Climate and Economic Justice Screening Tool
Technical Support Document
Version 2.0

December 2024

White House Council on Environmental Quality

Contents

I.   Introduction

The Climate and Economic Justice Screening Tool (CEJST or tool) is used by federal agencies to identify disadvantaged communities facing particular climate, environmental, and economic burdens. These are communities that are overburdened by pollution and marginalized by underinvestment in infrastructure and other basic services. This tool was first created as a component of the Biden-Harris Administration's historic commitment to advancing environmental justice.[1]

In November 2022, the White House Council on Environmental Quality (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 in February 2022 with support from the U.S. Digital Service (USDS), and in collaboration with other federal agencies and departments. A beta version of the tool was released 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 and USDS 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 federal agencies. 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, the Memorandum on Tribal Consultation and Strengthening Nation-to-Nation Relationships, and Executive Order 13175 on *Consultation and Coordination With Indian Tribal Governments*.

In December 2024, CEQ launched version 2.0 of the CEJST. This version was informed by updated data, recommendations from the WHEJAC, responses to a public survey of CEJST users, feedback from community engagement across the nation, federal agency expertise, and consultation with the White House Environmental Justice Interagency Council (IAC).

Updates to the tool in version 2.0 of the CEJST include:
- Grandfathering disadvantaged census tracts from version 1.0 of the CEJST into version 2.0 of the CEJST to enable federal agencies to consider communities identified as disadvantaged in either version of the tool to be disadvantaged;
- Methodological change to include additional low income areas within the U.S. Territories;
- Ability to search by census tract ID number; and
- Various technical fixes.

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, if they are on the lands of Federally Recognized Tribes, if they were

---

[1] In Executive Order 14008 on *Tackling the Climate Crisis at Home and Abroad*, President Biden directed the White House Council on Environmental Quality to develop a geospatial mapping tool to identify disadvantaged communities that face burdens.

identified as disadvantaged in version 1.0 of the tool, or if they are in certain U.S. Territories and meet the low income threshold only. 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 2.0 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. Additional information on the decision to identify disadvantaged communities at the census tract level can be found in Section III.B.

In version 2.0 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. This is consistent with the methodology used in version 1.0 of the tool. In addition, census tracts that are completely surrounded by disadvantaged communities are also considered disadvantaged if they meet an adjusted low income threshold. The tool now identifies as disadvantaged some additional census tracts that have a water boundary. The portion of the census tract along the body of water is excluded from the calculation of what is considered completely surrounded by disadvantaged census tracts.
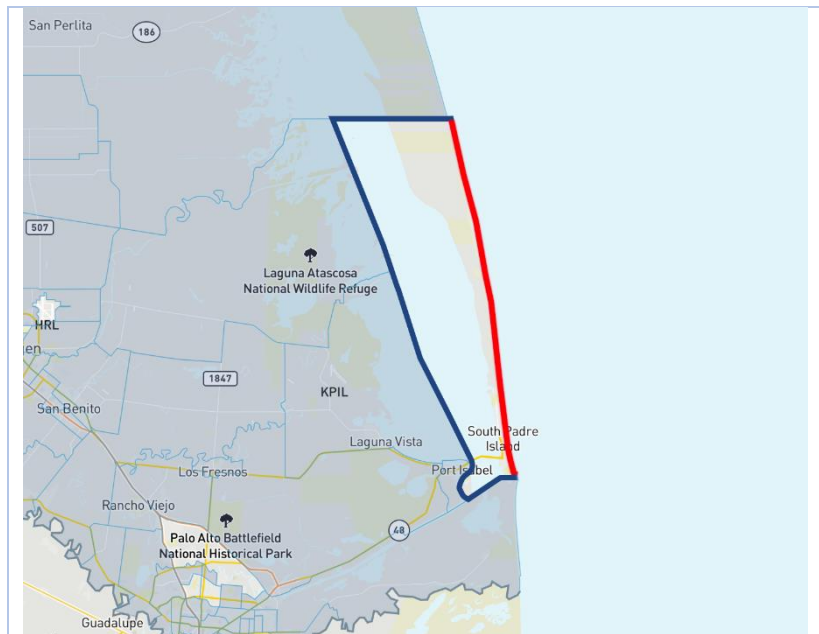
Figure 1. In this example census tract, the part of the boundary shown in red is a water boundary, which is excluded from the calculation of what is considered 100 percent surrounded by disadvantaged communities in version 2.0 of the tool. The census tract is now considered completely surrounded by disadvantaged census tracts.

In addition, the tool identifies communities as disadvantaged in American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands if the census tract

meets the 65th percentile low income threshold. Lastly, census tracts that were identified as disadvantaged in version 1.0, but longer meets the methodology in version 2.0, are "grandfathered" in as a disadvantaged community, as shown in the CEJST spreadsheet.

Table 1 provides an overview of the methodology and identifies the methodological changes made between the different versions of the tool.

Table 1. Overview of methodology used in version 2.0 of the CEJST
(Items marked as **NEW** are changes made in version 2.0)

Communities are considered disadvantaged in version 2.0 of the tool if:
- they are in census tracts that meet the thresholds for at least one of the tool's categories of burden;
- they are on land within the boundaries of Federally Recognized Tribes;
- they are in a census tract ID that was identified as disadvantaged in version 1.0 **(NEW)**; or
- they are in census tracts in certain U.S. Territories that meet the low income threshold only. **(NEW)**.

Census tracts that are completely surrounded by disadvantaged communities are also considered disadvantaged if they meet an adjusted low income threshold (≥ 50th percentile). **Note:** Tracts that are disadvantaged because they are "grandfathered" are not included in the calculation of completely surrounded.

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 OR 5. Projected wildfire risk ≥ 90th percentile	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 OR 2. Housing cost ≥ 90th percentile OR 3. Lack of green space ≥ 90th percentile OR 4. Lack of indoor plumbing ≥ 90th percentile OR 5. Lead paint ≥ 90th percentile	Low income*
Legacy pollution	1. Abandoned mine land present = Yes OR 2. Formerly Used Defense Site (FUDS) present = Yes 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 OR 3. Traffic proximity and volume ≥ 90th percentile	Low income*
Water and wastewater	1. Underground storage tanks and releases ≥ 90th percentile 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	Less than high school education > 10%

* Low Income = 65th percentile or above for percentage of people in households whose income is less than or equal to twice the federal poverty level.[2] The low income data for all 50 states, the District of Columbia, and Puerto Rico is sourced from the American Community Survey (2018-2022). The low income data for American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands is sourced from the Decennial Census Survey (2020).

B. Thresholds Approach for Burdens

1. Categories of Burdens

This section provides additional detail on the methodology for census tracts that are identified as disadvantaged because they meet the thresholds for at least one of the tool's categories of burden. 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 meets the burden thresholds, and is at or above the 50th percentile for low income, is also considered disadvantaged. **Note:** Tracts that are disadvantaged because they are "grandfathered" are not included in the calculation of completely surrounded.

The indicators of burdens used in the CEJST are informed by Executive Order 14008, which 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."[3] To achieve this goal, the CEJST burdens are grouped into categories that were informed by core federal environmental and economic justice investment focus areas.[4] The eight categories of burdens in the CEJST are: climate change, energy, health, housing, legacy pollution, transportation, water and wastewater, and workforce development.

---

[2] This does not include people in college dormitories, institutional group quarters (such as prisons and nursing homes), military barracks, and those in living situations without conventional housing (other than shelters).

[3] Section 219, Executive Order 14008, *Tackling the Climate Crisis at Home and Abroad* (January 27, 2001).

[4] OMB Memorandum M-21-28, *Interim Implementation Guidance for the Justice40 Initiative* (July 20, 2021).

Socioeconomic Burdens

In versions 1.0 and 2.0 of the CEJST, the low income burden for the continental United States and Puerto Rico is calculated by excluding college students. In version 2.0 of the tool, improvements were made to this calculation using the U.S. Census Bureau's 2015-2019 5-year American Community Survey (ACS) table (B14006) to more accurately characterize college students.[5] Thus, the low income burden is the percentage of a census tract's population in households where household income is at or below 200 percent of the federal poverty level, not including college students. Because table B14006 has data for 100 percent of the federal poverty level, the 200 percent level was interpolated, or estimated. These data are made available in the tool's downloadable files.

The workforce development category in versions 1.0 and 2.0 of the CEJST has one socioeconomic indicator, which is that fewer than 10 percent 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.

Categories of Burden

i. Climate Change

The burdens in the climate change category aim to measure expected agriculture value, building value, and population loss due to climate-related natural hazards, and projected wildfire and flood risk. 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. Data comes from the Federal Emergency Management Agency's (FEMA) National Risk Index, the First Street Foundation's Climate Risk dataset, and the U.S. Census Bureau's American Community Survey.

ii. Energy

The burdens in the energy category aim to measure the energy cost as well as energy-related pollution, including ambient concentrations of $PM_{2.5}$ in the air, 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. Data comes from the U.S. Department of Energy's (DOE) LEAD Score, the U.S. Environmental Protection Agency's (EPA) EJScreen, and the U.S. Census Bureau's American Community Survey.

iii. Health

The burdens in the health category 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

---

[5] The ACS dataset from the U.S. Census indicates people enrolled in undergraduate, graduate, or professional programs, people living in military barracks or institutional group quarters (such as prisons and nursing homes), people in living situations without conventional housing (other than shelters), or unrelated individuals under age 15 (such as foster children). This applies only to the contiguous U.S. and Puerto Rico, which are included in the ACS dataset.

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. Data come from the U.S. Centers for Disease Control and Prevention's (CDC) PLACES, the CDC's USALEEP, and the U.S. Census Bureau's American Community Survey.

iv.   Housing

The burdens in the housing category 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 structures without indoor plumbing or kitchens within a census tract. 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.[6] In addition, census tracts must meet the threshold for the low income burden, which is at or above the 65th percentile. Data come from the U.S. Department of Housing and Urban Development's (HUD) Comprehensive Housing Affordability Strategy, National Community Reinvestment Coalition (using data from maps created by the Home Owners' Loan Corporation between 1935 and 1940),[7] the Trust for Public Land and American Forestry's percent developed imperviousness (CONUS), and the U.S. Census Bureau's American Community Survey.

v.   Legacy Pollution

The burdens in the legacy pollution category aim to measure how much legacy, current, and potential pollution a census tract has through abandoned mines, Formerly Used Defense Sites (FUDS), proximity to hazardous waste facilitates, Risk Management Plan (RMP) facilities, and Superfund sites (otherwise known as National Priorities List (NPL). 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 or FUDS. In addition, census tracts must meet the threshold for the low income burden, which is at or above the 65th percentile. Data come from the Bureau of Land Management's (BLM) Abandoned Mine Land Inventory at the U.S. Department of the Interior (DOI), U.S. Army Corps of Engineers's (USACE) Formerly Used Defense Sites Inventory, EPA's EJScreen, and the U.S. Census Bureau's American Community Survey.

vi.   Transportation

The burdens in the transportation category aim to measure the transportation-related pollution, transportation barriers,[8] and traffic-related noise and proximity within a census tract. 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

---

[6] The tool uses the National Community Reinvestment Coalition's methodology for converting boundaries in the Home Owners' Loan Corporation's (HOLC) maps to census tracts. Previously, the historic underinvestment burden was not shown for tracts that were not included in the original HOLC maps because there is no underlying data. Now in version 2.0 of the tool, tracts where historic underinvestment did not occur will show the burden as not available.

[7] 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.

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

is at or above the 65th percentile. Data comes from the EPA's EJScreen, the U.S. Department of Transportation's (DOT) mapping tool of Transportation Disadvantaged Communities, and the U.S. Census Bureau's American Community Survey.

vii.   Water and Wastewater

The burdens in the water and wastewater category aim to measure the census tract's proximity to toxicity-weighted wastewater discharges and underground storage tanks that may leak. 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. Data come from the EPA's EJScreen, the EPA's Underground Storage Tank Finder via EJScreen, and the U.S. Census Bureau's American Community Survey.

viii.   Workforce Development

The burdens in the workforce development category aim to identify census tracts that would benefit from greater workforce development. This is in line with the 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 the percentage of a census tract's population in households where the household income is at or below 100 percent of the federal poverty level. To be considered disadvantaged in the workforce development 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 of the socioeconomic indicator, which is that fewer than 10 percent of people ages 25 or older have a high school education (i.e., graduated with a high school diploma).[9] Data come from the U.S. Census Bureau's American Community Survey (2015-2019) for all U.S. states, the District of Columbia, and Puerto Rico. Data come from the Decennial Census (2020) for American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands.

2. Communities Surrounded by Other Disadvantaged Tracts

Census tracts that are completely (i.e., 100 percent) surrounded by disadvantaged communities are identified as disadvantaged if they meet an adjusted low income threshold (at or above the 50th percentile). For tracts that have a water boundary (i.e., islands or coastal tracts), the water boundary is not included in the calculation of whether the tract is 100 percent surrounded by disadvantaged tracts. This is a change made in version 2.0 of the tool. The methodology for completely surrounded tracts 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

---

[9] As noted above, the higher education non-enrollment indicator used in the beta version of the CEJST is not used in 1.0 or 2.0 versions; 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. It was removed based on expert feedback.

surrounded by other disadvantaged tracts, then it is statistically likely that this tract was excluded as a result of data aberrations.

The following examples explain how the methodology works for tracts that are completely surrounded by disadvantaged communities.[10] Assume that these census tracts do not contain the lands of Federally Recognized Tribes.

- Assume there is a census tract that is completely surrounded by census tracts that are considered disadvantaged. This census 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 the census tract is surrounded by disadvantaged census tracts, and it is above the 50th percentile for low income, the census tract is identified as disadvantaged in version 2.0 of the CEJST.
- Assume there is a census 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. However, the census tract is surrounded by disadvantaged census 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 2.0 of the CEJST.
- Assume there is a census tract that is only surrounded by disadvantaged communities on 75 percent of its borders. This census tract does not otherwise meet any of the thresholds of the burdens in the version 2.0 methodology discussed above. This census tract is not identified as disadvantaged because it is not 100 percent surrounded by other disadvantaged census tracts.

3. Percentiles for Normalization

The tool ranks most of the burdens using percentiles. Percentiles show how much burden each census tract experiences when compared to other census 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 census tract has a higher measurement on this variable than 77 percent of all census tracts. Percentiles help to normalize data measured in different units and on different scales and enables the tool to use data for all regions and types of areas in the United States 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 version 2.0 of the tool have different scales. For instance, $PM_{2.5}$ is measured in micrograms of particulate matter per cubic meter of air. It generally ranges from about 2 or 3 to 300 micrograms per cubic meter. 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 on that indicator (e.g., a community is exposed to pollution). In order to achieve

---

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

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 of 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 is a bimodal distribution (e.g., a lot of "very safe air" neighborhoods 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. 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 of 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 United States. 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 percent 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 percent of the federal poverty line) is at or above 65th percentile. As noted above, version 2.0 of the CEJST uses a 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 is identified as disadvantaged if, in addition to meeting other criteria in the category, 10 percent or more of the percent of adults 25 or older living in the census tract did not graduate high school. Approximately 50 percent of census tracts in the CEJST have populations where 10 percent 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. For example, if a census

tract contains at least one Formerly Used Defense Site (FUDS) or an abandoned mine, 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 2.0 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 California's CalEnviroScreen and Maryland's EJ Screening Tool. In addition, EPA's EJScreen and the U.S. 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 tract that meets an environmental, climate, or other burden only:** Assume a census tract is in the 95th percentile for proximity to Superfund (or NPL) sites, but at the 40th percentile for low income. Even though the census tract 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 means the census tract would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tract that meets 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. However, the census tract 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 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 census 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. The census tract would not be considered a disadvantaged community, and would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tract that meets both the environmental, climate, or other burden and the related socioeconomic burden:** Assume a census tract is in the 92nd percentile for proximity to Superfund (or NPL) sites and the 87th percentile for low income. The census tract is disadvantaged because it exceeds the threshold 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.

Below are three examples for the thresholds used in the workforce development category offer additional insight:

- **Census tract that meets an environmental, climate, or other burden 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 percent of people have a high school diploma. This census tract would not be considered disadvantaged and would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tract that meets 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. 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 tract that meets both an environmental, climate, or other burden and the socioeconomic burden:** Assume a census tract is in the 92nd percentile for unemployment rate and that 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. Federally Recognized Tribes

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

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, the 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 census tract are identified as disadvantaged. The side panel that appears for each census tract on the tool's map states the percentage of the census tract that is comprised of the lands of Tribes. The tool also explains why the census tract is identified as partially disadvantaged.

D. Addressing Missing Data

For tracts that are missing income information, the low income indicator is estimated in the tool for the continental United States and Puerto Rico. 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 2.0 of the CEJST. Income is not estimated for tracts with unknown populations.

E. Bug Fixes

Version 2.0 of the CEJST also fixes two bugs that were present in version 1.0 of the tool. In version 2.0 of the tool, water boundaries are now excluded from the calculations to identify census tracts that are completely surrounded by disadvantaged census tracts. Additionally, the historic underinvestment indicator now shows in the side panel whether or not the tract is in an area that was formerly redlined.

F. Grandfathering

If a census tract no longer meets the criteria for the methodology identifying a disadvantaged community due to the update to the CEJST, that census tract will still be "grandfathered" in as a disadvantaged community for at least the next two years. This helps to reduce disruption and to support a smooth transition for agencies, applicants, and the public as new versions of the tool are released. The tool's instructions have more information about time-limited grandfathering. The CEJST spreadsheet will display all current and "grandfathered" disadvantaged communities.

III. Data

A. Data Overview

Version 2.0 of the CEJST includes data from a wide variety of sources, including the U.S. Census Bureau, the U.S. Centers for Disease Control and Prevention (CDC), the U.S. 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 National Community Reinvestment Coalition (using data from the Home Owners' 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 recommendations from federal agencies, environmental justice data experts, the White House Environmental Justice Advisory Council (WHEJAC), the National Academies of Sciences, Engineering, and Medicine (NASEM), and the public. Each dataset included meets the following relevance, availability, and quality requirements:

1. **Relevance to the goals** of 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, the District of Columbia, and the five U.S. Territories of Puerto Rico, American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands. Datasets that include all 50 states and the District of Columbia, 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 the District of Columbia. 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 Home Owners' Loan Corporation (HOLC) maps because there is no underlying data.
4. **Data are 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 are only available at site- or point-level, the tool assigns latitude/longitude coordinates to individual census tracts.

The datasets used in version 2.0 of the CEJST are summarized in the following table. 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 following table.

Source (date)	Dataset	Variables included or calculated	Relevant categories	Used for

Census (2018-2022)*	American Community Survey (ACS)	• High school education (CEJST calculates)*^	All categories	* All U.S. states, the District of Columbia, and Puerto Rico[11]
Census (2020)^	Decennial Census Survey	• 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)		^ American Samoa, Guam, the Commonwealth of 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 (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 and the District of Columbia
DOI/BIA (2018)	Land Area Representation	• Lands of Federally Recognized Tribes, including the locations of Alaska Native Villages and landless Tribes (NEW)	N/A	Federally Recognized Tribes, including Alaska Native villages and landless Tribes
DOI (2017)	Abandoned Mine Lands Inventory System	• Abandoned mine land	Legacy pollution	All U.S. states and the District of Columbia
DOT (2022)	Transportation Access Disadvantage	• Transportation barriers	Transportation	All U.S. states and the District of Columbia

---

[11] Linguistic isolation is not included for Puerto Rico.

EPA (varies, 2014-2021)	EJScreen	• Diesel particulate matter exposure[*] • PM$_{2.5}$ 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[*] • Wastewater discharge[*]	Climate change, 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 • Projected wildfire risk	Climate change	All contiguous U.S. states and the District of Columbia
HOLC (1935-1940) / NCRC (2020)	Historic Redlining Scores[12]	• Tracts determined to historically have been subject to redlining by the Home Owners' Loan Corporation between 1935-1940	Housing	Metro areas of the United States that were graded by HOLC
HUD (2014-2018)	Comprehensive Housing Affordability Strategy (CHAS)	• Housing cost (CEJST calculates) • Lack of indoor plumbing	Housing	All U.S. states, the District of Columbia, and Puerto Rico
MRLC consortium (2019)[13]	Percent Developed Imperviousness (CONUS)	• Lack of greenspace (CEJST calculates)	Housing	All contiguous U.S. states and the District of Columbia

---

[12] 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.

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

USACE (2019)	Formerly Used Defense Sites	• Formerly Used Defense Sites	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. However, it was decided that adequate data could not be reliably reported at this level at this time. The U.S. Census Bureau's American Community Survey has data available down to the census block group level. Block groups are usually between 600 and 3,000 people. These estimates at the census 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 census block group level may not be. For example, in order to report data at the census block group level, the 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), $PM_{2.5}$, and ozone estimates. For these datasets, EJScreen assigns each census block group the NATA, $PM_{2.5}$, 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 also 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.[14] 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.

Version 2.0 of the CEJST uses census tract boundaries for statistical areas from 2010. The CEJST is an iterative tool, and it is anticipated that future versions will be able to use the most current census tract boundaries and available datasets.

---

[14] 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 U.S. Census Bureau's hierarchy of geographies: https://mcdc.missouri.edu/applications/geocorr2018.html.

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, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands. Many nationally-consistent datasets are not currently available for the U.S. Territories, which is further discussed in section III.E below.

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.[15] The discussion of each dataset below describes when each approach was used.

1. U.S. Census Bureau's American Community Survey

The U.S. Census Bureau's American Community Survey 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 (2018-2022) includes data for all 50 states, the District of Columbia, and Puerto Rico for all of these variables.

Variables 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 percent 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 U.S. 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 2.0 of the CEJST 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.

Variables calculated by the CEJST before use in the tool:
- **Low income:** The first step in the development of the low income indicator is to calculate the percent of the population within a given census tract living at or below 200 percent of the federal poverty line. Next, the number of adults ages 15 and higher enrolled in higher education full time is extracted from the U.S. Census Bureau data. 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 percent of the federal poverty line.

---

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

Because this variable is critical to the tool, if data are 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.

- **High school education:** The high school education indicator 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 homes.
- **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 are 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 is not included as an indicator to identify disadvantaged communities in Puerto Rico in version 2.0 of the CEJST.

2. Decennial Census Survey

The Decennial Census Survey (2020) includes data for American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands. This dataset was added in version 2.0.

Variables used directly in the CEJST, without additional calculations:
- **Low income**: The low income variable is used in American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands without any adjustments.

3. U.S. Census Bureau's TIGER Files

Census tract boundaries are provided by the U.S. Census Bureau'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 are available for public download and consumption at https://data.census.gov/.

4. CDC's PLACES

PLACES is a public, interactive website launched by the U.S. Centers for Disease Control and Prevention (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. These data are 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.

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

The U.S. Small-Area Life Expectancy Estimates Project (USALEEP) – created by the U.S. Centers for Disease Control and Prevention (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 burdens used in the CEJST. These data are 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 version 2.0 of the CEJST, 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.

6. DOE's LEAD Tool

The U.S. Department of Energy's (DOE) 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. These data are 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.

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

The Land Area Representation (LAR) dataset at the U.S. Department of the Interior's (DOI) Bureau of Indian Affairs (BIA) designates land areas for Federally Recognized Tribes. 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.

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

The U.S. Department of the Interior's (DOI) Abandoned Mine Land Inventory System (e-AMLIS) from 2017 is 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:
- **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.

9. DOT's Transportation Access Disadvantage category

The U.S. Department of Transportation's (DOT) 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). 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.

10. EPA's EJScreen

EJScreen is an environmental justice mapping and screening tool developed by the U.S. 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. The data are 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 that are part of diesel exhaust in the air, as compiled by 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.
- **PM$_{2.5}$ 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 PM$_{2.5}$ emissions include power plants and industrial facilities.
- **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 the nearest facility 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 (or 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 the nearest facility 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, and flood and wildfire risk. Compiled from EPA's UST Finder from 2021.

11. FEMA's National Risk Index (NRI)

The Federal Emergency Management Agency's (FEMA) 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 2014-2021 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 and environmental burden indicators used in the CEJST. While the NRI for Natural Hazards produces a single risk index score, 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.

12. 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 are available for all U.S. states, the District of Columbia, and Puerto Rico. The projected wildfire risk data are 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 structures are at risk of floods occurring in the next thirty years from tides, rain, riverine and storm surges, or a 26 percent risk total over the 30-year time horizon. The risk is defined as an annualized 1 percent 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.

13. 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.

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 HOLC between 1935 and 1940. The tool uses the NCRC'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 census tracts that were not included in the original HOLC maps because there is no underlying data.

14. HUD's Comprehensive Housing Affordability Strategy
The U.S. Department of Housing and Urban Development's (HUD) 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. These data are 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 percent of the HUD Area Median Family Income by county and are paying greater than 30 percent 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).
- **Lack of indoor plumbing:** Percentage of homes in a census tract that do not have either indoor plumbing or a kitchen.

15. 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's (NLCD) 2019 Percent Developed Imperviousness (CONUS) dataset. The NLCD dataset was developed by the U.S. Geological Survey in partnership with several federal agencies.

---

[16] Nelson, Robert K., Winling, LaDale, Marciano, Richard, Connolly, Nathan, 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.

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 Land 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.

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

The U.S. Army Corps of Engineers's (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

The development team received and reviewed many suggestions for datasets from federal agencies, environmental justice data experts, the WHEJAC, and other stakeholders.

Many of the suggested datasets are not currently suitable for inclusion in a narrowly-targeted resource allocation methodology. Some suggested data are 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-Harris Administration. Some of the datasets that were considered are listed below. To our best knowledge, 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 U.S. Environmental Protection Agency's (EPA) Facility Registry Service (FRS) significantly underrepresent 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 (USDA) may have relevant data, as demonstrated by its 2012 Census of Agriculture. These data are not available at the census tract level.
- **Birth outcomes, such as preterm births or low birth weights:** These data are not currently publicly available at the census tract level.
- **Medicaid claims data:** These data are 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 the U.S. Department of Housing and Urban Development's (HUD) Neighborhood Stabilization Program.
- **Pesticide use data:** These data are 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 section 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. The linguistic isolation indicator was removed for Puerto Rico based on feedback received during the beta period.

American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands: For these U.S. Territories, the tool uses the following data:
- **Standalone threshold:** low income. (**NEW** for American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands)
- **Workforce development category:** unemployment, poverty, low median income, and high school education.

The CEJST uses a slightly different methodology to calculate the relevant percentiles for American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands and. To ensure that the percentiles used for the workforce development category make the appropriate comparisons, the CEJST combines the measures from these U.S. Territories with their corresponding measures from the 2010 American Community Survey in the 50 states, the District of Columbia, and Puerto Rico.[18]

The tool also makes an adjustment to the methodology for calculating the area median income percentile for American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands. Area median income is calculated by dividing the median income of the census tract by the median income of the entire territory. The 2020 Decennial Census 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 American Samoa, Guam, Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands, but it still achieves the same result of selecting roughly 10 percent of communities, with the usual distributional changes.

Due to lack of data for the U.S. Territories among the climate and environmental burden categories, the tool's methodology has been updated so that tracts in American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, or the U.S. Virgin Islands are considered disadvantaged if they meet the tool's methodology (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

---

[18] As noted earlier, data from the U.S. Census Bureau's American Community Survey (2018-2022) is otherwise used for the U.S. states, the District of Columbia, and Puerto Rico.

within the boundaries of Federally Recognized Tribes) or if the tract meets the 65th low income threshold only. As noted above, the Decennial Census is the source of the low income dataset.

IV. User Interface

A. Enhanced Side Panels

Version 2.0 of the CEJST features several new enhancements to the tool's interface. The sidebar of the map for each census tract includes additional information regarding formerly disadvantaged areas.

The side panel now explains if a census tract or portions of a tract are formerly disadvantaged. It will state either "Yes," "No," or "Partially" (which is explained in the Section II.F above).

B. Opportunities to Provide Feedback and Get Involved

The tool continues to provide numerous ways for the public to provide feedback on the tool.

- Users can find census tracts that they are familiar with and send feedback about those specific tracts. To access this survey, users can click the yellow "Send feedback" button in the side panel of any tract. That feedback is used to inform future versions of the tool. CEQ is especially interested in any information that may help to better reflect the realities of climate and economic burdens for local communities or census tracts.
- Users can also submit data sources or ideas for consideration by clicking the yellow button that reads "Share data sources with CEQ" near the top right of the page.
- In addition, users can take a short survey to help improve the experience of using the tool by clicking the yellow button that reads "Help improve the tool" near the bottom right of the page.

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.

C. Data Downloads

Version 2.0 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.
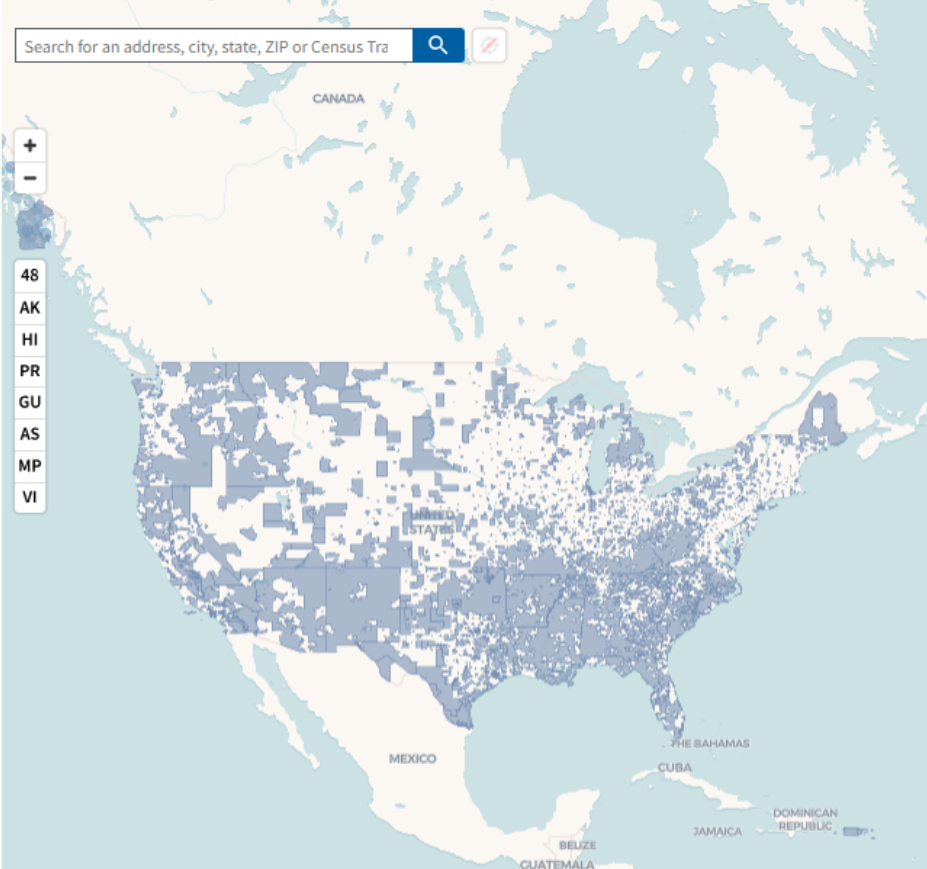
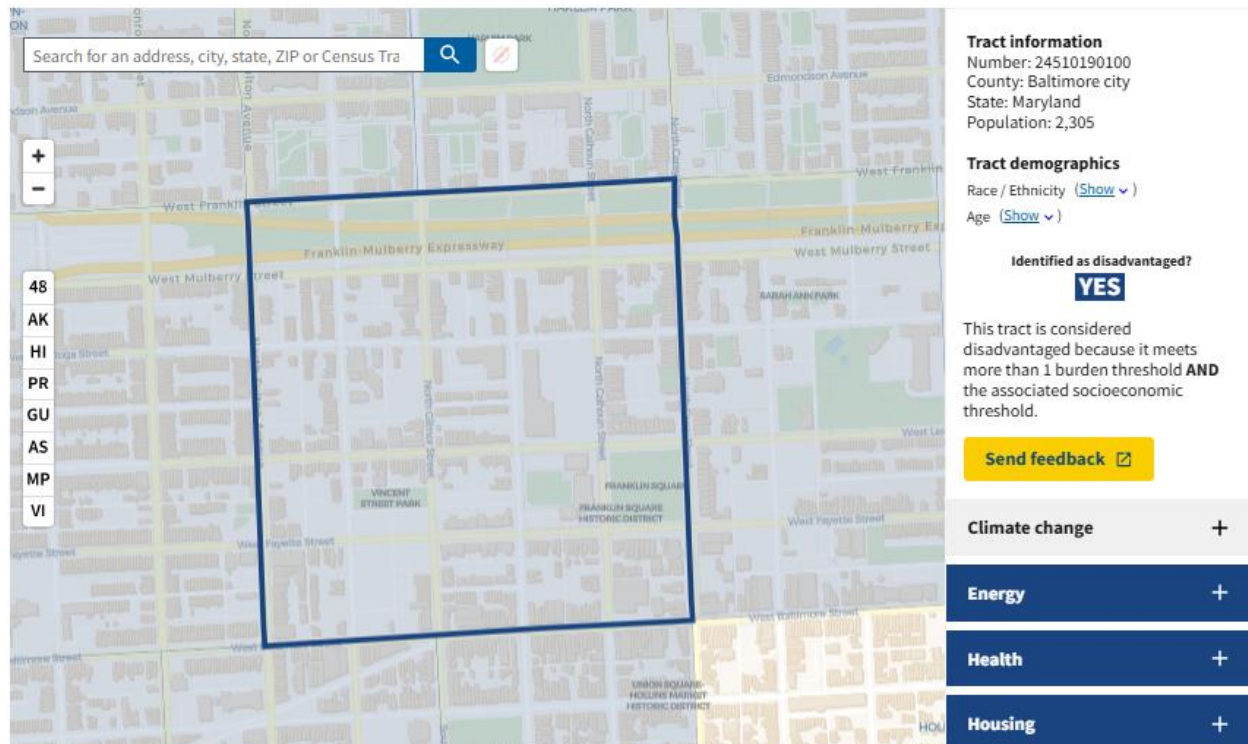Figure 2. A screenshot of the tool's explore the map page showing the lower 48 U.S. states.

Figure 3. A screenshot of a census tract identified as a disadvantaged community by the CEJST in version 2.0 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, and version 2.0 of the CEJST aims to further improve on version 1.0 of the tool in its identification of disadvantaged communities. CEQ plans to continue to update the CEJST periodically to continue to reflect the most current data on climate, environmental, and economic burdens affecting disadvantaged communities.