Climate and Economic Justice Screening Tool

Technical Support Document

Public Beta, Version 0.1

April 2022

White House Council on Environmental Quality

U.S. Digital Service

Contents

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 addressing inequity and inequality. 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 are marginalized, underserved, and overburdened by pollution. The tool will provide important information for the [Justice40 Initiative](#), which, for the first time in history, aims to provide 40 percent of the overall benefits of certain Federal investments in seven key areas to disadvantaged communities. These areas are: climate change, clean energy and energy efficiency, clean transit, affordable and sustainable housing, training and workforce development, reduction and remediation of legacy pollution, and the development of critical clean water and wastewater infrastructure.

This task of delivering the benefits of hundreds of Federal programs to communities that are marginalized, underserved, and overburdened by pollution requires fundamental and sweeping reforms to the ways in which the Federal government as a whole operates. For far too long, communities across the country have faced environmental injustices, such as bearing the brunt of toxic pollution, enduring underinvestment in infrastructure and critical services, and suffering disproportionate impacts from climate change. Since his first day in office, President Biden has committed to using every lever at his disposal to address these inequities, advance environmental justice, and spur economic opportunity, including through the Justice40 Initiative. To meet the Justice40 goal, the Administration is transforming hundreds of Federal programs across the whole government to ensure that disadvantaged communities receive the benefits of new and existing investments. The CEJST was developed with these Federal resource allocation purposes in mind. By helping agencies to identify disadvantaged communities, the CEJST seeks to fulfill the promise of the Justice40 Initiative.

CEQ launched a beta—or draft—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 the public and from Tribal Nations. The beta version of the CEJST website is available at https://screeningtool.geoplatform.gov. It features a user-friendly, searchable map[1] that identifies disadvantaged communities across all 50 states, the District of Columbia (D.C.), and the U.S. territories, to the extent that the data for the U.S. territories is available. The CEJST methodology uses environmental and climate indicators, together with socioeconomic indicators, to identify disadvantaged communities at the U.S. census tract level, which are geographic areas that contain approximately 4,000 people. The CEJST website also enables users to download a draft list of disadvantaged communities as a spreadsheet that includes the relevant data for all of the indicators.

This technical support document describes the data and methodology used in the beta version of the CEJST and is organized as follows. Section II provides an overview of the methodology and data used in

---

[1] U.S. census tracts that are identified as disadvantaged in the beta version of the CEJST are highlighted on the map in a transparent blue-gray shade that enables the user to also view street names, water bodies, and other relevant geographic details within each tract. The color blue was selected to aid readers with visual impairment; however, due to the translucency of the color, the shade may appear closer to gray on some browsers. Section II.C contains screenshots from the beta version of the CEJST.

the beta version of CEJST. In short, the tool utilizes publicly-available, nationally-consistent datasets to identify disadvantaged communities for prioritization in Federal resource allocation decisions. The methodology is based on a thresholds approach where a community qualifies as disadvantaged (1) if the U.S. census tract is above the threshold for one or more environmental or climate indicators and (2) the U.S. census tract is above the threshold for the socioeconomic indicators. The beta version of the CEJST includes data from a wide variety of sources, including the U.S. Census Bureau, the U.S. Environmental Protection Agency (EPA), the Centers for Disease Control and Prevention (CDC), the U.S. Department of Transportation (DOT), the U.S. Department of Energy (DOE), the Federal Emergency Management Agency (FEMA), and the U.S. Department of Housing and Urban Development (HUD). Section II also includes screenshots from the beta version of the CEJST.

Section III provides additional details on the beta version of the CEJST. Section III.A more fully describes the methodology, with three sub-sections outlining the use of (1) indicators; (2) thresholds and cut-off values; and (3) percentiles for normalization. Section III.B presents more information on the data used, with four sub-sections explaining (1) the selection of U.S. census tracts as the geographical unit; (2) the datasets used; (3) the datasets not currently eligible for inclusion; and (4) the challenges in identifying data for the U.S. territories.

Section IV concludes by describing ways to get involved, such as by providing feedback on the beta version of the CEJST or joining the open source community. In addition, leaders of federally recognized Tribes have been invited to engage in the Tribal consultation process. The beta version of the CEJST will be updated based on feedback received during the public beta period. Thereafter, the CEJST will be continuously updated and refined based on feedback and as new datasets and research become available.

II.   Overview of methodology and data
A. Methodology overview

As required by Executive Order 14008, the CEJST seeks to identify disadvantaged communities that are "marginalized and overburdened by pollution and underinvestment in housing, transportation, water and wastewater infrastructure, and health care."[2] To achieve this goal, the beta version of the CEJST uses a thresholds approach with categories that were informed by the areas of investment covered by the Justice40 Initiative. The eight categories in the beta version of the CEJST are: climate change, clean energy and energy efficiency, clean transit, affordable and sustainable housing, reduction and remediation of legacy pollution, critical clean water and wastewater infrastructure, health burdens, and training and workforce development.[3] Table 1, which is located at the end of this section, shows the indicators for each of the eight categories.

---

[2] Section 219, Executive Order 14008, Tackling the Climate Crisis at Home and Abroad (January 27, 2001).
[3] As set forth in OMB Memorandum M-21-28 (July 20, 2021) on "Interim Implementation Guidance for the Justice40 Initiative," the covered areas of investment are: climate change, clean energy and energy efficiency, clean transportation, affordable and sustainable housing, training and workforce development (related to climate,

In the beta version of the CEJST, a community qualifies as disadvantaged if (1) the U.S. census tract is above the threshold for one or more environmental or climate indicators,[4] and (2) the U.S. census tract is above the threshold for the socioeconomic indicators. For example, each of the tool's thresholds for the reduction and remediation of legacy pollution category uses (1) an environmental indicator related to proximity to pollution sites, and (2) socioeconomic indicators related to income and current enrollment in higher education. A U.S. census tract must exceed the threshold for both indicators to be considered disadvantaged. Below, each set of environmental or climate indicators and socioeconomic indicators is referred to as a single combined threshold. Additional information on the decision to identify disadvantaged communities at the U.S. census tract level can be found in the section "Census tract as the geographic unit," Section III.B.1.

Most indicators that feed into each combined threshold are percentiles. 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 reconcile data measured in different units and on different scales, so that the tool can use data relevant to 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 and make it possible to identify the relative burden that each census tract experiences. More details about the use of percentiles for normalization can be found in Section III.A.3.

Each indicator has a cutoff value as a threshold. To be considered disadvantaged, a census tract must exceed the cutoff values for both (1) an environmental or climate indicator, and (2) the corresponding socioeconomic indicators. For example, to exceed one of the combined reduction and remediation of legacy pollution thresholds, a tract must (1) meet an environmental or climate indicator (i.e., be closer to hazardous waste facilities, to Risk Management Plan (RMP) facilities, or to National Priorities List (NPL or Superfund) sites than 90% of all census tracts), and (2) meet the socioeconomic indicators designed to identify low income communities (i.e., have more households living at or below 200% of the Federal poverty level than 65% of all census tracts, and have 80% or more of the population over 15 not currently enrolled in higher education[5]).

---

natural disasters, environment, clean energy, clean transportation, housing, water and wastewater infrastructure, and legacy pollution reduction, including in energy communities), remediation and reduction of legacy pollution, and critical clean water and waste infrastructure.

[4] Environmental or climate criteria encompass environment, climate, health, or other socioeconomic burdens. In particular, the training and workforce development category reflects socioeconomic burdens experienced by marginalized, underserved, and overburdened communities. It is also the only complete threshold across all 50 states, D.C., Puerto Rico, the Northern Mariana Islands, the U.S. Virgin Islands, American Samoa, and Guam.

[5] The beta version of the CEJST attempts to accurately identify low income communities. College students often report low incomes to the U.S. Census because they do not work full-time, even though they may not necessarily be disadvantaged. Requiring that over 80% of the census tract's population not currently be enrolled in higher education helps ensure that communities with many students, such as those in college towns or university districts, are not be categorized as disadvantaged simply because many students live there.

The beta version of the CEJST uses thresholds rather than indexing, which enables the tool to account for regional, state, and other geographic differences across the U.S., including between urban and rural areas. Each threshold is measured independently and the thresholds do not compete with each other. In other words, adding a new threshold to the methodology will add some communities to the definition of disadvantaged without having to remove other communities. For example, in the climate change threshold category, coastal flooding and wildfires are both included among the environmental or climate indicators, so tracts that experience a high rate of coastal flooding, a high rate of wildfires, or both, are represented. More detail on the use of thresholds and cutoff values, including the cutoff values for the two indicators that use percentages, can be found in Section III.A.2.

The eight combined threshold categories and their indicators are shown in the table below. Additional details about each indicator can be found in Section III.A.1.

Table 1. Overview of methodology in the beta version of the CEJST

Category	Environmental or climate indicators	Socioeconomic indicators
Climate change	1. Expected agriculture loss rate is at or above the 90th percentile, 2. Expected building loss rate is at or above the 90th percentile, or 3. Expected population loss rate is at or above the 90th percentile	1. Low income (at or above 65th percentile for households living at or below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Clean energy and energy efficiency	1. Energy burden is at or above the 90th percentile, or 2. Particulate matter (PM) 2.5 in the air is at or above the 90th percentile	1. Low income (at or above 65th percentile for households living at or below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Clean transit	1. Diesel particulate matter exposure is at or above the 90th percentile, or 2. Traffic proximity and volume is at or above the 90th percentile	1. Low income (at or above 65th percentile for households living at or below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Affordable and sustainable	1. Percent of housing units built pre-1960 (used as an indicator of	1. Low income (at or above 65th percentile for households living at or

Category	Environmental or climate indicators	Socioeconomic indicators
housing	potential lead paint exposure) is at or above the 90th percentile AND median home value is at or less than the 90th percentile, or 2. Housing cost burden is at or above the 90th percentile	below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Reduction and remediation of legacy pollution	1. Proximity to hazardous waste facilities is at or above the 90th percentile, or 2. Proximity to Risk Management Plan (RMP) facilities is at or above the 90th percentile, or 3. Proximity to National Priorities List (NPL or Superfund) sites is at or above the 90th percentile	1. Low income (at or above 65th percentile for households living at or below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Critical clean water and wastewater infrastructure	1. Wastewater discharge is at or above the 90th percentile	1. Low income (at or above 65th percentile for households living at or below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Health burdens	1. Asthma is at or above the 90th percentile, or 2. Diabetes is at or above the 90th percentile, or 3. Heart disease is at or above the 90th percentile, or 4. Low life expectancy is at or above the 90th percentile	1. Low income (at or above 65th percentile for households living at or below 200% of the Federal poverty level), and 2. Over 80% of individuals 15 or older are not currently enrolled in higher education
Training and workforce development	1. Low median income as a percentage of area median income is at or above the 90th percentile, or 2. Linguistic isolation is at or above the 90th percentile, or 3. Unemployment is at or above the	1. More than 10% of adults 25 or older have not attained a high school degree, and 2. Over 80% of people 15 or older are not currently enrolled in higher education

Category	Environmental or climate indicators	Socioeconomic indicators
	90th percentile, or 4. The percentage of people living at or below 100% Federal poverty line is at or above the 90th percentile	

B. Data overview

The beta version of the CEJST includes data from a wide variety of sources, including the U.S. Census Bureau, EPA, CDC, DOT, DOE, FEMA, and HUD. The datasets selected were chosen after consideration of more than 150 datasets and ideas for datasets suggested by Federal agencies, environmental justice data experts, and the White House Environmental Justice Advisory Council (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 beta version of the tool includes comprehensive datasets related to climate, environmental, energy, and economic justice.
2. **Publicly-available:** Because the CEJST and its methodology are publicly-available, the data it uses must also be publicly-available. Confidential 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 (see Section III.B.4). In light of this, datasets were only selected if, at a minimum, they were comprehensive and consistently accurate for all 50 states and D.C. Effort was also made to include datasets that are relevant to both rural and urban areas.
4. **Data exists at the census tract level**: The beta version of the CEJST identifies disadvantaged communities at the census tract level because that 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. Additional information on the decision to use data at the census tract level can be found in Section III.B.1.

The datasets used in the beta version of the CEJST are summarized in the table below. As noted in the table, certain variables are used directly in the beta version of the CEJST, without additional calculations, whereas others are calculated by the beta version of the CEJST before use in the tool; more detail on how the datasets are used can be found in Section III.B.2.

Table 2. Overview of data in the beta version of the CEJST

Dataset	Source (date)	Variables included or calculated	Relevant categories
U.S. Small-area Life Expectancy Estimates Project (USALEEP)	CDC (2010-2015)	● Low life expectancy	Health burdens
American Community Survey (ACS)	U.S. Census (2015-2019)[6]	● Low income ● Unemployment ● Median home value ● Lead paint (housing stock age used as a proxy) ● Poverty (CEJST calculates) ● High school degree attainment rate (CEJST calculates) ● Linguistic isolation (CEJST calculates) ● Higher education non-enrollment	All thresholds
PLACES: Local Data for Better Health	CDC (2016-2019)	● Heart disease ● Asthma ● Diabetes	Health burdens
Low-Income Energy Affordability Data (LEAD) score	DOE (2018)	● Energy burden	Clean energy and energy efficiency
EJScreen	EPA (varies, 2014-2020)	● Traffic proximity and volume ● Diesel particulate matter exposure ● PM2.5 ● Wastewater discharge ● Proximity to Risk Management Plan (RMP) facilities ● Proximity to National Priorities List (NPL) sites ● Proximity to hazardous waste facilities	● Clean energy and energy efficiency ● Clean transit ● Reduction and remediation of legacy pollution ● Critical clean water and

---

[6] Note that the years of data available for the U.S. territories are different. See Section III.B.4 for more detail.

Dataset	Source (date)	Variables included or calculated	Relevant categories
			wastewater infrastructure
Comprehensive Housing Affordability Strategy (CHAS)	HUD (2014-2018)	● Housing cost burden (CEJST calculates)	Affordable and sustainable housing
National Risk Index	FEMA (2014-2021)	● Expected population loss rate (CEJST calculates) ● Expected building loss rate (CEJST calculates) ● Expected agriculture loss rate (CEJST calculates)	Climate change
TIGER files	U.S. Census (2010)	● Geographies of census tract boundaries	N/A

C. Screenshots of the beta version of the CEJST

This section contains two screenshots of a census tract from Baltimore, Maryland that is identified as disadvantaged by the beta version of the CEJST. These images are from the "Explore the Tool" webpage of the CEJST, which is available at https://screeningtool.geoplatform.gov/en/#3/33.47/-97.5.

Figure 1: A screenshot of a census tract identified as a disadvantaged community by the beta version of the CEJST, shown in the version of the CEJST at the time of the publication of this support document. The sidebar of the map's user interface is "closed."
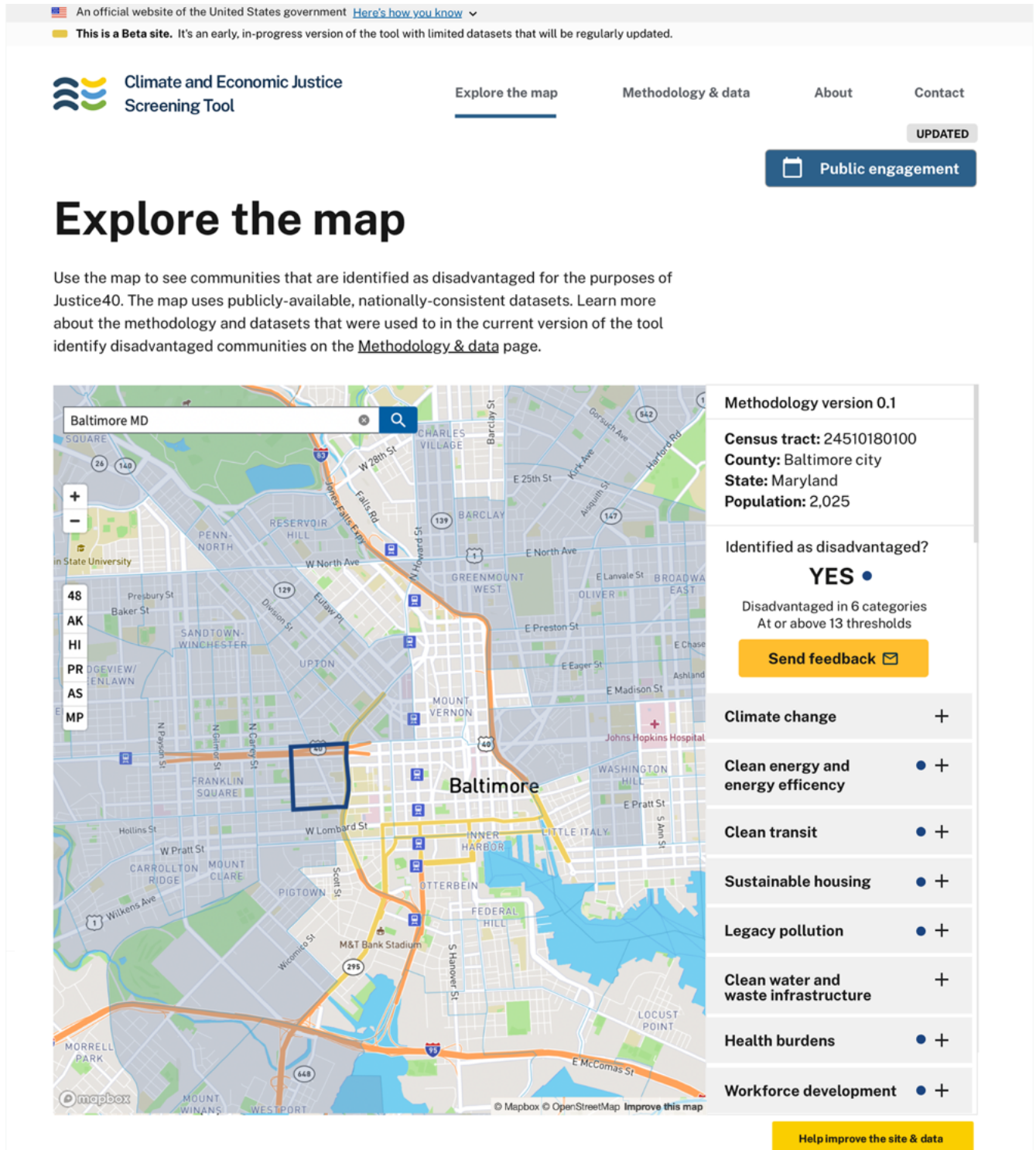
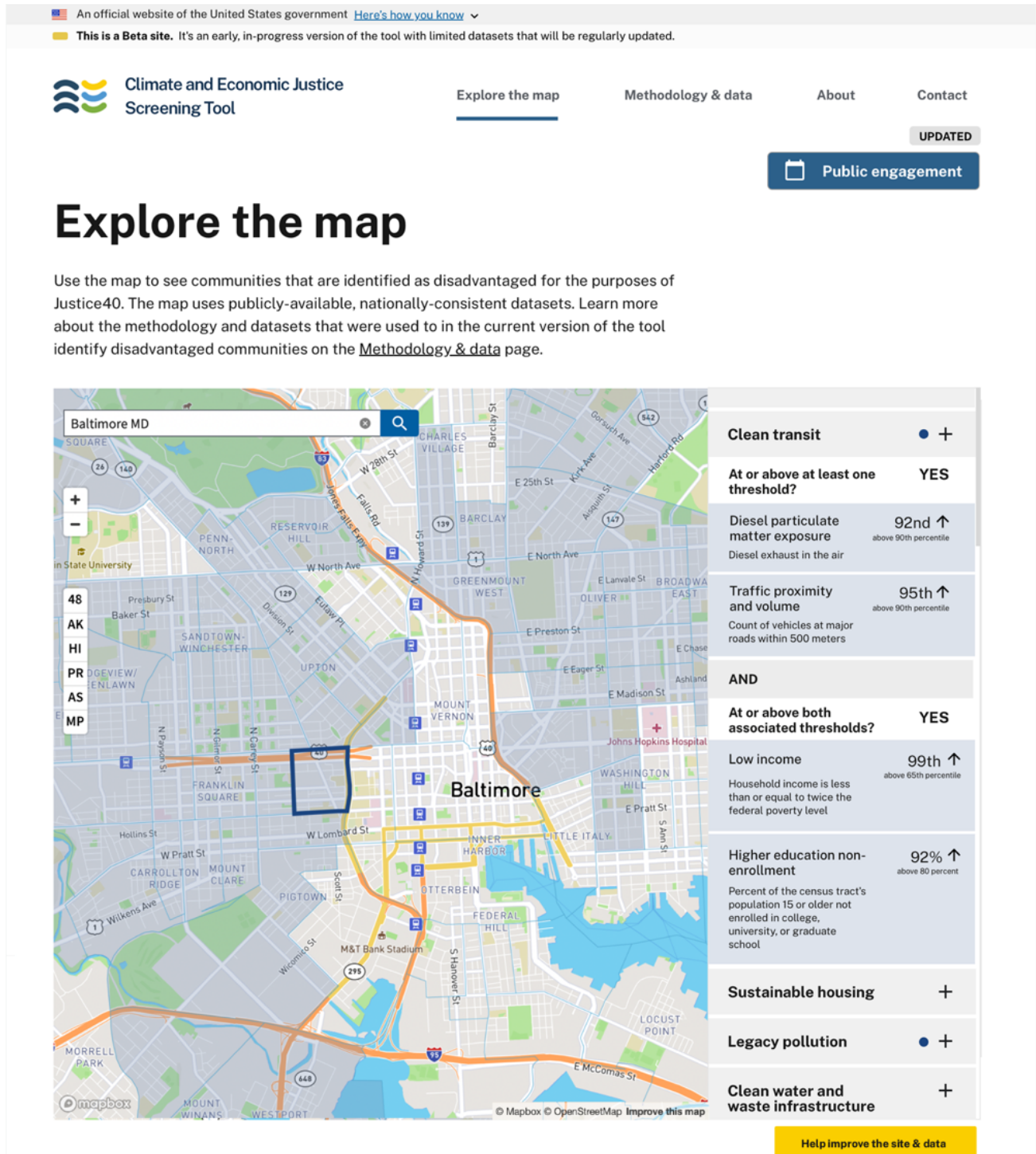**Figure 2:** A screenshot of a census tract identified as a disadvantaged community by the beta version of the CEJST, shown in the version of the CEJST at the time of the publication of this support document. The sidebar of the map's user interface is "open," revealing specific data elements about the census tract. In the full website, more data are visible as the user scrolls through the sidebar.

III.   Details on the methodology and data

A. Methodology

This section provides detailed information on the methodology used in the beta version of the CEJST, with three sub-sections outlining the use of (1) indicators; (2) thresholds and cut-off values; and (3) percentiles for normalization.

1. Indicators

The indicators used in the beta version of the CEJST are informed by the goal set forth in Executive Order 14008 "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."[7] In addition, the indicators are in threshold categories that track with the Justice40 covered areas of investment, which are climate change, clean energy and energy efficiency, clean transit, affordable and sustainable housing, training and workforce development (related to climate, natural disasters, environment, clean energy, clean transportation, housing, water and wastewater infrastructure, and legacy pollution reduction, including in energy communities), reduction and remediation of legacy pollution, and critical clean water and wastewater infrastructure.[8]

i.   Climate change

The environmental or climate burden indicators aim to measure expected agriculture value, building value, and population loss due to natural hazards that have some link to climate change within a census tract. The two socioeconomic indicators, household income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data comes from FEMA's National Risk Index and the U.S. Census's ACS.

ii.   Clean energy and energy efficiency

The environmental or climate burden indicators aim to measure the energy burden as well as energy-related pollution within a census tract. The two socioeconomic indicators, household income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data comes from DOE's LEAD Score, EPA's EJScreen, and the U.S. Census's ACS.

iii.   Clean transit

The environmental or climate burden indicators aim to measure the transportation-related pollution, noise, and proximity within a census tract. The two socioeconomic indicators, household income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data comes from EPA's EJScreen and the U.S. Census's ACS.

iv.   Affordable and sustainable housing

The environmental or climate burden indicators aim to measure the housing cost burden and the degree of lead paint exposure in housing within a census tract. The two socioeconomic indicators, household

---

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

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

income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data come from HUD's CHAS and the U.S. Census's ACS.

v. Reduction and remediation of legacy pollution

The environmental or climate indicators aim to measure how much legacy pollution, current pollution, and potential pollution through hazardous waste, NPL sites (otherwise known as Superfund), and RMP facilities within a census tract. The two socioeconomic indicators, household income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data come from EPA's EJScreen and the U.S. Census's ACS.

vi. Critical clean water and wastewater infrastructure

The environmental or climate burden indicator aims to measure the census tract's proximity to toxicity-weighted wastewater discharges. The two socioeconomic indicators, household income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data come from EPA's EJScreen and the U.S. Census's ACS.

vii. Health burdens

The environmental or climate indicators aim to identify areas facing high rates of asthma, diabetes, and heart disease within a census tract. The two socioeconomic indicators, household income and the share of the population not currently enrolled in higher education, together aim to identify census tracts that are low income. Data come from CDC's PLACES, CDC's USALEEP, and the U.S. Census's ACS.

viii. Training and workforce development

The environmental or climate indicators aim to identify census tracts that would benefit from greater workforce development, in line with the Justice40 Initiative's goal of promoting training and workforce development related to climate, clean energy, and other related categories. These indicators are low median income as a percentage of area median income, linguistic isolation, 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. The socioeconomic indicators identify areas with limited educational attainment: low high school degree attainment rate combined with a high share of the population not currently enrolled in higher education. The training and workforce development category thresholds are the only sources of data in the beta version of the CEJST that are currently available for some of the U.S. territories (see Section III.B.4 for more detail). Data come from the U.S. Census's ACS and the Decennial Census.

2. Thresholds and cutoff values

The methodology for the beta version of the CEJST uses a thresholds approach rather than indexing to identify disadvantaged census tracts. In contrast, some state-based environmental justice screening tools, such as [CalEnviroScreen](#) and [Maryland EJScreen Mapper](#), use averaging and multiplication to combine normalized data into a single index score, otherwise known as indexing. Similarly, the EPA's [EJScreen](#) also uses an index. The project development team carefully considered whether to develop the beta version of the CEJST as an index. Given the available datasets, however, it was determined that a thresholds approach would better enable the tool to identify disadvantaged communities in rural and

urban areas across the United States, as well as in the U.S. territories. The thresholds approach addressed concerns that certain areas of the nation and territories would be penalized in an index simply because they that had less available data.

Cutoff values are used to identify which census tracts exceed a given combined threshold and, thus, are identified as disadvantaged. Cutoffs were selected based on (1) how many census tracts would exceed the cutoff value; and (2) how easily understood the cutoff value would be, with a preference for round numbers.

The cutoff value for all environmental or climate indicators that use percentiles is the 90th percentile, which results in the selection of the top 10% of tracts for each indicator. This allows the beta version of the CEJST to identify census tracts with the greatest environmental and climate burdens. The cutoff for the low income indicator (<200% of the Federal poverty line) is 65th percentile, to reflect census tracts that exhibit some degree of socioeconomic need.

There are two cutoffs based on the raw percent value rather than percentile, which are exceptions to the rule that cutoffs are based on percentile. First, the beta version of the CEJST only identifies a community as disadvantaged if, in addition to meeting other criteria for each category, 80% or more of individuals above 15 years of age living in the census tract are not currently attending higher education (college, university, or graduate school). This indicator helps to identify low income communities by screening out those with many students, such as parts of a college town or university district, which may be categorized as low income simply because they have many students who do not work full-time. Over 96% of census tracts have more than 80% of the population not currently enrolled in higher education; in other words, less than 4% of tracts are screened out.

Consider, for example, a census tract immediately outside of University of California, Berkeley that has many residential college students. This tract is identified as low income because many full-time college students do not work full-time. Therefore, the average income of the census tracts immediately surrounding the university may be artificially lower due to the fact that many of the residents are higher education students. The beta version of the CEJST currently uses higher education enrollment to remove these tracts from being automatically considered disadvantaged.

Second, in the training and workforce development category, a community would not be identified as disadvantaged unless, 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 have not attained a high school degree. Approximately fifty percent of census tracts in the CEJST have populations where 10% or more of adults aged 25 or older do not have high school degrees.

i.   Examples of thresholds approach

The indicators for thresholds work together to identify disadvantaged census tracts. Again, the thresholds approach works to identify disadvantaged communities that trigger *both* environmental and socioeconomic indicators.

Consider three illustrative examples from the reduction and remediation of legacy pollution category:

- **Census tracts that meet the environmental or climate indicator only**: Assume a census tract is in the 95th percentile for proximity to NPL sites, but 40% of its residents are current college students. Even though it meets an environmental or climate burden indicator (proximity to Superfund sites), it is not considered a disadvantaged community for the reduction and remediation of legacy pollution threshold category because it does not meet the socioeconomic indicator (because of its high share of higher education students). It would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tracts that meet the socioeconomic indicator only**: Assume a census tract is in the 83rd percentile for low income and contains 7% higher education students. It satisfies the socioeconomic indicator for the reduction and remediation of legacy pollution threshold category. But, if the same census tract is at the 85th percentile for proximity to hazardous waste facilities, the 87th percentile for proximity to NPL sites, and the 86th percentile for proximity to RMP facilities, it does not satisfy any of the environmental or climate indicators for the threshold category, because none of these variables meet the cutoff of 90th percentile. Even though it meets both socioeconomic burden indicators (low income and low share of higher education students), because it does not meet the environmental or climate indicators (proximity to hazardous waste, proximity to NPL sites, and proximity to RMP), it is not considered a disadvantaged community. It would not be highlighted in a blue-gray shade on the tool's website map. (It is worth noting that as part of a public [Request for Information](#) on the beta version of the CEJST, CEQ has specifically solicited public input on how the tool's methodology could incorporate a cumulative impacts approach that quantitatively measures the combined adverse factors that contribute to the conditions that Justice40 is intending to address.)
- **Census tracts that meet both the environmental or climate and the socioeconomic indicators:** Assume a census tract is in the 92nd percentile for proximity to NPL sites, is in the 87th percentile for low income, and only has 5% of its population currently enrolled in higher education. The census tract is labeled as disadvantaged because it satisfies both the environmental and the socioeconomic indicators for reduction and remediation of legacy pollution threshold category. It would be highlighted the color blue-gray on the tool's website map. Note that it does not have to satisfy all of the other environmental or climate indicators to exceed a single threshold within the threshold category.

Three illustrative examples for the thresholds used in the training and workforce development category offer another perspective. This category provides the only currently available data in the CEJST that can be used to identify disadvantaged communities in the U.S. territories of Puerto Rico, American Samoa, and the Northern Mariana Islands:

- **Census tracts that meet the environmental or climate indicator only:** Assume a census tract is in the 95th percentile for unemployment rate, but 60% of its residents are not currently enrolled in higher education. Even though the census tract meets the unemployment rate indicator, it is not identified as disadvantaged because less than 80% of its residents are not currently enrolled in higher education; in other words, this appears to be a community with a lot of students, such

as part of a college town or university district. This tract would not be highlighted in a blue-gray shade on the tool's website map.

- **Census tracts that meet the socioeconomic indicator only:** Assume a census tract has an 83% high school degree attainment rate and has 93% of the population not enrolled in higher education. It satisfies the socioeconomic indicators for the training and 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 an English-speaking person over the age of 14, and 86th percentile for unemployment, it does not satisfy any of the environmental or climate indicators for the threshold category, because none of these variables meet the 90th percentile cutoff. Even though the census tract meets the socioeconomic burden indicators (low high school degree attainment rate and more than 80% of people 15 or older not currently enrolled in higher education), because it does not meet the environmental or climate indicator (median income, poverty, no English-speaking person over 14, and unemployment), it is not considered a disadvantaged community. It would not be highlighted in a blue-gray shade on the tool's website map.
- **Census tracts that meet both the environmental or climate and the socioeconomic indicators:** Assume a census tract is in the 92nd percentile for unemployment rate, has an 85% high school degree attainment rate, and a population where 95% are not currently enrolled in higher education. The census tract is labeled as disadvantaged because it satisfies both the environmental and the socioeconomic indicators for the training and workforce development category. It would be highlighted in a blue-gray shade on the tool's website map. Note that the census tract does not have to satisfy all of the other environmental or climate indicators to exceed the threshold category and be considered disadvantaged.

3. Percentiles for normalization

The beta version of the CEJST converts most variables to percentiles in order to normalize them and present them on a single scale. The variables used in the beta version of the CEJST naturally have very different scales. For instance, PM2.5 is measured in micrograms of particulate matter per cubic meter of air, which 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 indicators in the beta version of the CEJST, a higher percentile generally refers to a community being more overburdened and/or underserved on that indicator, i.e., a community is exposed to pollution. In order to achieve this directional uniformity, the percentiles for some indicators are calculated in reverse so that the data can be displayed consistently. For instance, for life expectancy in years, greater life expectancy is a sign of a healthy community. If life expectancy percentiles were included directly in the tool, it would create confusion by requiring users to remember that "low" values for this particular metric indicate high burden. Instead, the beta version of the CEJST uses the percentile measurement for *low* life expectancy, so that a high percentile value indicates a high amount of health burden.

The advantages and disadvantages of different methods for normalization were carefully considered before a decision was made to use percentiles in the beta version of 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.

B. Data

This section presents additional information on the data used in the beta version of the CEJST, with four sub-sections explaining (1) the selection of U.S. census tracts as the geographical unit; (2) the datasets used; (3) the datasets not currently eligible for inclusion; and (4) the challenges in identifying data for the U.S. territories.

1. Census tract as the geographic unit

The beta version of the CEJST identifies disadvantaged communities at the level of U.S. census tracts, which 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 precise targeting of benefits, but not so small that it introduces statistical unreliability in some of the data elements used in the tool. The decision to use census tracts was made after careful consideration of other alternatives.

The project development team initially examined whether it was possible for the unit of analysis to be U.S. census block groups, but ultimately decided that adequate data could not at this time be reliably reported at this level. The U.S. Census's Decennial Census and ACS have data available down to the census block group level, which are usually between 600 and 3,000 people. However, these estimates at the census block group level are statistically unreliable (i.e., they have high margins of error) in communities where sampling was limited. During testing of the beta version of the CEJST, the use of census data at the block level resulted in some spurious communities being identified as disadvantaged communities 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 EPA's EJScreen tool makes certain adjustments for certain variables 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 variables, EJScreen assigns each census block group the NATA, PM2.5, or ozone score of the census tract in which it is located; in other words, every census block group within the census tract is given the same exact score. When this same approach was tried for all the datasets used in the beta version of the CEJST, it generated significant statistical unreliability.

The project 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, for example, are reported at the level of U.S. counties, some of which—like Los Angeles County in California—contain millions of people. Other datasets are available at the zip code level. Some zip codes, however, can contain as many as 120,000 people. Further, zip codes do not fit neatly inside of the U.S. Census Bureau's hierarchy of geographies.[9] 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 crosswalks between zip codes and census tracts do exist, the disadvantages of using data reported at the zip code level outweighed the benefits.

2. Datasets used

This section describes how datasets were used in the beta version of the CEJST. 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. However, many datasets are not currently available for the U.S. territories. Unless otherwise noted below, the datasets listed below only include comprehensive and consistent data for the 50 states and D.C. Additional information on the challenges of identifying reliable data for the U.S. territories is discussed below in Section III.B.

The beta version of 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. The discussion of each dataset below describes when each approach was used. In addition, the discussion notes when the percentile is calculated in reverse order to address the fact that high values of that indicator indicate low burdens.

i. U.S. 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, and over 3.5 million households submit completed surveys each year. The American Community Survey collects data for all 50 states, D.C., and Puerto Rico.

Variables used directly in the beta version of CEJST, without additional calculations:
● **Low median income:** Low median income is defined as low median income of the given census tract as a percent of the area's median income. Area median income identifies the midpoint of a given area's income distribution, and accounts for regional differences in cost of living.

---

[9] 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.

- Unlike most of the other datasets, high values of this indicator indicate low burdens. For percentile calculations, the percentile is calculated in reverse order, so the tract with the highest value is at the 0th percentile.
- **Low income:** The low income variable reports what percentage of a tract's population consist of households with income at or below 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.
  - Unlike most of the other datasets, high values of this indicator indicate low burdens. For percentile calculations, the percentile is calculated in reverse order, so the tract with the highest value is at the 0th percentile.
- **Unemployment rates:** Unemployment rates report the percent of unemployed people in the civilian labor force. Unemployment is a crucial indicator of socioeconomic status.
- **Lead paint proxy:** Lead paint and lead-containing dust are pathways for significant exposure for many Americans, and elevated blood lead levels are considered an important environmental health issue. In the absence of data directly measuring lead paint exposure, the beta version of the CEJST uses the percent of housing units built before 1960 (in combination with low median home values to assess lead paint exposure and exclude historic mansions) as a proxy.
- **Low median home value:** The beta version of the CEJST uses low median home value of owner-occupied housing units from the ACS. Home value is the estimated price a home would sell if it went on the market today. Median home prices in the ACS data are based on survey respondents' answers to a subjective question asking for the dollar value of their home.
  - Unlike most of the other datasets, high values of this indicator indicate low burdens. For percentile calculations, the percentile is calculated in reverse order, so the tract with the highest value is at the 0th percentile.

Variables calculated by the beta version of CEJST before use in the tool:
For information on which specific fieldnames to include, please see the public GitHub repository for this project: https://github.com/usds/justice40-tool.

- **Poverty:** The ACS does not directly provide a measure of the percent of individuals living in households who are at or below 200% of the Federal poverty line for that year. Poverty is calculated by adding together households in different percentage brackets of the Federal poverty line.
- **High school degree attainment:** High school degree attainment rate is the percentage of adults over 25 years old within a given census tract who possess a high school diploma. The ACS does not directly provide a measure of high school degree attainment rate. This is calculated by adding together all measures of educational achievement up until less than high school graduation.
- **Linguistic isolation (households without an English-speaking person over the age 14):** Linguistic isolation refers to households that responded to ACS survey questions on language by stating all persons over the age 14 speak English less than "very well." The ACS does not directly provide a measure of linguistic isolation. This is calculated by adding together all measures of individually limited-English speaking households together.

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

The U.S. Small-Area Life Expectancy Estimates Project (USALEEP), created by the 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 beta version of the CEJST.

Variables used directly in the beta version of the CEJST, without additional calculations:
- **Low life expectancy:** Average number of years of life a person who has attained a given age can expect to live.
    - Unlike most of the other datasets used in the beta version of the CEJST, high values of this indicator indicate low burdens. For percentile calculations, the percentile is calculated in reverse order, so the tract with the highest value is at the 0th percentile.

iii.   CDC's PLACES

PLACES is a public, interactive website launched by the 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 beta version of the CEJST uses data from the 2016-2019 PLACES dataset.

PLACES contains data relevant to the health burden indicators used in the beta version of the CEJST.

Variables used directly in beta version of 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 ensures that survey responses match the population characteristics of each census tract.
- **Diabetes:** 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 ensures that survey responses match the population characteristics of each census tract.

iv.   DOE's LEAD Score

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 beta version of the CEJST, 2018 LEAD Score data are used to calculate energy burden.

Variables used directly in beta version of the CEJST, without additional calculations:
- **Energy burden:** Annual energy cost divided by household income, in U.S. dollars. Based on LEAD Tool data, the average energy burden for low income households is three times higher than for non-low income households, contributing to high energy insecurity.

v.   EPA's EJScreen

EJScreen is an environmental justice mapping and screening tool developed by the 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. Currently, EJScreen includes 11 environmental indicators and 6 demographic indicators. The beta version of the CEJST uses data from the 2020 EJScreen dataset.

EJScreen contains metrics relevant to the environmental burden indicators used in the beta version of the CEJST. These data comes from a myriad of sources and is compiled by the EPA.

Variables used directly in beta version of CEJST, without additional calculations:
- **Traffic proximity and volume:** Count of vehicles at major roads within 500 meters, divided by distance in meters, and 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 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.
- **Wastewater discharge:** Modeled toxic concentrations at stream segments within 600 meters, divided by distance in kilometers, and compiled from Risk-Screening Environmental Indicators (RSEI) Model from 2020.
- **Proximity to Risk Management Plan (RMP) facilities:** Count of potential chemical accident management plan facilities within 5 kilometers (or nearest one beyond 5 kilometers), divided by distance in kilometers, compiled from EPA's RMP database from 2020. These facilities are mandated by the Clean Air Act to file risk management plans because they handle substances with significant environmental and public health risks.
- **Proximity to National Priorities List (NPL) sites:** Count of proposed or listed NPL (Superfund) 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.

vi.   HUD's Comprehensive Housing Affordability Strategy

HUD's Comprehensive Housing Affordability Strategy (CHAS) serves as the strategic guide for housing and community development activities funded by HUD grants. CHAS data include information on housing problems and needs of low income households.

CHAS contains data relevant to the housing burden indicators used in the beta version of the CEJST. In the beta version of the CEJST, housing cost burden is measured as the percentage of households in a census tract that are both earning less than 80% of HUD Area Median Family Income by county and are paying more than 30% of their income to housing costs.

Variables calculated by the beta version of the CEJST before use in the tool:
For information on which specific fieldnames are included, please see the public github repository for this project, https://github.com/usds/justice40-tool.

- **Housing burden:** 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).

vii.   FEMA's National Risk Index (NRI)

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 beta version of the CEJST takes indicators from the 2014-2020 FEMA NRI dataset.

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

Variables calculated by the beta version of the CEJST before use in the tool:
For information on which specific fieldnames are included, please see the public github repository for this project: https://github.com/usds/justice40-tool.
- **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 beta version of 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 indicator, and the indicator is computed as the 90th percentile of tracts with agricultural value.

viii.   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, as well as housing unit count and population for a given year. These data are available for public download and consumption at https://data.census.gov/.

The beta version of the CEJST uses the 2010 tract boundaries because they match the data used in the tool. The U.S. Census updated these tract boundaries in 2020 and CEQ is evaluating whether to use these data in future updates.

Please note that 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.

3. Datasets not currently eligible for inclusion

The project development team compiled over 150 datasets and ideas for datasets from federal agencies, environmental justice data experts, and recommendations on the CEJST from the WHEJAC.

Many of the suggested datasets are not currently suitable for inclusion in a narrowly-targeted resource allocation methodology because they do not contain data at the census tract level. CEQ, with the support of the U.S. Digital Service, will work 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, and by collaborating in new data-gathering efforts that are 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 beta version of the CEJST.

- Proximity to concentrated animal feeding operations (CAFOs): The data for these sites in EPA's Facility Registry Service (FRS) significantly under-represent the reality of CAFOs across the U.S. and many of the addresses on the data could be improved to be more accurate. Similarly, the U.S. Department of Agriculture may have relevant data, as demonstrated by their 2012 Census of Agriculture; however, these data are not available at the census tract level.
- Coal ash dump sites: The project development team is working to identify a consistent source of information.

- Safe Drinking Water Information System (SDWIS): This dataset is not yet mapped between system names (and their health-based violations) and system geographic boundaries / census tracts.
- 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 HUD's Neighborhood Stabilization Program. The project development team is looking for more recent data.
- Pesticide use data: These data are not currently publicly-available at the census tract level.

CEQ plans to assess other datasets that are being suggested in response to the request for public comment (see Section IV for more detail). The CEJST will be continually updated as new data becomes available.

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

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

Puerto Rico: The U.S. Census Bureau's ACS data from 2015-2019 are used for all fields in the Training and Workforce category to identify disadvantaged communities in Puerto Rico. This is the only category for which data determined to be reliable exists.

Guam and the U.S. Virgin Islands: The beta version of the CEJST does not identify disadvantaged communities in Guam and the U.S. Virgin Islands. The available data and methodology for Guam and the U.S. Virgin Islands requires further investigation, and work is still underway to identify disadvantaged communities and update the CEJST accordingly for these two island territories.

American Samoa and the Northern Mariana Islands: The Decennial Census data from 2010 are used to identify disadvantaged communities in American Samoa and the Northern Mariana Islands. The beta version of the CEJST uses the following data indicators for these territories: unemployment, poverty, area median income, and high school degree achievement rate fields in the training and workforce development category.

The beta version of the CEJST uses a slightly different methodology to calculate the relevant percentiles for American Samoa and the Northern Mariana Islands because the relevant data are from the 2010 Decennial Census, which is not used for the other regions. To ensure that the percentiles used for the training and workforce development indicators make the appropriate comparisons, the beta version of the CEJST combines the measures from American Samoa and the Northern Mariana Islands with their corresponding measures from the 2010 ACS in the 50 states, D.C., and Puerto Rico in order to calculate the appropriate percentiles. (In general, the methodology in the beta version of the CEJST looks at the 90th percentile of a measure not only within a single state/territory or region, but also across the entire United States). The 2010 ACS data was used for the 50 states, D.C., and Puerto Rico because it contained

the relevant data; in contrast, the 2010 Decennial Census does not contain these data for the 50 states, D.C., and Puerto Rico. The methodology then uses this 90th percentile cutoff comparing 2010 data across the United States to highlight communities as disadvantaged in American Samoa and the Northern Mariana Islands.

Finally, the methodology also makes an adjustment to the methodology for calculating the area median income percentile for American Samoa and the Northern Mariana Islands. Area median income is calculated by dividing the median income of the census tract as reported in the 2010 Decennial Census by the median income of the entire territory as reported in the 2010 Decennial Census. Low income as a percentage of area median income is then converted into a percentile for American Samoa and the Northern Mariana Islands, and any tracts in the 90th percentile or above (that also meet the high school degree achievement criterion) are considered disadvantaged. In contrast to the area median income methodology used elsewhere, the percentiles within American Samoa and the Northern Mariana Islands 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 and the Northern Mariana Islands (i.e. Decennial Census versus ACS) but it still achieves the same result of selecting roughly 10% of communities, with the usual distributional changes.

IV. Get involved

A. Opportunities to provide feedback

Members of the public can explore the beta version of the map and data and provide feedback. This feedback will be used to help CEQ update and refine the tool. To provide feedback on the tool or raise support issues, please email Screeningtool-Support@omb.eop.gov, or fill out the survey at https://www.surveymonkey.com/r/cejst-survey, which is also linked on the CEJST.

CEQ has also published a Request for Information in the Federal Register to solicit input and feedback. The Request for Information can be found at https://www.federalregister.gov/documents/2022/02/23/2022-03920/climate-and-economic-justice-screening-tool-beta-version.

B. Open source codebase

All of the programming code for building the beta version of the CEJST is available at https://github.com/usds/justice40-tool. It is mostly written in Python and JavaScript (TypeScript). The 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 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 by suggesting "pull requests" to the codebase for consideration by CEQ. There is currently a regular meeting for members of the public interested in the CEJST open source development; please see the Github repository for details.

C. Tribal consultation

CEQ is engaging in consultation and coordination with Tribal Nations on the beta version of 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](). Leaders of Federally recognized Tribes should have received an email communication with relevant details from CEQ or the White House Office of Intergovernmental Affairs.

V.   Next steps

A new version of the CEJST, along with an updated technical support document, will be released after the feedback from the public comment period and Tribal Nations consultation has been reviewed. The CEJST will be continually updated as new data and research become available.

[end of document]