, New Chlamydia Cases, Syphilis Prevalence (primary & secondary syphilis), **Syphilis in Newborns** (congenital syphilis), and **New Gonorrhea Cases** were reported from the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention Atlas Plus of the CDC. Crude rates per 100,000 population were calculated for each STD (except congenital syphilis). Each rate was calculated by dividing the total incidence or prevalence for the calendar year by the population for that calendar year and then multiplying the result by 100,000. Congenital syphilis is reported per 100,000 live births.

HIV-Related Deaths was calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

LIFE EXPECTANCY AND DEATHS

Deaths

Deaths from All Causes was calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Premature Death (age < 75) was calculated using the Multiple Causes of Mortality restricted-use micro-data files from the National Vital Statistics System (NVSS). Years of potential life lost from all causes of death is derived using Dranger and Remington's[2] method and is age-adjusted and reported per 100,000 population. Weights for age-adjusting premature deaths were calculated using the US 2000 standardized population but were adjusted to include only the population aged 74 and younger. For details, please see Appendix 2.

Life Expectancy at Birth

Life Expectancy was calculated as the life expectancy from birth using the Multiple Causes of Mortality restricted-use micro-data files from the National Vital Statistics System (NVSS). Life expectancies were calculated as five-year moving averages (**Table 3**, shows how the data were aggregated for each period). The death incidences were calculated for all members at city level except for Las Vegas, NV and Louisville, KY where county level data were used. The population counts were obtained from ACS-1-year estimates. Abridged life tables and life expectancies were calculated using R Software Demo Tools package. To overcome the problem of unavailability of the population data for under 1-year old by race and sex strata, population "graduation" into single age groups was done for race-sex specific population using the formula by Sprague[3] and implemented using R package Demo Tools. Life expectancy estimates based on less than 700 deaths were suppressed.

Table 3: The years aggregated for the numerators and the denominators for Life Expectancy

"year" label	Numerator: Deaths					Denominator: Population				
2014	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
2015	2011	2012	2013	2014	2015	2011	2012	2013	2014	2015
2016	2012	2013	2014	2015	2016	2012	2013	2014	2015	2016
2017	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017
2018	2014	2015	2016	2017	2018	2014	2015	2016	2017	2018
2019	2015	2016	2017	2018	2019	2015	2016	2017	2018	2019
2020	2016	2017	2018	2019	2020	2016	2017	2018	2019	2019*

2019*- Population estimates in the 2020 ACS-1-year surveys were unreliable. 2019 ACS-1-year population estimates were used as proxy in accordance with the recommendations by the Census Bureau.

MATERNAL AND CHILD HEALTH

Births

Teen Births and **Low Birthweight** were calculated from National Vital Statistics System (NVSS) Natality files. For details, please refer to Appendix 2. For the teen birth rate, the numerator is the number of live births to mothers aged 15 to 19 years and denominator is the total female population aged 15 to 19 from the American Community Survey 1-year population estimates. Low birthweight is defined as the percentage of babies born under 2,500 grams. The denominator is the total number of live births.

Deaths

Infant Deaths and **Maternal Deaths** were calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Birth Control

Teen Birth Control were reported as received from the YRBS and is percent of teens who did not use any method to prevent pregnancy during last sexual intercourse among students who were currently sexually active. For more details, please see Appendix 2.

MENTAL HEALTH AND SUBSTANCE USE

Mental Health

Adult Mental Distress is reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-

PLACES) based primarily on the Behavioral Risk Factor Surveillance System. The numerator is respondents aged ≥18 years who report their mental health being "not good" for 14 or more days during the past 30 days. The denominator is all respondents aged ≥18 years. For more details, please see Appendix 2.

Suicide mortality rates were calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Electronic Bullying, **School Bullying**, **Teen Mental Distress** and **Teen Suicidal Ideation** were reported as received from the YRBS. Electronic Bullying is the percent of high school students who recounted being bullied through texting, Instagram, Facebook, or other social media, during the 12 months before the survey. School Bullying is the percent of high school students who recounted being bullied on school property in the past year. Teen Mental Distress is the percent of high school students who felt sad or hopeless almost every day for 2 weeks or more in a row so that they stopped doing some usual activities, during the 12 months before the survey. Teen Suicidal Ideation is the percent of high school students who planned about how they would attempt suicide during the 12 months before the survey. The denominator for all metrics were all respondents. For more details, please see Appendix 2.

Substance Use

Opioid Overdose Deaths were calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Adult Binge Drinking and **Adult Smoking** were reported as received from the Centers for Disease Control and Prevention's Population Level Analysis and Community Estimates (CDC-PLACES) based primarily on the Behavioral Risk Factor Surveillance System. Percentage of adults who binge drank is based on the Behavioral Risk Factor Surveillance System question about how many drinks a person had on one occasion in the past 30 days. Women who answered "four" and men who answered "five" are considered binge drinkers. Adult Smoking is percent of respondents aged ≥18 years who report having smoked ≥100 cigarettes in their lifetime and currently smoke every day or some days. Respondents who refused to answer were excluded from the analysis. For more details, please see Appendix 2.

Teen Alcohol, **Teen Smoking**, and **Teen Marijuana** were reported as received from the YRBS. Teen Alcohol is the percent of high school students who currently drank at least one drink of alcohol, on at least 1 day during the 30 days before the survey. Teen Smoking is the percent of high school students who currently reported smoking cigarettes on at least 1 day during the 30 days before the survey. Teen Marijuana is the percent of high school students who currently used marijuana also called grass, pot, or weed,

one or more times during the 30 days before the survey. The denominator for all metrics were all respondents. For more details, please see Appendix 2.

PHYSICAL AND BUILT ENVIRONMENT

Air Quality

Poor Air Quality and **Hazardous Air Quality** are calculated from the Air Quality Index (calculated for four major air pollutants regulated by the Clean Air Act: ground level ozone, particle pollution, carbon monoxide, and sulfur dioxide) value published by the Environmental Protection Agency. According to EPA classification, days with AQI >50 may be of health concern for sensitive groups and days with AQI >100 are considered unhealthy for all people. The numerator is the number of days when AQI was >50 or >100 and the denominator is number of days for which AQI is recorded.

Food Access

Limited Supermarket Access is calculated from indicators in the Food Access Research Atlas of the U.S. Department of Agriculture. Food Access Research Atlas includes several indicators to measure food access at the census tract level. The data were aggregated to county level. For limited healthy supermarket access, the numerator is the number of low-income people who did not live within half mile of supermarket and the denominator is the total number of low-income people in the county. Low-income is defined as people with income below 200% of Federal Poverty level.

Park Access

Green Space Access, Investment in Parks and **City Park System** as measured by ParkScore® were reported as received from the ParkScore database of the Trust for Public Land. The Trust for Public Land developed the ParkScore® index for the 100 largest cities in the US, to measure how well cities are meeting the need for parks. Cities are awarded points based on analysis of four important characteristics of an effective park system: acreage, investment, amenities, and access. Green Space Access is percent of residents living within a 10-minute walk of green spaces which include: publicly owned local, state, and national parks, trails, and open space; school parks with a joint-use agreement with the local government; and privately owned parks that are managed for full public use. Investment in Park is based on total spending per resident and is the sum of all public spending, nonprofit spending, and volunteer hours. ParkScore® index is the most comprehensive rating system ever developed to measure how well the 100 largest U.S. cities are meeting the need for parks. It is scaled from 0 – 100 with higher score representing better performance.

Active Transportation

Bikeability and **Walkability** were reported as received from the Redfin Corporation. Bike Score® is a patented measure of whether an area is good for biking. For a given location, a Bike Score is calculated by measuring bike infrastructure (lanes, trails, etc.), hills, destinations and road connectivity, and the number of bike commuters. Walk Score® measures the walkability of any address using a patented system. For each address, Walk Score® analyzes hundreds of walking routes to nearby amenities. Points are awarded based on the distance to amenities in each category. Walk Score® also measures pedestrian friendliness by analyzing population density and road metrics such as block length and intersection density

Walking to Work and **Riding Bike to Work** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B08137. The numerator is who reported walking or riding a bicycle to work respectively. The denominator is all workers 16+ years of age who did not work at home.

Transportation

Public Transit Access were reported as received from the Redfin Corporation. Transit Score® is a patented measure of how well a location is served by public transit. Transit Score is based on data released in a standard format by public transit agencies. To calculate a Transit Score®, a "usefulness" value to nearby transit routes based on the frequency, type of route (rail, bus, etc.), and distance to the nearest stop on the route is assigned. The "usefulness" of all nearby routes is summed and normalized to a score between 0-100.

Lack of Car were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B25044. The numerator is number of housing units (both owner and renter occupied) with no vehicles available. The denominator is all occupied housing units.

Public Transportation Use and **Drives Alone to Work** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B08137. The numerator is who reported using public transport, or those who drive to work alone in a car truck, or van. The denominator is all workers 16+ years of age who did not work at home.

Longer Driving Commute Time was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B08012. The numerator is those who reported commute time to work to be more than 30 minutes. The denominator is all workers 16+ years of age who did not work at home.

POISONING

Lead Poisoning

Child Lead Levels >5 mcg and **Child Lead Levels >10mcg** data for most states were collected from the Childhood Lead Poisoning Prevention Program of the CDC. The data for cities in Texas, California, and New York City were requested from the Childhood Lead Prevention Program division of the Department of Health. The estimates from these data sources are valid for comparison with data collected from the CDC. Two separate cut-offs were used; 5 ug/dl, the current CDC standard and 10 ug/dl, which was the previous standard used to define children who had unacceptably high blood lead levels. The numerator is the number of children with elevated blood levels and the denominator is the number of children under 6 who were screened.

Child Lead Testing is the number of children under 6 out of whole population under 6 who were tested for elevated blood lead level. The population data is obtained from the annual intercensal estimates for the most recent U.S. Census data. This metric is included to aid in the interpretation of the result of elevated blood lead level since screening is not mandated for children under 6 in all cities. The volume of children screened affects the percent of children detected with elevated blood lead level.

Housing Lead Risk were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B25034. The numerator is the housing units built before 1950. The denominator is all housing units.

DEMOGRAPHICS

Population

Single-parent Families was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B11005. The numerator is the number of households with a child below 18 with female only or male only householder (either family or non-family households). The denominator is all households.

Population Density per square mile was calculated using the weighted population from the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables (table B01003) and the land area is acquired from the gazetteer files published by the Census Bureau. The numerator is the total population, and the denominator is the land area.

Seniors and **Children** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated for total population.

These are calculated using table B01001 which gives the population counts by sex and age groups and reported as percent.

Race/Ethnicity

Minority Population were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B03002. This was calculated by subtracting the percentage of people who reported being non-Hispanic white from 100. The denominator is total population.

Language and Nativity

Primarily Speak English, Primarily Speak Chinese, and **Primarily Speak Spanish** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table C16001. The numerator are the people who reported speaking English, Spanish or Chinese primarily at home respectively. The denominator is population 5 years and older.

Foreign Born Population was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B05003. The numerator is all people who reported being born in a foreign country. The denominator is total population.

SOCIAL AND ECONOMIC FACTORS

Housing

Severe Housing Problems was calculated using data from the Comprehensive Housing Affordability Strategy of the Department of Housing and Urban development. This is measured for all occupied housing units. It is the percent of households with severe housing problem, namely if they have one or more of these four problems: lacks complete kitchen facilities, lacks complete plumbing facilities, overcrowding, and excess cost burden.

Vacant Housing Units, **Owner Occupied Housing**, **Renters vs. Owners** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated for housing units. Vacant Housing Units is the percent of total housing units that are vacant (table B25002). Owner Occupied Housing is the percent of occupied housing units that are occupied by the owner (table B25003). Renters vs. Owners is the ratio of renter occupied to owner occupied housing units among all occupied housing units (table B25003).

Homelessness metrics were calculated using data from the Point-in-Time (PIT) survey conducted by the Continuum of Care (CoC) Homeless Assistance Programs of the U.S. Department of Housing and Urban Development (HUD). The PIT is an unduplicated 1-night estimate (usually done in January) of both 'sheltered' and 'unsheltered' homeless populations, conducted according to HUD protocols.

'Sheltered' refers to people who are staying in emergency shelters or transitional housing programs. 'Unsheltered' refers to are people whose primary nighttime residence is a public or private place not designated for, or ordinarily used as, a regular sleeping accommodation for people (for example, the streets, vehicles, or parks).

For the metric "Homelessness, Non-Whites", the numerator is total who reported their race to be Black or African American, Asian, American Indian, or Alaska Native, Native Hawaiian or Other Pacific Islander, and Multiple Races. The denominator is total unhoused population. For the metric "Homelessness, Children" the numerator is all unhoused population who were less than 18 years old. The denominator is total unhoused population. For the metric "Homelessness and Vacant Housing", the numerator is all unhoused population. The denominator is total number of vacant rental units (table B25004) from the American Community Survey (ACS) 5-year estimate.

Education

Preschool Enrollment was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated for population over 3 years of age from table B14003. The numerator is 3- and 4-year-olds enrolled in school (both public and private schools), the denominator is the population of 3- and 4-year-old age group.

College Graduates were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B15002. The numerator is people who reported having educational attainment of bachelor's degree or higher. The denominator is population over 25 years of age.

Income

Poverty in All Ages and **Poverty in Children** were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated for population for whom the poverty status was determined from table B17001. For Poverty in All Ages, the numerator is population of all ages who reported family income for the past 12 months to be below the federal poverty level. The denominator is total population for whom the poverty status was determined. Poverty in Children, the numerator is population below 18 years of age who reported family income in the past 12 months to be below poverty level and the denominator is population below 18 years for whom the poverty status was determined.

Poverty and Near Poverty in All Ages was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated for population for whom the poverty status was determined (table C17002). The numerator is all whose ratio of income to federal poverty level was less than 200%.

Per-capita Household Income was reported from table B19301 of the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate. This was reported for total population.

Households with Higher-Incomes were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B19001. The numerator is households whose reported income in past year was greater than $50,000. The denominator is all households.

Income-related

Public Assistance was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B19058. The numerator is households who received cash public assistance or Food Stamps/SNAP. The denominator is all households.

Unemployment was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table B23001. The numerator is people in civilian workforce who reported being currently unemployed. The denominator is civilians 16+ years of age.

Service Workers was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from table C24010. The numerator is people who reported being employed in service or labor occupations. The denominator is civilians in workforce 16+ years of age.

Excessive Housing Cost was calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate tables and is calculated from tables B25074 and B25091. The numerator is all renter who reported the rent, or owner-occupied housing units who reported the mortgage to be >35% of the household income. The denominator is all housing units

Income Inequality

Household Income Inequality were calculated using the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate and is calculated from table B19001. The numerator are households whose annual income in < $25,000 or > $100,000. The denominator is all households.

Income Inequality expressed as GINI index was reported from table B19083 of the U.S. Census Bureau's American Community Survey (ACS) 5-year estimate. This was reported for all households.

RACIAL SEGREGATION INDICES

Racial Segregation, White and non-White; Racial Segregation, White and Black; Racial Segregation, White and Asian; and **Racial Segregation, White and Hispanic**

were assessed using the index of dissimilarity. The metric reflects racial segregation across neighborhoods (census tracts) in the city (census places).

The source data are tract-level demographic data from the 5-year American Community Survey (ACS). Tracts were assigned to cities (census places) based on the population allocation factor published by Missouri Census Data Center of the University of Missouri[1]. ACS surveys used were for the time periods 2010-2014, and 2015-2019.

We chose to show segregation using the index of dissimilarity because it among the most widely reported segregation metrics. The index reflects the evenness with which two groups (for example, Black and white residents) are distributed across the component geographic areas (census tracts) within a larger area (census places). The index ranges from 0 (complete integration) to 100 (complete segregation). The index is smallest when majority and minority populations are evenly distributed.

Generally, segregation of 50% or higher is considered 'high' segregation.

We include racial segregation in our platform because it is widely considered a fundamental cause of racial disparities in health in the United States. As described by Williams et al, and others, poor health is more common in places that are segregated from socioeconomic mobility and where there is concentrated exposure to social and physical hazards that harm health. CDC's Healthy People 2030 cites residential segregation as an example of how structural discrimination harms health (see their Social and Community Context domain).

As has been explained in numerous publications (academic and other publications), the persistence of high racial residential segregation rates across US cities is due to a long history and continuation of tactics that selectively restrict occupancy in neighborhoods reserved for whites. Tactics from the past and present include restricted covenants, redlining communities of color, intimidation, violence, and selective use of law enforcement and mass incarceration.

VIOLENCE AND INJURY

Crime Incidents

Violent Crime rates are reported as received from the Uniform crime reporting Program of the Federal Bureau of Investigation (F.B.I.). Under the program all city law enforcement agencies submit offense data to the FBI. Violent crime incidences included are murder, aggravated assault, robbery, and rape. The population estimate was computed by the FBI based on the 2010 decennial population counts and the rate of growth for the individual city. Crude rate per 100,000 population is reported.

Homicides mortality rates were calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Deaths

Injury Deaths was calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Firearm Deaths was calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

Police Killings and **Racial Disparity in Police Killings** are reported from the Mapping Police Violence database which records information on killings by police nationwide since 2013. This information has been sourced both from official police reports in some states (e.g., California, Texas etc.) and from the Fatal Encounters database, a nationwide impartial crowdsourced database on police killings. Police Killing is defined as a case where a person dies as a result of being shot, beaten, restrained, intentionally hit by a police vehicle, pepper sprayed, tasered, or otherwise harmed by police officers, whether on-duty or off-duty. Only killings where the person was unarmed has been considered. The numerator is the sum of all killings during the period and the denominator is the American Community Survey 5-year population estimates for 2015-2019. Following how the source dataset defines and labels these data, we refer to the

The rate is reported as crude annual rate per 1,000,000 population. The population estimates used were from the American Community Survey 5-year population estimates for 2015-2019. Racial Disparity in Police Killing is calculated as the difference in annual police killing rate between Black and White races.

Motor Vehicle Deaths mortality rates were calculated from the National Vital Statistics System (NVSS) Multiple Cause of Death data files. For details, please see Appendix 2 and detailed ICD 10 codes are given in **Table 6**.

School Violence

Weapons in School was reported as received from the YRBS. It is the percentage of high school students who carried a weapon such as a gun, knife, or club, on at least 1 day during the 30 days before the survey. The denominator is all respondents. For more details, please see Appendix 2.

Fighting in School was reported as received from the YRBS. It is the percentage of high school students who were involved in a physical fight one or more times during the 12 months before the survey. The denominator is all respondents. For more details, please see Appendix 2.

GLOSSARY OF TERMS

The following terms are defined by the CDC's Principles of Epidemiology in Public Health Practice, Third Edition: An Introduction to Applied Epidemiology and Biostatistics and the United States Cancer Statistics.

Rate – A rate is a measure of the frequency with which an event occurs in a defined population over a specified period of time.

Crude rate – The total number of cases of a particular disease or condition over the total population size for a given period of time. As the crude rate is influenced by the underlying age distribution of the state's population, cities will often report an age-adjusted rate.

Age-adjusted rate – Uses a standard population, generally the 2000 U.S. standard population, that is based on that year's population age groups. Using direct standardization, these populations by age group serve as weights for calculating the age-adjusted rate. This ensures differences in rates are not due to different age distributions of the populations.

APPENDICES

Appendix 1. List of data sources

Table 2: Data aggregation and period for which data is available from each data source.

Source	Multiple/ Single year	Time Period
American Community Survey, U.S. Census Bureau	Multiple	2010-2014, 2015-2019
Bike Score®, Redfin Corporation	Single	2020
Childhood Lead Poisoning Prevention Program, CDC	Single	2012 to 2017~
Economic Research Service, U.S. Department of Agriculture	Multiple	2015, 2019
Mapping Police Violence	Multiple	2013-2020
National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention AtlasPlus, CDC	Single	2010 to 2019
National Vital Statistics System (NVSS), CDC	Multiple	2010 to 2020
ParkScore®, The Trust for Public Land	Single	2015 to 2020
Population Level Analysis and Community Estimates (CDC-PLACES)	Multiple	2014 to 2019
Transit Score®, Redfin Corporation	Single	2020
U.S. Environmental Protection Agency	Single	2010 to 2020
Uniform Crime Reporting, FBI	Single	2010 to 2020~
Walk Score®, Redfin Corporation	Single	2020
Youth Risk Behavior Surveillance System, CDC	Single	2011, 2013, 2015, 2017, 2019~
~ available for most cities		

Appendix 2. Details on data sources.

AMERICA COMMUNITY SURVEY (ACS)

American Community Survey (ACS) 5-year estimate summary files were summarized at "summary level 160" for Places and analytic variables were created. Margins of error were not taken into account in the calculations. Places FIPS codes which is unique for city (7 digit- STATE FIPS+ Place FIPS) were used to isolate the data for the jurisdictions. Two non-overlapping 5-year surveys were used. The data for years 2010 to 2014 are from ACS 5-year survey 2010-2014. The data for years 2015 to 2019 are from ACS 5-year survey 2015-2019.

YOUTH RISK BEHAVIOR SURVEILLANCE SYSTEM (YRBS)

The Youth Risk Behavior Surveillance System (YRBS) prioritizes and monitors six health-risk behaviors among youths and young adults: (1) behaviors that contribute to unintentional injuries and violence; (2) sexual behaviors that contribute to human immunodeficiency virus (HIV) infection, other sexually transmitted diseases, and unintended pregnancy; (3) tobacco use; (4) alcohol and other drug use; (5) unhealthy dietary behaviors; (6) physical

inactivity. Data sources of YRBS include ongoing surveys as well as one-time national surveys, special population surveys, and methods studies. More specifically, YRBS data is comprised of a national school-based survey conducted by the Centers for Disease Control and Prevention (CDC). Additionally, education and health agencies conduct school-based state, territorial and freely associated states, tribal, and large urban school district surveys. All surveys are conducted biennially and include samples of students in grades 9-12. YRBS employs a two-stage, cluster sample design to produce representative samples of students in grades 9-12 for all state, territorial and freely associated states, tribal, and large urban school districts. A weight is applied to all student records to adjust for student nonresponse rates and the distribution of students by grade, sex, and race/ethnicity in each jurisdiction. The National Survey distributed by the CDC utilizes a three-stage, cluster sample design to obtain a nationally representative sample of the target population: students in grades 9-12 in the 50 states and the District of Columbia. U.S. territories are excluded from the sampling frame. Overall estimates and estimates for sex, grade, race/ethnicity, grade by sex, and race/ethnicity by sex subgroups meet the standard of producing estimates within ±5% at a 95% confidence interval level. Grade by race/ethnicity subgroups are accurate within ±5% at a 90% confidence interval level. For all surveys, students who refuse to participate are not replaced.

CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC)-PLACES

Small area estimation (SAE) of measures is generated through a multilevel statistical modeling framework. The CDC links geocoded health surveys and high spatial resolution population demographic and socioeconomic data through an innovative peer-reviewed multi-level regression and poststratification (MRP) approach. The MRP approach accounts for associations between individual health outcomes and characteristics, spatial contexts, and factors at multiple levels (i.e., state and county). Moreover, MRP predicts individual disease risk and health behaviors in a multi-level modeling framework in addition to estimating the geographic distributions of population disease burden and health behaviors. The MRP approach provides flexibility and produces modeled estimates of the prevalence for each indicator at the census tract and city level. Both internal and external validation studies conducted by the CDC confirm the strong consistency between MRP model-based SAE's and direct Behavior Risk Factor Surveillance System survey estimates at county and state levels. Primary data sources for this project include CDC Behavioral Risk Factor Surveillance System, the Census 2010 population, and the American Community Survey estimates.

NATIONAL VITAL STATISTICS SYSTEM (NVSS)

General Notes:

Restricted-use micro-data files for Natality and Multiple Causes of Mortality from the NVSS were used for calculation of these metrics. These files have city FIPS codes identifier for residents of cities with population size ≥ 100,000.

Geography: Deaths are assigned to the reported city of residence of the deceased (based on city FIPS code). Births are assigned to the city of residence reported by the mother.

Rates are reported for all cities except for Las Vegas, NV and Louisville, KY. For these cities, the data are reported at county level for all metrics (except for Infant Deaths, Maternal Deaths, Prenatal Care and Low Birth Weight).

Categorizing race/ethnicity: Race/ethnicity categories selected were limited by the race categorization available from the NVSS data and the population data from the Census. Race classifications used are Hispanic, White non-Hispanic, Black Hispanic and non-Hispanic, Asian/Pacific Islander Hispanic and non-Hispanic. For mortality, the race is classified according to the race of the deceased recorded in the death certificate. For the natality measures the race is classified according to the race of the mother.

Years: The rates were calculated as three-year moving averages (except 2010 and 2020 where two-year averages are used) for all metrics except for infant mortality and maternal mortality. The middle year is used as the representative year.

The **Table 4** & **5** shows the years aggregated for the numerators and the denominators.

Table 4: The years aggregated for the numerators and the denominators for mortality metrics.

"year" label	Numerator: Deaths or Births			Denominator: Population/Live Births		
2010	--	2010	2011	--	2010	2011
2011	2010	2011	2012	2010	2011	2012
2012	2011	2012	2013	2011	2012	2013
2013	2012	2013	2014	2012	2013	2014
2014	2013	2014	2015	2013	2014	2015
2015	2014	2015	2016	2014	2015	2016
2016	2015	2016	2017	2015	2016	2017
2017	2016	2017	2018	2016	2017	2018
2018	2017	2018	2019	2017	2018	2019
2019	2018	2019	2020	2018	2019	2019*
2020	2019	2020	--	2019	2019*	--

2019*- Population estimates in the 2020 ACS-1-year surveys were unreliable. 2019 ACS-1-year population estimates were used as proxy in accordance with the recommendations by the Census Bureau.

For infant deaths, rates were calculated as five-year moving averages. The years aggregated are as follows.

Table 5: The years aggregated for the numerators and the denominators for infant mortality.

"year" label	Numerator: Deaths					Denominator: Live Births				
	Mortality files years used					Natality files years used				
2014	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
2015	2011	2012	2013	2014	2015	2011	2012	2013	2014	2015
2016	2012	2013	2014	2015	2016	2012	2013	2014	2015	2016
2017	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017
2018	2014	2015	2016	2017	2018	2014	2015	2016	2017	2018
2019	2015	2016	2017	2018	2019	2015	2016	2017	2018	2019
2020	2016	2017	2018	2019	2020	2016	2017	2018	2019	2020

For **maternal deaths**, due to the concern of small number of death events, maternal mortality was calculated for the whole 11-year period 2010-2020.

Censoring Values: In accordance with the data use agreement with the National Center for Health Statistics (NCHS), if total count of events were less than 10 in any category, the values are suppressed. Data for some race strata were also suppressed based on internal criteria.

Prenatal care was censored for Phoenix, AZ (for years 2010 to 2014), Boston, MA (for years 2010 and 2011), Charlotte, NC (for years 2010 and 2011), Minneapolis, MN (for years 2010 and 2011) because they did not report prenatal care in these years.

Mortality Rates

Multiple Causes of Mortality restricted-use micro-data files from the NVSS were used for calculation of mortality rates.

Causes of death were classified in accordance with World Health Organization (WHO) regulations based on current revision of the International Statistical Classification of Diseases and Related Health Problems (ICD 10)[4]. The table of ICD 10 codes for different causes of mortality are given in **Table 6**.

Table 6: List of ICD 10 codes used for Mortality

Metrics	ICD-10 Codes
All Cancer Deaths	C00-C97
Breast Cancer Deaths (calculated for females)	C50
Lung Cancer Deaths	C34
Cardiovascular Disease Deaths	I00-I78
Heart Disease Deaths	I00-I09, I11, I13, I20-I51
Diabetes Deaths	E10-E14
Pneumonia or Influenza Deaths	J09-J18
HIV-Related Deaths	B20-B24

Metrics	ICD-10 Codes
Deaths from All Causes	All ICD-10 codes included
Drug Overdose Deaths	F11.0-F11.5, F11.7-F11.9, F12.0-F12.5, F12.7-F12.9, F13.0-F13.5, F13.7-F13.9, F14.0-F14.5, F14.7-F14.9, F15.0-F15.5, F15.7-F15.9, F16.0-F16.5, F16.7-F16.9, F17.0, F17.3-F17.5, F17.7-F17.9, F18.0-F18.5, F18.7-F18.9, F19.0-F19.5, F19.7-F19.9, X40-X44, X60-X65, X85,Y10-Y14
Opioid Overdose Deaths	Opioid overdose deaths are identified by the presence of any of the following Multiple Cause of Death (MCOD) codes in addition to the Underlying Cause of Death (UCD) codes X40-X44, X60-X65, X85, Y10-Y14: T40.0(opium); T40.1(heroin); T40.2 (natural opioid analgesics); T40.3 (methadone); T40.4 (other synthetic); or T40.6 (other and unspecified narcotics).
Suicide	X60-X84 and Y87
Homicides	X85-Y09, Y87.1
Injury Deaths	V01-X59, Y85-Y86
Firearm-Related Deaths	W32-W34, X72-X74, X93-X95, Y22-Y24, Y35.0
Motor Vehicle Deaths	V02-V04, V09.0, V09.2, V12-V14, V19.0-V19.2, V19.4-V19.6, V20-V79, V80.3-V80.5, V81.0-V81.1, V82.0-V82.1, V83-V86, V87.0-V87.8, V88.0-V88.8, V89.0, V89.2
Maternal Deaths	A34, O00-O95, O98, O99
Covid-19 Deaths	U07.1

Population: American Community Survey 1-year population estimates were used as population denominators for mortality rates. Age-group specific total-population, female, and male population values were obtained from tables B01001 (all races), B01001B (Black), B01001D&E (Asian/PI), B01001H (non-Hispanic White), and B01001I (Hispanic).

Infant deaths and maternal deaths are reported per number of live births. Number of live births are calculated from the Natality restricted-use micro-data files from the NVSS.

Weights: The mortality rates (except for infant deaths and maternal deaths) are age-adjusted by the direct method of age adjustment using the US 2000 standard population. All deaths where age was missing, unknown, or not stated are excluded from the analysis.

Metric Calculation:

Age-adjusted death rate =

$$\sum \frac{\text{Deaths from specific cause for age-group}}{\text{Total population for age-group}} * \text{age-group weight} * 100{,}000$$

$$\text{Infant Mortality Rate} = \frac{\text{Total infant deaths}}{\text{Total live births}} * 1000$$

Maternal Mortality Rate = $\frac{\text{Total maternal deaths}}{\text{Total live births}} * 100{,}000$

Birth Outcome

Natality restricted-use micro-data files from the NVSS were used for calculation of birth outcome. Birth outcome metrics are not age-adjusted. Birth weight, age of mother and prenatal care information was given in the birth records.

Population: Low birth weight and prenatal care are reported per number of live births. Number of live births are also calculated from the Natality restricted-use micro-data files from the NVSS. Teen birth is reported per unit female population aged 15-19 years. American Community Survey 1-year population estimates were used for female from tables B01001 (all races), B01001B (Black), B01001D&E (Asian/PI), B01001H (non-Hispanic White), and B01001I (Hispanic).

Metric Calculation:

Low Birthweight = $\frac{\text{All births-weights <2500 grams}}{\text{Total live births}} * 100$

Prenatal Care = $\frac{\text{All where prenatal care began in first 3 months}}{\text{Total live births}} * 100$

Teen Births = $\frac{\text{Births where age of mother is 15-19}}{\text{Female population 15-19 years}} * 1{,}000$

Appendix 3. References

1. Missouri Census Data Center. (n.d.). Retrieved September 29, 2021, from https://mcdc.missouri.edu/applications/geocorr.html.
2. Dranger E, Remington P. YPLL: A summary measure of premature mortality used in measuring the health of communities 2004
3. Sprague, Thomas B. "Explanation of a new formula for interpolation." Journal of the Institute of Actuaries 22.4 (1880): 270-285.
4. National Center for Health Statistics. Compressed Mortality File, 1999-2011 (machine readable data file and documentation, CD-ROM Series 20, No. 2Q) as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Hyattsville, Maryland. 2014.