Cross Cutting
Health Disparities

GUIDE FOR SUB-COUNTY ASSESSMENT OF LIFE EXPECTANCY (SCALE)

Improving the Capacity of States and Local Health Departments to Understand Life Expectancy Disparities

CSTE
COUNCIL OF STATE AND TERRITORIAL EPIDEMIOLOGISTS

Version 2.2
February 2019

Table of Contents

1.0. Executive Summary

As communities become increasingly data savvy, residents want to know what health conditions are in their neighborhood, not at a state, county, or national level. However, many local health jurisdictions have lacked the capacity to examine data below a county level, even though county and state level data hide some stark disparities. Examining neighborhood level estimates of Life Expectancy (LE) at birth in the context of known behavioral, social, and environmental risk and protective factors has been effective at reaching the community and gaining resources for communities. This **Sub-County Assessment of Life Expectancy** (SCALE) Guide is intended to serve as a resource for public health practitioners and their partners who are working to identify, measure, and understand growing and persistent community level health disparities and to catalyze collective actions to address the underlying causes of these disparities. The SCALE project goal is to improve the capacity of states and local health departments to calculate sub-county level LE estimates.

LE at birth is defined as the estimated number of years a newborn can expect to live if current age-specific death rates in that population remained the same over time [1]. This measure is particularly useful for examining community-level disparities because it reflects the impact of major illnesses and injuries and their underlying causes, enables direct comparisons across geographies and time, and is simpler and more intuitive to the public and policy makers than are other measures of death (e.g., standardized mortality ratios, age-adjusted mortality rates, and years of potential life lost) [2 -7].

Scaling these efforts across the US can inform future research and focus attention of policy makers, legislators, and the public on underlying conditions that are immediately actionable. To advance this initiative, the Council of State and Territorial Epidemiologists (CSTE), Centers for Disease Control and Prevention (CDC), six states (Florida, Massachusetts, Maine, New York, Washington, and Wisconsin) and two local (Los Angeles County and Public Health, Seattle & King County) health departments reviewed existing literature and methods, identified software tools, and developed this draft Guide.

The Guide is arranged to provide the background rationale for sub-county methods used in SCALE; explain how, what and where to find data for LE calculations; share an existing tool; and show how State and Local Health Departments utilized the process and what outcomes were experienced.

Calculation of LE can enable the following future public health practice and research applications:

1. Identify and monitor community hot spots of health disparities.
2. Visually examine the degree to which LE and associated contributing factors vary across populations and geographies.
3. Raise public awareness about the importance of place-based factors in creating health and health disparities including those not traditionally associated with public health (i.e., education, housing, transportation, community development, and employment).
4. Facilitate research on the relative contribution of specific behavioral, social, and environmental factors in creating health.
5. Catalyze multisector collaborations and empower communities to more effectively address upstream factors, reduce disparities, and improve community health.

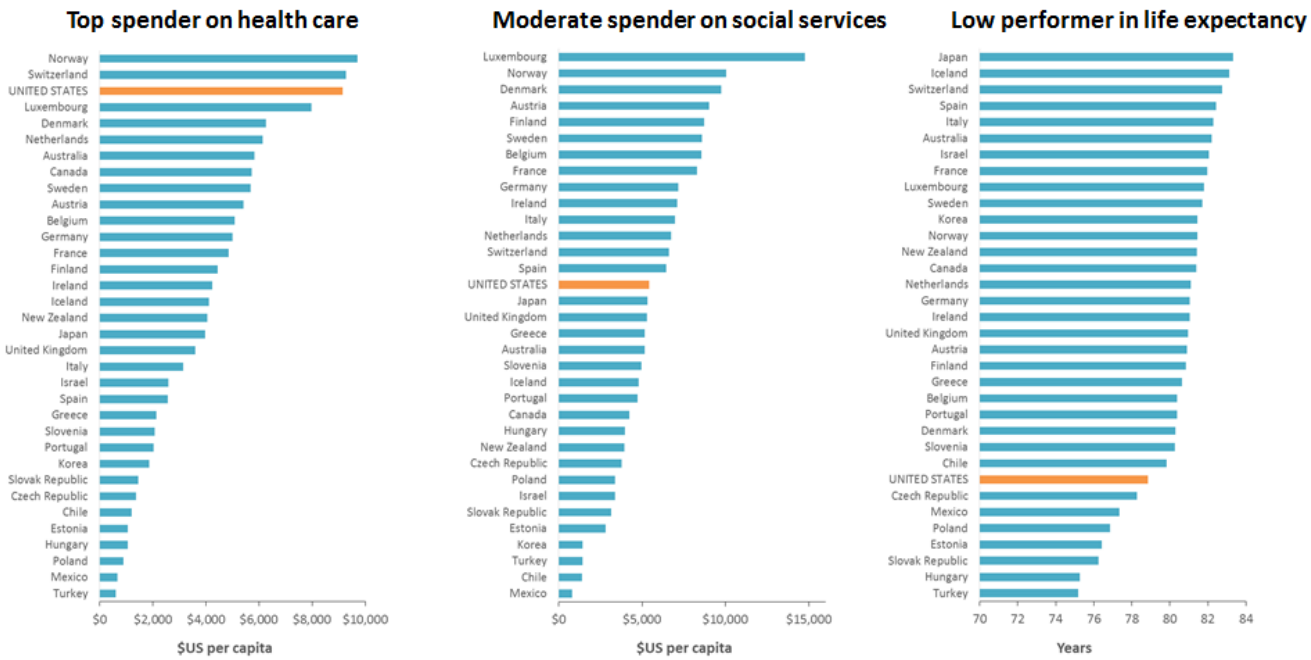
2.0. Introduction

LE is a measure that holds many people's attention, and is a concept that is easy for community to grasp. LE at birth is defined as the estimated number of years a newborn can expect to live if current age-specific death rates in that population remained the same over time [1], and is particularly useful for examining community-level disparities because it reflects the impact of major illnesses and injuries and their underlying causes, enables direct comparisons across geographies and time, and is simpler and more intuitive to the public and policy makers than are other measures of death (e.g., standardized mortality ratios, age-adjusted mortality rates, and years of potential life lost) [2 -7]. All states in the U.S. are required to routinely and systematically report death, and information from the death certificates (race/ethnicity, age, and a geographic identifier such as address, city, or ZIP code) can readily be used to calculate reliable and comparable LE estimates. Communities can galvanize around seeing a gap of 15 or 20 years between the longest lived and the shortest lived areas. In many cases, local public health may already have created calculations of LE at the county or state level, as those are often straightforward and may not require a detailed look at the underlying data going into the calculation. When looking at smaller geographies, however, there are a number of issues to keep in mind, particularly as numbers get small.

2.1. Disparities in Life Expectancy at the Country, County, and Local Levels

Disparities in LE estimates between the U.S. and other countries in the context of health expenditures have attracted increasing attention during the past few years. In 2010, the U.S. ranked 40th for male and 39th for female life expectancy at birth among 187 countries [8] [9], even though the U.S. spends almost twice as much per capita on health care than does any other country (Figure 1) [9][10]. All Americans, even the most educated, affluent, and well-insured, live sicker lives than those in other developed countries [11 -14]. More disturbing, the gap appears to be widening. Comparisons of historical trends of life expectancy between the U.S. and other countries found that, since ranking seventh in life expectancy during the 1950s, the U.S. has dropped more than 25 places, with the most rapid relative declines occurring during the past three decades among women [12] [15]. According to a 2012 Annual Review of Public Health article, this lag in U.S. health status results from "structural factors related to inequality and conditions of early life" [9].

Figure 1: Health Care and Social Services Spending versus Life Expectancy, by Country



Sources: Health expenditures per capita, 2013 and life expectancy at birth, 2013 (World Bank); Social expenditures per capita, 2011 (OECD)

[1]

Results of multiple studies suggest the primary driver of the poor relative national level performance of the U.S. on several public health indicators, including those seen in Figure 1, is profound disparities in LE at birth across U.S. counties [15-17]. For example, LE in some top performing counties— females in Marin County, California (85.0 years) and Montgomery County, Maryland (84.9 years) and males in Fairfax County, Virginia (81.7 years), and Gunnison County, Colorado (81.7 years)—is comparable with LE in countries where populations live the longest including Japan and Switzerland. In contrast, LE estimates for males in McDowell County, West Virginia (63.9 years), and Bolivar County, Mississippi (65.0), and for females in Perry County, Kentucky (72.7), and Tunica County, Mississippi (73.4), were lower than estimates for Algeria, Bangladesh, and Nicaragua [18].

The U.S. LE remains significantly lower than the longest lived countries. In 2009, the U.S LE was 78.2 years, compared to 81.8 in the longest lived countries. The current U.S. LE is equivalent to the LE that the 10 longest lived countries had in 1990. Extrapolating from this finding, it would take 19 years of improvement for the U.S. to catch up to the current LE in the longest lived countries. Similar comparisons demonstrate that disparities among counties when compared to the longest lived countries were even greater than between nations (Figure 3). LE ranged from 16 years *ahead* of the longest lived countries to more than 50 years behind [18]. If current trends hold, the worst performing counties don't have much of a chance to catch up to the best performing countries OR counties, as the top performing counties have seen steady gains over time, whereas LE in the worst performing counties have stagnated over the past 25 years [19].

[1] http://data.worldbank.org/, accessed July 2015

Figure 2: Historic and Projected Life Expectancy of the Longest-lived Countries, by Year, 1950 to 2050[2]

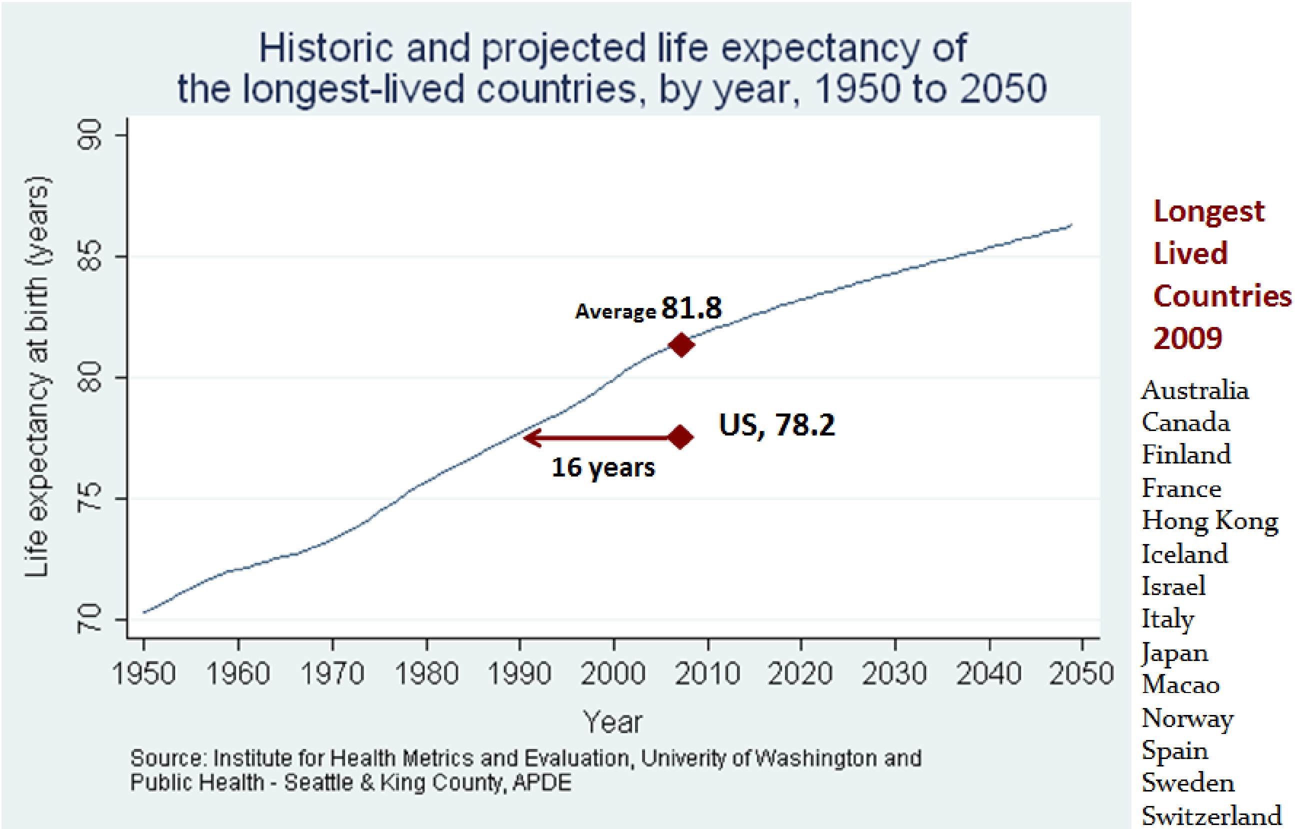
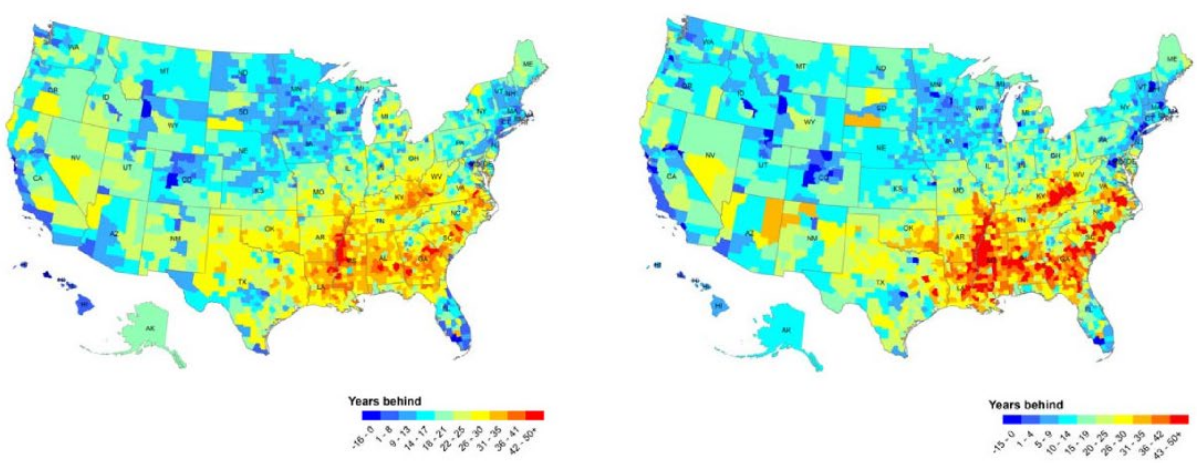


Figure 3: Life Expectancy by U.S. County and by the 10 Countries with the Highest Life Expectancy[18]



Females

Males

As seen in the comparisons above, the national estimates mask county-level disparities in LE. Recent evidence suggests that county-level LE measures are masking similar magnitudes of disparities at the sub-county level, even in counties that perform well overall on other measures of premature death.

[2] Presentation to King County Board of Health, Assessment, Policy Development & Evaluation, 5/2013

For example, researchers reported that the incidence of premature death in Boston was 1.39 times higher (95% CI 1.09–1.78) for persons living in census tracts where >20% of the population had incomes below the federal poverty level than it was for census tracts where <5% of the population lived in poverty. Similarly, the results of a study examining health disparities in 77 communities within Chicago found LE estimates varied by more than 15 years, ranging from 68.2 to 83.3 years [20]. As demonstrated in these examples, considering the geographic context of premature death has enormous potential for identifying local concentrated areas (or "hot spots") of health disparities and facilitating research on the role of local area factors including housing, education, employment opportunities, environmental conditions, behavioral factors, and access to health care and material goods that impact social disparities in health [21].

> RIGHT HERE IN KING COUNTY WE HAVE SOME OF THE BIGGEST HEALTH DISPARITIES YOU'LL FIND ANYWHERE. PEOPLE WHO LIVE IN CERTAIN KING COUNTY NEIGHBORHOODS ENJOY SOME OF THE LONGEST AND HEALTHIEST LIVES OF PEOPLE ANYWHERE, WHILE JUST A BIKE RIDE AWAY LIFE EXPECTANCY AND QUALITY OF LIFE IS MUCH MORE SIMILAR TO DEVELOPING COUNTRIES. THERE IS PLENTY THAT CAN BE DONE, AND MANY WAYS TO BE PART OF THE SOLUTION BOTH LOCALLY AND GLOBALLY.
>
> David Fleming, Former Director of Public Health Seattle King County

2.2. Interest in and Need for Life Expectancy Estimates

These large magnitudes of community-level disparities have caught the attention of U.S. legislators. For example, Senator Bernie Sanders, who chaired the Senate Subcommittee on Primary Health and Aging, held a congressional hearing in 2013 titled, "Dying Young: Why Your Social and Economic Status May Be a Death Sentence in America." The hearing included testimony from physician and research experts on health, economic, and educational factors that contribute to disparities in LE. At the hearing's conclusion, Senator Sanders cited poorer parts of the U.S., including some rural counties and inner-city neighborhoods, noting, "In many ways, the stress of poverty is a death sentence, which results in significantly shorter life expectancy. Parts of Boston and Baltimore have a lower life expectancy than Ethiopia and Sudan."

In addition to the increased legislator and public attention to community-level health disparities, several recent developments have increased the demand for assessing and improving local population health. First, the voluntary public health accreditation standards[3], launched in 2011, require a comprehensive community health assessment and community improvement plan as prerequisites for state and local health departments seeking accreditation. Second, Section 9007 of the Affordable Care Act requires that the >3,000 nonprofit hospitals across the U.S. complete a community health needs assessment every 3 years and adopt an implementation strategy to meet identified needs in order to continue to meet tax-exemption status. One specific requirement of the accreditation standards and the Internal Revenue Service regulations is identification of and engagement with community

[3] http://www.phaboard.org/

members or their representatives from populations experiencing health disparities within their jurisdictions. Finally, Public Health 3.0 [4] challenges local health to serve as a chief health strategist, looking at social determinants of health, and examining data at a local level, with community context.

3.0. Examples from the Field, Prior to SCALE

Several health departments have successfully used sub-county estimates of LE at birth to identify and explore local hot spots of health disparities, to raise public awareness, and to catalyze multisector partnerships and collective actions. In Sections 3.1 and 3.2, we present case studies from two such health departments, Public Health—Seattle & King County and the Los Angeles County Department of Public Health to provide examples of the promise and utility of sub-county life expectancy estimates for informing public health action. In addition, the Robert Wood Johnson Foundation (RWJF) has produced a series of story maps to catalyze conversations about the inequity in LE. We provide a brief summary of the RWJF effort in Section 3.3

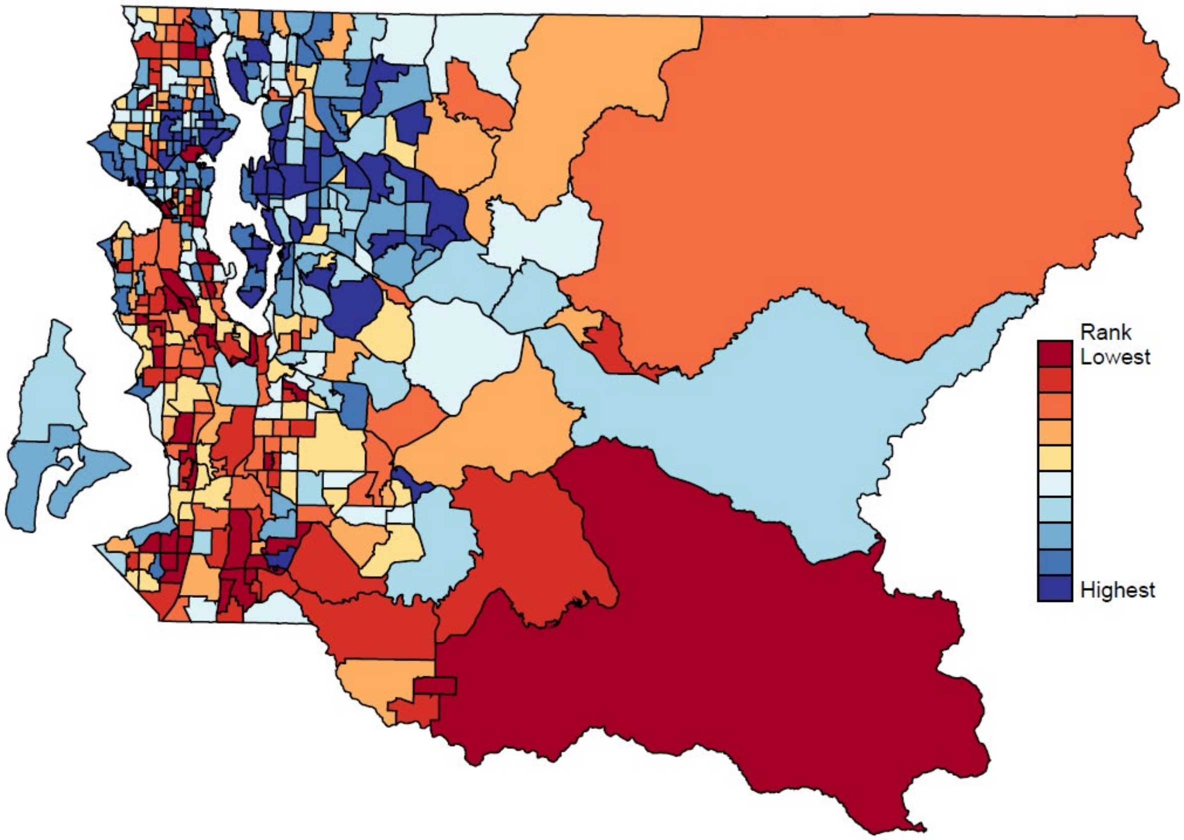
3.1. Public Health, Seattle & King County

King County, home to 2.1 million residents and 39 cities, is the most populous county in Washington State. Home to businesses such as Microsoft, Amazon, Weyhauser, and sporting other technology hubs, King County health outcomes and risk behaviors tend to compare favorably to other counties in the US; however, health equity work has shown the high performing county rate masks large disparities in health [27]. In 2012, Public Health, Seattle & King County (PHSKC) calculated LE at a census tract level for the 398 tracts in King County to begin examining place-based disparities. The 5-year estimates showed a range of 25 years (after suppression of unreliable rates), with a low of 72 years and a high of 96 years. The overall King County LE was 81.6 years.

These findings generated questions from local leaders and community members how additional health behaviors and health outcomes would look at a similar geography, and PHSKC embarked on a small area estimation project that showed a consistent pattern of disparities.[23] This work was presented in 2013 at a Federal Reserve Bank meeting, culminating in a place-based initiative called Communities of Opportunity (COO),[5] a public-private partnership with the Seattle Foundation and Living Cities, and is a cross-divisional initiative with PHSKC and the Department of Community and Human Services in King County. COO's goal is to create greater health, social, economic and racial equity in King County so that all people thrive and prosper, regardless of race or place, with a focus on economic, health, housing, and community metrics. The project is rooted in the community, using a collective impact framework that allows the community to shape their own solutions. The COO project is also aligned within the King County Accountable Community of Health.[6]

Figure 4. Life Expectancy at Birth, King County Census Tracts, 2008-2012

[4] DeSalvo, K. Public Health 3.0: Time for an Upgrade., AJPH 106(4), pp. 621–622
[5] http://www.kingcounty.gov/elected/executive/health-human-services-transformation/coo.aspx, last checked 2/2017
[6] http://www.kingcounty.gov/elected/executive/health-human-services-transformation/ach.aspx

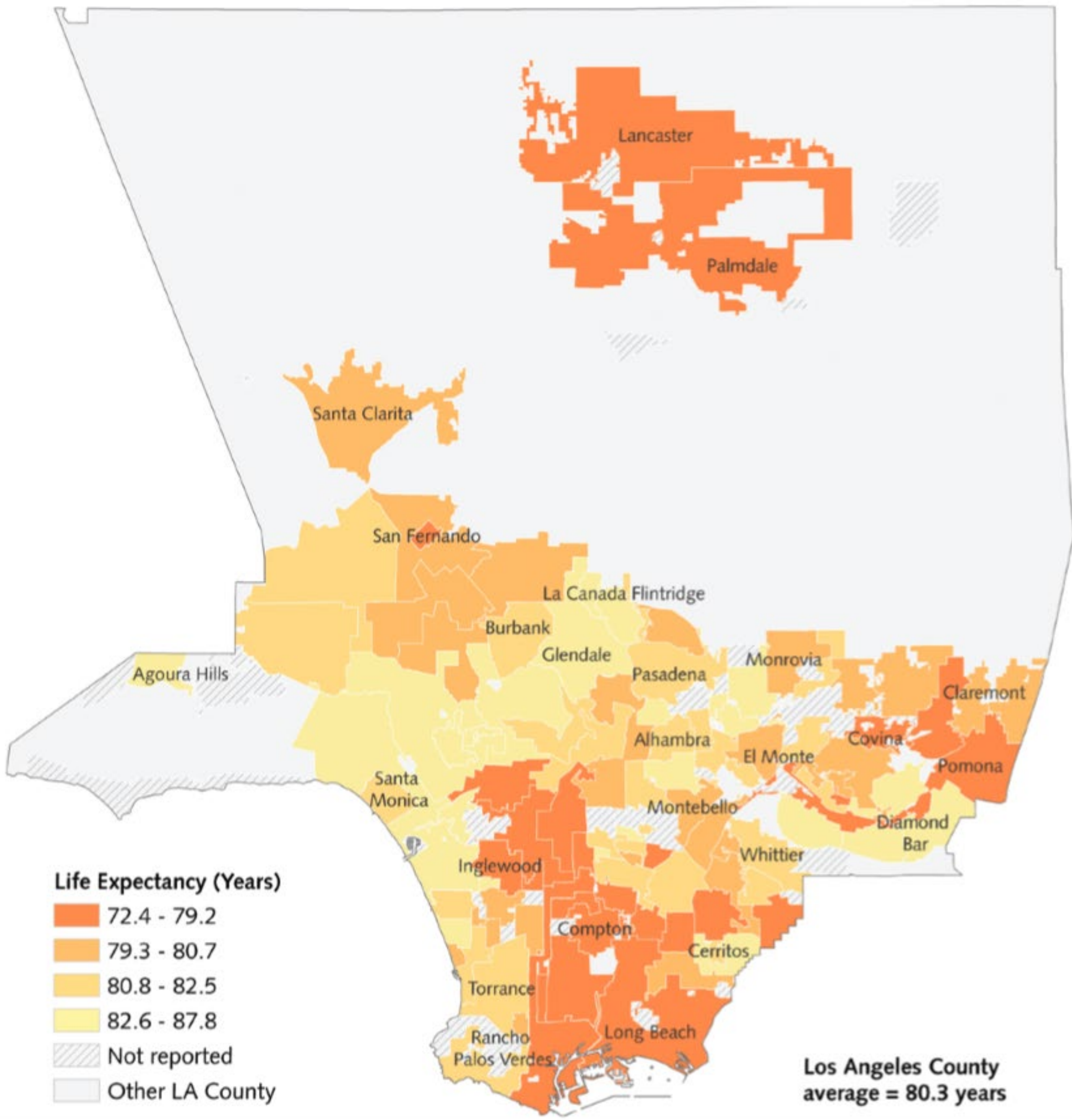


3.2. Los Angeles County

In 2009, the Los Angeles County Department of Public Health (LACDPH) calculated LE at birth for 103 cities and communities within the County.[2] In earlier analyses, large and persistent disparities in LE had been observed, and the LACDPH recognized that there was a need to bring increased attention to addressing the underlying social and physical determinants of health in order to make progress in narrowing these disparities. The County's cities and unincorporated communities were viewed as important partners in this effort, and it was hoped that examining LE at the city and community level would bring increased attention and engagement.

LE across cities and communities varied widely, ranging from 72.4 years to 87.6 years, and was strongly correlated with community-level economic hardship. Cities/communities were ranked by LE and by economic hardship, and this information was published in a report that was published and broadly disseminated to the general public, city mayors, councilmembers, city planners, and as other public health stakeholders. The information resulted in increased engagement with communities, local governments, policymakers, city planners, and other sectors; it also increased recognition of the important impacts of the physical and social environments on health.

Figure 5. Life Expectancy at Birth in LA County Neighborhoods



3.3 Robert Wood Johnson Foundation LE work

Similarly, RWJF has generated public attention on community-level health disparities by funding Virginia Commonwealth University to create a series of maps showing LE in 20 U.S. cities. These maps depict dramatic differences in LE in several jurisdictions. For example, the project demonstrated disparities by as much as 25 years within New Orleans (Figure 6).

Figure 6: Metro Map: New Orleans, Louisiana.[7] The average LE for babies born to mothers in New Orleans varies by as much as 25 years just a few miles apart.

[7] http://www.rwjf.org/en/library/articles-and-news/2015/09/city-maps.html

RWJF also calls out that gaps in health across neighborhoods stem from multiple factors[8]:

Education and income: Communities with weak tax bases cannot support high-quality schools and jobs are often scarce in neighborhoods with struggling economies.

Unsafe or unhealthy housing exposes residents to allergens and other hazards like overcrowding.

Stores and restaurants selling **unhealthy food** may outnumber markets with fresh produce or restaurants with nutritious food.

Opportunities to exercise, walk, or cycle may be limited, and some neighborhoods are unsafe for children to play outside.

Proximity to toxic agents such as highways, factories, or other sources of may expose residents to pollutants.

Access to primary care doctors and good hospitals may be limited.

Unreliable or expensive public transit can isolate residents from good jobs, health and child care, and social services.

Residential segregation and features that isolate communities (e.g., highways) can limit social cohesion, stifle economic growth, and perpetuate cycles of poverty.

[8] https://societyhealth.vcu.edu/work/the-projects/mapping-life-expectancy.html

4.0. SCALE Project

To address the needs articulated in Section 2.0 and to expand the work demonstrated by PHSKC, Los Angeles County, and other jurisdictions, the Centers for Disease Control and Prevention (CDC) provided funds to the Council of State and Territorial Epidemiologists (CSTE) in October 2014 to engage several state and local health departments in a multi-year project. This project, entitled the Sub-County Assessment of Life Expectancy (SCALE) Project, engaged several jurisdictions, developed and pilot tested resources to support state and local health departments throughout the US in calculating sub-county LE estimates, and provided training and presentation on the work. In this section, we describe the goals and key activities that comprise the SCALE Project.

4.1. SCALE Project Goal

The SCALE project goal is to improve the capacity of states and local health departments to calculate LE at birth estimates below the county level. Calculation of LE estimates in many jurisdictions throughout the US can enable the following future public health practice and research applications:

1. Identify and monitor community hot spots of health disparities.
2. Visually examine the degree to which LE and associated contributing factors vary across populations and geographic locations.
3. Raise public awareness about the importance of place-based factors in creating health and health disparities including those not traditionally associated with public health (i.e., education, housing, transportation, community development, and employment)
4. Facilitate research on the relative contributions of specific behavioral, social, and environmental factors to life expectancy.
5. Catalyze multisector collaborations and empowered communities to more effectively address upstream determinants of health, reduce disparities, and improve community health.

4.2. SCALE Project Activities

The SCALE project aims to improve the capacity of state and local health departments to calculate sub-county LE estimates by developing and disseminating easy-to-use resources, and identifying and sharing lessons learned from the project through evaluation activities (See Table 1). SCALE was a multi-phase effort in which the work of each phase builds upon the last. The project commenced in January 2015 when CSTE and CDC invited six state health departments (Florida, Massachusetts, Maine, New York, Washington, and Wisconsin) and two local health departments (LACDPH and PHSKC) with previous experience in small area analysis to participate in the first phase of the project.

During Phase I, the eight jurisdictions engaged in collaborative efforts to identify via literature review, test, and suggest methods for calculating sub-county LE, and produced a guidance document (this "Guide") to share lessons learned and decisions point guidance with other jurisdictions. These eight jurisdictions are the core Workgroup, which meets on a regular basis to discuss current updates to LE work, additional methods, and to provide TA to incoming jurisdictions.

In Phase II, 25 additional state and local health departments of varying sizes joined the SCALE project (Figure 7). These jurisdictions pilot tested the draft Guide produced during Phase I and assessing the extent to which the methods and associated tool identified in Phase I (i.e., the SEPHO tool, see Section 5.0) met their needs in calculating sub-county LE estimates.

In addition, states and localities from Phase I and a subset of Phase II participants engaged in discussions and activities to address several methodological questions raised during Phase I, such as how to address cells with zero deaths and how to treat areas where a large proportion of the population lives in group quarters. Initial efforts to identify best practices for mapping sub-county LE estimates drawing upon expertise from CDC's Geospatial Research, Analysis, and Services Program (GRASP) were also undertaken during Phase II.

Additional states and localities joined Phase III to continue testing and refining methodologies identified or developed in Phases I and II. In addition, during this final phase, the Workgroup addressed next steps after generating LE, discussed and made recommendations for visualizing sub-county LE, and provided tips for communicating about these estimates with various target audiences. Phase III also included a roundtable meeting with LE experts from Public Health England, who shared their lessons learned through many years of small area LE estimation. Lessons learned from Phase III and recommendations from these efforts will be included in future versions of this Guide.

Table 1. SCALE: Project Phases[9]

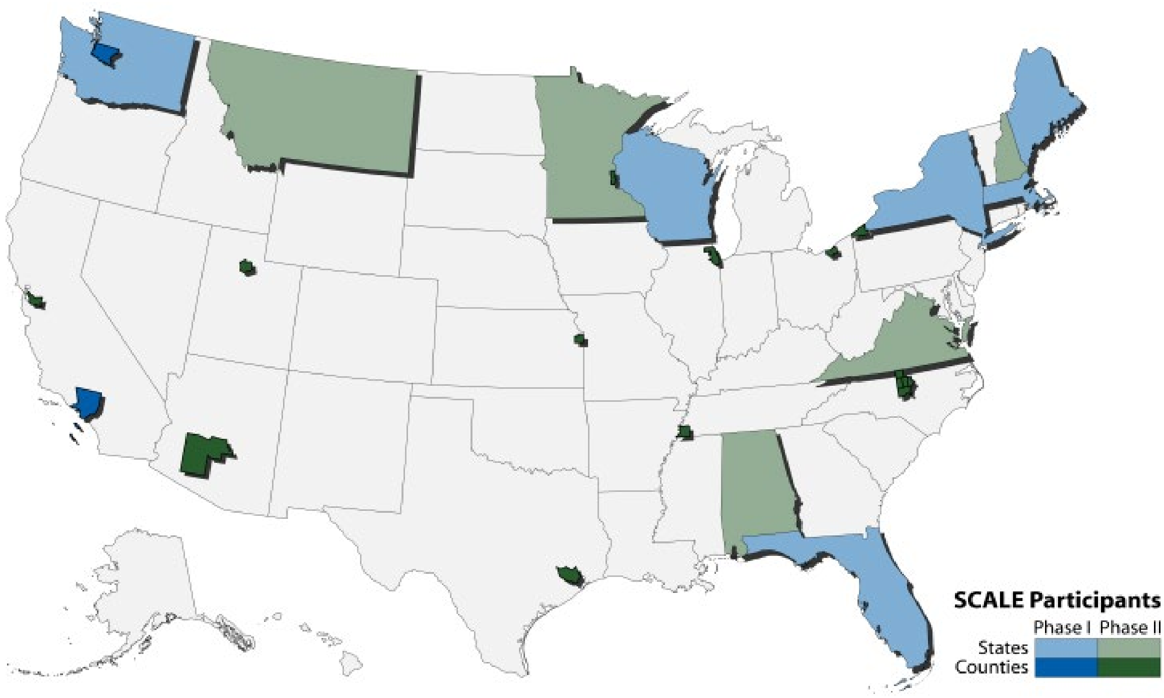
Phase I	Conducted a literature review to understand the approaches, available parameters, and lessons learned from previous efforts associated with constructing small-area LE estimates.
	Reviewed common approaches used in the literature for calculating direct small-area LE estimates and arriving at initial decisions about methodology.
	Identified other existing tools for calculating LE that might easily be adopted/adapted (SEPHO).
	Compared calculations produced by SEPHO tool with other methodologies for generating LE estimates (SAS and STATA code from previous LE efforts), refined approach.
	Developed evaluation plan for Phase II.
	Products include: (1) Draft Guide for state/local health departments with SEPHO tool as approach used, (2) Sub-county estimates for Phase I states/localities, (3) 2015 CSTE conference presentation, (4) Evaluation plan
Phase II	Recruit and orient new states/localities to methodology and general project purpose/approach.
	Implement evaluation; New state/localities pilot test draft materials from Phase I and provide feedback through the evaluation.
	States/localities assess potential refinements in methodology to expand geographic coverage by performing several sensitivity analyses.
	In collaboration with the Geospatial Research, Analysis, and Services Program (GRASP) the Agency for Toxic Substances and Disease Registry (ATSDR), identify methods for visualizing LE using direct estimates of LE.
	Engage expert panel to develop initial recommendations for visualizing and communicating about LE estimates.

[9] Paper, in progress

	Anticipated products include: (1) Revised tools for estimating LE, (2) Revised/updated Guide, (3) Recommendations regarding visualization and messaging, (4) 2016 CSTE conference presentations, (5) Evaluation findings, (6) Manuscript(s)
Phase III	Recruit and orient new states/localities to methodology and general project purpose/approach.
	Anticipated products include: (1) Revised/updated Guide, (2) Additional tools for LE calculation or consideration, (3) 20167 CSTE conference presentations, (4) Evaluation findings

LE: Life expectancy, SEPHO: South East Public Health Observatory, CSTE: Council of State and Territorial Epidemiologists, SCALE: Sub-County Assessment of Life Expectancy

Figure 7. States and Local Health Departments participating in SCALE



5.0. Review of Approaches for Calculating Life Expectancy

Definition of Life Expectancy

Life expectancy (LE) is a summary mortality measure often used to describe the overall health status of a population. For any given population, LE can be calculated at any age (e.g., birth, age 50 years, age 65 years). The SCALE project focuses on LE at birth, which is defined as the estimated number of years a newborn can expect to live if current age-specific death rates in that population remained the same over time [1].

In the US, LE is a commonly used summary indicator of population health and health disparities. Because all states require deaths to be routinely and systematically reported, information from the death certificates (race/ethnicity, age, and a geographic identifier such as address, city, or ZIP code) can readily be used to calculate reliable and comparable LE estimates.

LE requires knowing how many people died and the population size

5.1 Overview of approaches considered to develop the life expectancy calculation

Multiple methods exist for estimating LE. These include methods based on stable population concepts, biological theories of aging, estimation of population by age, regression equation methods that exploit the relationship between LE and other demographic indices, construction of abridged life tables and methods that combine traditional complete life table construction techniques with smoothing or graduation methods [22] [24]. After literature review and group discussion, the SCALE Workgroup decided to use an abridged life table method, and the Chiang II calculations. Details are provided below.

Life Tables

A life table shows the probabilities of a member of a particular age dying before their next birthday. In the US, two types of life tables are used: the cohort (or generation) life table and the period (or current) life table. The cohort life table is based on age-specific death rates observed through consecutive calendar years and reflects the mortality experience of an *actual* cohort from birth until no one from the group is alive [26]. The period life table represents the mortality experience of a *hypothetical* birth cohort if it experienced throughout its entire life the mortality conditions of the period of interest. The period life table can be considered “a snapshot of current mortality experience and shows the long-range implications of a set of age-specific death rates that prevailed in a given year” [26].

CDC’s National Center for Health Statistics publishes complete period life tables annually.[10] Given the routine nature and availability of a nationally published period life table, the Workgroup settled on using this as the basis for Phase I of the SCALE project.

Abridged Life Table

A *complete life table* contains data for every year of age, whereas an *abridged life table* typically contains data by 5- or 10-year age intervals. The SCALE project suggests using an abridged life table with 5-year age intervals except for the first interval, which is set at 0–1 year, and the last interval, which is defined as 85+ years. SCALE recommends using an abridged table for smaller geographic areas

[10] http://www.cdc.gov/nchs/products/life_tables.htm.

in recognition that as populations are smaller in sub-county geographies, single age categories may have no deaths recorded during that time period, resulting in many zero cells, and biasing the estimates.. It's important to separate the infant deaths (before age 1) from the 1-4 age category, as infants who die have a much higher rate of death in the first 28 days, compared to other ages, which have a more normal distribution of deaths across the year and age-group. Grouping infant deaths with the 1-4 category will result in a higher LE estimate and may dilute health disparities experienced across different populations. The abridged life table method can be used for any geographic area, including census tracts, ZIP codes, city boundaries, or other geopolitical units.

Adjusted Chiang II Methods

The long-established Chiang method for estimating LE by using a period (current) life table has been widely used internationally [7] [27]. The Chiang method and its variations assume that deaths are spread evenly throughout each age period, except for persons <1 year of age, for whom deaths are highly skewed toward the first 28 days of life. For all other age groups, Chiang assumes a 0.5 age interval; the <1 group is valued at 0.1. One major concern about the Chiang method involves a miscalculation in standard error in age groups for which no deaths occur. One way to prevent zero deaths in a single year of age is to collapse age categories, which is why the abridged life table is suggested. In addition, researchers have developed alternative methods to address this issue, including the Chiang II method and the adjusted Chiang II method [10].

The adjusted Chiang II method was proposed to modify the assumption in the Chiang I method of a zero variance for the final age band. The adjusted Chiang II method uses the formula by Silcocks to adjust for variance in the final age band [3]. The SCALE recommendations use the adjusted Chiang II methods.

5.2 Addressing small-area methodologic issues

Calculation of LE at the county or state level is typically straightforward, given generally large population numerators and denominators, and may not require a detailed look at the underlying data going into the calculation. County and states are geographies routinely assigned, and there are often population estimates available at a state or county level. When looking at smaller geographies, however, there are a number of issues that must be considered, as small numbers in the geography or age group (deaths or population) or how unequally distributed data (deaths or population) may influence the LE calculations.

Small Populations/Minimum Population Size

Several authors have examined the impact of small populations on life expectancy. Variations on the suggested minimum population size range from 3,750 to 7,000. Additional information can be found in Appendix A. Some researchers have suggested that life table estimates overestimate LE for populations less than 5,000 years of life at risk [1] [28] [29]. For the SCALE Phase I project, several Workgroup members assessed the minimum population size for SCALE applications by generating estimates for all tracts, and examined standard error and special conditions of the tract to assess whether a unique feature of the tract (e.g., nursing home residents, incarcerated populations, university students) affected the LE estimate. Others used an R-based tool to aggregate tracts to the minimum population size. One Workgroup member excluded tracts that had a high percentage (>=50%) living in group quarters, tracts that had no land area (were only water), and military bases. In one Workgroup case, tracts remained too small and they used Minor Civil Divisions (MCDs).

Because death data rarely come with assigned geocodes, jurisdictions may need to use address information to geocode the death to tract or other geography (see Section 6, below). Once the geographic assignment is made, numerators and denominators can be evaluated to see whether aggregation of tracts or another higher level geography is helpful to present stable rates.

Standard error and confidence intervals

Because LE can be a tightly grouped dataset, determining a meaningful difference between LE ranges is important. Most of the reviewed papers did not discuss a specific standard error. The SCALE Phase I Workgroup recommends using a standard error of ±2, based on literature review and after analysis of calculated LE for local jurisdictions. The Workgroup suggests considering suppressing LEs with a large standard error.

Zero Cells

Based on the population, population distribution, and death rates, zero cells can occur, even when grouping ages and across years, especially when small geographic units, such as census tract, are used. A zero death count gives an estimate of zero for an age interval, which causes underestimation of the variation; the more zeros and the more underestimation that occur, the larger the underestimation of standard error [25]. Several corrections have been suggested to address the concern, including small substitutions of values for zero and expected numbers of deaths [1] [7] [26] [28]. The biggest concern is when there is a zero count in the <1 or the oldest age band, as those have the largest impact on rates. At this time, SCALE Workgroup view each of the options as acceptable, depending on the jurisdiction's choices. New York State chose to use an R-based tool to group tracts to hit a threshold of 60 deaths to avoid zero cells.

Age 85+ Year Category

A death count of zero in the 85+ year age category would strongly affect LE because the Chiang calculation would give the cohort an infinite survival, raising the LE estimate and standard error. To address this, some options include: replacing the zero cells with a national death rate for the country or a national age-specific death rate for the country [7] [28], or increasing the last age band to 75+. Most Workgroup participants did not have issues with zero cells in the 85+ age category given the 5 year aggregation. Those who did tended to use replace the zero cell with the overall 85+ mortality rate.

Age 90+ Year Category – section in new LE tool

Public Health England revised their upper age band for their LE estimates. They noted that in 2016, 40% of deaths in England occur over age 85, including more than half of all female deaths. When looking at the impact, the results were more similar to the single year life table, which meant that the 90+ age band gave a more robust/accurate estimate (lower variance).[11] The newest version of the PHE tool includes this age band.

Population

Routinely generated population estimates are provided from such entities as the National Center for Health Statistics (NCHS), but only on a county-wide level. Some jurisdictions may have access to local population small area estimates, which could be used for this project. Workgroup participants who did not have access to sub-county population estimates used 2010 Census data as a mid-point for LE. This is still a recommendation for small area estimates, if there are no other local or updated sources for

[11] Personal discussion from March 2018 roundtable.

the data. MA, FL, WA and PHSKC have generated local small area estimates and used these as population estimates for their LE work.

The American Community Survey (ACS), the Census Bureau's annual household survey, does produce demographic characteristic distribution at a census tract level but is not recommended for use as a population denominator. The primary purpose of the ACS is to measure changes in a community's socioeconomic characteristics based on a small sample of households surveyed every month. The Census recommends ACS for gauging trends over time and for comparing characteristics across areas, but specify that it lacks the precision for population estimates. Many of the tracts have very high coefficients of variation, which indicate the lack of precision. In addition, ACS provides summary data for aggregated age groups, including 0-4; they do not include results for the <1 population. One Phase I Workgroup member compared their LE results using the 2010 Census population vs the 5-year ACS data, and showed that the ACS data markedly changed results.[12]

Special populations

5.3 Methods Selected for SCALE

The Phase I SCALE Workgroup chose the adjusted Chiang II method because 1) it adjusts for the variance in the final age band and 2) it addresses age bands with zero deaths. Research shows simulation results suggested that use of the adjusted Chiang II method provides the closest approximations to reference LE and standard errors by not imputing any values into age groups with zero deaths [1], with the exception of the oldest age group.

An environmental scan of existing LE tools by the Workgroup led to discovery of an existing Life Expectancy Tool, created by the South East England Public Health Observatory (SEPHO) group (http://www.sepho.org.uk/viewResource.aspx?id=6626). This LE tool, which also uses an adjusted Chiang II method, is well documented. The Workgroup extensively tested and validated the results against SAS® (SAS Institute Inc. Cary, NC, USA.) programming used by LACDPH Stata (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX, USA: StataCorp LP.) programs used by PHSKC, and an in-house Excel tool for calculation of LE created by NYS. Given the extensive documentation of the SEPHO tool, similarities of, ease of use, and accessibility, the Workgroup unanimously recommended it as a tool for SCALE. SAS and Stata tools are available upon request.

6.0. Acquiring and formatting data for SCALE

6.1. Acquiring Data

For jurisdictions that have not been involved in previous efforts to perform small-area analyses, it is not uncommon to spend several months (between six to nine months) acquiring data for the purpose of calculating sub-county LE estimates. One jurisdiction in Phase I needed to have an Institutional Review Board review their request for data. Connecting with the data providers and the stakeholders who could use the small area LE is beneficial in helping to gain access to the data as well as potential

[12] Personal communication, SCALE Workgroup. Data to be released after paper publication.

technical assistance with various aspects of the process, including cleaning and formatting data, geocoding, selecting an appropriate small-area, and statistical analysis.

6.1.1. Data sources, necessary variables, and formatting

Two types of data are needed: (1) death certificate data and (2) population estimates. The death certificate data and the population data will need to be generated at the same geographic level (census tract, ZIP code, city). Most **death certificate** data will not come assigned to a census tract or neighborhood level, but more state and local health departments are starting to geocode. See Section XX for considerations about which geographic area might be optimal for use. One of the most essential decisions involves the level of geography to present the data. Phase I participants chose census tract, aggregations of census tracts, ZIP code, and a Minor Civil Division. There are different factors driving the decision – for example, Florida chose to use ZIP code as they have a number of other health indicators also mapped at the ZIP code level. Maine tried a variety of geographies until they hit one that worked for their population and data reliability requirements. New York State used aggregations of census tracts (and suppression of tracts that were primarily group quarters). Several of the jurisdictions chose to use census tract for ease of incorporating social determinants of health from the American Community Survey.

Phase I and Phase II Workgroup participants typically used between 5 and 10 years of mortality data to be able to generate small area estimates. For those who used 5 years of data, they requested 2008-2012 data.

Population data may come from internally created numbers, or from the Census. Many Phase I participants began with the 5 year period spanning the 2010 Census, multiplying the Censal counts times 5 to generate a population estimate. A few had their own locally generated population estimates.

Both data sources should be broken down into 19 age groups (<1, 1–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85+) for the purpose of this analysis. This will be fed into an Excel spreadsheet tool, so the age group aggregation should follow this set-up.

6.1.2. Formal data agreements and approvals to consider

Mortality data can be obtained from a statewide department of health[13] or a vital registration office.[14] State limitations, statutes, and practices differ with respect to the requirements for acquiring mortality data however in some cases this includes a requirement for a data sharing agreement (DSA) or a memorandum of understanding (MOU). In rare instances, some jurisdictions may also need to gain the approval of an Institutional Review Board (IRB). Localities new to the project may want to first do a quick internet or intranet search to see if previous work has already been created such as an MOU or DSA used by another county that can be repurposed. See Appendix B for examples of MOUs or DSAs used by SCALE participants.

---

[13] https://www.cdc.gov/stltpublichealth/sitesgovernance/index.html

[14] https://www.cdc.gov/nchs/nvss/deaths.htm

It is important to forge relationships to help with data sharing. Analysts will need to make a number of decisions about how to calculate and present the data, and having stakeholder involvement in this process. See the Figure 12 for a flowchart summary walkthrough of suggested steps. At a minimum, **date of birth, gender, date of death, cause of death (primary and underlying), street address, ZIP code, city, and death certificate number** should be requested. Data users might want to consider asking for geocode (tract or latitude/longitude), a measure of geocode quality, and other demographic or socioeconomic pieces for additional analysis, such as race/ethnicity, occupation, etc. Date of birth rather than an age or age range is useful, especially for the <1 age group. Phase I and Phase II participants reported having had <1 "age" be recorded as a month, which the tools converted to a year, making calculations incorrect.

Some of these variables are not used to calculate LE but may be useful in pursuing analysis of LE disparities. Including them in the initial data request prevents needing to go back for another request. It's unlikely that jurisdictions will be able to use these direct estimates to examine the granular cause of death in small areas, but some Phase II participants were able to show impacts at a county or state level by collapsing categories of death (e.g. all cancers, all heart disease) and to look at population and geographical patterns.

Another question to address when requesting data from a vital stats provider is to know whether they have a reciprocal agreement to get data for residents of their state that die in another state. (See Section 8.2 – border areas for more discussion). If they do not, jurisdictions will need to decide if they want to consider approaching the other states, as this process may be time consuming and would be best initiated at the outset of the project.

One caveat: The standard death certificate form changed in 2004, and jurisdictions adopted it at different times. To date (2018), not all jurisdictions have adopted the new certificate, and some may have switched certificate forms during the 5-year period of interest. It's worth a discussion with the vital statistics registrar or others who work with the data to understand the impact of the certificate change on the data collected.

Once the data agreements are established (if necessary) and analysts have become more facile with the data, it allows for ease in multiple iterations of the process. An example of a boilerplate DSA is located in the Appendix B.

Questions to consider when acquiring data for sub-county life expectancy

- √ What elements can be included without a DSA or MOU?
- √ Has another jurisdiction or organizational unit created a successful DSA or MOU that can be adapted?
- √ Are the data geocoded?
- √ Have other programs or units worked with geocoded mortality data? Might you be able to leverage this work?
- √ Will you need to work with other jurisdictions to get information about geographies along the border of your geography, such as counties bordering another state or a ZIP code that crosses county or state boundaries?
- √ Are there existing requirements around suppression or censorship of unreliable numbers?
- √ How many years of data might be needed for sufficient sample size?
- √ Could this data benefit other priorities in your LHJ?
- √ Will you want other demographic information for analysis (e.g. race/ethnicity, education, occupation)?

6.2. Geocoding Mortality Data

Geocoding is the process by which descriptors (e.g., address, city, ZIP code, province) are assigned a place on a map, also known as a geospatial location, creating a georeferenced dataset. Although this section is not designed to be a comprehensive tutorial, it will provide some links for additional information and best practices.[15] Recent assessments of state health department epidemiology capacity indicate that more than 50% do not routinely geocode their data.[30] As a result, it will be important many jurisdictions that wish to calculate sub-county LE will need to geocode their mortality data as part of this process. In this section we offer some tips to make this process smoother than it may be otherwise.

6.2.1. Geocoding defined

The Workgroup suggests the initial task of geocoding is to assign address information from a death certificate to a 2010 census tract. Because this project is examining geographic disparities in life expectancy, geocoding data is essential to examination of sub-county LE estimates.

Some jurisdictions may already have access to geocoded data when the data are requested from the source; others may need to perform this step. Even when geocoded data are available, it is important to understand how the geocoding process was accomplished, and the attendant levels of accuracy resulting from the method used. For example, some automatic geocoders might assign a centroid (a center point) of a city or a ZIP code to an address that cannot be matched to a street location. This can artificially inflate the number of deaths that are occurring in that tract, as the individuals are being assigned to the numerator but are not in the denominator. If data come already geocoded, one recommendation would be to request a match score and match type, which will provide the data user

[15] http://naaccr.org/LinkClick.aspx?fileticket=ZKekacM8k_IQ0%3d&tabid=239&mid=699

with appropriate information in deciding how to use the data. The user can manually match the centroid or missing values, they can assign based on an overlay, or they can randomly distribute non-matching cases.

If data do span a Censal year, data users should also verify that any geocoded data are assigned to the same census geography. One Phase I participant reported needing to re-geocode the 2008 and 2009 data to 2010 Census boundaries.

6.2.2. Software for batch geocoding

Should a participant need to geocode the death data, there are several different software options available for batch geocoding. Batch geocoding occurs when the datasets are processed through a software package and assigned to a geolocation automatically. Any geographic information system (GIS) will have batch geocoding options. Following are several software options that may be valuable to explore in more detail:

- **ArcGIS (from ESRI)** . States that participate in the CDC Building GIS Capacity for Chronic Disease Surveillance (http://www.cdc.gov/dhdsp/programs/gis_training/index.htm) should have access to this software. A geocoding tutorial for this software is accessible on the Web: http://help.arcgis.com/en/arcgisdesktop/10.0/pdf/geocoding-tutorial.pdf.
- **Google Earth.** Google Earth has geocoding capability. To use the full functionality of this tool, it is helpful for the user to know HTML and JavaScript. Users must sign up for Google Earth and obtain an API key, which is what allows a user to connect and geocode using Google Earth. A tutorial is available. There is a paid, unlimited version and a free version, which is limited to 2,500 geocodes a day. http://www.drew.edu/ess/wp-content/uploads/sites/82/Tutorial-7-Geocoding-with-Batch-Geocode-and-Google-Earth.pdf
- **Statistical analysis packages.** Some analytic software packages, such as SAS and R, also have geocoding tools. For example, SAS/GRAPH uses a proc geocode (http://support.sas.com/documentation/cdl/en/graphref/63022/HTML/default/viewer.htm#a003121448.htm) and R has a package on CRAN https://www.rdocumentation.org/packages/ggmap/versions/2.6.1/topics/geocode . The default R package uses the Google Earth API, but other geocoding reference data also can be used.

6.2.3. After batch geocoding

No software can match all addresses to a location, as there are errors in the address file, the street file for matching, or addition of news roads that aren't yet incorporated into the GIS software. Records that don't geocode are called exceptions. Exceptions can result from an incorrect street number, misspelling of a street, incorrect post office box, name of building instead of street, incorrect directional mismatch between ZIP code and street, or an incorrect underlying street layer as some examples. Rural areas may have route boxes that also are difficult to assign. These exceptions should be manually reviewed and matched when possible, as most exceptions do not happen randomly. Sometimes local GIS addresses will be better than a national one; some may just require spelling corrections. If a ZIP code is part of the record, and that ZIP falls entirely within a tract, that tract can be assigned to the record. A final geocoding match of ≥90% is ideal, and the higher it can be, the better. Most Phase I participants achieved a greater than 95% match score.

Addresses that remain unmatched after review can be handled one of two ways. They can be treated as missing, if there is non-differential classification (i.e., if the unmatched cases are similar to the matched cases in terms of demographics), they can be randomly assigned by having them distributed across all tracts, or geocodes can be imputed, based on demographic factor similarity between the case and the population.

6.3. Preparing the data

The population and death data need to be arranged in the 19 age groups listed above in Microsoft Excel or CSV format. This fits the abridged life table format. See Figures A and B as examples. Remember, if using the population estimates from the Census, multiply by 5 (or the number of years used to aggregate). Sort your data by geography (GeoID in this example). Note the number of rows for geographies – the numerator and denominator should have the same number of rows. This will be used in the SEPHO tool.

Figure A: Example of numerator/death data, 5 year aggregate[16]

GeoID	<1	1-4	5-9	10-14	15-19	20-24	25-29	…	70-74	75-79	80-84	85+
G01	32	39	3	3	11	13	22	…	141	163	196	488
G02	41	13	6	6	18	31	43	…	190	290	228	514
G03	27	32	2	2	2	2	16	…	145	161	194	518
G04	39	6	3	3	7	16	33	…	196	261	265	588

Figure B. Example of denominator/population data; 2010 population multiplied by 5

GeoID	<1	1-4	5-9	10-14	15-19	20-24	25-29	…	70-74	75-79	80-84	85+
G01	9005	10805	35125	42150	33890	40670	54140	…	25315	14715	17660	4810
G02	7865	25190	35330	29655	40070	36795	47215	…	19440	25335	11475	5300
G03	5840	7010	25820	30985	25845	31010	45860	…	32050	18260	21910	4915
G04	8780	27305	36505	27210	36635	39430	54615	…	18330	19005	9990	5665

There should be two spreadsheets: one each for the number of deaths and the number of population. In each spreadsheet, each row represents a small area, and each column represents an age group.

[16] Calculating LE for Small Areas, T. Talbot, CSTE 2016

Each row should also have a geographic identifier (city, ZIP, census tract). Those using the SEPHO tool described in Section 7 will see this referred to as an **area code**. This geographic identifier is required to calculate LE for individual areas.

Questions to consider with respect to preparing the data

- √ Are there geographic areas with populations less than 5000? Users may want to flag these areas for evaluation of the result.
- √ Are there cells with no deaths in the <1 age category or 85+?
- √ What does the distribution of deaths look like across tracts? Might tracts need to be grouped together for reliable results?
- √ Will I need to request data from another jurisdiction that abuts my boundaries?
- √ What is the quality of the geocodes?

7.0. Selecting a Software: Available Options

Workgroup participants scanned the available toolsets, and this Guide presents the 3 most common options. All Phase I and Phase II members used the SEPHO tool (See Section 5 for more information), from the South East England Public Health Observatory group, who have been developing small area LE for many years. This tool was verified and tested for accuracy. Given the simple nature of needing no special statistical software, the ease of cutting and pasting, this tool has a low bar to entry for ease of use. It also has many sub-tabs that allow the user to see exactly what is going on in each calculation and can be useful in helping to discover why some results may be different than expected. https://fingertips.phe.org.uk/documents/PHE%20Life%20Expectancy%20Calculator.xlsm

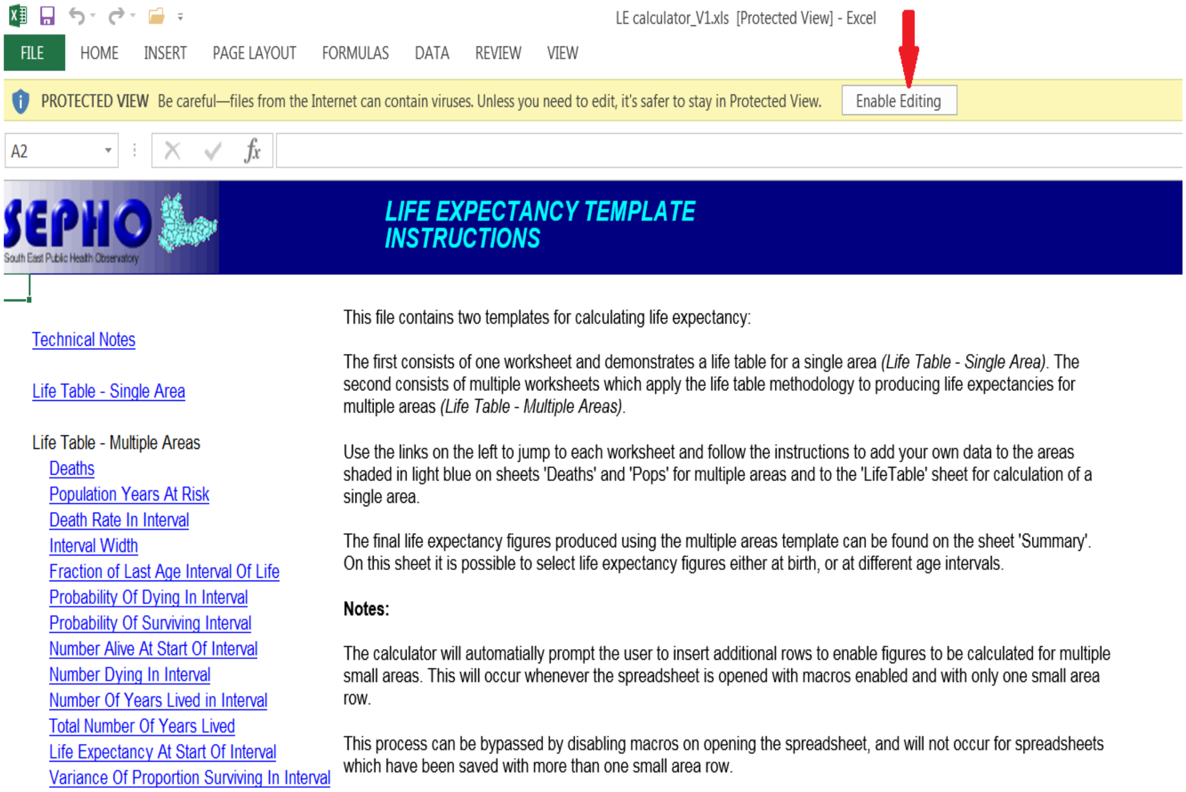
7.1 Using the SEPHO Excel tool

The numerator and population data will be copied from Excel or the CSV format and pasted in the SEPHO Excel tool, which contains macros that will calculate the LE result. This is a simple, easy to use tool that was verified for accuracy against statistical programs. (See section 7.2 for more details). This will involve a download from a website, and works best with Excel version 10 or higher. Phase I, Phase II, and Phase III SCALE participants used the SEPHO tool for simplicity. Regardless of the tool used, the steps in preparing data are the same for the calculations.

SEPHO tool Download

1. Download the SEPHO tool from https://fingertips.phe.org.uk/documents/PHE%20Life%20Expectancy%20Calculator.xlsm
2. Open the file (Life Expectancy calculator_V1.xls)
3. If this is the first time you are opening the file, depending on the security setting of your Microsoft Office, you might need to click "Enable Editing" (see Figure 8 for an example)

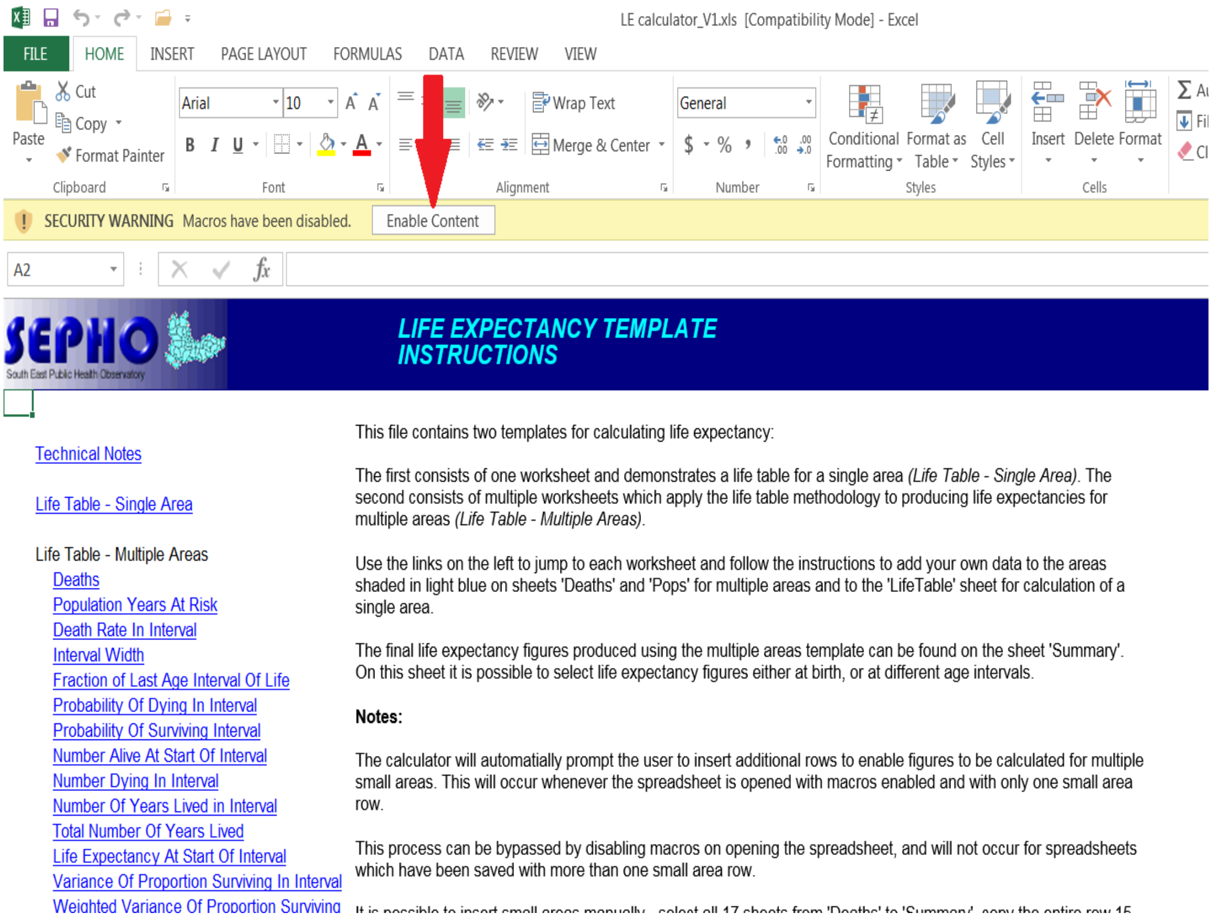
Figure 8. SEPHO tool, highlighting the "Enable Editing" button.



4. Again, depending on the security settings of your Microsoft Excel, you might need to enable the Macro by clicking “Enable Content” (where the red arrow points).

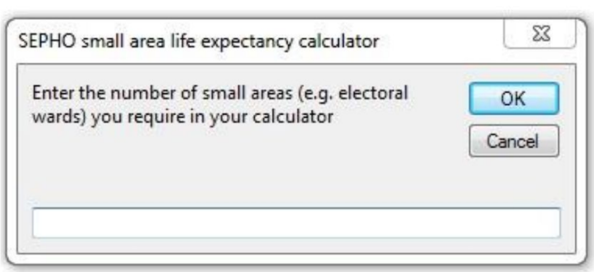
Figure 9. SEPHO tool, showing where to click to enable the macro.

5. Save the enabled version, and then save again with a new name to keep an unedited copy.



6. Click on "Deaths" under Life Table – Single Area.
7. A dialog box titled "SEPHO small area life expectancy calculator" will pop out requesting that the user "[E]nter the number of small areas (e.g., electoral wards) you require in your calculator." Put in the number of small areas in the box below – this should match the number of geographic units (tract, ZIP MCD, etc) being used in the calculation. This should be the number of rows observed in step 6.3 Make sure your data (numerator and denominator) are sorted by geography.

   Figure 10. SEPHO calculator pop up

   This will create the appropriate number of rows.
8. Copy the numbers of death by age group to the "Deaths" spreadsheet (if the spreadsheet is not shown, click the right arrow on the bottom left corner to show the "Deaths" spreadsheet).Copy the number of population by age group to the "Pops." Data should be sorted by geography so that the order matches in the Deaths and Pops page.

Figure 11. Spreadsheet example.

Area Code	Area Name	Age Group																		Total
		Under 1	1-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+	All Ages
Code	Standard/Nat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Code	Small area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note that the column "Area code" is required for calculating life expectancy and standard errors at the individual area level. This is just the geographic identifier from the input data set. If the area code is not provided, the life expectancy and standard error will be calculated for the aggregation but not for the individual areas.

9. Results will be displayed on the spreadsheet "Summary." Intermediate results are displayed as well with Life Expectancy at Start of Age interval displayed on spreadsheet "e." Make sure the table is set to "Birth" to compute LE at birth. Users can change this to look at the impact of mortality rates on other age groups as well, but that is beyond the scope of the SCALE Guide.
10. To calculate life table for a single area, copy and paste your data on to the spreadsheet "LifeTable," results will be displayed on columns "P, U,V,W" for the point estimate of the life expectancy at the start of the age interval, sample standard error, lower bound of the 95% confidence interval, and upper bound of the 95% confidence interval.
11. Save the Excel Workbook under a different name.
12. Sophisticated users can modify the life table tab if they prefer to use a different life table. For example, the upper age band can be modified to be 90+ or 95+ or the age group bands can be increased from 5 to 10 (keeping <1 separate).

More details are available at the archives SEPHO maintains http://webarchive.nationalarchives.gov.uk/20160701122411/http://www.sepho.org.uk/viewresource.aspx?id=8943 or on the CSTE page. https://fingertips.phe.org.uk/documents/PHE%20Life%20Expectancy%20Calculator.xlsm

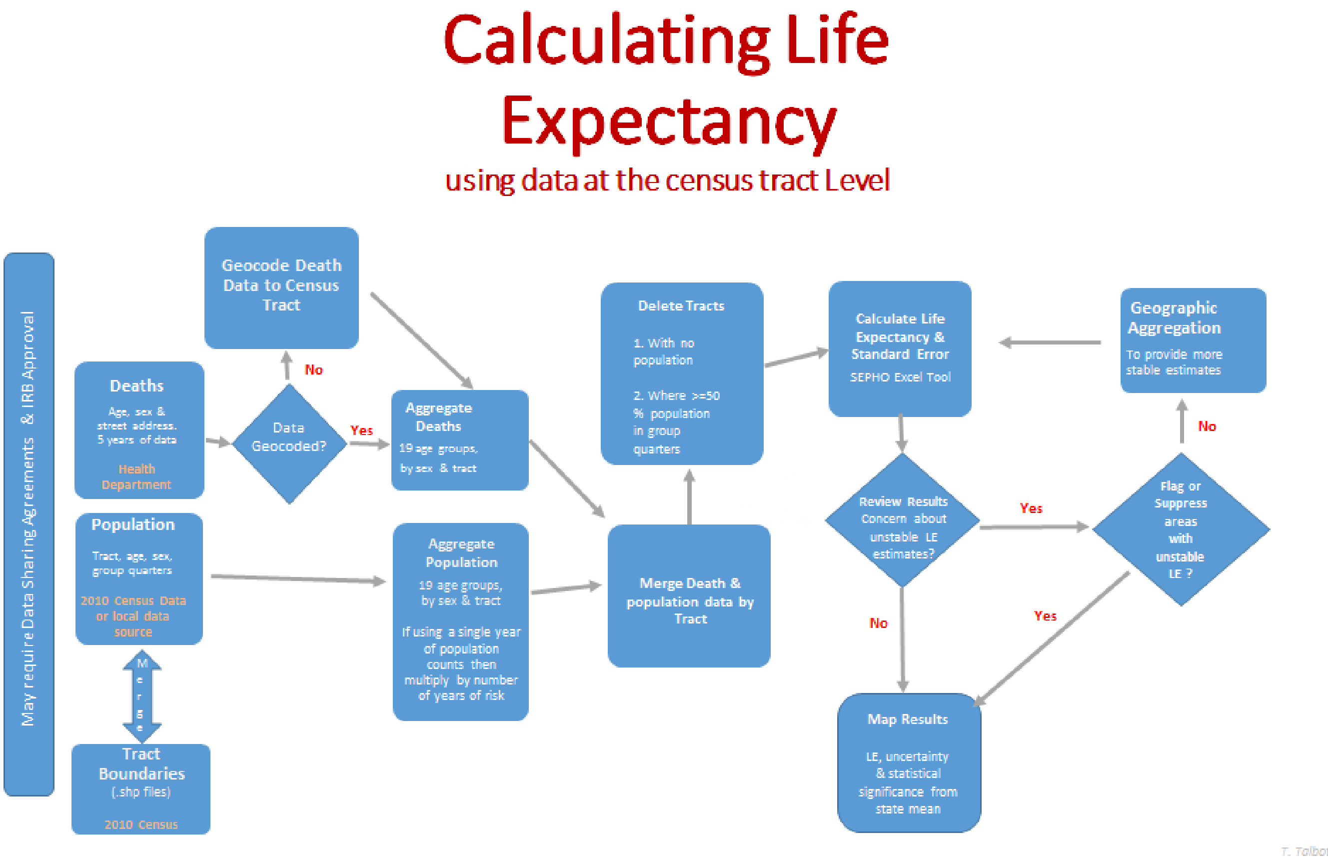
7.2 SAS or Stata option

For a jurisdiction wishing to use a SAS- or Stata-based tool for calculating life expectancy, please contact CSTE, who will connect you with a SCALE workgroup member. Numerator and denominator data need to be created as described above for the SEPHO tool.

7.3 Flowchart

The overall workflow will look similar to the flowchart below, created by a Phase I Workgroup member. Most Workgroup members needed to run through several iterations to reach a final product that was ready for analysis, display, and distribution, as they discovered more about the data with which they were working.

Figure 12. A flowchart of activities for calculating LE.



8.0 Interpreting the Findings

8.1 Results

Numerator: Of the Phase I participants, four used death data from 2008-2012; two used more recent data (2009-2013), and one jurisdiction with more sparse populations used 10 years of data (2001-2010) to generate sub-county LE.

Denominator: Four jurisdictions used 2010 Census population estimates while three had local population estimates. All participants found the SEPHO tool easy to use, and customizable for the specific type of geographies pertinent to their locality.

Each jurisdiction examined the output of the numerator, denominator, and the standard errors, and made some decisions about what to present. One Phase I participant (Florida) initially used ZIP codes as it aligned with other sub-county health and behavioral indicators that were already produced. Maine needed to group 10 years of data and provide them at the Minor Civil Division (MCD), which is a Census designation for a civil township, precinct, or magisterial district. The other jurisdictions produced census tract estimates.

Many jurisdictions found that mapping the data after the first pass through the SEPHO tool led to ask more questions (see section 8.2). The Workgroup discussed sensitivity analyses, what to consider unreliable estimates, how to handle unreliable estimates, and special considerations around specific geographic types.

Phase II: Most Phase II participants were adopting recommendations from the Phase I group. The majority of participants (85%) were able to get data and complete calculations. Of these, most had shared internally (91%) and 7% had shared with audiences external to the health department.

Given the high rate of success for Workgroup member, the ease of these tools in developing sub-county LE can easily be adopted in most jurisdictions.

8.2 Special considerations

Areas with unexpectedly high or low LE

Border areas

One question to ask the data provider is if the vital statistics department has a reciprocal agreement with other states. This means that a resident of State A dies in another state, State B, that State B would send a death certificate to State A. Not having this agreement means that LE might be artificially inflated in areas that are close to borders, especially ones that have a hospital or nursing facility on one side of the border, as individuals might be seen in the hospital or nursing facility in State B, die there, and State A would not know. Border issues typically arise between county, country, and state lines. If there is not a reciprocal agreement, it may be worth contacting the other jurisdiction to see what options might be available to establish one, or present the data with caveats for those areas.

Small populations:

In some cases, tracts (or other sub-county geography) may not meet the recommended minimum size of 5,000 person years at risk due to a small number of people residing in that geography. Others may have only a few deaths. Workgroup participants handled small populations in one of several ways. Some suppressed the data. Some calculated the LE, and if it had a reasonable standard error, the data were presented. Others grouped additional years of data or aggregated geographies together. For example, NYS aggregated tracts to get to 60 deaths, using their R-based Geographic Aggregation Tool. [17] In addition, there are some tracts with unique qualities that make calculation of LE challenging.

Unique tracts:

Examples of unique tracts includes ones that are primarily assigned to an area that has few residents, such as a prison, an airport or ones that contain a large number of university students or a hospital or nursing facility.

Tracts assigned specifically to an airport or to a prison tend to have very small populations, if any are included in the area at all. University populations, prisons and hospitals may have deaths recorded at that location, but the problem is that the person resides somewhere else and isn't captured in the population at risk. Nursing homes may also show a spike in the number of deaths in that tract as the facility could be put on the death certificate as the residence.

[17] http://www.albany.edu/faculty/ttalbot/GAT/

One jurisdiction identified these using data from the Census or ACS, excluding tracts that have more than 50% of the population living in group quarters. Deaths can still occur in those geographies, but the individuals would have an official residence somewhere else, inflating the LE estimate for that geography. Other members of the Workgroup suggested adopting this standard.

Standard errors and confidence intervals

Standard error (and hence confidence limits) increase as population decreases. Literature suggests that a population of 5,000 life years at risk produces an 'acceptable' standard error of +/– 2 years (or a 95% confidence limit of +/– 4 years). Phase I and Phase II participants found that even with 5,000 person years at risk, there were still small areas with a standard error outside the range of 2. Workgroup participants handled it in a variety of ways – some suppressed SEs less than 2, some presented with caution notes.

Impact of Migration on LE

Populations that may be highly mobile or areas that have large numbers of influxing younger, healthy people, may see changes in life expectancy. This is an issue that the SCALE Workgroup discussed but has not examined in-depth. However, it remains important to realize that LE data and trends must be examined in the local context to understand what can be driving those changes.

9.0 Using the LE estimates

Policy makers and community members often ask what is driving the geographic disparities that are seen. The National Environmental Public Health Tracking Network (NEPHT) has produced list of useful related social-economic indicators which are important when looking at health outcomes such as LE. Virginia also computed Healthy Life Expectancy.

Florida

Florida was able to roll out the LE estimates into their environmental public health tracking network portal. While estimates were initially at a ZIP code level, they continue to work towards census tract estimates. In addition to presentations to policy makers, other environmental health, local areas, and CSTE, they correlated LE to an economic hardship index (see Section 13) within a specific county. These results were featured on the local news and community involvement has begun.

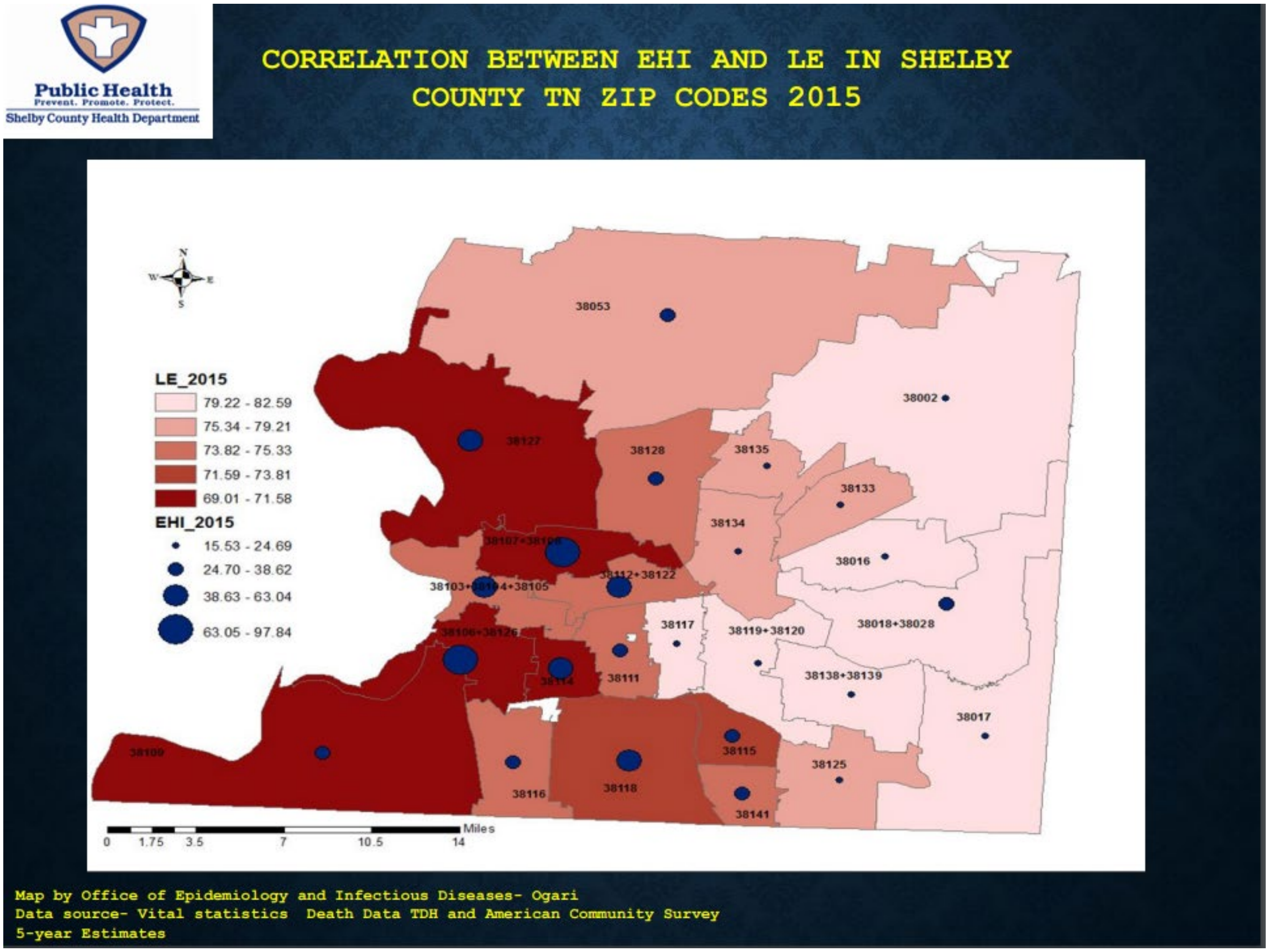
NYS

New York State released an article discussing their SCALE experience, including an in-depth geocoding process, discussing the aggregation methods, and methodological challenges they encountered along the way.

Shelby County, TN

Shelby County, TN showed a variation in LE from 71.6 to 82.6 years; a gap of more than a decade. A similar gap was seen when comparing whites to Black residents. After presenting this data to policy makers, the policy makers wanted to understand what might be driving these differences. Using the economic hardship index, they showed that the areas with the most economic hardship also had the lowest LE. By looking at each individual component, they were able to identify areas that were influencing the outcomes, and examined specific chronic diseases that are leading causes of death in Shelby County.[18]

Figure 13. EHI and LE in Shelby County, TN by ZIP code



10.0 Mapping and Display of LE results

Phase I and Phase II participants who had successfully generated LE were surveyed in August 2016, to get a sense of how they were choosing to display their results. Most participants were using static maps generated in mapping software, although a few had some interactive maps on the internet. Some participants were mapping other social determinants of health alongside the LE estimates, including education, poverty, lack of health insurance, and risk factor or behavioral data.

Even in creating static maps, no gold standard has yet emerged on the best way to present the data. All participants created a classified thematic or choropleth map where shades represent a range of values. Each tract falls into a specific "bin" or range. There was a wide variety of methods in how the map cutpoints were created.

Equal interval:

[18] https://cdn.ymaws.com/www.cste.org/resource/resmgr/pdfs/pdfs2/Moving_Upstream-Economic_Har.pdf, last accessed 12/2018

Each class consists of an equal data interval along the dispersion (from the high to the low point) of the data. Intervals are determined by dividing the range of all your data by the number of classes desired. Equal intervals are recommended if the data is distributed in a rectangular shape or if classification steps are nearly equal in size. The major disadvantage of this method is that class limits fail to reveal the distribution of the data along the number line. There may be classes that remain blank, which of course is not particularly meaningful on a map.

Natural Breaks:

Groups (or clusters) the data around different classes, with the goal of minimizing the deviation in each class while maximizing the difference between the other groups. In other words, it minimizes value differences between the data within the same class and emphasize the differences between the classes.

A disadvantage of this method is that class limits may vary from one map-maker to another due to the author's subjective class definition (Good graphic way of determining natural group of similar values by searching for significant depressions in frequency distribution. Minor troughs can be misleading and may yield poorly defined class boundaries.

Quantiles:

This is a rank order approach, using an equal number of observations into each class, generated by taking the total observations and dividing by the desired number of classes. Each class is approximately equally represented. However, it may introduce an apparent pattern of dissimilarity if there is none, overweight the data, or there may be gaps between the observations.

Standard Deviation:

This method measure how the value is distributed along a dispersion graph, and it show the standard deviation from the statistical mean of our dataset. The resulting classes reveal the frequency of elements in each class. Standard deviation may be useful to show statistical significance or to direct attention to the high and low valued. It's typically best for a standard normal distribution. One downside is it doesn't list the exact LE rate, and may be difficult for a layperson to understand.

Table 2.

Methods employed to generate legends (n=22)	
Natural breaks Jenks	6
Quantile	4
Defined interval	1
Standard deviation	1
Geometric interval	1
Other*	3

*included manual classification into bins and difference from statewide average.

Another display/mapping issue includes determining how to represent unstable rates. Some chose to grey out or crosshatch unreliable data. Some suppressed; others aggregated sub-county units until they did not need to suppress the data.

Table 3.

Methods employed to indicate areas with high SE (n=12)*	
Greying out	5
Crosshatch	3
Suppressed	3
Aggregated	2

*Participants indicated multiple methods were employed

11.0 Economic hardship index: incorporating SDOH with LE

There are a number of indices that bring together elements of the Social Determinants of Health (SDOH) that may lend itself well to starting to explain the variation and inequalities seen with LE.

Economic Hardship Index

The Economic Hardship Index (EHI) was developed in 1976 by Nathan and Adams, and focuses on 6 key social determinants, all of which have data available from the American Community Survey. They include: *unemployment* (for populations over the age of 16 years); *education* (the % of the population over 25 years of age without a high school diploma); *per capita income level* (the amount of money earned per person); *poverty* (% below the federal poverty level); *crowded housing* (housing units with more than one person per room), and *dependency* (population under 18 or over 64 years of age). Shelby County, TN was the first SCALE participant to use this index with LE, and it was presented as part of the CSTE Annual Conference Workshop in 2017.

Examples of EHI, relationship to health, and variations in included variables

San Antonio[19]

In 2017, a researcher mapped EHI in San Antonio, TX based on these six metrics:

- Percent of population age 25 and over below the poverty level;
- Per capita income;
- Unemployment levels for those 16 and over;
- Educational attainment for those 25 and over, which equates to having a high school diploma or a GED;
- Crowding level, which is the percentage of housing units with more than one occupant per room; and
- Dependency amount, the (combined) percentage of the population under age 18 and above age 65.

She found a “crescent of comfort” in comparison to a concentration of “more difficult living” as well as strong correlation between % below poverty and low educational attainment. Among the 4 highest

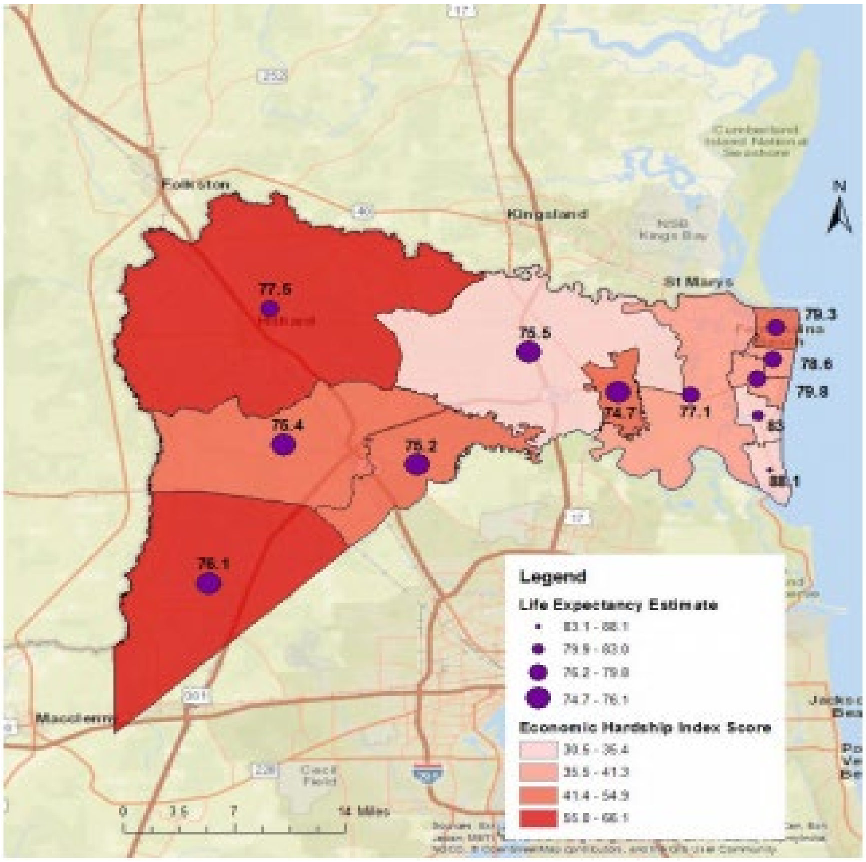
[19] https://www.huffingtonpost.com/entry/san-antonio-and-the-geography-of-poverty_us_59cbcc8ce4b02ba6621ff986, last accessed 12/2018

hardship ZIP codes, they had worse poor health outcomes, low educational attainment, high poverty, and low housing values.

Nassau County, FL[20]

An analysis showed higher LE is associated with less economic hardship. LE in high and low economic areas differ by almost 14 years.

Figure 14. EHI and LE in Nassau County, FL.

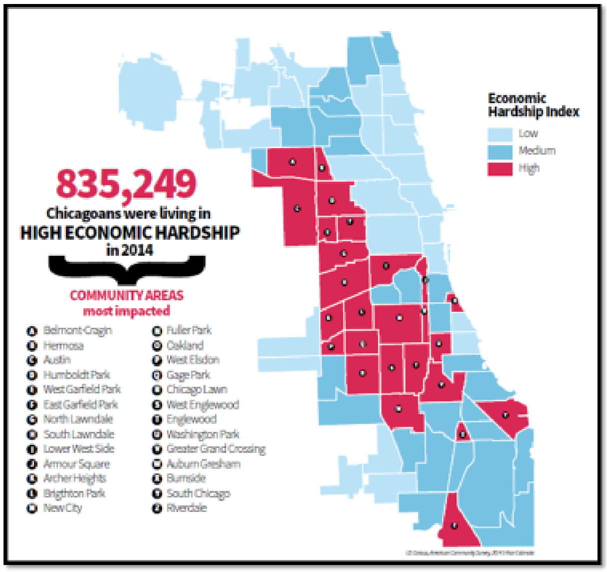


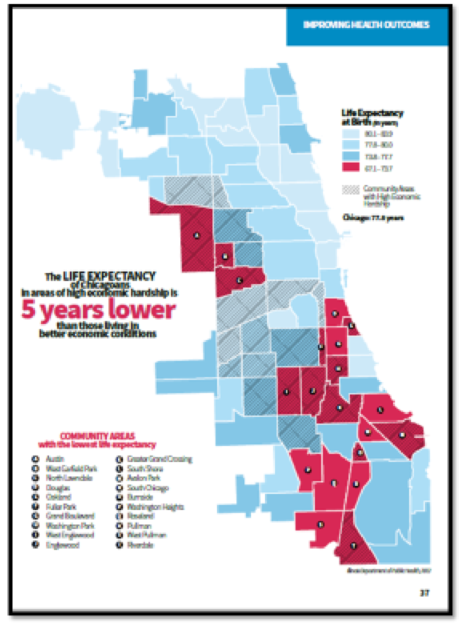
[20] http://www.thecountyinsider.com/uploads/2/4/7/0/24700962/final_press_release_-__april_6__public_health_week__right_to_health.pdf, last accessed 12/2018

[Chicago, IL][21]

Chicago found that the LE in high economic hardship areas was 5 years less than those in better economic conditions. They also compared other health indicators and found a similar pattern.

Figure 15. EHI and LE in Chicago, IL.





See Appendix C for more examples.

Steps to calculate EHI:

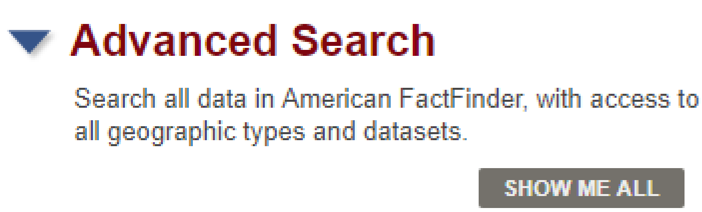
If using the ACS data, it is available in two forms: 1 year and 5 year. Most sub-county data such as the ZIP code/ZCTA[22] or census tract data uses the 5 year aggregations. Users can select the EHI timeframe that matches the LE data.

This example uses data from the 2011-2015 ACS dataset.

1. Navigate to the American Community Survey FactFinder website.

https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t

If you get there by a web search, on the start page, make sure to click “Advanced Search”



[21] https://greatcities.uic.edu/2016/09/19/economic-hardship-index-shows-stark-inequality-across-chicago/, last accessed 12/2018

[22] ZCTA: Zip Code Tabulation Area, which are approximations of ZIP codes as the boundaries were established in 2010

Once on the page, click “*geographies*” on the side bar, and then the “*name*” tab in the box (Figure X)

Figure 16. Navigating FactFinder start page for Geographies and Name tab.

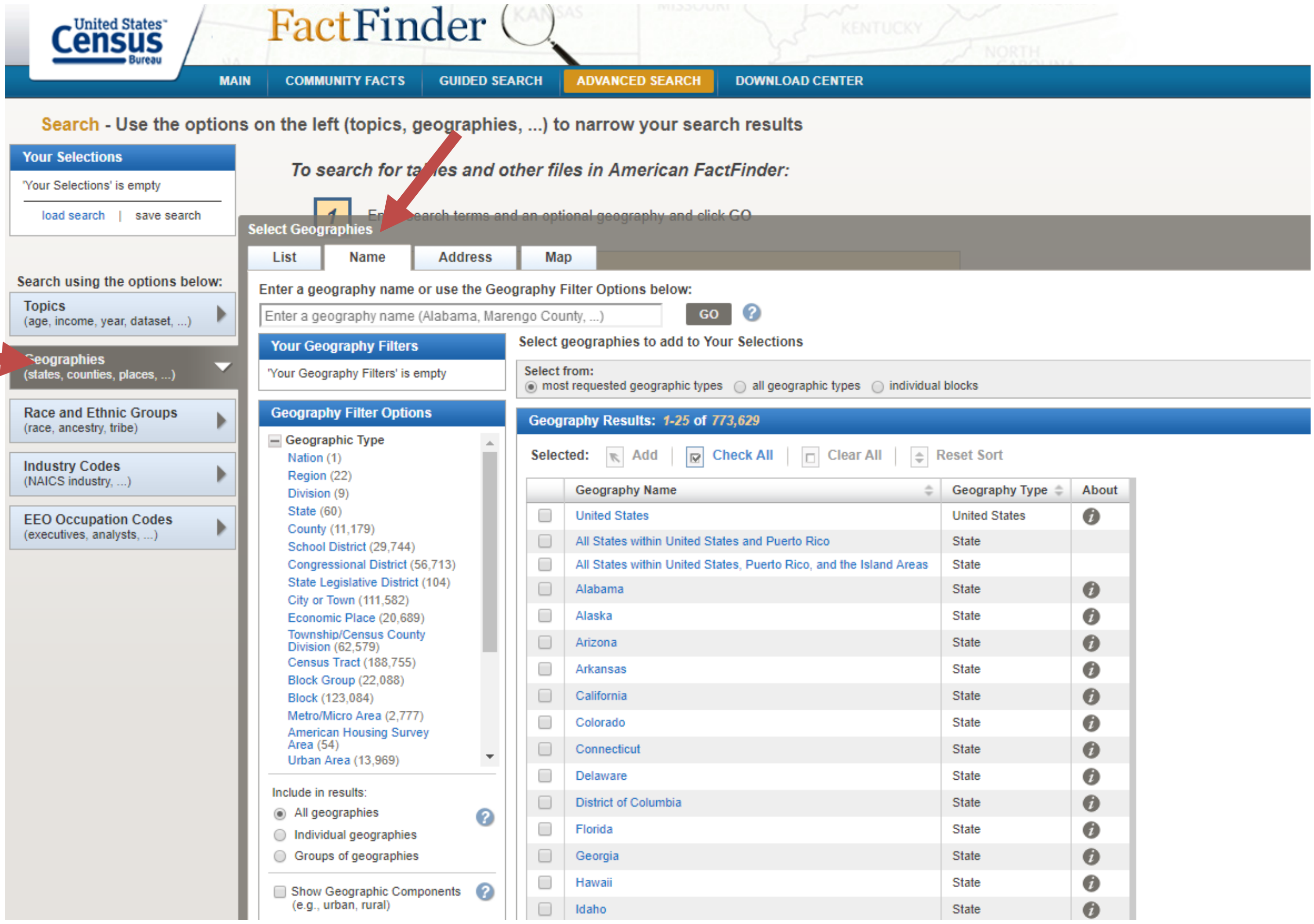
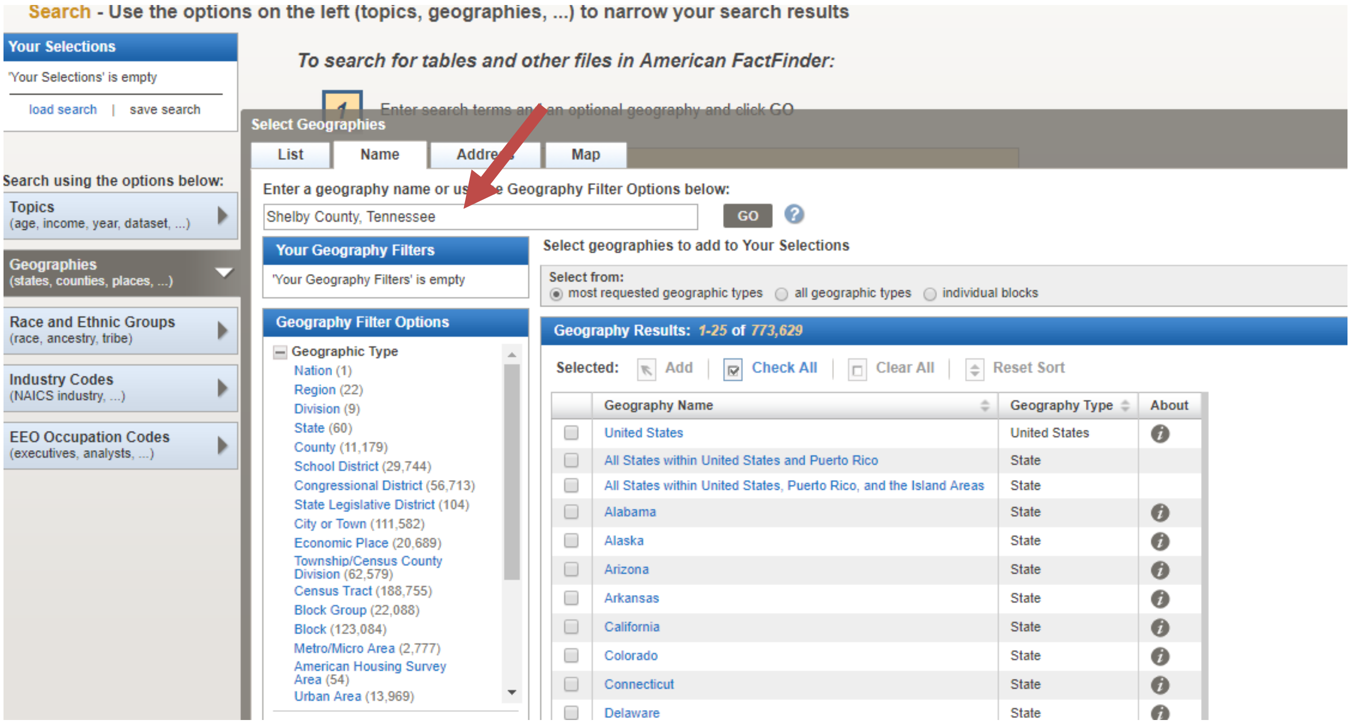
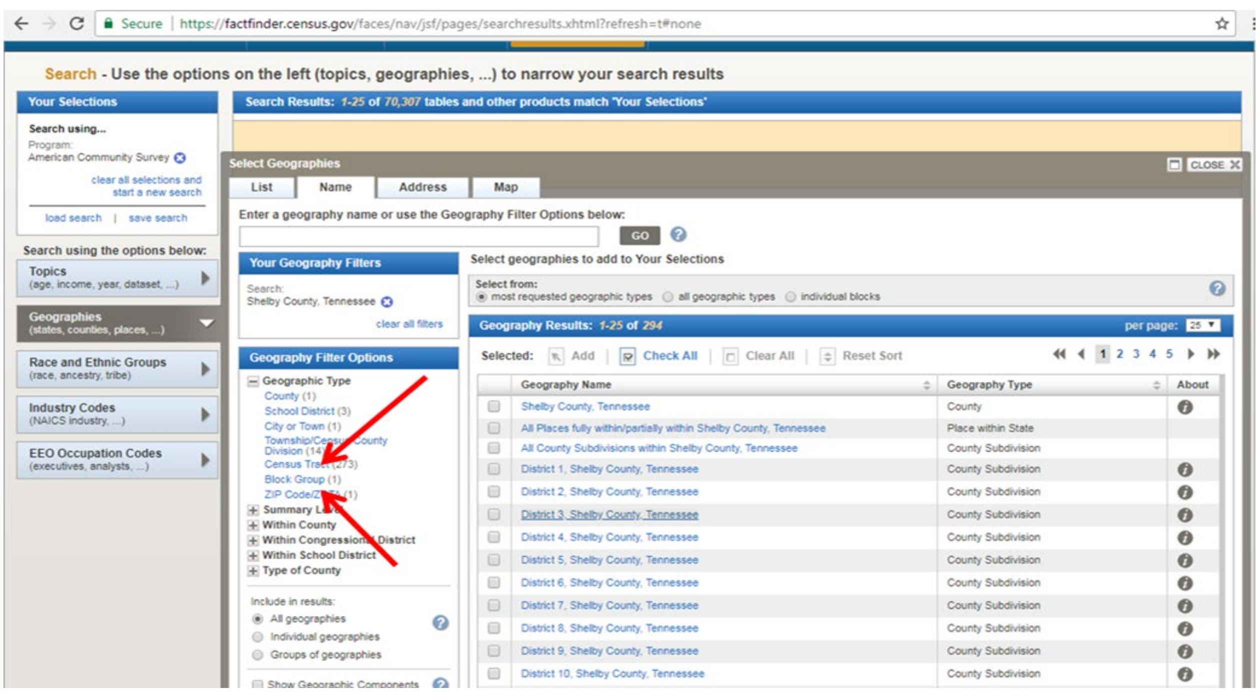


Figure 17. Choosing *Names* tab on the FactFinder site.

2. Enter the name of the specific geography: it could be a state, a city, or a county. This example uses Shelby County, TN.



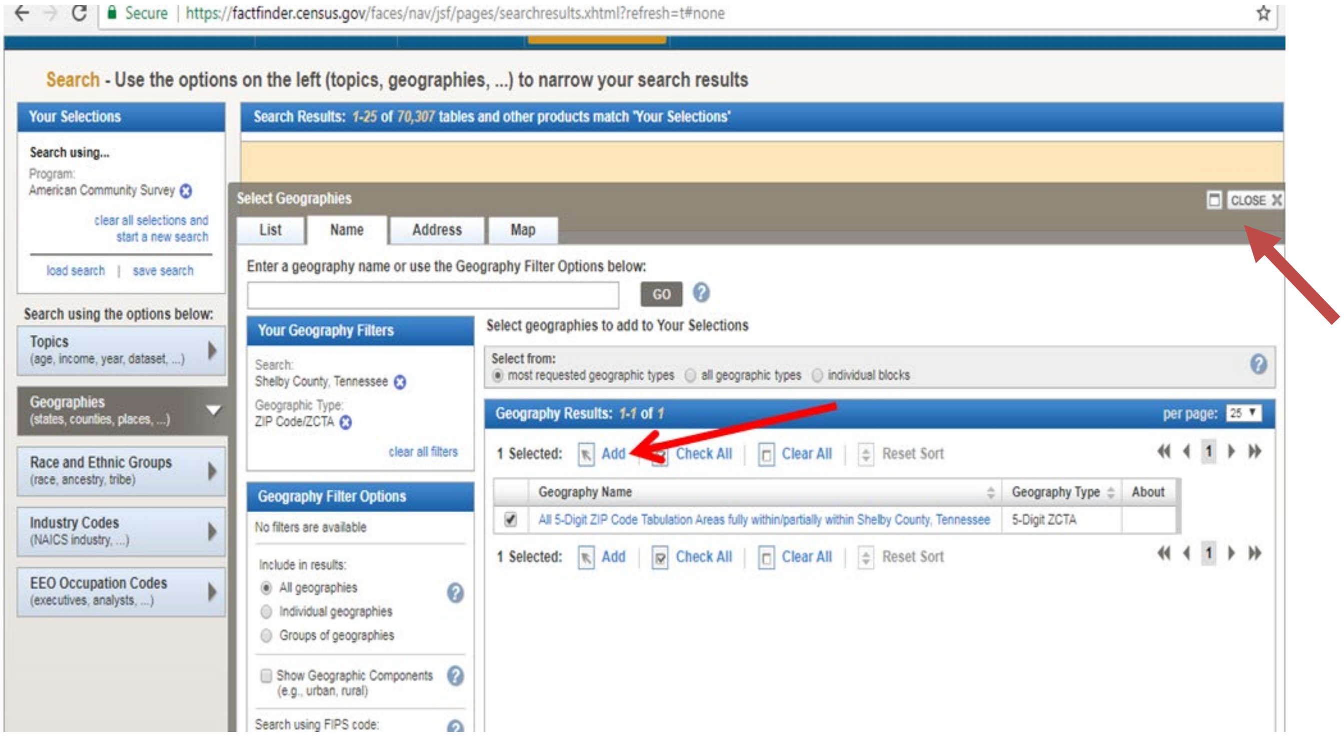
3. Under the *Geography Filter* options, select "census tract" if the LE data is at the tract level, or "ZIP Code/ZCTA[23]" if the data is at a ZIP code level. You will see the box that says "Your Geography Filters" update with the geographies you pick.



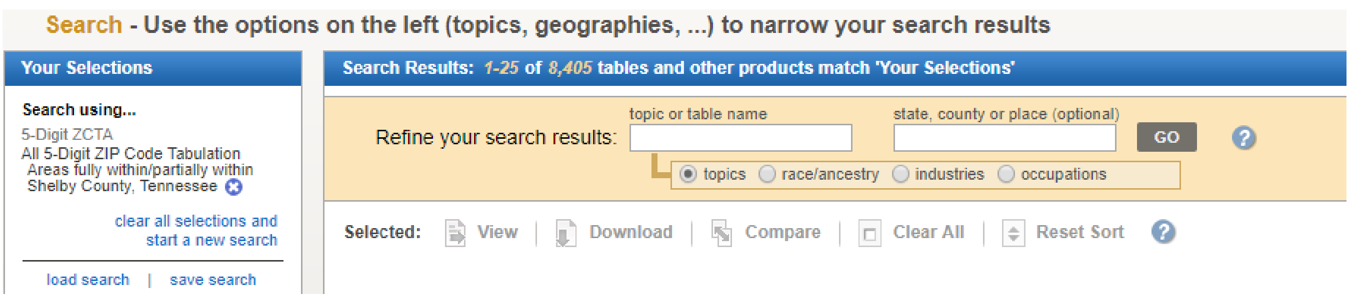
4. Once the geography has been checked, click *add to your selection*, then click the *close* button 

.

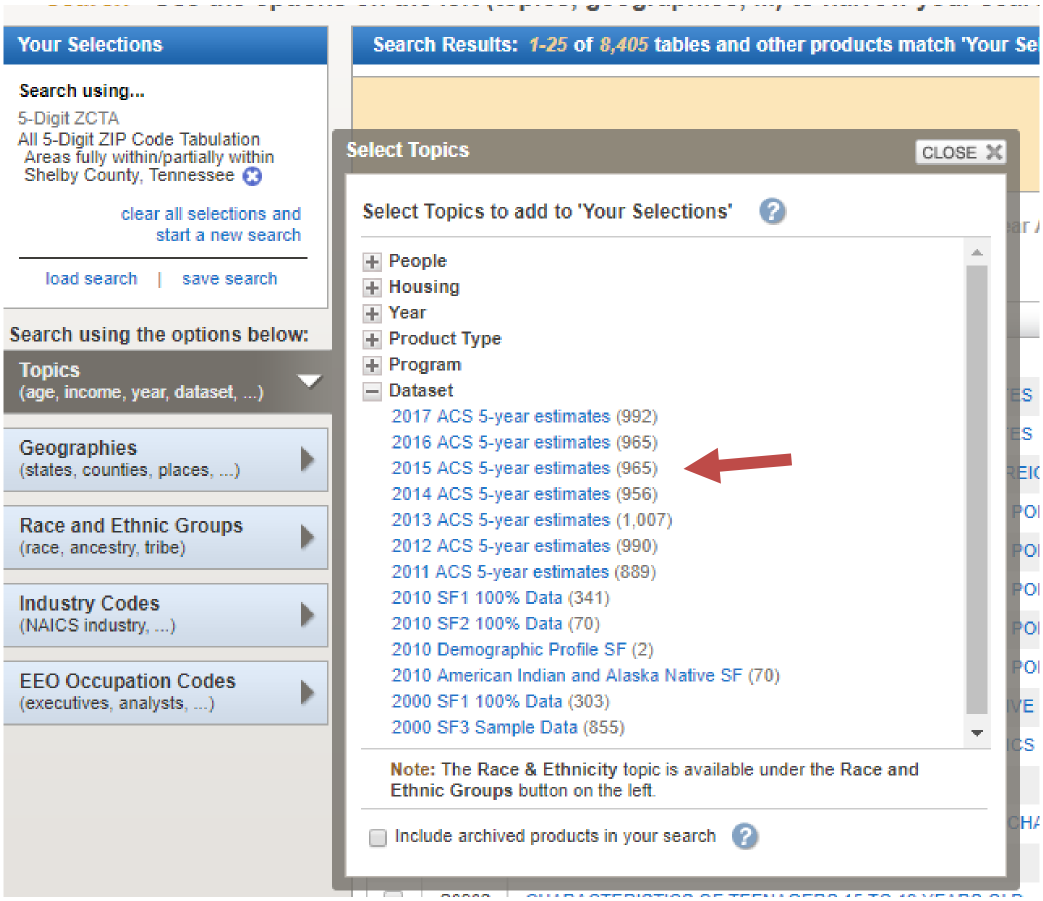
[23] ZCTA: Zip Code Tabulation Area, which are approximations of ZIP codes as the boundaries were established in 2010.



After closing the selection box, you should see the “Your Selections” box in the top-left hand box update with your choices.



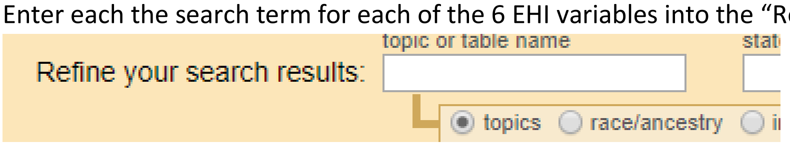
5. Select the dataset you want to use for the ACS variables. For the most part, you will be using ACS 5-year aggregated data for the ZCTA or tract geography. Wherever possible, try to align the dates for the ACS and for the data used. In this example, we will be using data from the 2015 5-year data.



Select 2015 ACS 5-year estimates; it will add to "your selections" and then click the 

button.

6. Enter each the search term for each of the 6 EHI variables into the "Refine your search results" box.



T

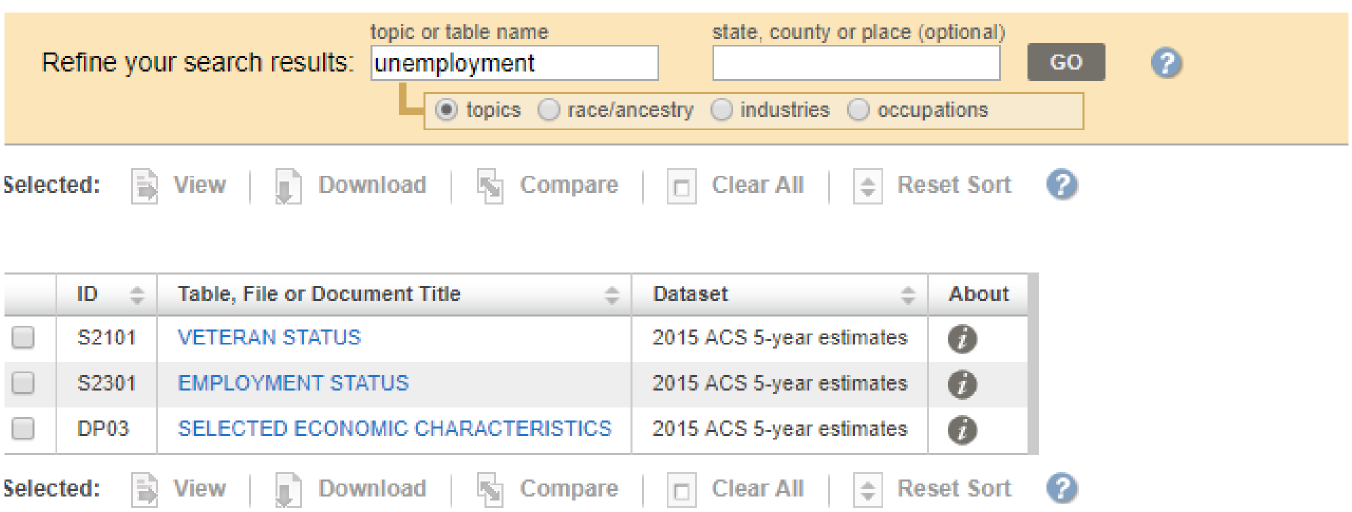
Table 6a has the topic name/concept, description from the Census and the table name and variable, along with any calculation notes.

Table 6a. Variables used in search; description and table and variable names.

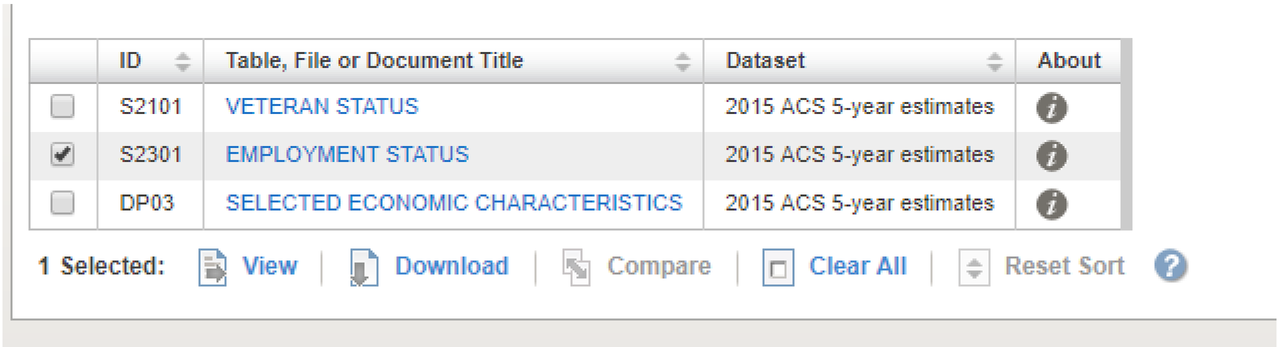
VARIABLE TOPIC	DESCRIPTION CENSUS	CENSUS TABLE & Variable Name	CALCULATION DESCRIPTION NOTES
Unemployment	Unemployment rate; Estimate; Population 16 years and over	ACS_5_YR_S2301 **Employment Status** HC04_EST_VC01	NA
Dependency	Total; Estimate; SELECTED AGE CATEGORIES - 18 years and over	ACS_5_YR_S0101 **Age and Sex** HC01_EST_VC25, HC01_EST_VC33	HC02_EST_VC25+ HCO2_EST_VC33

	Total; Estimate; SELECTED AGE CATEGORIES - 65 years and over		
No HS Diploma population 25 years and over	Percent; Estimate; Percent high school graduate or higher	ACS_5_YR_S1501 **EDUCATIONAL ATTAINMENT** HC02_EST_VC17	100 – (HC02_EST_VC17)
Poverty	Percent below poverty level; Estimate; Population for whom poverty status is determined	ACS_5_YR_S1701 **POVERTY STATUS IN THE PAST 12 MONTHS** HC03_EST_VC01	NA
Crowded Housing	Percent; OCCUPANTS PER ROOM - Occupied housing units - 1.00 or less)	ACS_5_YR_DP04 **SELECTED HOUSING CHARACTERISTICS** HC03_VC113	100 – (HC03_VC113)
Per capita income in the past 12 months	Estimate; Per capita income in the past 12 months (in 2016 inflation-adjusted dollars)	ACS_5_YR_B19301 **PER CAPITA INCOME IN THE PAST 12 MONTHS** HD01_VD01	NA

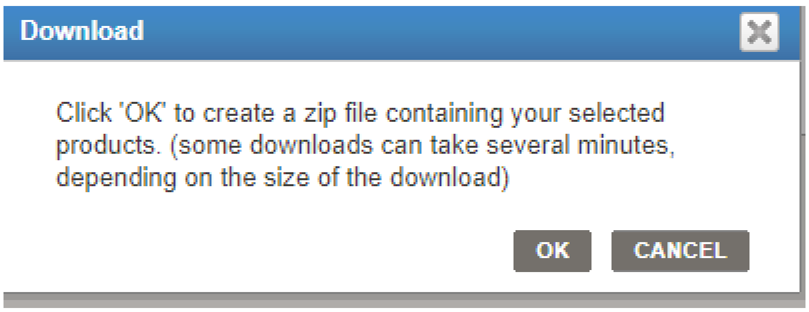
For example, enter “unemployment” into the box under “table or topic name” and click “go”.



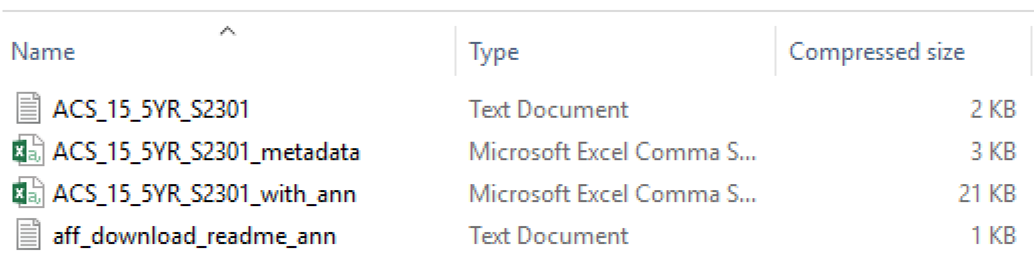
7. Selection options will be populated with 3 tables. Using Table X above, check the table ID (S2301), then select that by clicking the checkbox next [Download] to it, then click to download the data.



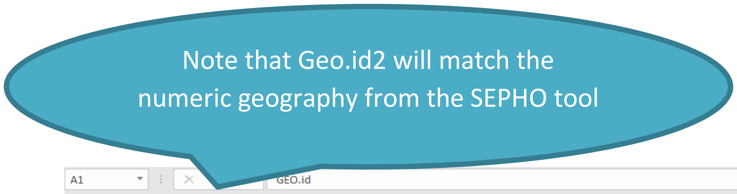
You will get a prompt to let you know it will create a .zip file. Click "OK"

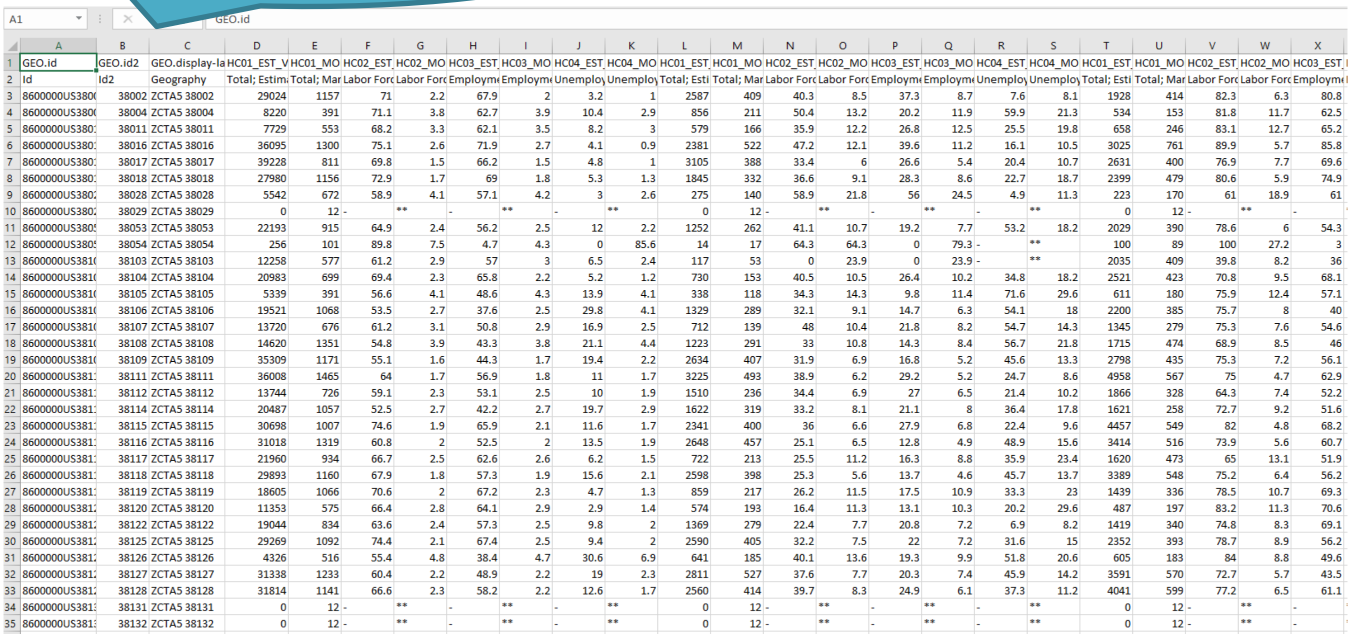


8. Open the .ZIP file, which likely saved to the \Download folder (on a Windows computer). Upon opening the .ZIP file, you will see the following files:



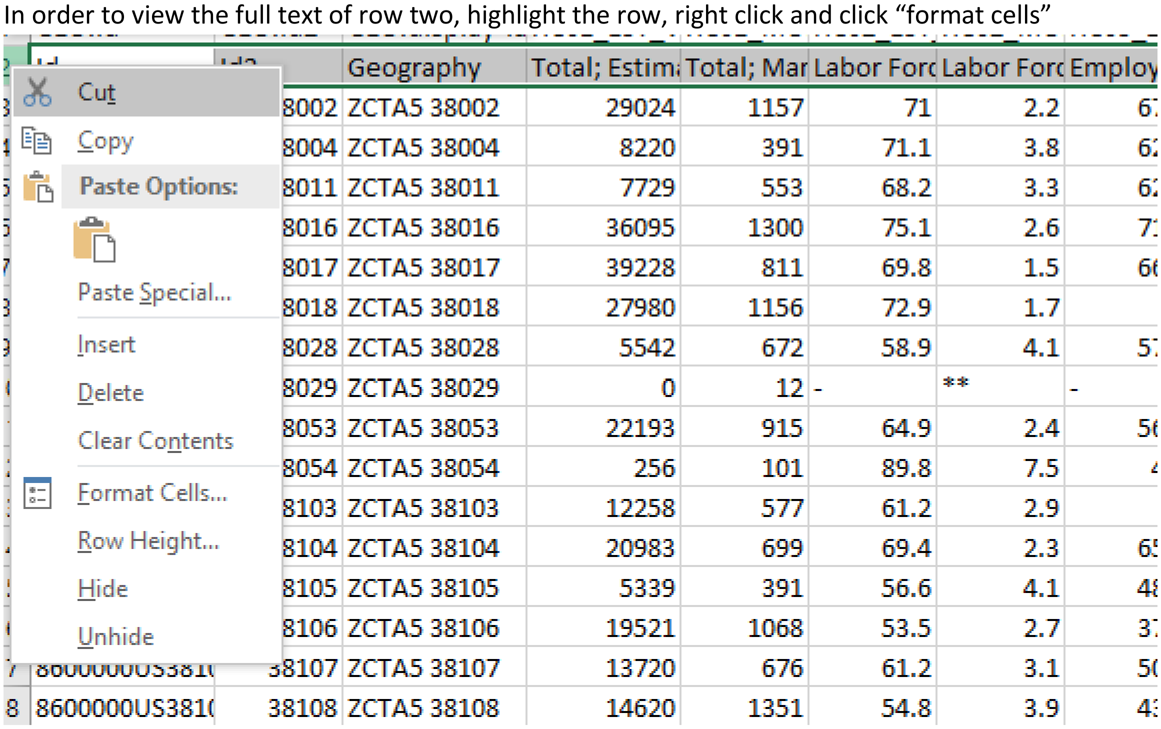
9. Open the ACS_15_5_YR_S2301_with_ann.xlsx, which contains the data. The spreadsheet will look similar.



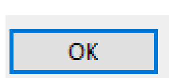


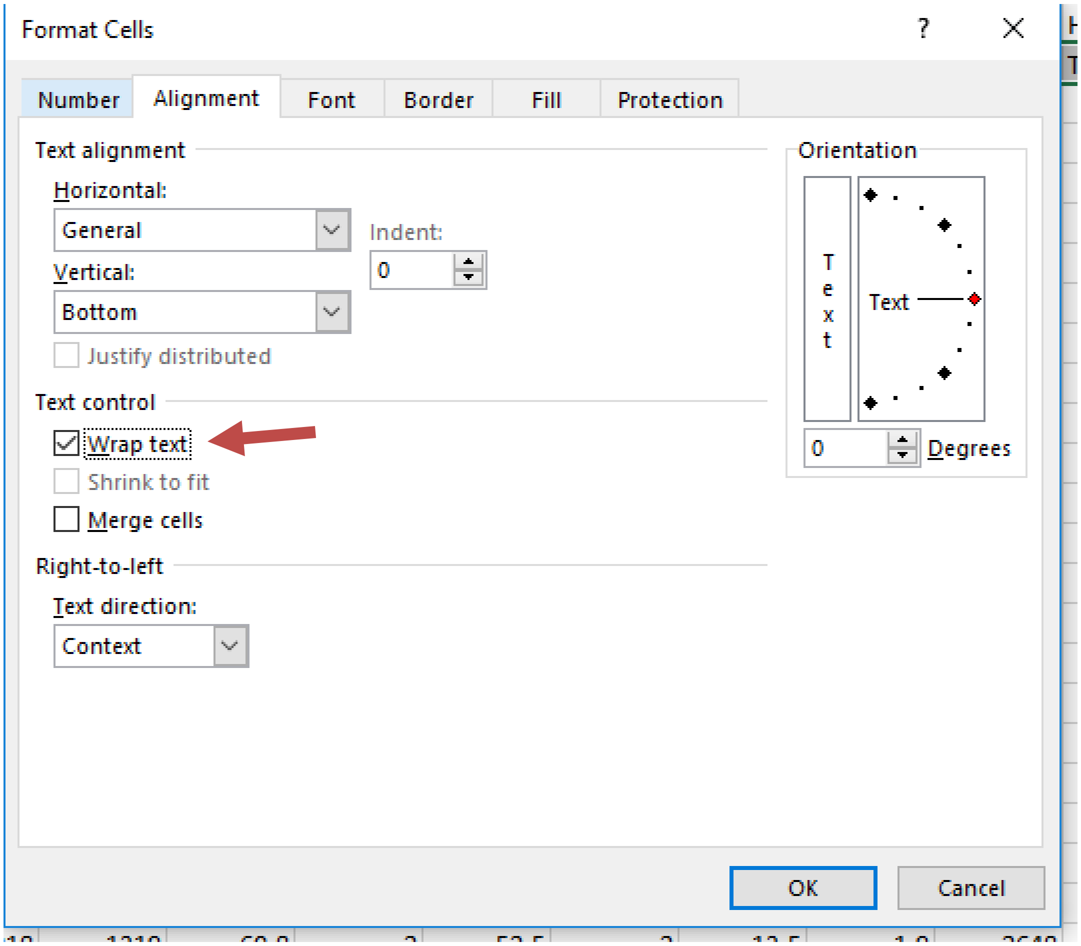
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	GEO.id	GEO.id2	GEO.display-la	HC01_EST_V	HC01_MO	HC02_EST_	HC02_MO	HC03_EST_	HC03_MO	HC04_EST_	HC04_MO	HC01_EST_	HC01_MO	HC02_EST_	HC02_MO	HC03_EST_	HC03_MO	HC04_EST_	HC04_MO	HC01_EST_	HC01_MO	HC02_EST_	HC02_MO	HC03_EST_
2	Id	Id2	Geography	Total; Estim	Total; Mar	Labor Forc	Labor Forc	Employme	Employme	Unemploy	Unemploy	Total; Esti	Total; Mar	Labor Forc	Labor Forc	Employme	Employme	Unemploy	Unemploy	Total; Esti	Total; Mar	Labor Forc	Labor Forc	Employme
3	8600000US380	38002	ZCTA5 38002	29024	1157	71	2.2	67.9	2	3.2	1	2587	409	40.3	8.5	37.3	8.7	7.6	8.1	1928	414	82.3	6.3	80.8
4	8600000US380	38004	ZCTA5 38004	8220	391	71.1	3.8	62.7	3.9	10.4	2.9	856	211	50.4	13.2	20.2	11.9	59.9	21.3	534	153	81.8	11.7	62.5
5	8600000US380	38011	ZCTA5 38011	7729	553	68.2	3.3	62.1	3.5	8.2	3	579	166	35.9	12.2	26.8	12.5	25.5	19.8	658	246	83.1	12.7	65.2
6	8600000US380	38016	ZCTA5 38016	36095	1300	75.1	2.6	71.9	2.7	4.1	0.9	2381	522	47.2	12.1	39.6	11.2	16.1	10.5	3025	761	89.9	5.7	85.8
7	8600000US380	38017	ZCTA5 38017	39228	811	69.8	1.5	66.2	1.5	4.8	1	3105	388	33.4	6	26.6	5.4	20.4	10.7	2631	400	76.9	7.7	69.6
8	8600000US380	38018	ZCTA5 38018	27980	1156	72.9	1.7	69	1.8	5.3	1.3	1845	332	36.6	9.1	28.3	8.6	22.7	18.7	2399	479	80.6	5.9	74.9
9	8600000US380	38028	ZCTA5 38028	5542	672	58.9	4.1	57.1	4.2	3	2.6	275	140	58.9	21.8	56	24.5	4.9	11.3	223	170	61	18.9	61
10	8600000US380	38029	ZCTA5 38029	0	12	-	**	-	**	-	**	0	12	-	**	-	**	-	**	0	12	-	**	-
11	8600000US380	38053	ZCTA5 38053	22193	915	64.9	2.4	56.2	2.5	12	2.2	1252	262	41.1	10.7	19.2	7.7	53.2	18.2	2029	390	78.6	6	54.3
12	8600000US380	38054	ZCTA5 38054	256	101	89.8	7.5	4.7	4.3	0	85.6	14	17	64.3	64.3	0	79.3	-	**	100	89	100	27.2	3
13	8600000US381	38103	ZCTA5 38103	12258	577	61.2	2.9	57	3	6.5	2.4	117	53	0	23.9	0	23.9	-	**	2035	409	39.8	8.2	36
14	8600000US381	38104	ZCTA5 38104	20983	699	69.4	2.3	65.8	2.2	5.2	1.2	730	153	40.5	10.5	26.4	10.2	34.8	18.2	2521	423	70.8	9.5	68.1
15	8600000US381	38105	ZCTA5 38105	5339	391	56.6	4.1	48.6	4.3	13.9	4.1	338	118	34.3	14.3	9.8	11.4	71.6	29.6	611	180	75.9	12.4	57.1
16	8600000US381	38106	ZCTA5 38106	19521	1068	53.5	2.7	37.6	2.5	29.8	4.1	1329	289	32.1	9.1	14.7	6.3	54.1	18	2200	385	75.7	8	40
17	8600000US381	38107	ZCTA5 38107	13720	676	61.2	3.1	50.8	2.9	16.9	2.5	712	139	48	10.4	21.8	8.2	54.7	14.3	1345	279	75.3	7.6	54.6
18	8600000US381	38108	ZCTA5 38108	14620	1351	54.8	3.9	43.3	3.8	21.1	4.4	1223	291	33	10.8	14.3	8.4	56.7	21.8	1715	474	68.9	8.5	46
19	8600000US381	38109	ZCTA5 38109	35309	1171	55.1	1.6	44.3	1.7	19.4	2.2	2634	407	31.9	6.9	16.8	5.2	45.6	13.3	2798	435	75.3	7.2	56.1
20	8600000US381	38111	ZCTA5 38111	36008	1465	64	1.7	56.9	1.8	11	1.7	3225	493	38.9	6.2	29.2	5.2	24.7	8.6	4958	567	75	4.7	62.9
21	8600000US381	38112	ZCTA5 38112	13744	726	59.1	2.3	53.1	2.5	10	1.9	1510	236	34.4	6.9	27	6.5	21.4	10.2	1866	328	64.3	7.4	52.2
22	8600000US381	38114	ZCTA5 38114	20487	1057	52.5	2.7	42.2	2.7	19.7	2.9	1622	319	33.2	8.1	21.1	8	36.4	17.8	1621	258	72.7	9.2	51.6
23	8600000US381	38115	ZCTA5 38115	30698	1007	74.6	1.9	65.9	2.1	11.6	1.7	2341	400	36	6.6	27.9	6.8	22.4	9.6	4457	549	82	4.8	68.2
24	8600000US381	38116	ZCTA5 38116	31018	1319	60.8	2	52.5	2	13.5	1.9	2648	457	25.1	6.5	12.8	4.9	48.9	15.6	3414	516	73.9	5.6	60.7
25	8600000US381	38117	ZCTA5 38117	21960	934	66.7	2.5	62.6	2.6	6.2	1.5	722	213	25.5	11.2	16.3	8.8	35.9	23.4	1620	473	65	13.1	51.9
26	8600000US381	38118	ZCTA5 38118	29893	1160	67.9	1.8	57.3	1.9	15.6	2.1	2598	398	25.3	5.6	13.7	4.6	45.7	13.7	3389	548	75.2	6.4	56.2
27	8600000US381	38119	ZCTA5 38119	18605	1066	70.6	2	67.2	2.3	4.7	1.3	859	217	26.2	11.5	17.5	10.9	33.3	23	1439	336	78.5	10.7	69.3
28	8600000US381	38120	ZCTA5 38120	11353	575	66.4	2.8	64.1	2.9	2.9	1.4	574	193	16.4	11.3	13.1	10.3	20.2	29.6	487	197	83.2	11.3	70.6
29	8600000US381	38122	ZCTA5 38122	19044	834	63.6	2.4	57.3	2.5	9.8	2	1369	279	22.4	7.7	20.8	7.2	6.9	8.2	1419	340	74.8	8.3	69.1
30	8600000US381	38125	ZCTA5 38125	29269	1092	74.4	2.1	67.4	2.5	9.4	2	2590	405	32.2	7.5	22	7.2	31.6	15	2352	393	78.7	8.9	56.2
31	8600000US381	38126	ZCTA5 38126	4326	516	55.4	4.8	38.4	4.7	30.6	6.9	641	185	40.1	13.6	19.3	9.9	51.8	20.6	605	183	84	8.8	49.6
32	8600000US381	38127	ZCTA5 38127	31338	1233	60.4	2.2	48.9	2.2	19	2.3	2811	527	37.6	7.7	20.3	7.4	45.9	14.2	3591	570	72.7	5.7	43.5
33	8600000US381	38128	ZCTA5 38128	31814	1141	66.6	2.3	58.2	2.2	12.6	1.7	2560	414	39.7	8.3	24.9	6.1	37.3	11.2	4041	599	77.2	6.5	61.1
34	8600000US381	38131	ZCTA5 38131	0	12	-	**	-	**	-	**	0	12	-	**	-	**	-	**	0	12	-	**	-
35	8600000US381	38132	ZCTA5 38132	0	12	-	**	-	**	-	**	0	12	-	**	-	**	-	**	0	12	-	**	-

The first row contains the variable names used by the census. Row 2 contains annotations for what the code means. In order to view the full text of row two, highlight the row, right click and click "format cells"



And choose "wrap text" and then click





The row will expand to see the text in the field.

10. Locate the variable listed in the table. As an example, for unemployment, locate the field **HC04_EST_VC01,** which will give you the % of the population older than 16 years that is unemployed. No additional calculations are needed. Copy this row and paste into a clean spreadsheet. Also copy the Geo.id2 and paste in the spreadsheet.
11. Repeat for the other variables. Note that, in some cases, calculations are needed. For example, in the dependency section, you will need to add 2 columns and sum up the counts of those <18 and those >64, which are variables **HC02_EST_VC25** and **HC02_EST_VC33.** This will give you the % of the population that falls into the "dependency" category.
12. Arrange the data in the data in the Excel spreadsheet, and add an EHI column. Use the following formulas to calculate EHI

<u>For the calculation of all variables except for income:</u>
X = ((Y-Ymin)/(Ymax—Ymin))*100
Where : X = standardized value of component variable (for example, unemployment rate) for each area to be computed.
Y = unstandardized value of component variable for each area
Ymin = the minimum value for Y across all area
Ymax = the maximum value for Y across all area

For example:
Zip Code A has unemployment rate of 35%
Minimum unemployment rate across all Zip Codes is 5%
Maximum unemployment rate across all Zip Codes is 70%

EHI for Unemployment (X) = ((35-5)/ (70-5))*100

For the calculation of Income:

X = ((Y-Ymax)/(Ymin—Ymax))*100
For example
Zip Code A has average income of $40, 000
Minimum income across all Zip Codes $15,000
Maximum income across all Zip Codes $100,000

EHI for income (X) = ((40,000-100,000)/ (15,000-100,000))*100

Remember that higher EHI indicates greater hardship.

13. Map EHI against LE and look for patterns. It's also worthwhile to run correlations between each of the 6 individual components to compare patterns. SCALE Workgroup participants found the relationships to be different in each locality.

12.0 Other LE Methods since SCALE

USALEEP project

In 2018, the National Center for Health Statistics, the Robert Wood Johnson Foundation (RWJF), and the National Association for Public Health Statistics and Information Systems (NAPHSIS) produced some small area LE called the US Small-area Life Expectancy Estimate Project (USALEEP). They use a very different methodology[24] from the SCALE project, including using 6 years worth of death and population data (2010-2015), using the Decennial Census population, and developing population estimates using American Community Survey survey-based estimates, and imputing rates in age groups that had zero death counts over the 6 year period. They limited analyses to tracts with more than 5,000 people over the 6-year period, and modeled estimates based on the tracts with sufficient death and population counts. SCALE Workgroup participants are evaluating the differences between the USALEEP and the direct estimates as created by SCALE. The Guide will be updated once more information is available.

Data can be found at the following location. https://www.cdc.gov/nchs/nvss/usaleep/usaleep.html

13.0 Summary

The SCALE project successfully piloted a simple way for local jurisdictions to calculate LE at sub-county areas. There are many lessons to be learned from the process, including how to obtain the correct data, using an appropriate tool, and the power of collaboration. Many of the Phase I and Phase II participants have been able to use their LE results locally and several have papers about the process in the works. Due to the time constraints around data acquisition, cleaning, and production, many of the Workgroup participants are still in active process of developing estimates. The appetite for this work is also important to note: each time a Workgroup participant has presented about this project, additional jurisdictions continued to sign on and request additional information, as there is a large hunger for being able to analyze sub-county data, LE in particular. Additional work around data visualization

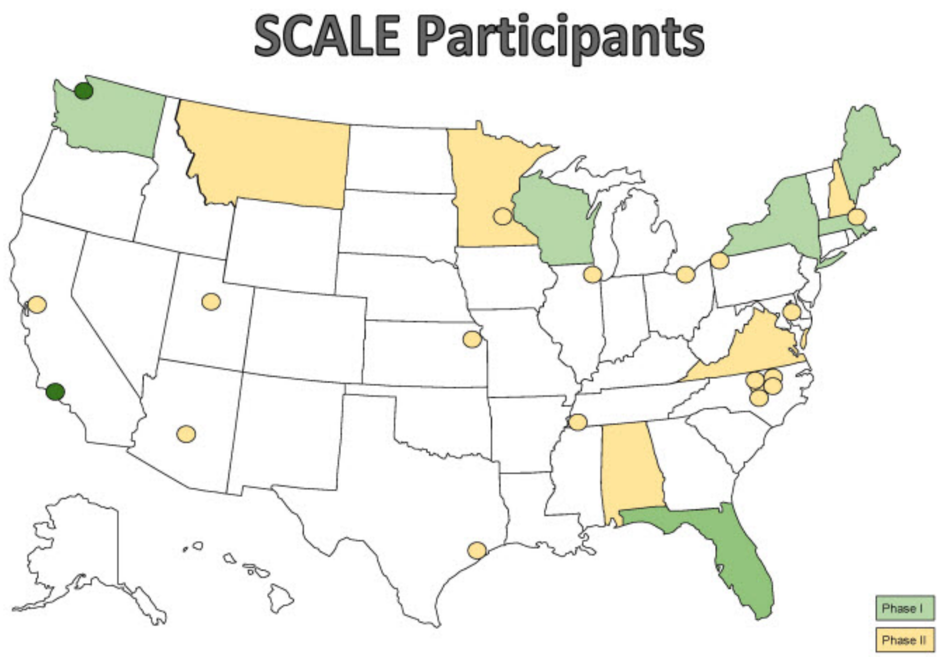
[24] Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski J. U.S. Small-area Life Expectancy Estimates Project: Methodology and Results Summary. National Center for Health Statistics. Vital Health Stat 2(181). 2018.

approaches, dissemination, and next steps are still on-going areas of discussion and development for the Workgroup. It is the Workgroup's hope that this interest will continue to be catalyzed and used for positive policy, system, and environmental changes in order to decrease health inequities.

Feedback about the guide or questions about the project may be directed to CSTE. http://www.cste.org/?page=SCALE&hhSearchTerms=%22scale%22

14.0 Acknowledgements

Many thanks to the CDC and CSTE staff who have supported this project, including Vickie Boothe, Jessica Wurster, Hayleigh McCall, and evaluator Leslie Fierro.



SCALE Phase I Participants

Florida
Los Angeles County, CA
Maine
Massachusetts
Washington
New York
Seattle & King County. WA
Wisconsin

SCALE Phase II Participants

Alabama
Maricopa Co., AZ
Alameda Co., CA
District of Columbia
Cook Co., IL
Alamance Co., NC
Caswell Co., NC
Chatham Co., NC
Durham Co., NC
Orange Co., NC
Johnson Co., KS
Erie Co., PA
Shelby Co., TN
Houston, TX
Montana
New Hampshire
Cleveland, OH
Metro Area Planning Council, MA
Washington Co., MN
Minnesota
Salt Lake Co., UT
Virginia

References

1. Toson, B. and A. Baker, *Life expectancy at birth: methodological options for small populations.* National statistics methodological series, 2003. **33**.
2. Los Angeles Department of Public Health, O.o.H.A.a.E., *Life Expectancy in LA County: How long do we live and why? A cities and community report.* . 2010.
3. Silcocks, P.B.S., D.A. Jenner, and R. Reza, *Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities.* Journal of Epidemiology and Community Health, 2001. **55**(1): p. 38-43.
4. Jonker, M.F., et al., *The impact of nursing homes on small-area life expectancies.* Health & Place, 2013. **19**(0): p. 25-32.
5. Harper, S., R. MacLehose, and J. Kaufman, *Trends in the black-white life expectancy gap among US states, 1990-2009.* Health affairs, 2014. **33**(8): p. 1375-82.
6. Adam Smith, J., S. Harper, and N. Auger, *Causes of widening life expectancy inequalities in Québec, Canada, 1989-2004.* Canadian Journal of Public Health, 2011. **102**(5): p. 375-81.
7. Stephens, A.S., et al., *Life expectancy estimation in small administrative areas with non-uniform population sizes: application to Australian New South Wales local government areas.* BMJ open, 2013. **3**(12): p. e003710-e003710.
8. Wang, H., et al., *Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010.* Lancet (London, England), 2012. **380**(9859): p. 2071-94.
9. Avendano, M. and I. Kawachi, *Why do Americans have shorter life expectancy and worse health than do people in other high-income countries?* Annual review of public health, 2014. **35**: p. 307-25.
10. Syme, S.L., *Reducing racial and social-class inequalities in health: the need for a new approach.* Health affairs, 2008. **27**(2): p. 456-9.
11. Banks, J., et al., *Disease and disadvantage in the United States and in England.* JAMA: the Journal of the American Medical Association, 2006. **295**(17): p. 2037-45.
12. Bezruchka, S., *The Hurrider I Go the Behinder I Get: The Deteriorating International Ranking of U.S. Health Status.* Annual review of public health, 2012. **33**(1): p. 157-173.
13. Squires, D., *The global slowdown in health care spending growth.* JAMA: the Journal of the American Medical Association, 2014. **312**(5): p. 485-6.
14. Woolf, S.H. and P. Braveman, *Where health disparities begin: the role of social and economic determinants--and why current policies may make matters worse.* Health Affairs, 2011. **30**(10): p. 1852-9.
15. Ezzati, M., et al., *The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the United States.* PLoS Medicine, 2008. **5**(4): p. e66.
16. Murray, C.J.L., et al., *Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States.* PLoS Medicine, 2006. **3**(9): p. e260.
17. Cullen, M., C. Cummins, and V. Fuchs, *Geographic and racial variation in premature mortality in the U.S.: analyzing the disparities.* PLoS ONE, 2012. **7**(4): p. e32930.
18. Wang, H., et al., *Left behind: widening disparities for males and females in US county life expectancy, 1985-2010.* Population Health Metrics, 2013. **11**(1): p. 8.
19. *Measuring the Risks and Causes of Premature Death: Summary of Workshops*. 2015, Washington DC: 2015 by the National Academy of Sciences.
20. Hunt, B., G. Tran, and S. Whitman, *Life Expectancy Varies in Local Communities in Chicago: Racial and Spatial Disparities and Correlates.* Journal of Racial and Ethnic Health Disparities, 2015: p. 1-9.

21. Chen, J., et al., *Mapping and measuring social disparities in premature mortality: the impact of census tract poverty within and across Boston neighborhoods, 1999-2001.* Journal of urban health, 2006. **83**(6): p. 1063-84.
22. King County Equity and Social Justice Annual Report, November 2014. http://www.kingcounty.gov/elected/executive/equity-social-justice.aspx, last checked 10/2015.
23. Song L, Mercer L, Wakefield J, Laurent A, Solet D. Using Small-Area Estimation to Calculate the Prevalence of Smoking by Subcounty Geographic Areas in King County, Washington, Behavioral Risk Factor Surveillance System, 2009–2013. Prev Chronic Dis 2016;13:150536.
24. Chiang, C.L., *A stochastic study of the life table and its applications. II. Sample variance of the observed expectation of life and other biometric functions.* Human biology, 1960. **32**: p. 221-238.
25. Bravo, JM and Malta, J. Estimating life expectancy in small population areas. United Nations Statistical Commission and Economic Commission for Europe. 23 April 2010.
26. Arias, E., *United States life tables, 2008.* National vital statistics reports, 2012. **61**(3): p. 1-64.
27. Chang, C.L., *Introduction to stochastic processes in biostatistics.* 1968.
28. Eayres, D. and E.S. Williams, *Evaluation of methodologies for small area life expectancy estimation.* Journal of epidemiology and community health, 2004. **58**(3): p. 243-9.
29. Scherbov, S. and D. Ediev, *Significance of life table estimates for small populations: Simulation-based study of estimation errors.* Demographic Research, 2011. **24**: p. 527-550.
30. Hadler, J.I. Lampkins, R., Lemmings, J., et al. Assessment of Epidemiology Capacity in State Health Departments—United States, 2013. MMWR Morb Mortal Wkly Rep 2015;64: 394-398.

Appendix A Summary of Peer-Reviewed Literature (to be added)

Appendix B Examples of MOU/DSA

Example 1

INFORMATION SHARING AGREEMENT

For

CONFIDENTIAL INFORMATION OR LIMITED DATASET(S)

Between
DATA PROVIDER
And
DATA RECIPIENT

This Agreement documents the conditions under which the DATA PROVIDER shares confidential information or limited Dataset(s) with other entities.

CONTACT INFORMATION FOR ENTITIES RECEIVING AND PROVIDING INFORMATION

	INFORMATION RECIPIENT	**INFORMATION ROVIDER**
Organization Name		
Business Contact Name		
Title		
Address		
Telephone #		
Email Address		
Fax #		
IT Security Contact		
Title		
Address		
Telephone #		
Email Address		
Privacy Contact Name		
Title		
Address		
Telephone #		
Email Address		
Data User Contact Name		
Telephone #		
Email Address		

I. DEFINITIONS:

Confidential Information means information that is protected from public disclosure by law. There are many state and federal laws that make different kinds of information confidential. LOCAL LAWS *the two most common are the Public Records Act, and the Healthcare Information Act.*

Limited Dataset means a data file that includes potentially identifiable information. A limited Dataset does not contain direct identifiers.

Potentially Identifiable Information means information that includes indirect identifiers which may permit linking an individual to that person's health care information. Examples of potentially identifiable information include:

birth dates,

admission, treatment or diagnosis dates,

healthcare facility codes

other data elements that may identify an individual. These vary depending on factors such as the geographical location and the rarity of a person's health condition, age or other characteristic.

Restricted Confidential Information means confidential information where especially strict handling requirements are dictated by statutes, regulations or contractual agreements. Violations may result in enhanced legal sanctions.

II. PURPOSE AND AUTHORITY/SCOPE OF AGREEMENT

PURPOSE: Data are used for routine community health assessment and surveillance. Analysis is provided for a variety of requestors, including the public, school districts, elected officials, students, and other departments within the agency. Only aggregate data are disclosed. Data may be stratified on race/ethnicity, income, gender, and age to examine disparities in the data and conduct small area analyses to look at geographic variations. The data are also used for departmental program evaluation.

Parties shall use the information described in this Agreement solely for the purpose stated this Agreement.

III. STATUTORY AUTHORITY TO SHARE INFORMATION

DATA PROVIDER statutory authority to disclose the confidential information or limited Dataset(s) identified in this agreement to the Information Recipient: LOCAL LAWS *(Disclosure of Vital Records for research purposes) and (disclosure of records without personal identifiers).*

Information Recipient's statutory authority to receive the confidential information or limited Dataset(s) identified in this Agreement: LOCAL LAWS

If the purpose is for research, has an Institutional Review Board (IRB) review and approval been received?

☐ Yes ☐ No

IV. PERIOD OF PERFORMANCE

This Agreement shall be effective from DATE1 through DATE2

V. DESCRIPTION OF INFORMATION

Information Provider will make available the following information under this agreement (Include the name of the database and a list of the data elements):

Death data

Geocoded death

The information described in this section is:

☐ Restricted Confidential Information

☐ Confidential Information

Potentially identifiable information

Any reference to information in this Agreement shall be the information as described in this Section.

VI. ACCESS TO INFORMATION

a. METHOD OF ACCESS/TRANSFER

DATA PROVIDER Web Application (indicate application name):

State Secure File Transfer Service Encrypted CD/DVD or other storage device

Other: (describe the methods for access/transfer)

Note: DATA PROVIDER IT Security Officer must approve “Other” prior to Agreement execution. IT Security Officer will send approval/denial directly to DATA PROVIDER Contracts Office and DATA PROVIDER Business Contact.

b. FREQUENCY OF ACCESS/TRANSFER

One time: DATA PROVIDER shall deliver information by_____(date) Repetitive: frequency or dates Annual files as available

As available within the period of performance stated in Section I.

OTHER PROVISIONS:

USE OF INFORMATION

The Information Recipient agrees to strictly limit use of information obtained or created under this Agreement to the purposes stated in the Agreement. For example, unless the Agreement specifies to the contrary the Information Recipient agrees not to:

link information received under this Agreement with any other information.

use information received under this Agreement to identify or contactindividuals.

The Information Recipient shall construe this clause to provide the maximum protection of the information that the law allows.

VII. SAFEGUARDING INFORMATION

a. CONFIDENTIALITY

Information Recipient agrees to:

limit access and use of the information:

To the minimum amount of information

The fewest people

For the least amount of time required to do the work.

Assure that all people with access to the information understand theirresponsibilities regarding it.

Assure that every person (e.g., employee or agent) with access to the information signs and dates the "Use and Disclosure of Confidential Information Form" (Example 1 Appendix A) before accessing the information.

Retain a copy of the signed and dated form as long as required in Data Disposition Section

The Information Recipient acknowledges the obligations in this section survive completion, cancellation, expiration or termination of this Agreement.

b. SECURITY

The Information Recipient assures that its security practices and safeguards meet STATE IT Security Standards:

For the purposes of this Agreement, compliance with the HIPAA Security Standard and the HITECH Act meets the ISB IT Security Standards.

The Information Recipient agrees to adhere to the Data Security Requirements in Example 1 Appendix B.

The Information Recipient further assures that it has taken steps necessary to prevent unauthorized access, use or modification of the information in any form.

c. BREACH NOTIFCATION

The Information Recipient shall notify the DATA PROVIDER IT Security Officer within two (2) business days of any suspected or actual breach of security or confidentiality of information covered by the Agreement.

Note: The DATA PROVIDER IT Security Officer must approve any changes to this section prior to Agreement execution. IT Security Officer will send approval/denial directly to DATA PROVIDER Contracts Office and DATA PROVIDER Business Contact.

d. RE-DISCLOSURE OF INFORMATION

Information Recipient agrees to not disclose in any manner all or part of the information identified in this Agreement except as the law requires, this Agreement permits, or with specific prior written permission by the DATA PROVIDER.

If the Information Recipient must comply with state or federal public record disclosure laws, and receives a records request where all or part of the information subject to this Agreement is responsive to the request: the Information Recipient will notify the DATA PROVIDER Privacy Officer of the request ten (10) business days prior to disclosing to the requestor. The notice must:

be in writing

include a copy of the request or some other writing that shows the:

date of the Information Recipient received the request

DATA PROVIDER records the Information Recipient believes are responsive to the request and the identity of the requestor, if known.

e. ATTRIBUTION REGARDING INFORMATION

Information Recipient agrees to cite "DATA PROVIDER" or other citation as specified, as the source of the information subject of this Agreement in all text, tables and references in reports, presentations and scientific papers. Other citation:

Information Recipient agrees to cite its organizational name as the source of interpretations, calculations or manipulations of the information subject of this Agreement.

VIII. REIMBURSEMENT TO DATA PROVIDER

Payment for services to create and provide the information is based on the actual expenses DATA PROVIDER incurs, including charges for research assistance when applicable.

Information Recipient agrees to pay DATA PROVIDER by check or account transfer within 30 calendar days of receiving the DATA PROVIDER invoice.

Upon expiration of the Agreement, any payment not already made shall be submitted within 30 days after the expiration date or the end of the fiscal year, which is earlier.

Charges for the services to create and provide the information are:

$______ or _______ No charge.

IX. DATA DISPOSITION

Unless otherwise directed in writing by the DATA PROVIDER Business Contact, at the end of this Agreement, or at the discretion and direction of DATA PROVIDER, the Information Recipient shall:

☐ Immediately destroy all copies of any data provided under this Agreement after it has been used for the purposes specified in the Agreement. Acceptable methods of destruction are described in Example 1 Appendix B. Upon completion, the Information Recipient shall submit the attached Certification of Data Disposition (Example 1 Appendix C) to the DATA PROVIDER Business Contact.

☐ Immediately return all copies of any data provided under this Agreement to the DATA PROVIDER Business Contact after the data has been used for the purposes specified in the Agreement, along with the attached Certification of Data Disposition (Example 1 Appendix C).

☐ Retain the data for the purposes stated herein for a period of time not to exceed

_______ (*e.g., one year, etc*.), after which Information Recipient shall destroy the data (as described below) and submit the attached Certification of Data Disposition (Example 1 Appendix C) to the DATA PROVIDER Business Contact.

☐ Other (Describe): Upon request by DATA PROVIDER program staff, at the end of the Agreement term, or when no longer needed, the Contractor shall return the Confidential Information or Sensitive Data to DATA PROVIDER information technology staff or the Contractor shall certify in writing that they employed a DATA PROVIDER approved method to destroy the information. The Contractor may obtain information regarding approved destruction methods from the DATA PROVIDER contact identified on the cover page of this Agreement.

X. AGREEMENT ALTERATIONS AND AMENDMENTS

This Agreement may be amended by mutual agreement of the parties. Such amendments shall not be binding unless they are in writing and signed by personnel authorized to bind each of the parties.

XI. CAUSE FOR IMMEDIATE TERMINATION

The Information Recipient acknowledges that unauthorized use or disclosure of the Information or any other violation of section VI may result in the immediate termination of this Agreement.

XII. CONFLICT OF INTEREST

The DATA PROVIDER may, by written notice to the Information Recipient:

Terminate the right of the Information Recipient to proceed under this Agreement if it is found, after due notice and examination by the Contracting Office that gratuities in the form of entertainment, gifts or otherwise were offered or given by the Information Recipient, or an agency or representative of the Information Recipient, to any officer or employee of the DATA PROVIDER, with a view towards securing this Agreement or securing favorable treatment with respect to the awarding or amending or the making of any determination with respect to this Agreement.

In the event this Agreement is terminated as provided in (a) above, the DATA PROVIDER shall be entitled to pursue the same remedies against the Information Recipient as it could pursue in the event of a breach of the Agreement by the Information Recipient. The rights and remedies of the DATA PROVIDER provided for in this section are in addition to any other rights and remedies provided by law. Any determination made by the Contracting Office under this clause shall be an issue and may be reviewed as provided in the "disputes" clause of this Agreement.

DISPUTES

Except as otherwise provided in this Agreement, when a genuine dispute arises between the DATA PROVIDER and the Information Recipient and it cannot be resolved, either party may submit a request for a dispute resolution to the Contracts and Procurement Unit. The parties agree that this resolution process shall precede any action in a judicial and quasi-judicial tribunal. A party's request for a dispute resolution must:

be in writing, state the disputed issues, state the relative positions of the parties, state the Information Recipient’s name, address, and his/her department Agreement number, and be mailed to the DATA PROVIDER Contracts and Procurement Unit, within thirty (30) calendar days after the party could reasonably be expected to have knowledge of the issue which he/she now disputes.

This dispute resolution process constitutes the sole administrative remedy available under this Agreement.

XIII. EXPOSURE TO DATA PROVIDER BUSINESS INFORMATION NOT OTHERWISE PROTECTED BY LAW AND UNRELATED TO CONTRACT WORK

During the course of this contract, the information recipient may inadvertently become aware of information unrelated to contract work. Information recipient will treat such information respectfully, recognizing DATA PROVIDER relies on public trust to conduct its work. This information may be hand written, typed, electronic, or verbal, and come from a variety of sources.

XIV. GOVERNANCE

This Agreement is entered into pursuant to and under the authority granted by the laws of the state of DATA PROVIDER and any applicable federal laws. The provisions of this Agreement shall be construed to conform to those laws.

In the event of an inconsistency in the terms of this Agreement, or between its terms and any applicable statute or rule, the inconsistency shall be resolved by giving precedence in the following order: applicable state and federal statutes and rules; any other provisions of the Agreement, including materials incorporated by reference.

XV. HOLD HARMLESS

Each party to this Agreement shall be solely responsible for the acts and omissions of its own officers, employees, and agents in the performance of this Agreement. Neither party to this Agreement will be responsible for the acts and omissions of entities or individuals not party to this Agreement. DATA PROVIDER and the Information Recipient shall cooperate in the defense of tort lawsuits, when possible.

XVI. LIMITATION OF AUTHORITY

Only the Authorized Signator for (DATA PROVIDER) (delegation to be made prior to action) shall have the express, implied, or apparent authority to alter, amend, modify, or waive any clause or condition of this Agreement on behalf of the DATA PROVIDER. No alteration, modification, or waiver of any clause or condition of this Agreement is effective or binding unless made in writing and signed by the Authorized Signator for DATA PROVIDER.

XVII. RIGHT OF INSPECTION

The Information Recipient shall provide the DATA PROVIDER and other authorized entities the right of access to its facilities at all reasonable times, in order to monitor and evaluate performance, compliance, and/or quality assurance under this Agreement on behalf of the DATA PROVIDER.

XVIII. RIGHTS IN INFORMATION

Information Recipient agrees to provide, if requested, copies of any research papers or reports prepared as a result of access to DATA PROVIDER information under this Agreement for DATA PROVIDER review prior to publishing or distributing.

In no event shall the Information Provider be liable for any damages, including, without limitation, damages resulting from lost information or lost profits or revenue, the costs of recovering such Information, the costs of substitute information, claims by third parties or for other similar costs, or any special, incidental, or consequential damages, arising out of the use of the information. The accuracy or reliability of the Information is not guaranteed or warranted in any way and the information Provider's disclaim liability of any kind whatsoever, including, without limitation, liability for quality, performance, merchantability and fitness for a particular purpose arising out of the use, or inability to use the information.

If checked, please submit the following:

copies of_______(insert list of items)_____ to the attention of: (insert name of DATA PROVIDER employee)

at_____(insert address to which material is sent)__.

XIX. SEVERABILITY

If any term or condition of this Agreement is held invalid, such invalidity shall not affect the validity of the other terms or conditions of this Agreement, provided, however, that the remaining terms and conditions can still fairly be given effect.

XX. SURVIVORSHIP

The terms and conditions contained in this Agreement which by their sense and context, are intended to survive the completion, cancellation, termination, or expiration of the Agreement shall survive.

XXI. TERMINATION

Either party may terminate this Agreement upon 30 days prior written notification to the other party. If this Agreement is so terminated, the parties shall be liable only for performance rendered or costs incurred in accordance with the terms of this Agreement prior to the effective date of termination.

XXII. WAIVER OF DEFAULT

This Agreement, or any term or condition, may be modified only by a written amendment signed by the Information Provider and the Information Recipient. Either party may propose an amendment.

Failure or delay on the part of either party to exercise any right, power, privilege or remedy provided under this Agreement shall not constitute a waiver. No provision of this Agreement may be waived by either party except in writing signed by the Information Provider or the Information Recipient.

XXIII. ALL WRITINGS CONTAINED HEREIN

This Agreement contains all the terms and conditions agreed upon by the parties. No other understandings, oral or otherwise, regarding the subject matter of this Agreement shall be deemed to exist or to bind any of the parties hereto.

IN WITNESS WHEREOF, the parties have executed this Agreement.

DATA PROVIDER	DATA RECIPIENT
______________________________	______________________________
Signature	Signature
______________________________	______________________________
Date	Date

EXAMPLE 1 APPENDIX A

USE AND DISCLOSURE OF CONFIDENTIAL INFORMATION

People with access to confidential information are responsible for understanding and following the laws, policies, procedures, and practices governing it. Below are key elements:

A. CONFIDENTIAL INFORMATION

Confidential information is information federal and state law protects from public disclosure. Examples of confidential information are social security numbers, and healthcare information that is identifiable to a specific person under LOCAL LAWS. The general public disclosure law identifying exemptions is LOCAL LAWS.

B. ACCESS AND USE OF CONFIDENTIAL INFORMATION

1. Access to confidential information must be limited to people whose work specifically requires that access to the information.

2. Use of confidential information is limited to purposes specified in sections V and VI of this Agreement.

C. DISCLOSURE OF CONFIDENTIAL INFORMATION

1. An Information Recipient may disclose an individual's confidential information received or created under this Agreement to that individual or that individual's personal representative consistent with law.

2. An Information Recipient may disclose an individual's confidential information, received or created under this Agreement only as permitted under section *VIII* -***Re-Disclosure of Information*** of the Agreement, and state and federal laws allow.

D. CONSEQUENCES OF UNAUTHORIZED USE OR DISCLOSURE

An Information Recipient's unauthorized use or disclosure of confidential information is the basis for the Information Provider immediately terminating the Agreement. The Information Recipient may also be subject to administrative, civil and criminal penalties identified in law.

E. ADDITIONAL DATA USE RESTRICTIONS: (if necessary)

Signature: ____

Date:____

DATA SECURITY REQUIREMENTS

Protection of Data

The Information Recipient agrees to store data on one or more of the following media and protect the data as described:

A. Hard disk drives

1. Data stored on local workstation hard disks: The data must be encrypted as described under *F. Data Storage on portable devices or media*. Encryption is not required when data classified as "Potentially Identifiable Information" is stored temporarily on local workstation hard disks. Temporary storage is thirty (30) days or less.

2. Access to the data will be restricted to authorized users by requiring logon to the local workstation using a unique user ID and Complex Password or other authentication mechanisms which provide equal or greater security, such as biometrics or smart cards. Accounts must lock after 5 unsuccessful access attempts and require administrator reset.

Complex Passwords are:

- At least 8 characters in length
- Contain at least three of the following character classes: uppercase letters, lowercase letters, numerals, special characters.
- Do not contain the user's name, user ID or any form of their full name
- Do not consist of a single complete dictionary word, but can include a passphrase
- Are changed at least every 120 days.

B. Network server disks

Data stored on hard disks mounted on network servers and made available through shared folders:

3. Access to the data will be restricted to authorized users through the use of access control lists which will grant access only after the authorized user has authenticated to the network.
   a. Authentication must occur using a unique user ID and Complex Password or other authentication mechanisms which provide equal or greater security, such as biometrics or smart cards. Accounts must lock after 5 unsuccessful access attempts, and require administrator reset.
4. Data on disks mounted on such servers must be located in a secured computer area, which is accessible only to authorized personnel with access controlled through use of a key, card key, combination lock, or comparable mechanism.
5. If the servers are not located in a secured computer area **or** if the data is classified as Confidential or Restricted Confidential, it must be encrypted as described under *F. data storage on portable devices or media.*

C. Optical discs (CDs or DVDs) in local workstation optical disc drives

a. Data provided on optical discs must be encrypted as described under *F. Data storage on portable devices or media*. When not in use for the purpose of this agreement, such discs must be locked in a drawer, cabinet or other physically secured container to which only authorized users have the key, combination or mechanism required to access the contents of the container.

D. Optical discs (CDs or DVDs) in drives or jukeboxes attached to servers

Data provided on optical discs must be encrypted as described under. *Data Storage on portable devices or media.*

6. Access to data on these discs will be restricted to authorized users through the use of access control lists which will grant access only after the authorized user has authenticated to the network.
7. Authentication must occur using a unique user ID and Complex Password or other authentication mechanisms which provide equal or greater security, such as biometrics or smart cards. Accounts must lock after 5 unsuccessful access attempts and require administrator reset.
8. Data on discs attached to such servers must be located in an area which is accessible only to authorized personnel, with access controlled through use of a key, card key, combination lock, or comparable mechanism.

E. Access via remote terminal/workstation over the State Governmental Network (SGN) or the Internet

9. When data is transferred between the Information Provider and the Information Recipient, access to the data will be controlled by the Information Provider, who will issue authentication credentials. Information Recipient will notify the Information Provider immediately whenever:
    - An authorized person in possession of such credentials is terminated or otherwise leaves the employ of the Information Recipient
    - Whenever a person's duties change such that the person no longer requires access to perform work for this agreement.

10. The data shall not be transferred or accessed over the Internet by the Information Recipient in any other manner unless specifically authorized within the terms of the Agreement.

    - If so authorized secure encryption protocols and multi-factor authentication mechanisms such as; hardware or software tokens, smart cards, digital certificates and biometrics, must be used.
    - During transmission the data must be encrypted using a key length of at least 128 bits. Industry standard mechanisms and algorithms, such as those validated by the National Institute of Standards and Technology (NIST) are required.

F. Data storage on portable devices or media

11. Portable devices include, but are not limited to; handhelds/PDAs, Ultramobile PCs, flash memory devices (e.g. USB flash drives, personal media players), portable hard disks, and laptop/notebook computers.

12. Portable media includes, but is not limited to; optical media (e.g. CDs, DVDs), magnetic media (e.g. floppy disks, tape, Zip or Jaz disks), or flash media (e.g. CompactFlash, SD, MMC).

13. The data shall not be stored by the Information Recipient on portable devices or media unless specifically authorized within the terms of this Agreement. If so authorized, the data shall be given the following protections:

    - Use industry standard encryption mechanisms validated by the National Institute of Standards and Technologies (NIST).
    - Encrypt the data with a key length of at least 128 bits
    - Control access to devices with a unique user ID and Complex Password or stronger authentication method such as a physical token or biometrics. Whenever technically possible accounts must lock after 5 unsuccessful access attempts and require administrator reset.
    - Manually lock devices whenever they are left unattended and set devices to lock automatically after a period of inactivity, if this feature is available. Maximum period of inactivity is 10 minutes.
    - Physically protect the portable device(s) and/or media by
        - Keeping them in locked storage when not in use
        - Using check-in/check-out procedures when they are shared, and
        - Taking frequent inventories

14. When being transported outside of a secure area, portable devices and media with data provided under this agreement must be under the physical control of Information Recipient staff with authorization to access the data.

G. Backup Media

Data may be backed up as part of Information Recipient's normal backup process provided that it is encrypted and the process includes secure storage and transport.

H. Paper documents

Any paper records must be protected by storing the records in a secure area which is only accessible to authorized personnel. When not in use, such records must be stored in a locked container, such as a file cabinet, locking drawer, or safe, to which only authorized persons have access.

I. Data Segregation

a. Data provided under this agreement must be segregated or otherwise distinguishable from all other data. This is to ensure that when no longer needed by the Information Recipient, all of the data can be identified for return or destruction. It also aids in determining whether the data has or may have been compromised in the event of a security breach.

b. When it is not feasible or practical to segregate the data from other data then all data, which it is commingled with the data provided under this agreement, must be protected as described in this exhibit

J. Data Disposition

If data destruction is required by the Agreement, the data shall be destroyed using one or more of the following methods:

Data stored on:	**Will be destroyed by:**
Server or workstation hard disks	Using a “wipe” utility which will overwrite the data at least three (3) times using either random or single character data, or Degaussing sufficiently to ensure that the data cannot be reconstructed, or Physically destroying the disk , or Delete the data and physically and logically secure data storage systems that continue to be used for the storage of confidential data to prevent any future access to stored information. One or more of the preceding methods must be performed before transfer or surplus of the systems or media containing the data.
Paper records	On-site shredding, pulping, or incineration, or Recycling through a contracted firm provided the contract with the recycler assures that the confidentiality of data will be protected. Paper documents containing confidential information requiring special handling (e.g. protected health information) On-site shredding, pulping, or incineration
Optical discs (e.g. CDs or DVDs)	Incineration, shredding, or completely defacing the readable surface with a course abrasive
Magnetic tape	Degaussing, incinerating or crosscut shredding Removable media (e.g. floppies, USB flash drives, portable hard disks, Zip or similar disks) Using a “wipe” utility which will overwrite the data at least three (3) times using either random or single character data Physically destroying the disk Degaussing magnetic media sufficiently to ensure that the data cannot be reconstructed

K. Notification of Compromise or Potential Compromise

The compromise or potential compromise of the data must be reported to the DATA PROVIDER IT Security Officer within two (2) business days of discovery.

EXAMPLE 1 APPENDIX C CERTIFICATION OF DATA DISPOSITION

Date of Disposition ____

☐ All copies of any Datasets related to agreement #_____________have been deleted from all data storage systems. These data storage systems continue to be used for the storage of confidential data and are physically and logically secured to prevent any future access to stored information. Before transfer or surplus, all data will be eradicated from these data storage systems to effectively prevent any future access to previously stored information.

☐ All copies of any Datasets related to agreement #_____________have been eradicated from all data storage systems to effectively prevent any future access to the previously stored information.

☐ All materials and computer media containing any data related to agreement#_____________ have been physically destroyed to prevent any future use of the materials and media.

☐ All paper copies of the information related to agreement #_____________have been destroyed on-site by cross cut shredding.

☐ All copies of any Datasets related to agreement #_____________that have not been disposed of in a manner described above, have been returned to DATA PROVIDER.

☐ Other

The data recipient hereby certifies, by signature below, that the data disposition requirements as provided in agreement #________________, Section C, item B Disposition of Information, have been fulfilled as indicated above.

__________________________________ ________________

Signature of data recipient Date

LIST OF DATA ELEMENTS TO BE RECEIVED

Field	Field Description	Position

<PAGE STRIPPED OF DATA>

EXAMPLE 1 APPENDIX E: Geocoding elements

GEOCODED FILES LIST OF DATA ELEMENTS

Description	Position

<PAGE STRIPPED OF DATA>

Appendix C: More examples of EHI

SAN ANTONIO 2017

After creating a series of dozens of maps about various population characteristics in San Antonio by ZIP Code, I used the original Brookings formula to create a comparable hardship index for San Antonio, then ran it by the Chicago researchers to make sure I had it right. The result is above, and it confirms what I had been seeing in map after map of population characteristics. The hardship index is based on these six metrics:

- Percent of population age 25 and over below the poverty level;
- Per capita income;
- Unemployment levels for those 16 and over;
- Educational attainment for those 25 and over, which equates to having a high school diploma or a GED;
- Crowding level, which is the percentage of housing units with more than one occupant per room; and
- Dependency amount, the (combined) percentage of the population under age 18 and above age 65.

LOS ANGELES (2014)

- Percentage of Housing Units with More than 1 Person per Room
- Percentage of Population with Income Less than the Federal Poverty Level
- Percentage of Unemployed Workers Age 16 and Over
- Percentage of Population Over 25 That Did Not Graduate High School
- Percentage of Population Over Age 65; & Percentage of Population Under Age 18
- Median Household Income

WISCONSIN (2015)

EHI is comprised of six measures:

- crowded housing (percentage of housing units with >1 person/room),
- poverty (percentage of households below the federal poverty level),
- unemployment (percentage of people ≥16 years who are unemployed),
- education (percentage of people ≥25 years without a high school education),
- dependency (percentage of population <18 or >64 years) and
- per capita income.

SMART COMMUNITIES An Update on Urban Hardship 2005

Richard Nathan and Charles Adams are the creators of the "Urban Hardship Index" which examines six key statistics in dozens of major metropolitan areas. The six key statistics are:

- Unemployment (those over age 16 who were unemployed)
- Dependency (percentage of the population over 64 or under 18 years of age)
- Education (percentage of population over age 25 with less than a high school education)
- Income level (per capita)

- Crowded housing (percent of occupied housing units with more than one person per room)
- Poverty (percent of people living below the federal poverty line)

AN UPDATE ON URBAN HARDSHIP Nelson A. Rockefeller Institute of Government 2004

The ***Intercity Hardship Index*** draws together six key factors:2

- Unemployment, defined as the percent of the civilian population over the age of 16 who were unemployed;
- Dependency, the percentage of the population that are under the age of 18 or over the age of 64;
- Education, the percentage of the population over the age of 25 who have less than a high school education;
- Income Level, the per capita income;
- Crowded Housing, measured by the percent of occupied housing units with more than one person per room; and
- Poverty, the percent of people living below the federal poverty level.

Healthy Places Index (HPI 2.0)
http://healthyplacesindex.org

Intercity Hardship Index

The Intercity Hardship Index, also known as the Urban Hardship Index and Economic Hardship Index is a composite of 6 indicators:

- Unemployment rate,
- Dependency (percent of the population aged >65 years or < 18 years),
- Low education attainment (the percentage of those over the age of 25 with less than a high school education),
- Per capita income,
- Housing overcrowding (>1 occupant/room),
- Poverty level (<100% of federal poverty).

Except for dependency, the 5 other indicators in the Hardship Index are part of the HPI.

The association between neighborhood economic hardship, the retail food environment, fast food intake, and obesity: findings from the Survey of the Health of Wisconsin 2015

The Economic Hardship Index (EHI) developed by the Rockefeller Institute of Government was derived from 2000 census data. The EHI measures the social and economic conditions of communities on a block group level using six indicators including:

- crowded housing (percent of housing units with more than 1 person per room),
- poverty (percent of households living below the federal poverty level),
- unemployment (percent of persons over the age of 16 years that are unemployed),
- education (percent of persons over the age of 25 years without a high school education),
- dependency (percent of population that is under age 18 or over age 64 years),
- income level (median per capita).

CHICAGO

Chicago's Vision Zero 2017

Economic Hardship Index (EHI) developed by the Rockefeller Institute of Government. The EHI measures the social and economic conditions of communities at the census tract level using six indicators including:

- crowded housing (percent of housing units with more than 1 person per room),
- poverty (percent of households living below the federal poverty level),
- unemployment (percent of persons over the age of 16 years that are unemployed),
- education (percent of persons over the age of 25 years without a high school education),
- dependency (percent of population that is under age 18 or over age 64 years),
- income level (median per capita).

FLORIDA

PUBLIC HEALTH WEEK 2018 - Ensuring the Right to Health

The Nassau County EHI score is the average of the following six variables that have been standardized on a scale from 0 to 100:
- Unemployment (over the age of 16 years)
- Education (over 25 years of age without a high school diploma)
- Per capita income level
- Poverty (?below the federal poverty level)
- Crowded housing (housing units with more than one person per room)
- Dependency (population under 18 or over 64 years of age)

SHELBY COUNTY

Topic to search for	EHI Variable	Description of variable
• Employment Status	Unemployment	the percent of the civilian population over the age of 16 who were unemployed
• Population under 18 years by age • Population 60 years and over	Dependency	the percentage of the population that are under the age of 18 or over the age of 64

• Educational Attainment for the population 25 years and over.	Education	the percentage of the population over the age of 25 who have less than a high school education
• Per capita income in the past 12 months	Income Level	the per capita income
• Selected housing characteristics (Occupants per room)	Crowded Housing	Measured by the percent of occupied housing units with more than one person per room
• Poverty status in the past 12 months	Poverty	The percent of people living below the federal poverty level