

Climate and Economic Justice Screening Tool

Technical Support Document

Version 1.0

November 2022

White House Council on Environmental Quality

U.S. Digital Service

Contents

I. Introduction 4
II. Methodology 5
A. Methodology overview 5
B. Thresholds Approach for Burdens 7
1. Categories of Burdens 7
2. Communities surrounded by other disadvantaged tracts 11
3. Percentiles for normalization 12
4. Thresholds and cutoff values 13
C. Lands of Federally Recognized Tribes 15
D. Addressing missing data 16
E. Bug Fixes 16
III. Data 16
A. Data overview 16
B. Census tracts 20
C. Datasets used 22
1. Census's American Community Survey (ACS) 22
2. Census TIGER files 24
3. CDC's PLACES 24
4. CDC's U.S. Small-Area Life Expectancy Estimates Project (USALEEP) 24
5. DOE's LEAD Tool 25
6. DOI/BIA's Land Area Representation (LAR) 25
7. DOI's Abandoned Mine Lands Inventory System (eAMLIS) 25
8. DOT's Transportation Access Disadvantage category 25
9. EPA's EJScreen 26
10. FEMA's National Risk Index (NRI) 27
11. First Street Foundation's Climate Risk Data Access 28
12. HOLC maps / NCRC historic redlining score 28
13. HUD's Comprehensive Housing Affordability Strategy 29
14. MRLC's NLCD Percent Developed Imperviousness (CONUS) 30

15. USACE's Formerly Used Defense Sites (FUDS) 30
D. Datasets not currently eligible for inclusion 30
E. Challenges in identifying data for the U.S. territories 31
IV. User Interface 32
A. Improved Side Panels 32
B. Additional Features 32
C. Opportunities to provide feedback and get involved 33
D. Data Downloads 33
V. Looking ahead 35

I. Introduction

The Climate and Economic Justice Screening Tool ("CEJST" or "tool") is a critical component of the Biden-Harris Administration's historic commitment to advancing environmental justice. In Executive Order 14008 on Tackling the Climate Crisis at Home and Abroad, President Biden directed the White House Council on Environmental Quality (CEQ) to develop a geospatial mapping tool to identify disadvantaged communities that face burdens. These communities have been marginalized by society, overburdened by pollution, and underserved by infrastructure and other basic services.

Federal agencies will use the CEJST for the Justice40 Initiative. It will help them identify disadvantaged communities that should benefit from the Justice40 Initiative. The Justice40 Initiative seeks to deliver 40% of the overall benefits of certain Federal investments to disadvantaged communities. These investments relate to seven areas: climate change; clean energy and energy efficiency; clean transit; affordable and sustainable housing; the remediation and reduction of legacy pollution; the development of critical clean water and wastewater infrastructure; and training and workforce development. This task of delivering the benefits of hundreds of Federal programs to disadvantaged communities is challenging. It requires fundamental and sweeping changes to the ways in which the whole Federal government operates.

On November 22, 2022, CEQ launched version 1.0 of the CEJST. Version 1.0 incorporates feedback that CEQ received on the beta—or draft—version of the tool. CEQ had previously launched the beta version of the tool on February 18, 2022 with support from the U.S. Digital Service, and in collaboration with other Federal agencies and departments. The tool was released in a beta version in order to solicit feedback from Federal agencies, Tribal Nations, state and local governments, the White House Environmental Justice Advisory Council (WHEJAC), key stakeholders, and the public. By the end of the public comment period, CEQ/USDS had received almost 3,000 comments on the tool. Feedback was gathered through a variety of methods during the beta period, including: a Request for Information posted in the Federal Register, which had a 90 day comment period; an email support inbox; a survey accessible via the CEJST; census tract level feedback that could be provided via the map, and listening sessions. CEQ and USDS also conducted three public trainings and two agency-only trainings on the tool. They also met one-on-one with nearly 20 agencies with Justice40 covered programs. In addition, CEQ held two Tribal consultations on the tool to provide Tribal Nations with meaningful opportunities for input, consistent with CEQ's Action Plan for Consultation and Coordination with Tribal Nations, President Biden's Memorandum on Tribal Consultation and Strengthening Nation-to-Nation Relationships, and Executive Order 13175 on Consultation and Coordination With Indian Tribal Governments.

The CEJST is available at https://screeningtool.geoplatform.gov. The tool uses publicly-available, nationally-consistent datasets to identify disadvantaged communities. The datasets are indicators of burdens that disadvantaged communities face. These burdens are related to climate change, environment, health, and economic opportunity. Communities are considered disadvantaged if they are in census tracts that meet the thresholds for at least one of the tool's categories of burden, or if they are on the lands of Federally Recognized Tribes. The CEJST features a user-friendly, searchable map of all 50

states, the District of Columbia, and the U.S. territories. The CEJST website also has data files, such as spreadsheets and shapefiles, available for download.

This technical support document describes the data and methodology used in version 1.0 of the CEJST. It is organized as follows: Section II describes the methodology; Section III describes the data sources; Section IV describes changes to the user interface; and Section V concludes with next steps.

II. Methodology

A. Methodology overview

The tool highlights disadvantaged census tracts across all 50 states, the District of Columbia, and the U.S. territories. Census tracts are small units of geography that generally contain between 1,200 to 8,000 people, with an average size of 4,000 people. Census tract boundaries for statistical areas are determined by the U.S. Census Bureau once every ten years. The tool utilizes the census tract boundaries from 2010. This was chosen because many of the data sources in the tool currently use the 2010 census boundaries. Additional information on the decision to identify disadvantaged communities at the census tract level can be found in Section III.B.

In version 1.0 of the CEJST, communities are considered disadvantaged if they are in census tracts that meet the thresholds for at least one of the tool's categories of burden, or if they are on land within the boundaries of Federally Recognized Tribes. In addition, census tracts that are completely surrounded by disadvantaged communities are also considered disadvantaged if they meet an adjusted low income threshold. Table 1 provides an overview of the methodology and identifies the methodological changes made between the beta version and version 1.0 of the tool.

Table 1. Overview of methodology used in version 1.0 of the CEJST

(Items marked as NEW are changes made from the beta version to version 1.0)

Communities are considered disadvantaged:

- if they are located in a census tract that meets the thresholds for at least one of the tool's categories of burden, or;
- if they are on land within the boundaries of Federally Recognized Tribes **(NEW)**

Census tracts that are completely surrounded by disadvantaged communities are also considered disadvantaged if they meet an adjusted low income threshold (≥ 50th percentile). **(NEW)**

Category	Environmental, climate, or other burdens	Socioeconomic burden
Climate change	1. Expected agriculture loss rate ≥ 90th percentile OR 2. Expected building loss rate ≥ 90th percentile OR 3. Expected population loss rate ≥ 90th percentile OR 4. Projected flood risk ≥ 90th percentile **(NEW)** OR 5. Projected wildfire risk ≥ 90th percentile **(NEW)**	Low income*
Energy	1. Energy cost ≥ 90th percentile OR 2. PM 2.5 in the air ≥ 90th percentile	Low income*
Health	1. Asthma ≥ 90th percentile OR 2. Diabetes ≥ 90th percentile OR 3. Heart disease ≥ 90th percentile OR 4. Low life expectancy ≥ 90th percentile	Low income*
Housing	1. Historic underinvestment = Yes **(NEW)** 2. Housing cost ≥ 90th percentile OR 3. Lack of green space ≥ 90th percentile **(NEW)** OR 4. Lack of indoor plumbing ≥ 90th percentile **(NEW)** OR 5. Lead paint ≥ 90th percentile	Low income*
Legacy pollution	1. Abandoned mine land present = Yes **(NEW)** OR 2. Formerly Used Defense Site (FUDS) present = Yes **(NEW)** OR 3. Proximity to hazardous waste facilities ≥ 90th percentile OR 4. Proximity to Superfund or National Priorities List (NPL) sites ≥ 90th percentile OR 5. Proximity to Risk Management Plan (RMP) sites ≥ 90th percentile	Low income*
Transportation	1. Diesel particulate matter ≥ 90th percentile OR 2. Transportation barriers ≥ 90th percentile **(NEW)** OR 3. Traffic proximity and volume ≥ 90th percentile	Low income*
Water and wastewater	1. Underground storage tanks and releases ≥ 90th percentile **(NEW)** OR 2. Wastewater discharge ≥ 90th percentile	Low income*
Workforce development	1. Linguistic isolation ≥ 90th percentile OR 2. Low median income ≥ 90th percentile OR 3. Poverty ≥ 90th percentile OR 4. Unemployment ≥ 90th percentile	High school education < 10%

* Low Income = 65th percentile or above for census tracts that have people in households whose income is less than or equal to twice the federal poverty level, not including students enrolled in higher education **(NEW method of calculation)**

B. Thresholds Approach for Burdens

1. Categories of Burdens

The tool uses datasets as indicators of burdens. The burdens are organized into categories. A community is highlighted as disadvantaged on the CEJST map if it is in a census tract that is:

(1) at or above the threshold for one or more environmental, climate, or other burdens, and
(2) at or above the threshold for an associated socioeconomic burden.

In addition, a census tract that is completely surrounded by disadvantaged communities that meet the burden thresholds—and is at or above the 50th percentile for low income—is also considered disadvantaged.

The indicators of burdens used in the CEJST are informed by Executive Order 14008. It states that the goal is "to secure environmental justice and spur economic opportunity for disadvantaged communities that have been historically marginalized and overburdened by pollution and underinvestment in housing, transportation, water and wastewater infrastructure, and health care."[1] To achieve this goal, the CEJST burdens are grouped into categories that were informed by Justice40 investment focus areas.[2] The eight categories of burdens in the CEJST are: climate change, energy, health, housing, legacy pollution, transportation, water and wastewater, and workforce development.[3] The CEJST methodology was developed to reflect what EO 14008 makes clear—that disadvantaged communities face numerous challenges because they have been marginalized by society, overburdened by pollution, and underserved by infrastructure and other key services.

As a result of feedback received during the public comment period, the methodology for calculating the low income burden has changed between the beta version and version 1.0 of the CEJST. The beta version had used two socioeconomic burdens in all categories except for workforce development: low income and higher education non-enrollment. In order to be identified as a disadvantaged community, the beta version of the CEJST had required that 80% of the census tract's population be not currently enrolled in higher education. This higher education non-enrollment indicator had sought to ensure that communities with many students, such as those in college towns or university districts, would not be categorized as disadvantaged simply because many students lived there. However, as a result of feedback received during the beta period, it became clear that this indicator needed to be changed. In version 1.0 of the CEJST, the low income burden is now calculated by first excluding the population of students enrolled in higher education. Thus, the low income burden is now the percentage of a census

[1] Section 219, Executive Order 14008, Tackling the Climate Crisis at Home and Abroad (January 27, 2001).
[2] OMB Memorandum M-21-28, Interim Implementation Guidance for the Justice40 Initiative (July 20, 2021).
[3] Because health burdens are so tightly correlated with environmental and climate burdens, the CEJST also includes a health category even though is not a Justice40 investment area. In addition, this categorization also enables CEQ to understand where additional nationwide data sources are needed. For example, the data team has been working to identify additional indicators for the water and wastewater quality category due to the lack of relevant data currently available at the census tract level.

tract's population in households where household income is at or below 200% of the Federal poverty level, not including students enrolled in higher education.

The workforce development category in version 1.0 of the CEJST only has one socioeconomic indicator, which is that fewer than 10% of people ages 25 or older have a high school education (i.e. graduated with a high school diploma). In the beta version of the CEJST, the higher education non-enrollment indicator was also used in the workforce development category. It measured the percentage of the census tract's population 15 or older that was not enrolled in college, university, or graduate school. As a result of feedback received, version 1.0 of the CEJST no longer uses the higher education non-enrollment indicator for the workforce development category.

i. Climate change

The burdens in the climate change category aim to measure expected agriculture value, building value, and population loss[4] due to climate-related natural hazards, as well as projected wildfire risk and projected flood risk due to climate change. These last two burdens (wildfire and flood risk) were added to version 1.0 of the CEJST in response to public feedback. To be identified as disadvantaged in the climate change category, communities must be in census tracts that are at or above the threshold for at least one of these burdens, i.e. at or above the 90th percentile. They must also meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education. Data comes from Federal Emergency Management Agency (FEMA)'s National Risk Index, the First Street Foundation's Climate Risk dataset, and the U.S. Census's American Community Survey.

ii. Energy

The burdens aim to measure the energy cost as well as energy-related pollution within a census tract. To be identified as disadvantaged in the energy category, communities must be in census tracts that are at or above the threshold for at least one of these burdens, i.e. at or above the 90th percentile. They must also meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education. Data comes from the Department of Energy (DOE)'s LEAD Score, the Environmental Protection Agency (EPA)'s EJScreen, and the U.S. Census's American Community Survey.

iii. Health

The burdens aim to identify areas facing high rates of asthma, diabetes, heart disease, and low life expectancy within a census tract. To be identified as disadvantaged in the health category, communities must be in census tracts that are at or above the threshold for at least one of these burdens, i.e. at or above the 90th percentile. They must also meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for

[4] In version 1.0, the population loss burden was only applied to census tracts with populations of 20 or greater.

calculating the low income burden that excludes students enrolled in higher education. Data come from the Centers for Disease Control (CDC)'s PLACES, the CDC's USALEEP, and the U.S. Census's American Community Survey.

iv. Housing

The burdens aim to measure the housing cost, the degree of lead paint exposure in housing, historic underinvestment due to redlining, lack of green space, and the share of homes without indoor plumbing or kitchens within a census tract. These last three burdens are new to version 1.0. To be identified as disadvantaged in the housing category, communities must be in census tracts that are at or above the threshold of at least one of these burdens. The threshold is set at or above the 90th percentile for all of the burdens with one exception. For the historic underinvestment burden, census tracts display "yes" for this burden when they have a score of 3.25 or more out of 4.[5] In addition, census tracts must meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education. Data come from Housing Urban Development (HUD)'s Comprehensive Housing Affordability Strategy, National Community Reinvestment Coalition (using data from maps created by the Home Owners Loan Corporation between 1935 and 1940),[6] the Trust for Public Lands and American Forestry's Percent developed Imperviousness (CONUS), and the U.S. Census's American Community Survey.

v. Legacy pollution

The burdens aim to measure how much legacy, current, and potential pollution a census tract has through proximity to hazardous waste, Superfund sites (otherwise known as National Priorities List (NPL)), Risk Management Plan (RMP) facilities, abandoned mine land, and Formerly Used Defense Sites (FUDS). To be identified as a disadvantaged community in the legacy pollution category, communities must be in census tracts that are at or above the threshold of at least one of these burdens. The threshold is set at or above the 90th percentile for all of the burdens with two exceptions. Census tracts display "yes" if they contain at least one abandoned mine land or FUDS, both of which are new burdens in version 1.0 of the tool. In addition, census tracts must meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education. Data come from Environmental Protection Agency (EPA)'s EJScreen, the Bureau of Land Management (BLM)'s Abandoned Mine Land Inventory at the Department of the Interior (DOI), U.S. Army Corps of

[5] The tool uses the National Community Reinvestment Coalition's methodology for converting boundaries in the Home Owners Loan Corporation (HOLC)'s maps to census tracts. The historic underinvestment burden is not available for tracts that were not included in the original HOLC maps because there is no underlying data.
[6] The full citation for the dataset is as follows: Meier, Helen C.S., and Mitchell, Bruce C. . Historic Redlining Scores for 2010 and 2020 US Census Tracts. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-10-15. https://doi.org/10.3886/E141121V2

Engineers (USACE)'s Formerly Used Defense Sites Inventory, and the U.S. Census's American Community Survey.

vi. Transportation

The burdens aim to measure the transportation-related pollution, transportation barriers, traffic related noise and proximity within a census tract. Transportation barriers is a new burden added to version 1.0 of the CEJST.[7] To be considered disadvantaged in the transportation category, communities must be in census tracts that are at or above the threshold for at least one of these burdens, i.e. at or above the 90th percentile. In addition, census tracts must meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education. Data comes from the Environmental Protection Agency (EPA)'s EJScreen, the Department of Transportation (DOT)'s mapping tool of Transportation Disadvantaged Communities, and the U.S. Census's American Community Survey.

vii. Water and wastewater

The burdens aim to measure the census tract's proximity to toxicity-weighted wastewater discharges and underground storage tanks that may leak, which is a new burden added to version 1.0 of the CEJST. To be considered disadvantaged in the water and wastewater category, communities must be in tracts that are at or above the threshold for at least one of these burdens, i.e. at or above the 90th percentile. In addition, census tracts must meet the threshold for the low income burden, which is at or above the 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education. Data come from the Environmental Protection Agency (EPA)'s EJScreen, the EPA's Underground Storage Tank Finder via EJScreen, and the U.S. Census's American Community Survey.

viii. Workforce development

The burdens aim to identify census tracts that would benefit from greater workforce development. This is in line with the Justice40 Initiative's goal of promoting training and workforce development related to climate, clean energy, and other related categories. These burdens are low median income as a percentage of area median income, percent of households in linguistic isolation, percent of the workforce experiencing unemployment, and percentage of a census tract's population in households where the household income is at or below 100% of the Federal poverty level. To be considered disadvantaged in the workforce development category, communities must be census tracts that are at or above the threshold for at least one of these burdens, i.e. at or above the 90th percentile. They must also meet the threshold of the socioeconomic indicator, which is that fewer than 10% of people ages 25 or older have a high school education (i.e. graduated with a high school diploma). As noted above, the higher education non-enrollment indicator used in the beta version of the CEJST is no longer used in version 1.0; that indicator had measured the percentage of the census tract's population 15 or older that was not enrolled in college, university, or graduate school. Data come from the U.S. Census's

[7] The transportation barriers indicator was only used in census tracts with populations of 20 or greater.

American Community Survey (2015-2019) for all U.S. states, the District of Columbia, and Puerto Rico. Data come from the American Community Survey (2010) for American Samoa, Guam, the Northern Mariana Islands, and the U.S. Virgin Islands.

2. Communities surrounded by other disadvantaged tracts

In response to feedback on the beta version of the tool, census tracts that are completely (i.e. 100 percent) surrounded by disadvantaged communities are now identified disadvantaged if they meet an adjusted low income threshold (at or above the 50th percentile).

This methodology change was informed by concerns about geospatially correlated errors in datasets, i.e. several datasets might not capture an accurate picture of the same location. For small geographic units like census tracts, these errors can have magnifying effects. Statistical theory, namely Tobler's Law, states that geographically close areas are more similar than areas that are far apart. In other words, adjacent census tracts generally exhibit similar characteristics. If a census tract is not considered disadvantaged, even though it is almost completely surrounded by other disadvantaged tracts, then it is statistically likely that this tract was excluded as a result of data aberrations. As the feedback received during the public comment of the beta period made clear, addressing concerns about census tracts surrounded by disadvantaged tracts is more than just statistics: being surrounded by disadvantaged communities is, in itself, a form of disadvantage. These communities may have a harder time accessing public services and may experience greater environmental burdens because they are surrounded by disadvantaged communities. This methodological change in version 1.0 of the CEJST attempts to address this problem.

The following examples explain how the methodology works for disadvantaged that are completely surrounded by disadvantaged communities.[8] Assume that these census tracts do not contain the lands of Federally Recognized Tribes.

- Assume there is a tract that is completely surrounded by tracts that are considered disadvantaged. This tract does not meet or exceed the cutoffs for any environmental, climate, or other burden. It also does not meet the cutoff for the low income burden. This means it is not considered disadvantaged by the baseline CEJST methodology. However, because it is surrounded by disadvantaged tracts and it is above the 50th percentile for low income, the census tract is identified as disadvantaged in version 1.0 of the CEJST.
- Assume there is a tract that is above the 90th percentile for several environmental, climate, or other burdens, but it is not above the 65th percentile for low income. But, it is surrounded by disadvantaged tracts. This means it only needs to be above the 50th percentile for low income. If it is, the census tract is identified as disadvantaged in version 1.0 of the CEJST.
- Assume there is a tract that is only surrounded by disadvantaged communities on 75 percent of its borders. This tract does not otherwise meet any of the thresholds of the burdens in the

[8] The calculation for being completely surrounded does not include borders of the tract that are water, such as oceans or rivers.

version 1.0 methodology discussed above. This tract is not be identified as disadvantaged because it is not 100 percent surrounded by other disadvantaged tracts.

3. Percentiles for normalization

The tool ranks most of the burdens using percentiles. Percentiles show how much burden each tract experiences when compared to other tracts. Percentiles reflect the distribution of a variable and run from the 0th percentile, meaning the lowest score on a given scale, to the 100th percentile, meaning the highest score on a given scale. A variable that lists a census tract as "77th percentile" means that this tract has a higher measurement on this variable than 77% of all census tracts. Percentiles help to normalize data measured in different units and on different scales. This means the tool can use data for all regions and types of areas in the U.S. and U.S. territories. Percentiles provide a useful way to describe and compare data that may be on different scales. They make it possible to identify the relative burden that each census tract experiences.

The CEJST converts most variables to percentiles in order to normalize them. This also makes it possible to present them on a single scale. The variables used in the CEJST 1.0 naturally have very different scales. For instance, PM2.5 is measured in micrograms of particulate matter per cubic meter of air. It generally ranges from about 2 or 3 to 300. In contrast, wastewater discharge is measured as toxicity-weighted stream concentrations at stream segments within 500 meters, divided by distance in kilometers. A percentile-based distribution can combine two such disparate variables into a single scale by converting them both to percentiles.

For burdens in the CEJST, a higher percentile generally refers to a community being more overburdened and/or underserved on that indicator, e.g., a community is exposed to pollution. In order to achieve this directional uniformity, the tool reverses the percentiles for some burdens. This change makes the data display consistent with other burdens. For instance, greater life expectancy is a sign of a healthy community. It would be confusing to require users to remember that "low" values for this particular metric indicate high burden. Instead the CEJST uses the percentile measurement for *low* life expectancy, so that a high percentile indicates a high amount of burden.

The advantages and disadvantages of different methods for data normalization were considered before a decision was made to use percentiles in the tool. One of the downsides of using percentiles is that there is no measurement of orders of non-linear magnitude. For instance, if there's a bimodal distribution (e.g., a lot of "very safe air" and a lot of "very bad air" neighborhoods), the gap between those modes will not be represented well by converting the data into a linear, percentile-based scale. The difference in impact of an indicator (e.g., the effect that air quality has on human breathing) between the 50th and 51st percentile may be a lot smaller than the difference in impact between the 90th and 91st percentile. However, other approaches to normalization also have notable downsides as well. For instance, z-scores are often inappropriate for data that is not normally distributed, and there is no perfect solution for this challenge. Compared to the other available options, the advantages to using percentiles outweighed the disadvantages. Percentiles can be interpreted and understood more easily

than many of the other methodologies for normalization. Moreover, indicators that have high impacts at the 90th percentile or higher were selected in order to reduce concerns about non-linear impacts.

4. Thresholds and cutoff values

The burdens in the CEJST use thresholds, or cutoffs, which enable the tool to account for regional, state, and other geographic differences across the U.S. This includes between urban and rural areas. Each threshold is measured independently. The thresholds do not work against each other. Adding a new threshold to the methodology will add some communities to the definition of disadvantaged without removing other communities. For example, in the climate change category, projected coastal flooding and wildfires are both included among the environmental or climate burdens. This means that tracts that are projected to experience a high rate of coastal flooding, a high rate of wildfires, or both, are represented.

Most of the CEJST burdens use percentiles as cutoff values. The cutoff value for the environmental, climate, and other burdens that use percentiles is at or above the 90th percentile. This results in the selection of the top 10% of tracts for each burden. This allows the CEJST to identify census tracts with the greatest environmental, climate, and other burdens. The cutoff for the low income indicator (<200% of the Federal poverty line) is at or above 65th percentile. As noted above, version 1.0 of the CEJST uses a new methodology for calculating the low income burden that excludes students enrolled in higher education.

There is one cutoff based on the raw percent value rather than percentile. In the workforce development category, a community would is identified as disadvantaged if, in addition to meeting other criteria in the category, 10% or more of the percent of adults 25 or older living in the census tract did not graduate high school. Approximately fifty percent of census tracts in the CEJST have populations where 10% or more of adults aged 25 or older did not graduate high school.

A few burdens in the CEJST use a simple yes/no (i.e. Boolean strings) for their cutoffs. A yes means that people in the census tract are experiencing the burden. If a census tract contains at least one Formerly Used Defense Sites (FUDS) or abandoned mine land, then the respective burden is represented as "yes." This means that the threshold has been met. The historic underinvestment burden also uses yes/no to represent whether the census tract was extensively redlined or not. As noted above, the tool uses the National Community Reinvestment Coalition's methodology for converting boundaries in redlining maps (i.e. maps created by Federal government's Home Owners' Loan Corporation between 1935 and 1940) into census tracts. Census tracts meet the threshold in version 1.0 of the CEJST when they have a score of 3.25 or more out of 4.

In contrast to the threshold approach used in the CEJST, some environmental justice screening tools use an indexing method, averaging and multiplying indicators to combine normalized data into a single index score. These approaches often incorporate a cumulative impacts approach by showing how adverse factors create additional climate, environmental, or other burdens when combined together. These include state-based tools such as CalEnviroScreen and Maryland's EJScreen Mapper. In addition, Environmental Protection Agency (EPA)'s EJScreen and the Department of Health and Human Services'

Environmental Justice Index are examples at the Federal level. The CEJST is an iterative tool and it is anticipated that a future version of the tool will better represent cumulative impacts.

i. Examples of thresholds approach

The burden thresholds work together to identify disadvantaged census tracts. The thresholds approach works to identify disadvantaged communities that are at or above the cutoff values for *both* environmental, climate, and other burdens and the related socioeconomic burdens.

Consider three examples from the legacy pollution category:

- **Census tracts that meet the environmental, climate, or other burden only**: Assume a census tract is in the 95th percentile for proximity to Superfund (or National Priorities List (NPL)) sites, but it is at the 40th percentile for low income. Even though it meets an environmental, climate or other burden's threshold (proximity to Superfund sites), it is not considered a disadvantaged community for the legacy pollution category because it does not meet the related socioeconomic indicator. This also means it would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tracts that meet the socioeconomic burden only**: Assume a census tract is in the 83rd percentile for low income. It meets the cutoff for the related socioeconomic burden for the legacy pollution category. But, it does not meet any of the environmental, climate or other burdens for the category. Assume, for example, that the census tract is only in the 85th percentile for proximity to hazardous waste facilities, the 87th percentile for proximity to Superfund (or National Priorities List (NPL)) sites, and the 86th percentile for proximity to RMP facilities; and, that it does not have a Formally Used Defense Site (FUDS) or an abandoned mine in the tract. Further assume that this tract is not completely surrounded by disadvantaged communities. Thus, even though it meets the related socioeconomic burden (low income), it does not meet the environmental, climate, or other burdens. It would not be considered a disadvantaged community, and would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tracts that meet both the environmental, climate, or other burdens and the related socioeconomic burdens:** Assume a census tract is in the 92nd percentile for proximity to Superfund (or National Priorities List (NPL)) sites and is in the 87th percentile for low income. The census tract is disadvantaged because it exceeds the cutoffs for both the environmental burden and the related socioeconomic burden for the legacy pollution category. It would be highlighted the color blue-gray on the tool's website map. Note: the census tract does not have to meet or exceed all of the other environmental, climate, or other burdens to exceed the threshold within the legacy pollution category.

Three examples for the thresholds used in the workforce development category offer more perspective. This category has some of the available CEJST data that can be used to identify disadvantaged communities in the U.S. territories:

- **Census tracts that meet the environmental, climate, or other burdens only:** Assume a census tract is in the 95th percentile for unemployment rate, but 55 percent of the population aged 25

years or older graduated from high school. The socioeconomic burden is not met because the threshold is set as fewer than 10% of people have a high school diploma. This tract would not be considered disadvantaged and would not be highlighted in a blue-gray shade on the tool's website map.

- **Census tracts that meet the socioeconomic burden only:** Assume a census tract has a population where 20 percent of the population aged 25 or older did not graduate from high school. This tract meets the cutoff for the socioeconomic burden for the workforce development category. But, if the same census tract is at the 85th percentile for low median income, the 87th percentile for percentage of people living at or below the poverty line, the 40th percentile for number of households without a person over the age of 14 who speaks English well, and 86th percentile for unemployment, it does not satisfy any of the environmental, climate, or other burdens for the category. None of these variables meet the 90th percentile cutoff. Even though the census tract meets the socioeconomic burden (high school education), it does not meet the environmental, climate, or other burdens (median income, poverty, no person over 14 who speaks English well, and unemployment). It would not be considered a disadvantaged community, and would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tracts that meet both the environmental, climate, or other burdens and the socioeconomic burdens:** Assume a census tract is in the 92nd percentile for unemployment rate and 20 percent of the population aged 25 or older did not graduate from high school. The census tract is labeled as disadvantaged because it satisfies at least one of the environmental, climate, or other burdens and also the socioeconomic burden for the workforce development category. It would be highlighted in a blue-gray shade on the tool's website map. Note: the census tract does not have to meet the cutoff for all of the other environmental, climate, or other burdens to be considered disadvantaged within the category.

C. Lands of Federally Recognized Tribes

To respect Tribal sovereignty and self-government and to fulfill Federal trust and treaty responsibilities to Tribal Nations, land within the boundaries of Federally Recognized Tribes are designated as disadvantaged on the map. Alaska Native Villages are also displayed as point locations. Census tracts that contain 99.5% or more of the lands of Tribes are identified as disadvantaged communities. When the lands of Tribes is displayed on the map, the name of the Tribe is also shown.

This decision was made after meaningful and robust consultation with Tribal Nations and in response to other feedback received during the public comment period on the beta version. This approach is also consistent with CEQ's Action Plan for Consultation and Coordination with Tribal Nations, President Biden's Memorandum on Tribal Consultation and Strengthening Nation-to-Nation Relationships, and Executive Order 13175 on Consultation and Coordination With Indian Tribal Governments.

Regardless of whether a census tract meets the threshold for at least one of the tool's categories of burdens or if it contains the lands of Tribes, it is shaded on the map in the same color. However, where

there is overlap (i.e. the tract meets the threshold for at least one of the tool's categories of burdens AND if it contains the lands of Tribes), the shade is darker due to the way in which the tool displays the overlapping layers. The different shading is a function of layering the two different methodologies on the map and should not be viewed as being significant for purposes of the CEJST; all areas that are shaded are disadvantaged communities for purposes of the CEJST.

In some instances, a census tract is partially comprised of the lands of Federally Recognized Tribes, but does not otherwise meet the threshold for at least one of the tool's categories of burdens. These tracts are considered to be partially disadvantaged because only the lands of Tribes in that tract are identified as disadvantaged communities. The side panel that appears for each census tract on the tool's map states the percentage of the tract that is comprised of the lands of Tribes. The tool also explains why the tract is identified as partially disadvantaged.

D. Addressing missing data

Following the release of the beta version of the CEJST, Federal agencies and members of the public called attention to certain census tracts that were missing income information. Approximately 800 census tracts were missing income data in the beta version of the CEJST. The missing income data meant that these tracts were automatically disqualified from being identified as disadvantaged communities. The version 1.0 of the CEJST addresses this problem for census tracts with known populations greater than 0.

For tracts that are missing income information, the low income indicator is now estimated in the tool. This estimate is done by taking the mean of the low income indicator for the contiguous tracts. If the contiguous tracts were also missing income values, then the average of the low income burden for the county or state is used. In order to compute the revised low income burden, the percent of the population currently enrolled in higher education is also estimated. This aligns with best practices for working with geospatial data and is a common statistical practice. These tracts will show the symbol for approximately (~) before the low income burden in version 1.0 of the CEJST.

E. Bug Fixes

Version 1.0 of the CEJST also fixes two bugs that were in the beta version. Life expectancy data that was missing for Maine and Wisconsin in the beta version has now been added to the tool. In addition, census tracts that are 100% water are not identified as disadvantaged communities on the map. However, these tracts will still be listed in the data downloads because they are included in the census data.

III. Data

A. Data overview

Version 1.0 of the CEJST includes data from a wide variety of sources, including the U.S. Army Corps of Engineers (USACE), the U.S. Census Bureau (Census), the U.S. Centers for Disease Control (CDC), the Department of Energy (DOE), the U.S. Department of Housing and Urban Development (HUD), the U.S.

Department of the Interior (DOI) (including the Bureau of Indian Affairs (BIA)), the U.S. Department of Transportation (DOT), the U.S. Environmental Protection Agency (EPA), the Federal Emergency Management Agency (FEMA), the First Street Foundation, the U.S. Department of Housing and Urban Development (HUD), the National Community Reinvestment Coalition (using data from the Home Owner's Loan Corporation), the Multi-Resolution Land Characteristics (MRLC) consortium, and the U.S. Army Corps of Engineers (USACE). The datasets were chosen after receiving hundreds of dataset ideas from Federal agencies, environmental justice data experts, and the WHEJAC. Each included dataset meets the following relevance, availability, and quality requirements:

1. **Relevance to the goals of Executive Order 14008 and the Justice40 Initiative:** To achieve the environmental justice goals set forth in Executive Order 14008 and ensure that the CEJST is relevant to the Justice40 Initiative, the tool includes comprehensive datasets related to climate, environmental, energy, and economic justice.
2. **Publicly available:** The CEJST and its methodology are publicly available. Confidential or proprietary data are not included in the tool.
3. **Comprehensive and consistently accurate:** CEQ sought to identify disadvantaged communities across the 50 states, D.C., and the five U.S. territories of Puerto Rico, American Samoa, the Northern Mariana Islands, Guam, and the U.S. Virgin Islands. Datasets that include all 50 states and D.C., as well as the five U.S. territories are preferred, but many datasets are not available for the U.S. territories. Datasets were only selected if, at a minimum, they were comprehensive and consistently accurate for all 50 states and D.C. The tool's datasets are relevant to both rural and urban areas. The historic underinvestment burden is not available for tracts that were not included in the original HOLC maps because there is no underlying data.
4. **Data is available either at the point-level or at the census tract level**: The CEJST identifies disadvantaged communities at the census tract level. This level is the smallest geographic unit for which publicly available, nationally consistent datasets can be consistently displayed on the tool. Datasets that provide information at the state or county level do not have the level of granularity needed to incorporate into the CEJST. The tool generally uses datasets available at the census tract level; for crucial data that is only available at site- or point-level, the tool assigns latitude/longitude coordinates to individual census tracts.

The datasets used in version 1.0 of the CEJST are summarized in the table below. Certain variables are used directly in the CEJST, without additional calculations. Others are calculated by the CEJST before use in the tool. More detail about how the datasets are used can be found in the table below.

Table 2. Overview of data used in version 1.0 of the CEJST

Source (date)	Dataset	Variables included or calculated	Relevant categories	Used for
Census (2015-2019)* Census (2010)^	American Community Survey (ACS)	● High school education (CEJST calculates)*^ ● Lead paint (CEJST calculates using housing age and median home value)* ● Linguistic isolation (CEJST calculates)* ● Low income* ● Low median income*^ ● Poverty (CEJST calculates)*^ ● Unemployment*^ ● Demographic information*^ (displayed in the map's side panel for each census tract) **(NEW)**	All categories	* All U.S. states, the District of Columbia, and Puerto Rico[9] ^ American Samoa, Guam, the Northern Mariana Islands, and the U.S. Virgin Islands
Census (2010)	TIGER files	● Geographies of census tract boundaries	N/A	All U.S. states, the District of Columbia, and U.S. territories
CDC (2016-2019)	PLACES: Local Data for Better Health	● Asthma ● Diabetes ● Heart disease	Health	All U.S. states and the District of Columbia
CDC (2010-2015)	U.S. Small-area Life Expectancy Estimates Project (USALEEP)	● Low life expectancy	Health	All U.S. states and the District of Columbia
DOE (2018)	Low-Income Energy Affordability Data (LEAD) Tool	● Energy cost	Energy	All U.S. states, the District of Columbia, and Puerto Rico

[9] Linguistic isolation was removed for Puerto Rico in version 1.0 of the CEJST is based on feedback received during the beta period.

Source (date)	Dataset	Variables included or calculated	Relevant categories	Used for
DOI/BIA (2018)	Land Area Representation	● Lands of Federally Recognized Tribes, including the locations of Alaska Native Villages **(NEW)**	N/A	Federally Recognized Tribes, including Alaska Native villages
DOI (2017)	Abandoned Mine Lands Inventory System	● Abandoned mine land **(NEW)**	Legacy Pollution	All U.S. states and the District of Columbia
DOT (2022)	Transportation Access Disadvantage	● Transportation barriers **(NEW)**	Transportation	All U.S. states and the District of Columbia
EPA (varies, 2014-2021)	EJScreen	● Diesel particulate matter exposure* ● PM2.5 in the air+ ● Proximity to hazardous waste facilities* ● Proximity to Risk Management Plan (RMP) facilities* ● Proximity to Superfund (or National Priorities List (NPL)) sites* ● Traffic proximity and volume* ● Underground storage tanks and releases* **(NEW)** ● Wastewater discharge*	● Energy ● Transportation ● Legacy pollution ● Water and wastewater	*All U.S. states, the District of Columbia, and Puerto Rico + All U.S. states and the District of Columbia
FEMA (2014-2021)	National Risk Index	● Expected agriculture loss rate (CEJST calculates) ● Expected building loss rate (CEJST calculates) ● Expected population loss rate (CEJST calculates)	Climate change	All U.S. states and the District of Columbia
First Street Foundation (2022)	Climate Risk Data Access	● Projected flood risk **(NEW)** ● Projected wildfire risk **(NEW)**	Climate Change	All contiguous U.S. states and the District of Columbia

Source (date)	Dataset	Variables included or calculated	Relevant categories	Used for
HOLC (1935-1940)/ NCRC (2021)	Historic Redlining Scores[10]	• Tracts determined to historically have been subject to redlining by the Home Owner's Loan Corporation between 1935-1940 **(NEW)**	Housing	Metro areas of U.S. that were graded by HOLC
HUD (2014-2018)	Comprehensive Housing Affordability Strategy (CHAS)	• Housing cost (CEJST calculates) • Lack of indoor plumbing **(NEW)**	Housing	All U.S. states, the District of Columbia, and Puerto Rico
MRLC consortium[11]	Percent Developed Imperviousness (CONUS)	• Lack of greenspace (CEJST calculates) **(NEW)**	Housing	All contiguous U.S. states and the District of Columbia
USACE (2019)	Formerly Used Defense Sites	• Formerly Used Defense Sites **(NEW)**	Legacy Pollution	All U.S. states and the District of Columbia

B. Census tracts

The CEJST identifies disadvantaged communities at the census tract level. Census tracts generally contain between 1,200 to 8,000 people, with an average size of 4,000 people. Selecting the appropriate geographic unit of analysis for a screening tool used for geographically-based resource allocation involves a tradeoff: the geographic area must be small enough for targeting benefits, but not so small that it introduces statistical unreliability in the tool. The decision to use census tracts was made after consideration of other alternatives such as census block groups.

The development team initially examined whether it was possible for the unit of analysis to be U.S. census block groups. It was decided that adequate data could not be reliably reported at this level at this time. The U.S. Census's American Community Survey have data available down to the census block group level. Block groups are usually between 600 and 3,000 people. These estimates at the census

[10] Meier, Helen C.S., and Mitchell, Bruce C. Historic Redlining Scores for 2010 and 2020 US Census Tracts. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-10-15. https://doi.org/10.3886/E141121V2

[11] The Trust for Public Lands and American Forests worked with the development team to compile and analyze the data for the lack of greenspace indicator.

block group level are statistically unreliable in communities where sampling is limited. This means that there are high margins of error. During testing of the beta version of the CEJST, the use of census data at the block level resulted in some communities being identified as disadvantaged due to unreliable data.

Limited nationally consistent environmental and climate data are reported at the census block group level. Even data that at first glance appear to be at the U.S. census block group level may not be. For example, in order to report data at the U.S. census block group level, the Environmental Protection Agency (EPA)'s EJScreen tool makes adjustments for data that are only actually reported at the census tract level, such as the National Air Toxics Assessment (NATA), PM2.5, and ozone estimates. For these datasets, EJScreen assigns each census block group the NATA, PM2.5, or ozone score of the census tract in which it is located. Every census block group within the census tract is given the same exact variable. When this same approach was used for all the CEJST datasets, it generated significant unreliability.

The development team considered other geographic units that were ultimately determined to be too large to be useful for the purpose of targeting benefits to disadvantaged communities. Many datasets are reported at the level of U.S. counties. Some counties like Los Angeles County in California contain millions of people. Other datasets are available at the zip code level. Some zip codes can contain as many as 120,000 people. Zip codes do not fit neatly inside of the U.S. Census Bureau's hierarchy of geographies.[12] If zip codes were used as the unit of analysis, data reported by census tract would not easily be able to be matched to zip codes. Although translations between zip codes and census tracts do exist, the disadvantages of using data reported at the zip code level outweigh the benefits.

Like the beta version, version 1.0 of the CEJST uses census tract boundaries for statistical areas from 2010.[13] This was chosen because many of the data sources in the tool currently use the 2010 census boundaries. Although the CEJST uses 2010 tract *boundaries*, much of the data in the CEJST comes from more recent years. For instance, proximity to hazardous waste facilities is calculated using data from 2020, but these data were collected using 2010 census tract boundaries. The 2010 boundaries are simply the geographic unit in which the data was collected, not the date of the data collection itself. As noted below, the Census TIGER files are used for these boundaries.

The U.S. Census Bureau updated the census tract boundaries for statistical areas in 2020. The CEJST is an iterative tool and it is likely that a future version will adopt the 2020 census tract boundaries.

[12] The Missouri Census Data Center's tool Geocorr 2018 was used to compare the data from census tracts against other geographic boundaries that do not cleanly fit inside the Census's hierarchy of geographies.
[13] The U.S. Census also makes small adjustments to the census tract boundaries for legally defined geographic areas on a yearly basis.

C. Datasets used

The goal of the CEJST is to identify disadvantaged communities across the 50 states, the District of Columbia, and the five U.S. territories of Puerto Rico, American Samoa, the Northern Mariana Islands, Guam, and the U.S. Virgin Islands. Many datasets are not currently available for the U.S. territories.

The CEJST uses some datasets directly, without any additional calculations. For other datasets, certain calculations have to be made prior to the data being included in the tool.[14] The discussion of each dataset below describes when each approach was used.

1. Census's American Community Survey (ACS)

The American Community Survey (ACS) is an annual survey that the U.S. Census Bureau administers to gain information about the nation and its people. The U.S. Census Bureau invites a randomized subset of 1 in 38 U.S. households to participate in the survey. Over 3.5 million households submit completed surveys each year. The American Community Survey (2015-2019) includes data for all 50 states, D.C., and Puerto Rico for all of these variables. The American Community Survey (2010) includes data for American Samoa, Guam, the Northern Mariana Islands, and the U.S. Virgin Islands.

Data used directly in the CEJST, without additional calculations:

- **Low median income:** Low median income is defined as a percent of the area's median income. Area median income identifies the midpoint of a given area's income distribution. It accounts for regional differences in cost of living.
    - Unlike most of the other datasets, high values of this indicator indicate low burdens. The tool reverses this percentile. The tract with the highest value is at the lowest or 0th percentile.
- **Poverty:** The American Community Survey does not directly provide a measure of the percent of individuals living in households who are at or below 100% of the Federal poverty line for a given year. Poverty is calculated by adding together households in different percentage brackets of the Federal poverty line. The Federal poverty level, also known as the "poverty line," is a measure of the minimum amount of income needed for households to pay for essentials. This level is issued every year by the Department of Health and Human Services.
- **Unemployment:** Unemployment reports the percent of unemployed people in the civilian labor force.
- **Demographic information**: The user interface of version 1.0 of the CEJST now displays demographic information (race/ethnicity and age) that is available through the American Community Survey. This information is provided to help inform the public about the tool and is not used in the methodology.

[14] For information on which specific fieldnames to include, please see the public GitHub repository for this project: https://github.com/usds/justice40-tool

Variables calculated by the CEJST before use in the tool:

- **Low income:** The low income variable requires calculations. First, the percent of the population within a given census tract living at or below 200% of the federal poverty line is calculated. Next, the percent of adults ages 15 and higher enrolled in higher education full time is calculated. The two numbers are subtracted to get the percentage of the census tract that is not enrolled full time in higher education and is also below 200% of the federal poverty line. Because this variable is critical to the tool, if it is missing it is estimated based on the average of the contiguous geographic neighboring census tracts. In the case of missing income and education data, the combined low income indicator is estimated. In the case of missing education data or income data only, only the missing variable is estimated.
    - Note: The way in which the low income indicator is calculated changed between the beta version and version 1.0 of the CEJST. The low income indicator now excludes the population of students enrolled in higher education. Version 1.0 of the CEJST no longer has a separate indicator for "higher education non-enrollment."
- **High school education:** High school education is the percentage of adults over 25 years old within a given census tract who have not graduated from high school. The American Community Survey does not provide a measure of high school diploma attainment. It is calculated by adding together all measures of educational achievement up until less than high school graduation.
- **Lead paint:** In the absence of data directly measuring lead paint exposure, lead paint is calculated in the tool via a proxy.
    - The CEJST uses the percent of housing units built before 1960, which indicates potential lead paint exposure.
    - Tracts with extremely high home values (i.e. median home values above the 90th percentile) that are less likely to face health risks from lead paint exposure are not included in the lead paint indicator. The CEJST uses low median home value of owner-occupied housing units from the American Community Survey. Home value is the estimated price a home would sell if it was put on the market. Median home prices in the American Community Survey data are based on survey respondents' answers to a subjective question asking for the dollar value of their home.
- **Linguistic isolation:** Linguistic isolation refers to households that responded to the American Community Survey questions on language by stating all persons over the age 14 speak English less than "very well." The American Community Survey does not provide a measure of linguistic isolation. It is calculated by adding together all measures of individually limited-English speaking households together.
    - Note: Linguistic isolation is not included in all places where data is available. Linguistic isolation does not function as an effective burden in some areas of the nation. Spanish is an official language of Puerto Rico. Thus, linguistic isolation was removed as an indicator to identify disadvantaged communities in Puerto Rico in version 1.0 of the CEJST. This change was made based on feedback during the tool's public beta period.

2. Census TIGER files

Census tract boundaries are provided by the U.S. Census's 2010 TIGER files. TIGER/Line shapefiles are an extract of selected geographic and cartographic information from the U.S. Census Master Address File / Topologically Integrated Geographic Bureau's Master Address File (MAF/TIGER) Database (MTDB). These files provide the data on the location of census blocks. They also provide housing unit counts and populations for a given year. This data is available for public download and consumption at https://data.census.gov/.

3. CDC's PLACES

PLACES is a public, interactive website launched by the Centers for Disease Control (CDC) in partnership with the Robert Wood Johnson Foundation and the CDC Foundation. PLACES has estimates of 29 health measures, including health risk behaviors, health outcomes, health status, and clinical preventative practices. The CEJST uses data from the 2016-2019 PLACES dataset. This data is collected for all U.S. states and the District of Columbia.

Variables used directly in the CEJST, without additional calculations:

- **Asthma:** Weighted percentage of people who answer "yes" to both of the following questions: (1) have you ever been told by a doctor, nurse, or other health professional that you have asthma?, and (2) do you still have asthma? The weighted percentage is calculated in PLACES ensures that survey responses match the population characteristics of each census tract.
- **Diabetes:** Weighted percentage of people ages 18 years and older who report being told by a health professional that they have diabetes, other than diabetes during pregnancy.
- **Heart disease:** Weighted percentage of people ages 18 years and older who report ever having been told by a health professional that they had angina or coronary heart disease. The weighted percentage is calculated in PLACES ensures that survey responses match the population characteristics of each census tract.

4. CDC's U.S. Small-Area Life Expectancy Estimates Project (USALEEP)

The U.S. Small-Area Life Expectancy Estimates Project (USALEEP), created by the Centers for Disease Control (CDC), the Robert Wood Johnson Foundation, and the National Association for Public Health Statistics and Information Systems, produces estimates of life expectancy at birth by census tract using data from 2010-2015. The USALEEP contains information on health burden used in the CEJST. This data is collected for all U.S. states and the District of Columbia.

Variables used directly in the CEJST, without additional calculations:

- **Low life expectancy:** Average number of years of life a person can expect to live.
    - Unlike most of the other datasets used in the CEJST 1.0, high values of this indicator indicate low burdens. The tool reverses this percentile. The tract with the highest value is at the lowest or 0^{th} percentile.

5. DOE's LEAD Tool

The Department of Energy (DOE)'s Low-Income Energy Affordability Data (LEAD) Tool is a web accessible, interactive platform that allows users to compare energy characteristics in low income households across the nation. It contains data on housing unit counts and average monthly housing electricity, gas, and other fuel expenditures. For the CEJST, 2018 LEAD Tool data are used to calculate energy cost. This data is collected for all U.S. states, the District of Columbia, and Puerto Rico.

Variables used directly in the CEJST, without additional calculations:

- **Energy cost:** Average household annual energy cost divided by the average household income, in U.S. dollars. Based on LEAD Tool data, the average energy cost for low income households is three times higher than for non-low income households. This contributes to high energy insecurity.

6. DOI/BIA's Land Area Representation (LAR)

The Land Area Representation (LAR) dataset at the Bureau of Indian Affairs (BIA), Department of the Interior (DOI) designates land areas for Federally Recognized Tribes. This dataset is new to version 1.0 of the CEJST. The LAR dataset that was provided by BIA also contained a points file that contained the locations of Alaska Native Villages. According to BIA:

> The LAR dataset is the BIA's official geospatial representation of Federal Indian land areas. It is used in its systems of records. The BIA LAR dataset's spatial accuracy and attribute information is continuously being updated, improved, and used as the single authoritative land area data for the BIA's mission. Since the LAR dataset was prepared for public release, it can be used by various organizations, agencies, units of government (i.e. Federal, state, county, and city), and other entities according to the restrictions on appropriate use which can be found in the metadata.

7. DOI's Abandoned Mine Lands Inventory System (eAMLIS)

The Department of the Interior's Abandoned Mine Land Inventory System (e-AMLIS) from 2017 is used. This dataset is new to version 1.0 of the CEJST. The data are available for all U.S. states and the District of Columbia.

Variables calculated by the CEJST before use in the tool:

- **Abandoned mine land**: Presence of an abandoned mine left by legacy coal mining operations. This is calculated by using the latitude and longitude coordinates to locate an abandoned mine within a census tract. If the value is greater than 1, then it is represented as "yes" in the tool.

8. DOT's Transportation Access Disadvantage category

The Department of Transportation (DOT)'s Transportation Access Disadvantage category identifies communities and places that spend more, and take longer, to get where they need to go. The dataset is from 2022. It is comprised of four variables: (1) percent of total population with a drive time to

employment greater than or equal to 30 minutes (Census, 2019); (2) percent of total population with no vehicle(s) available (CDC, 2018); (3) a composite index of economic and built environment characteristics representing the extent to which the location is not supportive to walking (EPA, 2014); and (4) transportation costs percentage income for the regional typical household (Census 2019, HUD 2019). This dataset is new to version 1.0 of the CEJST. The data are available for all U.S. states and the District of Columbia.

Variables used directly in the CEJST, without additional calculations:

- **Transportation barriers**: Average relative cost and time spent on transportation relative to all other tracts.

9. EPA's EJScreen

EJScreen is an environmental justice mapping and screening tool developed by the Environmental Protection Agency (EPA). It utilizes a nationally consistent combination of demographic and environmental indicators to highlight geographic areas that contain vulnerable populations with disproportionately high environmental burdens.[15] The data is used for U.S. states, the District of Columbia, and Puerto Rico.

EJScreen contains metrics relevant to the environmental burdens used in the CEJST. These data come from a myriad of sources and are compiled by the EPA.

Variables used directly in the CEJST, without additional calculations:

- **Traffic proximity and volume:** Count of vehicles at major roads within 500 meters, divided by distance in meters. This is compiled from DOT traffic data from 2017.
- **Diesel particulate matter exposure:** Mixture of particles part of diesel exhaust in the air, and compiled from NATA using 2014 data. Air toxics, or hazardous air pollutants, are known or expected to cause cancer or other serious health effects such as birth defects**.**
- **PM2.5 in the air:** Fine inhalable particles with diameters that are generally 2.5 micrometers and smaller, compiled from the Office of Air and Radiation (OAR) fusion of model and monitor data from 2017 as compiled by EPA's EJScreen, sourced from NATA and DOT traffic data. Common sources of PM2.5 emissions include power plants and industrial facilities.

[15] In the fall of 2022, EPA released EJScreen2.1, which contained additional environmental indicators for the U.S. territories, as well as updated data for lead paint, Superfund proximity, RMP facility proximity, hazard waste proximity, and underground storage tanks. Version 1.0 of the CEJST was able to incorporate underground storage tank data (2021). However, because EJScreen2.1 uses 2020 census boundaries, it was not possible to incorporate in the latest 2022 EJScreen data into the CEJST without compromising the integrity of the data. As noted earlier, the CEJST uses the census tract boundaries from 2010 because many of the data sources in the tool use those boundaries. It is anticipated that a future version of the CEJST will use the 2020 census tract boundaries and then be able to include this data.

- **Wastewater discharge:** Modeled toxic concentrations at stream segments within 600 meters, divided by distance in kilometers. This is compiled from the Risk-Screening Environmental Indicators (RSEI) Model from 2020.
- **Proximity to Risk Management Plan (RMP) facilities:** Facilities that use extremely hazardous substances are required under the Clean Air Act to develop a Risk Management Plan that identifies the potential effects of a chemical accident, identifies steps the facility is taking to prevent an accident, and spells out emergency response procedures should an accident occur. This indicator counts RMP facilities within 5 kilometers (or nearest one beyond 5 kilometers), divided by distance in kilometers, compiled from EPA's RMP database from 2020.
- **Proximity to Superfund (or National Priorities List (NPL)) sites:** Count of proposed or listed Superfund (NPL) sites within 5 kilometers (or nearest outside of 5 kilometers), each divided by distance in kilometers, and compiled from EPA's Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Database from 2020.
- **Proximity to hazardous waste facilities:** Count of hazardous waste facilities (Treatment, Storage, and Disposal facilities, and Large Quantity Generators) within 5 kilometers (or nearest one beyond 5 kilometers), divided by distance in kilometers, compiled from Treatment, Storage, and Disposal Facilities (TSDF) data calculated from EPA's Resource Conservation and Recovery Act (RCRA) Info Database from 2020.
- **Underground storage tanks (USTs) and releases:** Formula of the density of leaking underground storage tanks and the number of all active storage tanks within 1,500 feet of census tract boundaries. These tanks are used to store petroleum or hazardous substances. EPA's UST Finder contains information about proximity of UST facilities to surface and groundwater public drinking water protection areas, private domestic wells, flood and wildfire risk. Compiled from EPA's UST Finder from 2021. This dataset is new to version 1.0 of the CEJST.

10. FEMA's National Risk Index (NRI)

The Federal Emergency Management Agency (FEMA)'s National Risk Index (NRI) is a dataset and interactive online mapping tool that illustrates communities in the United States at risk for 18 natural hazards. The CEJST takes indicators from the 2014-2020 FEMA NRI dataset.

Variables calculated by the CEJST before use in the tool:

FEMA's National Risk Index includes information relevant to the climate change environmental burden indicators used in the CEJST. The NRI for Natural Hazards does produce a single risk index score, but it is not used by the CEJST. Instead, after consultation with FEMA, the CEJST calculates risk in population, building value, and agricultural value due to 14 natural disasters that have been linked to climate change. These are: avalanche, coastal flooding, cold wave, drought, hail, heat wave, hurricane, ice storm, landslide, riverine flooding, strong wind, tornado, wildfire, and winter weather.

- **Population loss rate:** Sum of the expected fatalities and expected injuries (which are counted as one-tenth of a fatality) in a census tract divided by total population in the tract to produce a per

capita rate. The Spatial Hazard Events and Losses Database for the United States and National Centers for Environmental Information (NCEI) report direct and indirect fatalities and injuries due to the 14 disasters the CEJST includes. To compute population loss rate, each injury is counted as one-tenth of a fatality. Both direct and indirect injuries and fatalities are counted as population loss.

- **Building value loss rate**: Sum of the building value at risk divided by total building value to get a wealth-neutral percent of building value at risk.
- **Agricultural value loss rate**: Sum of the agricultural value at risk divided by total agricultural value to get a wealth-neutral percent of agricultural value at risk. Because agricultural value in most census tracts is low and rates of agricultural value loss are low even among tracts most at risk, only tracts with at least $408,000 of agricultural value (about the 10th percentile for rural tracts) are eligible for this burden. The burden is computed as the 90th percentile of tracts with agricultural value.

11. First Street Foundation's Climate Risk Data Access

The First Street Foundation is a non-profit organization that "makes its flood, wildfire, and extreme heat risk statistical data available at the census tract, zip code, county, congressional district, and state levels for non-commercial users in an effort to maximize the use of [its] data for public good." The CEJST uses the Climate Risk Data Access from 2022. The First Street Foundation projected flood risk data is available for all U.S. states, the District of Columbia, and Puerto Rico. The projected wildfire risk data is available for the contiguous U.S. states and the District of Columbia.

Variables used directly in the CEJST, without additional calculations:

- **Projected flood risk**: A high precision, climate-adjusted model that projects flood risk for properties in the future. The dataset calculates how many properties are at risk of floods occurring in the next thirty years from tides, rain, riverine and storm surges, or a 26% risk total over the 30-year time horizon. The risk is defined as an annualized 1% chance. The tool calculates tract-level risk as the share of properties meeting the risk threshold. The risk does not consider property value.
- **Projected wildfire risk**: A 30-meter resolution model projecting the wildfire exposure for any specific location in the contiguous U.S., today and with future climate change. The risk of wildfire is calculated from inputs associated with fire fuels, weather, human influence, and fire movement. The risk does not consider property value.

12. HOLC maps / NCRC historic redlining score

Between 1935 and 1940, the Home Owners Loan Corporation (HOLC), a now defunct Federal agency, created redlining maps for over 200 cities, focusing on cities with populations greater than 40,000. The color-coded maps gave neighborhoods grades based on detailed risk-based characteristics, such as the neighborhood's quality of housing, the recent history of sale and rent values, as well as its demographic composition. These grades affected the ability of homeowners in those neighborhoods to access credit, which resulted in historic underinvestment. Using digitized HOLC maps made available by the Mapping

Inequality Project,[16] researchers affiliated with the National Community Reinvestment Coalition (NCRC) created a dataset assigning each metro area census tract a single historic redlining score.[17] This score represents how much of the tract was redlined in the HOLC maps. The NCRC researchers gave a score of 4 to a formerly redlined neighborhood that was graded D ("hazardous") in a HOLC map. In contrast, neighborhoods graded A ("best"), B ("desirable") and C ("declining") were assigned the scores 1, 2 and 3, respectively. By geographically matching these former neighborhoods with Census tracts, the dataset creates a composite score for each census tract. The NCRC researchers used a cutoff of 3.25 to determine the communities that were subject to the most redlining. This 2021 dataset is new to version 1.0 of the CEJST.

Variables used directly in the CEJST, without additional calculations:

- **Historic Underinvestment**: Census tracts that experienced historic underinvestment based on redlining maps created by the federal government's Home Owners' Loan Corporation (HOLC) between 1935 and 1940. The tool uses the National Community Reinvestment Coalition's methodology for converting boundaries in the HOLC maps to census tracts. Census tracts meet the threshold when they have a score of 3.25 or more out of 4. The historic underinvestment burden is not available for tracts that were not included in the original HOLC maps because there is no underlying data.

13. HUD's Comprehensive Housing Affordability Strategy

Housing and Urban Development (HUD)'s Comprehensive Housing Affordability Strategy (CHAS) serves as the strategic guide for housing and community development activities funded by HUD grants. CHAS data (2014-2018) include information on housing problems and needs of low income households. This data is available for all the U.S. states, the District of Columbia, and Puerto Rico.

Variables calculated by the CEJST before use in the tool:

- **Housing cost:** Percentage of households in a census tract that are earning less than 80% of HUD Area Median Family Income by county and are paying greater than 30% of their income to housing costs. The calculations for this are drawn from Table 8 of the Comprehensive Housing Affordability Strategy. This methodology was developed in collaboration with HUD.
    - *To compute this variable, calculate the following:* (# of Owner Occupied Units Meeting Criteria + # of Renter Occupied Units Meeting Criteria) / (Total # of Owner Occupied Units + Total # of Renter Occupied Units - # of Owner Occupied Units with HUD Area Median Family Income (HAMFI) Not Computed - # of Renter Occupied Units with HAMFI Not Computed).

---

[16] Robert K. Nelson, LaDale Winling, Richard Marciano, Nathan Connolly, et al., "Mapping Inequality," American Panorama, ed. Robert K. Nelson and Edward L. Ayers, accessed November 20, 2022, https://dsl.richmond.edu/panorama/redlining.

[17] Meier, Helen C.S., and Mitchell, Bruce C. Historic Redlining Scores for 2010 and 2020 US Census Tracts. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-10-15. https://doi.org/10.3886/E141121V2

- **Lack of indoor plumbing**: Percentage of homes in a census tract that do not have either an indoor plumbing or kitchen.

14. MRLC's NLCD Percent Developed Imperviousness (CONUS)

The data for the lack of greenspace indicator comes from the Multi-Resolution Land Characteristics (MRLC) consortium, which makes available the National Land Cover Database (NLCD)'s 2019 Percent Developed Imperviousness (CONUS) dataset. The NLCD dataset was developed by the U.S. Geological Survey in partnership with several Federal agencies.

Variables calculated by the CEJST before use in the tool:

- **Lack of greenspace**: Share of land with developed surfaces covered with artificial materials like concrete or pavement, excluding crop land used for agricultural purposes. Places that lack green space are also known as nature-deprived. The Trust for Public Lands and American Forests are two non-profit organizations with data expertise in nature-deprived communities. They provided extensive help in compiling and analyzing the data used in this indicator.

15. USACE's Formerly Used Defense Sites (FUDS)

The U.S. Army Corps of Engineers (USACE) Formerly Used Defense Sites (FUDS) data from 2019 are used. The data are available for all U.S. states and the District of Columbia.

Variables calculated by the CEJST before use in the tool:

- **Formerly Used Defense Sites**: Properties that were owned, leased, or possessed by the United States, under the jurisdiction of the Secretary of Defense prior to October 1986. This is calculated by using the latitude and longitude coordinates to locate the FUDS within a census tract. If the value is greater than 1, then it is represented as "yes" in the tool.

D. Datasets not currently eligible for inclusion

During the public beta period, the development team received and reviewed many suggestions for datasets from federal agencies, environmental justice data experts, WHEJAC, and other stakeholders. Ultimately, nine new datasets were included in version 1.0 of the CEJST. As discussed above, these are: projected flood risk; projected wildfire risk; historic underinvestment due to redlining; lack of green space; lack of plumbing; abandoned mine lands; formerly used defense sites; transportation barriers; and underground storage tanks and releases.

Many of the suggested datasets are not currently suitable for inclusion in a narrowly-targeted resource allocation methodology. The data is not at the census tract level. The development team is working to make additional data available at a census tract level by working with the responsible parties for existing data sources to release data with more granularity. New data-gathering efforts are also underway in the Biden Administration.

Some of the datasets that were considered are listed below. These datasets are not currently eligible for inclusion in the CEJST.

- **Proximity to concentrated animal feeding operations (CAFOs):** The data for these sites in the Environmental Protection Agency (EPA)'s Facility Registry Service (FRS) significantly under-represent the reality of CAFOs across the nation. Many of the addresses in the data could be improved for accuracy. Similarly, the U.S. Department of Agriculture may have relevant data, as demonstrated by their 2012 Census of Agriculture. These data are not available at the census tract level.
- **Coal ash dump sites:** The development team is working to identify a consistent source of information.
- **Safe Drinking Water Information System (SDWIS):** This data is not yet mapped between system names (and their health-based violations) and census tracts.
- **Birth outcomes, such as preterm births or low birth weights:** This data is not currently publicly available at the census tract level.
- **Medicaid claims data:** This data is not currently publicly available at the census tract level.
- **Foreclosure rates:** Some of the highest quality data at the census tract level was last updated in 2009 in Housing and Urban Development (HUD)'s Neighborhood Stabilization Program. The development team is looking for more recent data.
- **Pesticide use data:** This data is not currently publicly available at the census tract level.

E. Challenges in identifying data for the U.S. territories

Not all data included in the CEJST are available for all U.S. territories. This description identifies the available data that was used in the CEJST.

Puerto Rico: The data used for Puerto Rico are from all relevant and available fields in the energy, housing, legacy pollution, transportation, and workforce development categories. The following data are used: low income, projected flood risk, energy cost, lack of indoor plumbing, lead paint, housing cost, proximity to hazardous waste facilities, proximity to Superfund or National Priorities List (NPL) sites, proximity to Risk Management Plan (RMP) facilities, diesel particulate matter exposure, traffic proximity and volume, underground storage tanks and releases, wastewater discharge, low median income, poverty, unemployment, and high school education. Linguistic isolation was removed for Puerto Rico based on feedback received during the beta period.

Guam, the U.S. Virgin Islands, American Samoa, and the Northern Mariana Islands: For these U.S. territories, the tool uses the following data: unemployment, poverty, low median income, and high school education. These burdens are in the workforce development category.

The CEJST uses a slightly different methodology to calculate the relevant percentiles for Guam, the U.S. Virgin Islands, American Samoa and the Northern Mariana Islands because the relevant data are from the 2010 American Community Survey, which is not used for the other regions. To ensure that the

percentiles used for the workforce development category make the appropriate comparisons, the CEJST combines the measures from these territories with their corresponding measures from the 2010 American Community Survey in the 50 states, D.C., and Puerto Rico.[18]

The methodology also makes an adjustment to the methodology for calculating the area median income percentile for Guam, the U.S. Virgin Islands, American Samoa and the Northern Mariana Islands. Area median income is calculated by dividing the median income of the census tract by the median income of the entire territory. The 2010 American Community Survey data are used in both sides of the comparison. Low median income as a percentage of area median income is then converted into a percentile for these territories, and any tracts in the 90th percentile or above are considered disadvantaged. In contrast to the area median income methodology used elsewhere, the percentiles within these territories are not calculated across the entire United States, but instead are calculated separately for each island area. This adjusted methodology for area median income accounts for the different source data available for Guam, the U.S. Virgin Islands, American Samoa and the Northern Mariana Islands, but it still achieves the same result of selecting roughly 10% of communities, with the usual distributional changes.

IV. User Interface

A. Improved Side Panels

Version 1.0 of the CEJST features several new enhancements to the tool's interface. First, the sidebar of the map for each census tract has been redesigned to include additional information. The side panel more clearly explains why a census tract is or is not considered a disadvantaged community. It will state either "Yes," "No" or "Partially" (which is explained in the Section II.C above). If at least one of the thresholds for a category of burden has been met, the entire category is highlighted in dark blue.[19]

The side panel of version 1.0 of the CEJST contains demographic information. Using data from the U.S. Census, information about the race/ethnicity and age breakdowns of each census tract is provided. This data is provided to help inform the public about the tool and is not used in the methodology.

B. Additional Features

Version 1.0 of the CEJST also contains additional features to enhance the user experience. The map now contains a "geolocation" feature that enables users to find their location with one click. The zooming in and out functionality has also been improved.

[18] As noted earlier, data from the U.S. Census's American Community Survey (2015-2019) is otherwise used for the U.S. States, the District of Columbia, and Puerto Rico.
[19] The display on older web browsers may look different.

C. Opportunities to provide feedback and get involved

The tool continues to provide numerous ways for the public to provide feedback on the tool. The sidebar of each census tract on the map has a "Send Feedback" button that allows users to send specific feedback about that tract. This is an important way to "groundtruth" the tool. Ground truthing helps to make sure that the tool better reflects the realities of people living in disadvantaged census tracts.

The tool also contains links to several different surveys. People can provide feedback about the tool, provide data sources and suggestions, or ask questions.

In addition, all of the programming code for building the CEJST is open source and available at https://github.com/usds/justice40-tool. It is mostly written in Python and JavaScript (TypeScript). Documentation for working with the codebase is currently available in both English and Spanish.

Those with interest in the data and analysis are encouraged to dive directly into the code and data. It is available through the open source codebase. Members of the public are also able to build the tool on their own computers and to make proposed updates to the tool. They can suggest "pull requests" to the codebase for consideration by CEQ. See the Github repository for details.

D. Data Downloads

Version 1.0 of the CEJST also contains numerous files available for download. These include spreadsheets that contain the list of disadvantaged communities. This list can be sorted or filtered by state, county, or other data field. There is also a shapefile that includes the data for all of the indicators. This shapefile can be uploaded into alternate mapping platforms such as Esri.

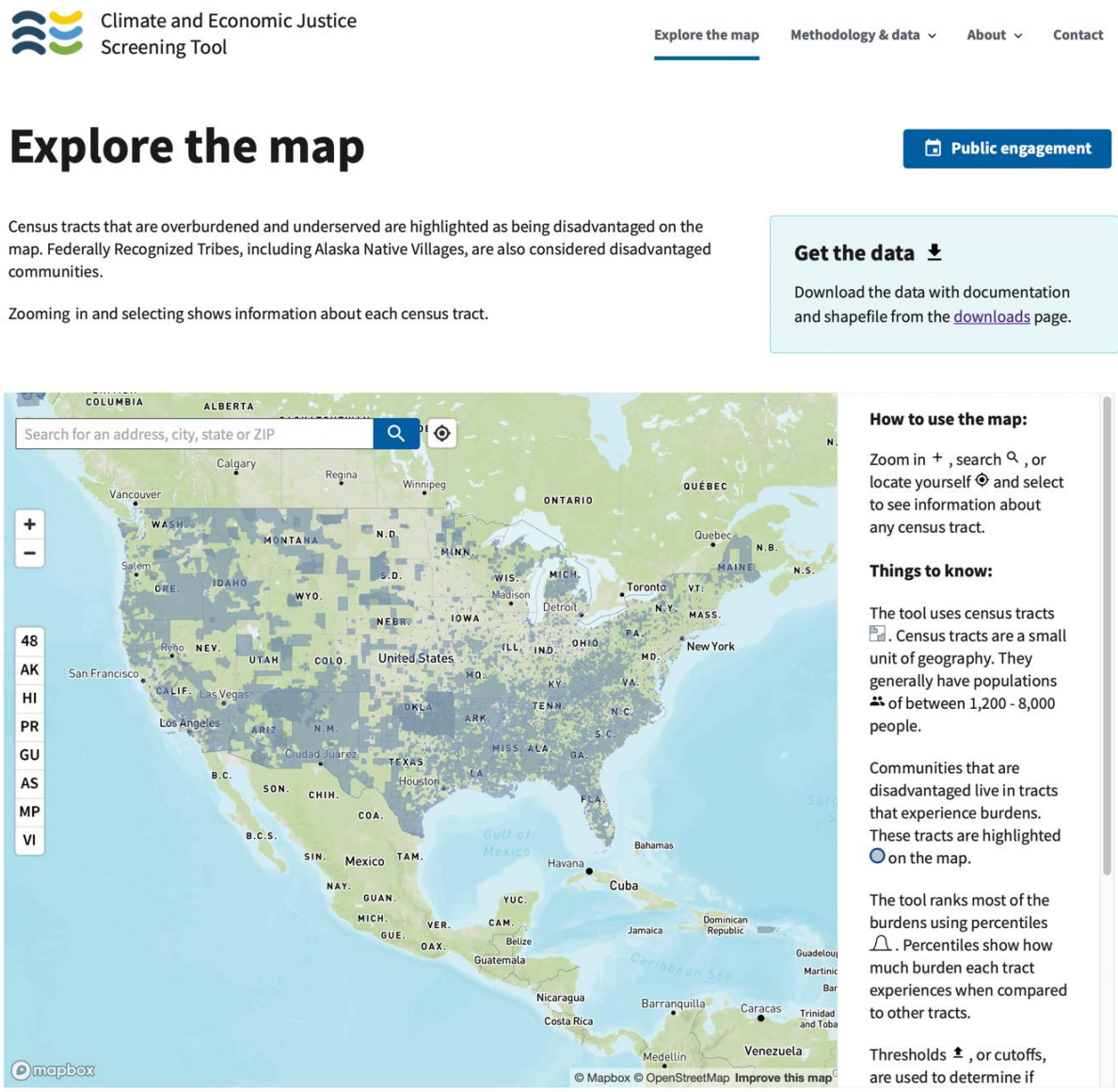


Figure 1: A screenshot of the tool's explore the map page showing the lower 48 U.S. states.

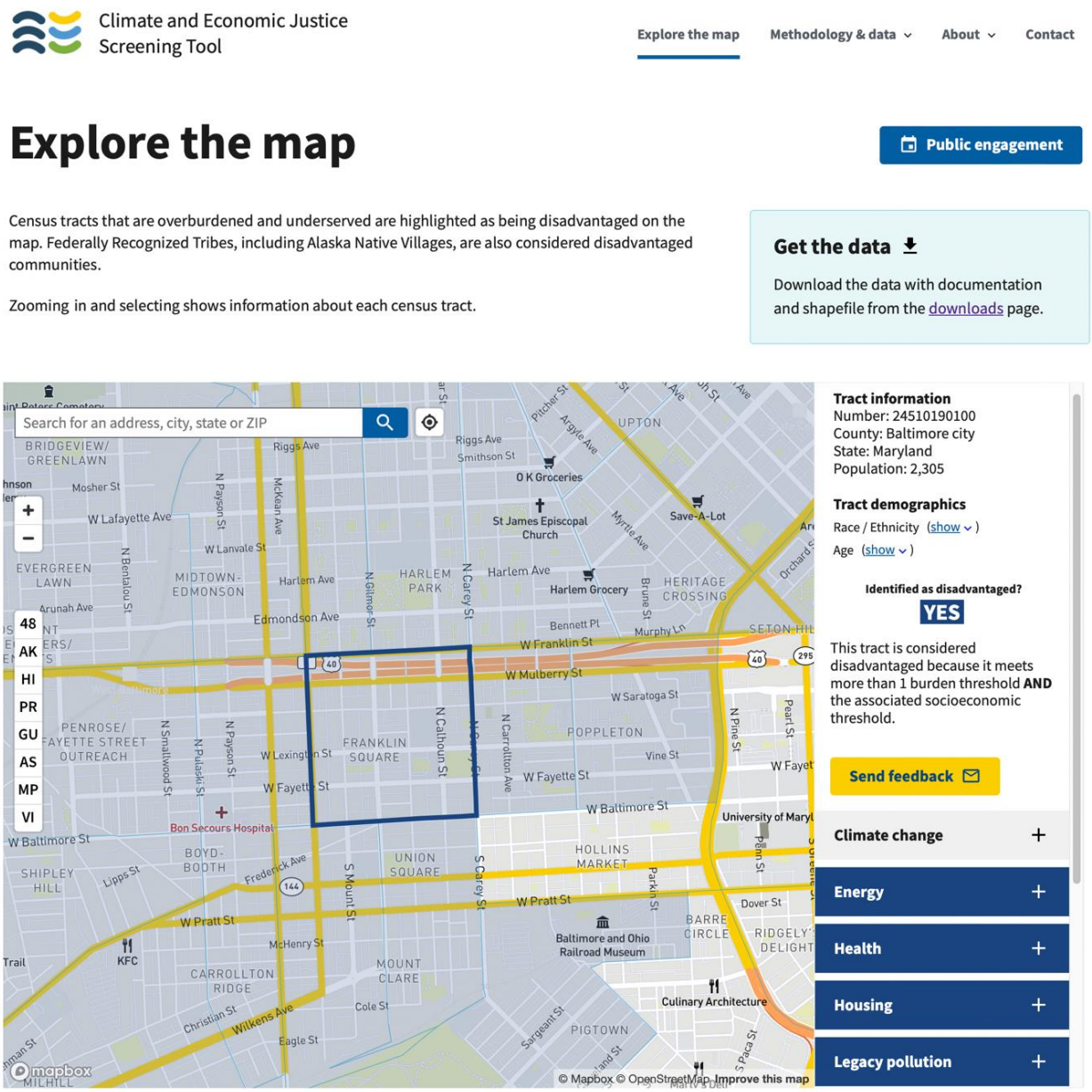


Figure 2: A screenshot of a census tract identified as a disadvantaged community by the CEJST in version 1.0 of the CEJST. The sidebar of the map's user interface is "open," revealing specific data elements about the census tract. On the website, more data become visible as the user scrolls through the sidebar.

V. Looking ahead

The CEJST is intended to be an iterative tool. CEQ will update the CEJST at least annually each year, as directed by Executive Order 14008. The annual update is expected to coincide with the start of each Federal Fiscal Year. Knowing the frequency and timeline for CEJST updates helps Federal agencies know

when the list of disadvantaged communities might be updated. It also helps members of the public and other stakeholders know how they can provide feedback in a timely manner. Version 1.0 of the CEJST improves on the beta version, and future updates will continue to refine the tool.

[end of document]