HealthLandscape

How HRSA is Leveraging ACS and CDC Data for Small Area Estimates and Geographic Visualizations of Substance Use Outcomes

Mark Carrozza, Anuradha Jetty, Jennifer Rankin,
and Jene Grandmont, American Academy of Family Physicians

American Community Survey Data Users Group
2019 DATA USERS CONFERENCE

What is the UDS Mapper?

- An online mapping tool developed to provide access to maps, data, and analysis using Uniform Data System (UDS) and other relevant data to visualize service area information for Health Center Program (HCP) awardees and look-alikes
- Compares HCP awardee and look-alike data to community/ population data and shows spatial relationships between the program, community attributes, and other resources

Data Mapped by ZCTA

- ZIP Code Tabulation Area - an approximation of ZIP Codes from the U.S. Census Bureau
  - 2010 U.S. Census Boundaries for ZCTAs
- UDS data, 2017
- UDS data are submitted to HRSA by HCP awardees and look-alikes every calendar year
  - Data are analyzed and prepared
  - Population demographics and health (various sources)

Calculated Layers in the UDS Mapper

- Penetration rates (% of target population going to health centers) and unmet need
  - Low-income population
  - Total population
  - Population by insurance categories
- Change in patients and population over time (2-year and 1-year)
  - Total patients and patients by insurance categories
  - Total population and population by insurance categories

Other Health Center Data in the UDS Mapper

- Health center characteristics
  - Health Center Count and Health Center Dominance
- Population not served by health centers (unmet need)
  - Low-income population
  - Population by insurance categories
- Patient mix
  - Patients by insurance categories

Population Data in the UDS Mapper

- Poverty level
  - Poverty (100% Federal Poverty Level (FPL))
  - Low-Income (200% FPL)
- Race and ethnicity
- Age
- Social environment
- Insurance status
- Uninsurance by income level
  - 138% FPL
  - 200% FPL
  - 138-400% FPL

Other Data in the UDS Mapper

- Non-geographic UDS data by organization (Information Card Deck)
- Population and MAP for MAT Indicators
  - American Community Survey
  - HRSA Area Resource File
  - Centers for Disease Control and Prevention (CDC) Vital Statistics
  - CDC Behavioral Risk Factor Surveillance System
  - Centers for Medicare and Medicaid Services
  - Substance Abuse and Mental Health Administration

Small Area Estimation - Background

- Local level health data typically not available through national surveys
- Demand for public health data for small geographies is on the rise
- Small areas - Estimates from national surveys
    - o large variance
    - o not reliable
    - o privacy and confidentiality issues
- Therefore, Small Area Estimation (SAE) is the key to develop local area level health measures

Small Area Estimation - Background

SAE methods can be broadly classified into

- Design-based – design-based distribution
- **Model-based** – estimation solely depends on the model specified
- Multilevel regression and poststratification (MRP) is a model-based SAE
- CDC researchers developed MRP to estimate census tract level health measures for 500 cities
- HealthLandscape and the Robert Graham Center - ZCTA level prevalence of select chronic diseases and health behaviors

Small Area Estimation – Data Sources

Variable Type	Data Source	Measure
Outcome	BRFSS/National database	Health Behavior/Chronic Disease
Independent	BRFSS/National database	Age (13 groups)
		Sex (male, female)
		Race/Ethnicity (8 categories)
	ACS	County level – Percent population <150% FPL
		Tract level – Percent population <150% FPL
	Census 2010	Block level population (for each of the 208 combinations)

Zhang et al., Am J Epidemiol. 2014;179: 1025–1033

Small Area Estimation - Measures

Age	Sex	Race/Ethnicity	Poverty	County	State
18-24	Male	NH, White	County level <150% FPL	Sampled	Sampled
25-29	Female	NH, Black	Census Tract <150% FPL		
30-34		American Indian/Alaska Native			
35-39		Asian			
40-44		Native Hawaiian /Pacific Islander			
45-49		Other, Single Race			
50-54		Two or More Races			
55-59		Hispanic			
60-64					
65-69					
70-74					
75-79					
80+					

Zhang et al., Am J Epidemiol. 2014;179: 1025–1033

Small Area Estimation - Methodology

Zhang et al., Am J Epidemiol. 2014;179: 1025–1033

Small Area Estimation - Methodology

First Step

- Multilevel logistic regression models for each outcome - 13 categories of age, sex, 8 groups race/ethnicity and county level poverty (<150% FPL)
- Used SAS Proc GLIMMIX
- Fixed effects for age, sex, race/ethnicity and county level poverty
- County and state random effects

Zhang et al., Am J Epidemiol. 2014;179: 1025–1033

Equation

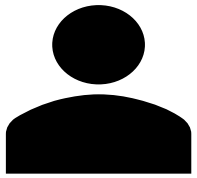
$$
\begin{aligned}
P^{b}_{ijkcs}(y_{ijkcs} &= 1) \\
&= \text{logit}^{-1}(\alpha_i + \beta_j + \gamma_k + x_b'\eta + \mu_c + \nu_s) \\
&= \frac{\exp(\alpha_i + \beta_j + \gamma_k + x_b'\eta + \mu_c + \nu_s)}{1 + \exp(\alpha_i + \beta_j + \gamma_k + x_b'\eta + \mu_c + \nu_s)}
\end{aligned}
$$

Level

State s

County c

Person

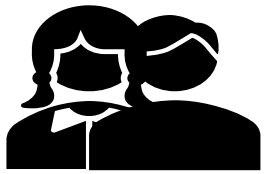


Measure

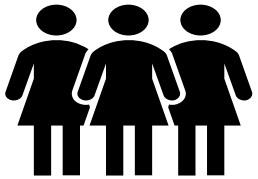
$

Poverty b

Age i



Gender j



Race/Ethnicity k

Zhang et al., Am J Epidemiol. 2014;179: 1025–1033

Small Area Estimation - Methodology

Second Step

- Bayesian Spatial Smoothing with adjacent county random effects for counties not sampled in BRFSS
- Merged this dataset with census tract level poverty (<150% FPL) and Census block level population for each of the 208 age, sex and race combinations

Third Step

- Solved the regression equation
- Post-stratification with census demographics
- Monte Carlo simulations (1000) Model-based

Validated against survey-based estimates (correlations)

Zhang et al., Am J Epidemiol. 2014;179: 1025–1033

Small Area Estimation - Purpose

- public health prevention program planning
- resource allocation
- health policy formulation
- health care decision-making and delivery

Small Area Estimation - Limitations

- Geographic context – is critical to evaluate the relationship between the health outcomes and demographic characteristics
- Area-level poverty indicator used and not the individual-level poverty
- Different independent variables may need to be included in the multilevel models to estimate different health outcomes
- Cannot track trends
- Cannot evaluate the effectiveness of intervention unless intervention variables included in the source survey

References

- Zhang, X, Holt, J B, Lu, H, Wheaton, A G, Ford, E S, Greenlund, K J and Croft, J B. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. Am J Epidemiol. 2014;179: 1025–1033. DOI: https://doi.org/10.1093/aje/kwu018
- Lehtonen R, Veijanen A. Design-based methods of estimation for domains and small areas. InHandbook of statistics 2009 Jan 1 (Vol. 29, pp. 219-249). Elsevier.
- Wang Y, Holt JB, Xu F, Zhang X, Dooley DP, Lu H, et al. Using 3 Health Surveys to Compare Multilevel Models for Small Area Estimation for Chronic Diseases and Health Behaviors. Prev Chronic Dis 2018; 15:180313. DOI: https://doi.org/10.5888/pcd15.180313.
- Berkowitz Z, Zhang X, Richards TB, Sabatino SA, Peipins LA, Holt J, White MC. Multilevel regression for small-area estimation of mammography use in the United States, 2014. Cancer Epidemiology and Prevention Biomarkers. 2019 Jan 1;28(1):32-40.

Population Indicators Tool

- Enables you to do cold spot analysis to find high-need areas based on data that are common indicators of health status
- About the data:
  - Total population (not the patient population)
  - Drawn from a number of different sources, at a variety of geographic levels, and are updated annually (if underlying data are updated)
  - See https://www.udsmapper.org/knowledge-base.cfm for the most current data source dates
  - All data are displayed by ZCTA
  - Data that are not available from the source at the ZCTA level were imputed based on U.S. Census population data

Population Indicators: Data and Sources (1)

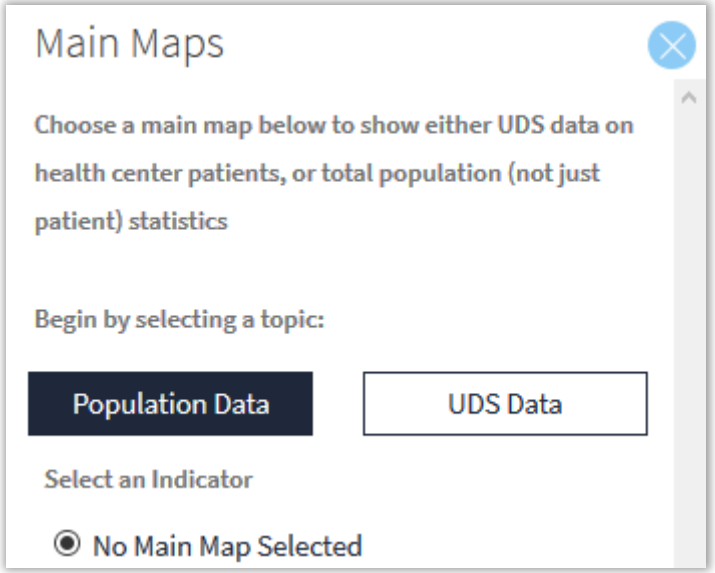
- ACS 5-year estimates, natively by ZCTA
  - % of Population in Poverty (at or below 100% FPL)
  - % of Population Low-Income (at or below 200% FPL)
  - % of Population Not Employed
  - % of Households with Limited English Proficiency
  - % of Population with Less Than High School Education
  - % of Population with a Disability
- ACS 1-year estimates, natively by county (imputed)
  - % of Population Uninsured

Population Indicators: Data and Sources (2)

- HRSA Area Resource File, natively by county (imputed)
  - Low Birth Weight Rate
- CDC WONDER, natively by county (imputed)
  - Age-Adjusted Mortality Rate
- CDC Behavioral Risk Factor Surveillance System (BRFSS), natively by county (imputed)
  - % of Adults Ever Told They Have Diabetes
  - % of Adults Ever Told They Have High Blood Pressure
  - % of Adults Who Are Obese
  - % of Adults with No Dental Visit in the Past Year
  - % of Adults Who Have Delayed or Not Sought Care Due to High Cost
  - % of Adults with No Usual Source of Care
  - % of Adults Who Ever Smoked and are Currently Smokers
  - % of Adults Who Binge Drink

Clear Colored Layers from the Map

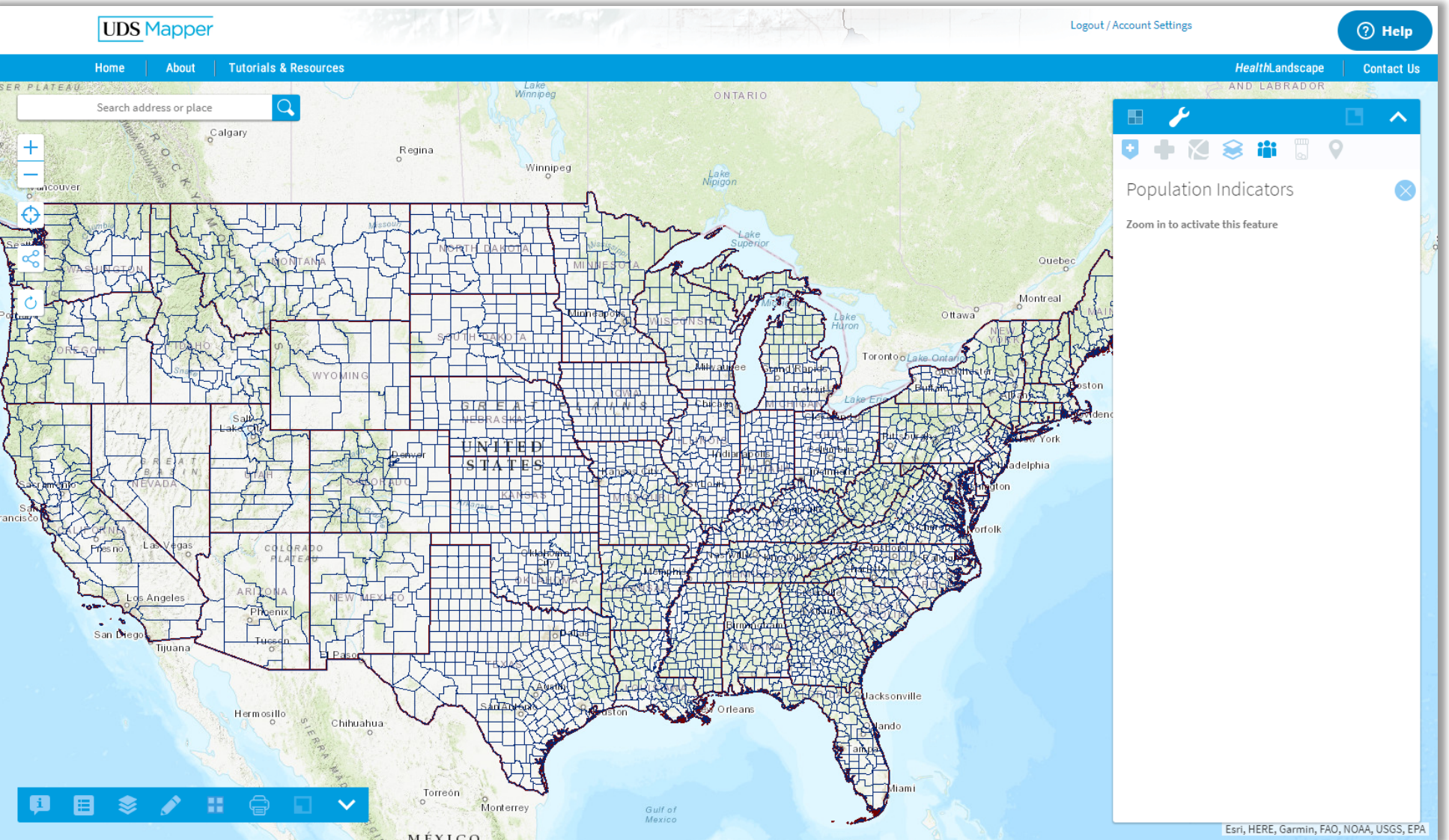
- Population Indicators are best visualized with a clear map background
- Before displaying Population Indicators data on the map, you should remove other colorful data layers including the Main Maps
  - Open the **Main Maps** tool, click **No Main Map Selected**



How to Open the Population Indicators Tool

- Click on the **Population Indicators** icon (looks like a group of people) in the Tool Box in the upper right corner of the map
- If you only see a blue bar with a few tool buttons, click the **white, downward-pointing arrow** to reveal the tools

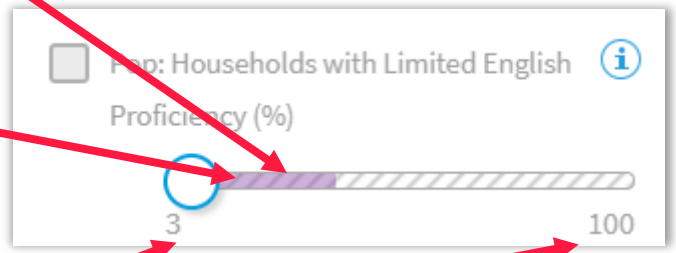
Zoom Level to Use the Population Indicators Tool



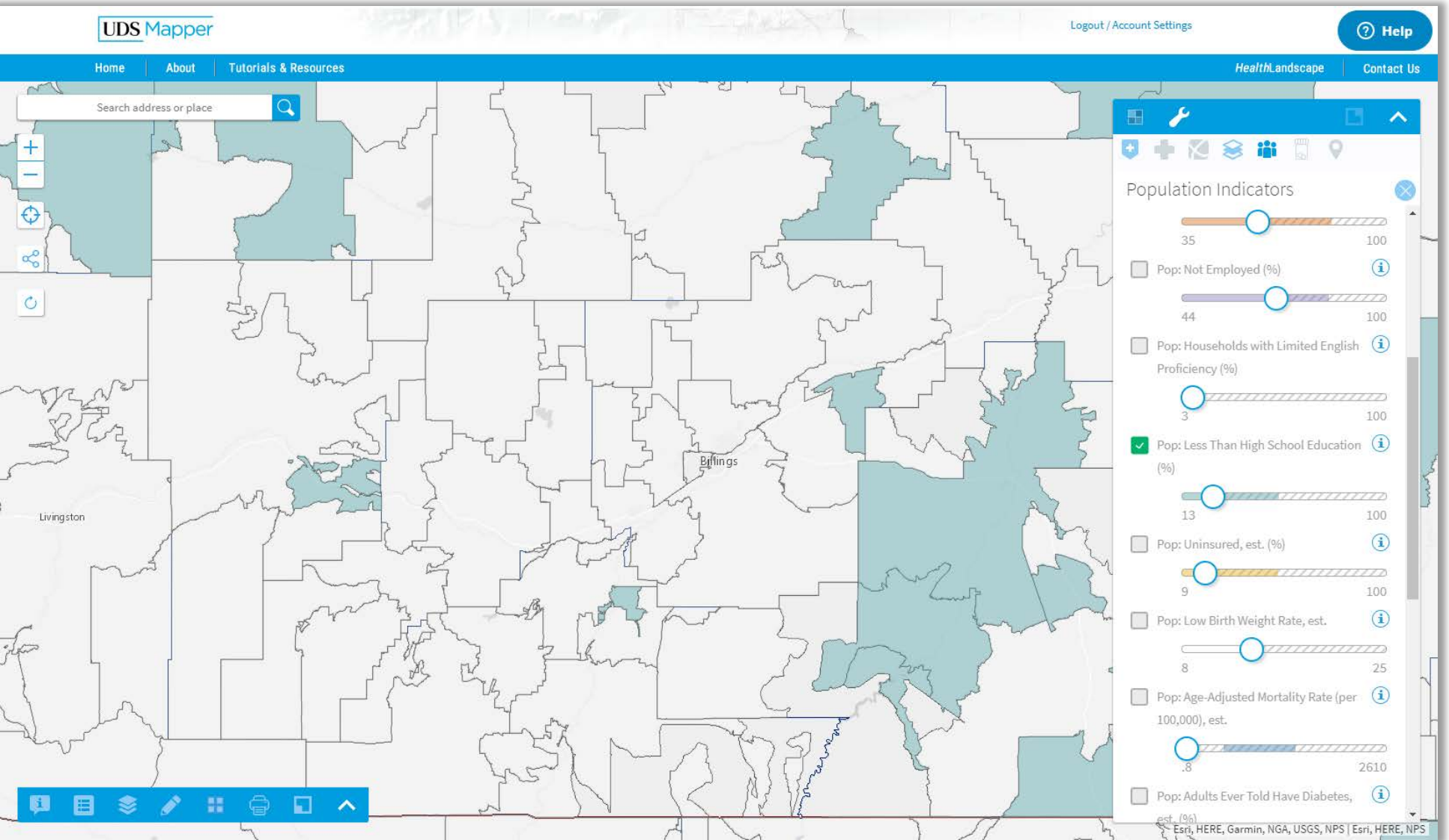
- You must be close in to the map to activate the indicators
- If starting at the default level, **zoom in** three times

National and Local Data Ranges in the Population Indicators

- The color within each slider shows the range of the data for the geography you are currently seeing on the screen and will change as you move around the country
- The hashmarks within each slider indicate the range of values that are displayed on the map
- The number range under the slider bar shows the value your indicator is set to (left number) and the maximum value of that dataset for the nation



Turn on a Population Indicator



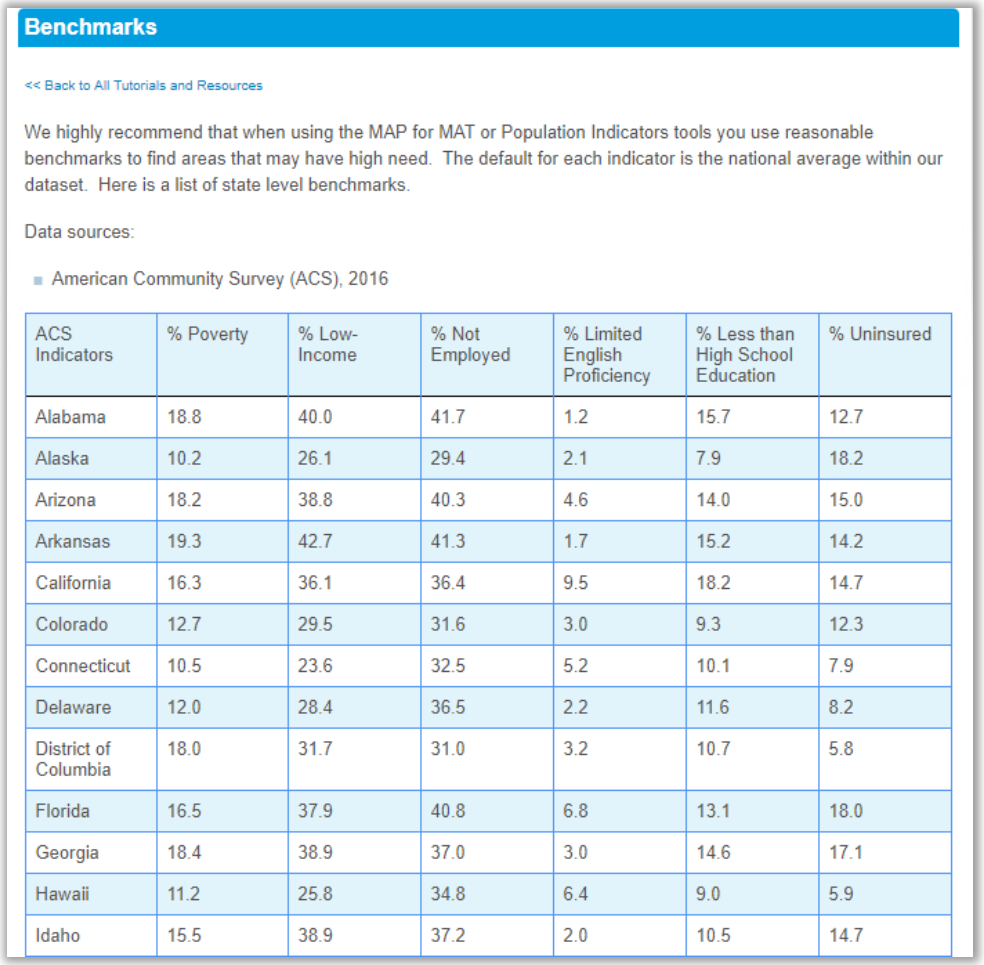
?

- Every indicator is set to the national average by default
- When you turn on an indicator, the colored areas are ZCTAs whose value for that indicator are equal to or greater than the national average, or show higher need

Use Reasonable Benchmarks

- Because you can set your own values in this tool, it is important to use reasonable benchmarks when looking for “high” need
  - By default, indicators are set to the national average
- Stating that an area has high need may be false – it must be compared to something tangible rather than just a user selected number

Where to Find Benchmarks



Benchmarks

<< Back to All Tutorials and Resources

We highly recommend that when using the MAP for MAT or Population Indicators tools you use reasonable benchmarks to find areas that may have high need. The default for each indicator is the national average within our dataset. Here is a list of state level benchmarks.

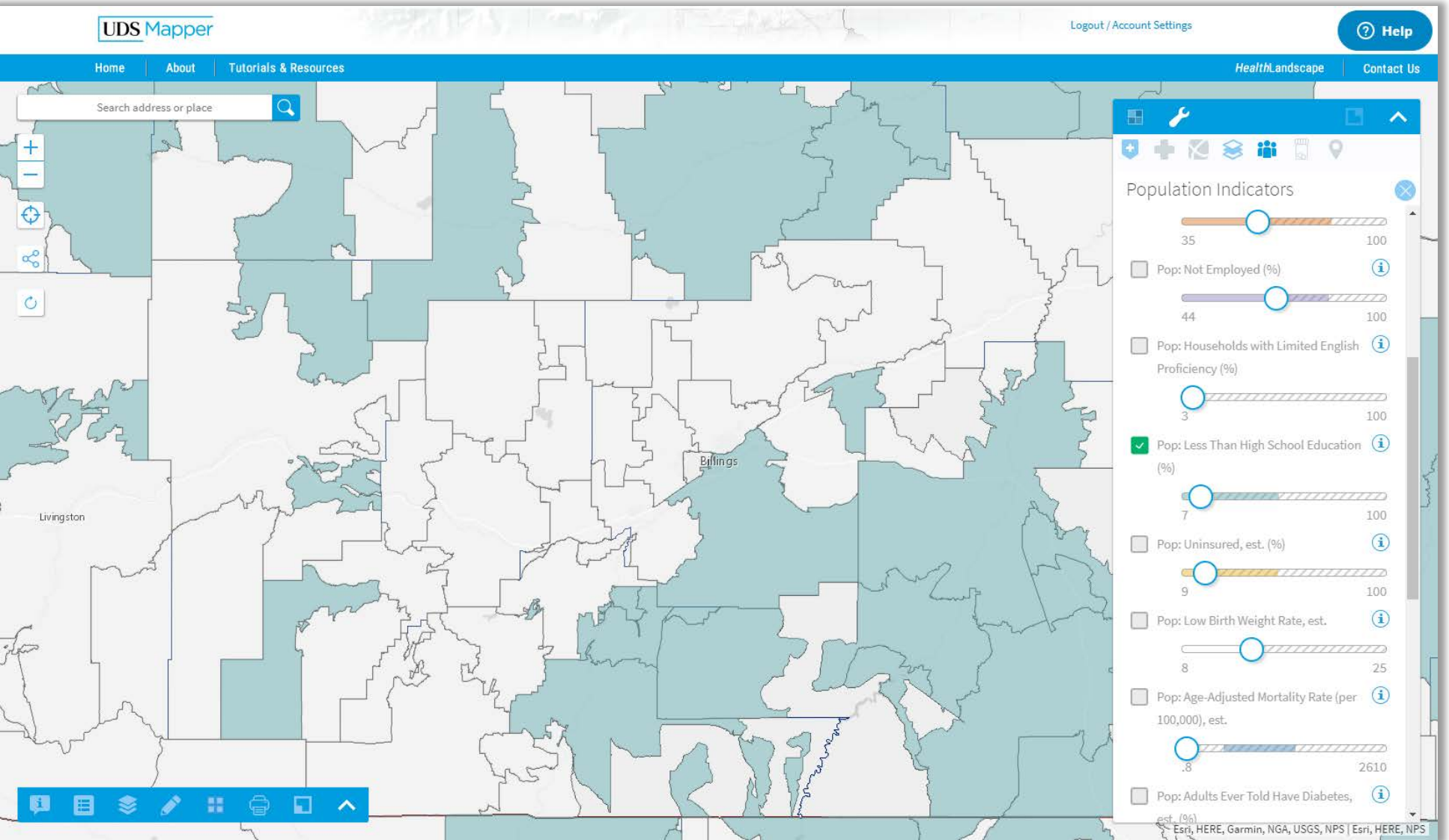
Data sources:

- American Community Survey (ACS), 2016

ACS Indicators	% Poverty	% Low-Income	% Not Employed	% Limited English Proficiency	% Less than High School Education	% Uninsured
Alabama	18.8	40.0	41.7	1.2	15.7	12.7
Alaska	10.2	26.1	29.4	2.1	7.9	18.2
Arizona	18.2	38.8	40.3	4.6	14.0	15.0
Arkansas	19.3	42.7	41.3	1.7	15.2	14.2
California	16.3	36.1	36.4	9.5	18.2	14.7
Colorado	12.7	29.5	31.6	3.0	9.3	12.3
Connecticut	10.5	23.6	32.5	5.2	10.1	7.9
Delaware	12.0	28.4	36.5	2.2	11.6	8.2
District of Columbia	18.0	31.7	31.0	3.2	10.7	5.8
Florida	16.5	37.9	40.8	6.8	13.1	18.0
Georgia	18.4	38.9	37.0	3.0	14.6	17.1
Hawaii	11.2	25.8	34.8	6.4	9.0	5.9
Idaho	15.5	38.9	37.2	2.0	10.5	14.7

- Benchmarks other than the national average include published targets (Healthy People 2020) or state or regional averages
- While you can use any benchmark, there is a table of state rates for each benchmark available on the UDS Mapper site https://www.udsmapper.org/benchmarks.cfm

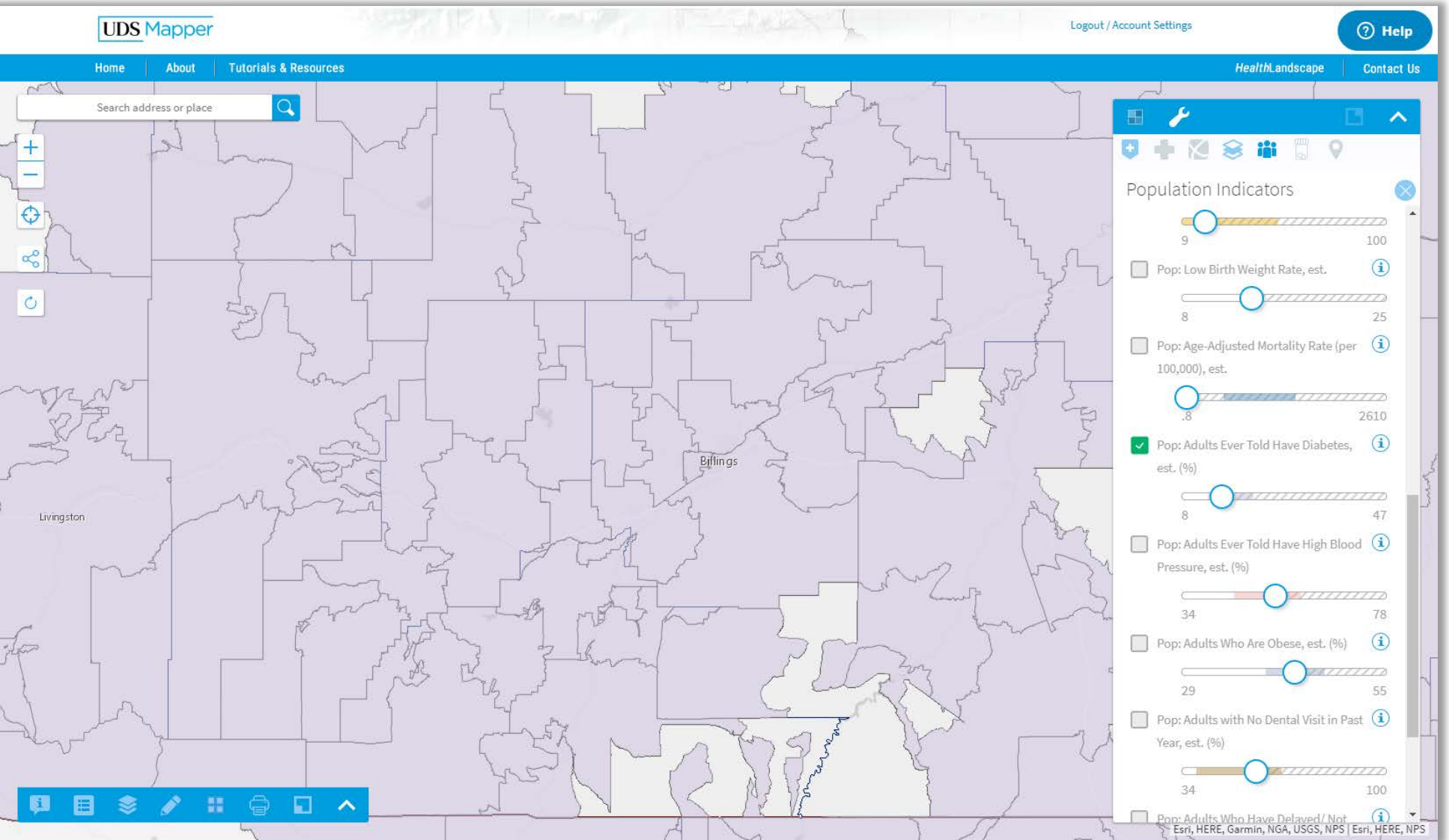
Population Indicator Benchmark (1)



Set appropriate benchmark

- As we see in the Benchmarks table, 7.2% of Montanans have less than a high school education
- People with this education level may have a hard time finding employment that offers health insurance

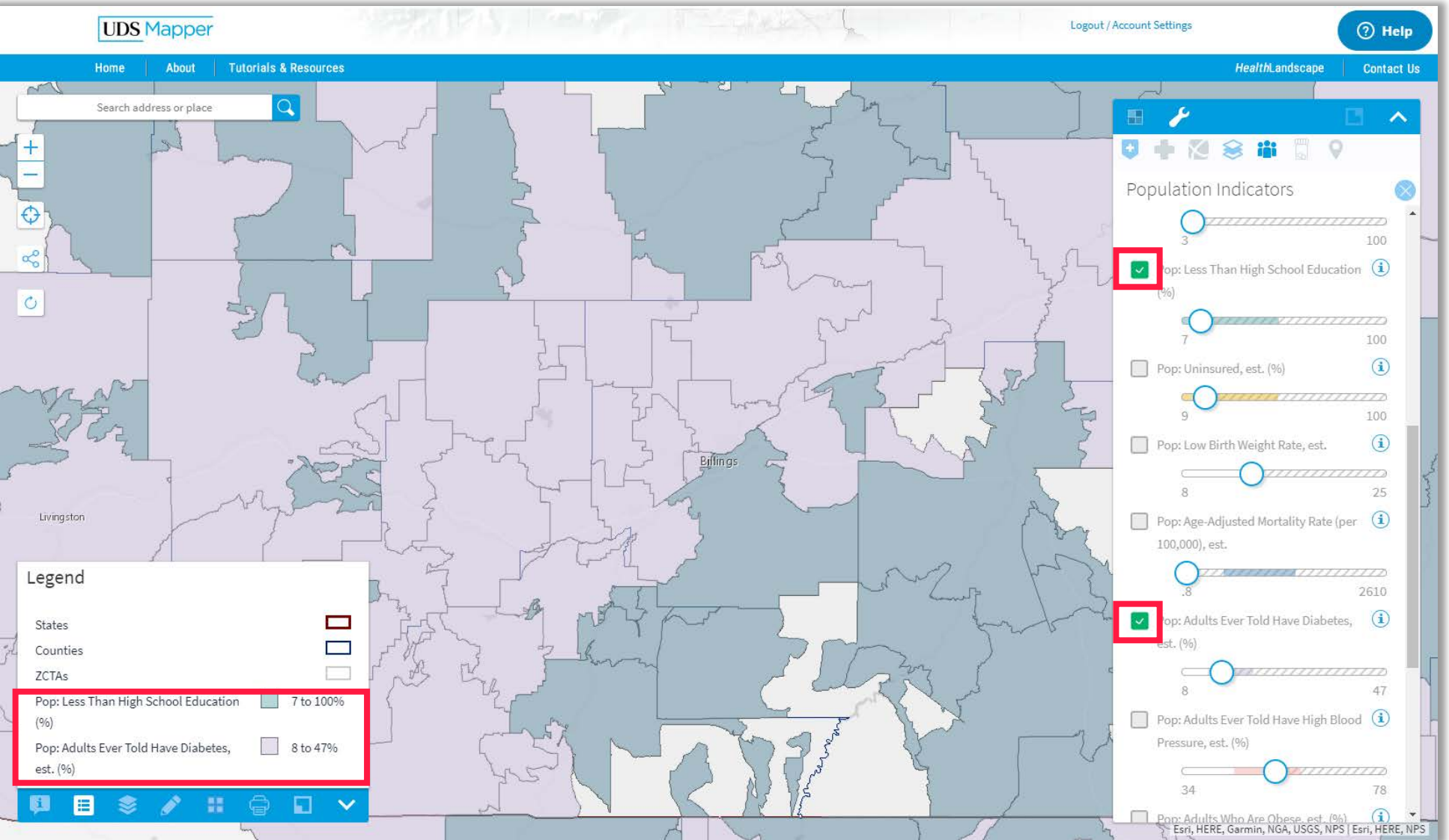
Population Indicator Benchmark (2)



Turn off education level indicator and turn on the diabetes indicator

- As we see in the Benchmarks table, Montana has a diabetes prevalence of 7.9%

Compare Indicators



Turn education level indicator back on to look for overlapping cold spots of need based on multiple indicators

- Blue areas: at least 7% of population with less than high school education
- Purple areas: at least 8% diabetes prevalence
- Blended areas: at least 7% of population less than high school education AND at least 8% diabetes prevalence
- Because blended colors do not appear in the legend, we recommend you not turn on more than two indicators at a time, as colors will blend and may become confusing

How to Download the Data

- The Population Indicators data are available to view in the **Data Table** and can be downloaded
  - The data will only be available for the ZCTAs that are selected in the Explore Service Area tool
- To visually determine the rate in a specific ZCTA, gradually move the slider button and note when the ZCTA becomes unfilled
  - For example, if you move the slider setting for % of Adults With No Usual Source of Care from 10 to 11, and you see a ZCTA become unfilled, you know that 10% of adults in that ZCTA have no usual source of care

Questions and Comments

Mark Carrozza

Mcarrozza@HealthLandscape.org

www.HealthLandscape.org

www.UDSMapper.org