TC Explorer Technical Methodology

3 January 2025

This technical methodology document is for the January 2025 release of the US Department of Transportation's Transportation Community (TC) Explorer. The webpage for the TC Explorer is: TC Explorer | US Department of Transportation.

Contents

Note on links: Throughout this document, links in roman type are links to external documents or websites, while links in *italic type* are to terms defined in the glossary, starting on page 61.

1. Introduction to the TC Explorer

The Transportation Community (TC) Explorer is an economic development tool that allows communities to better identify transportation investments that can benefit communities, including rural and tribal communities. When done right, transportation policy can transform economies, connect people to opportunities, and empower underserved communities to build generational wealth for the future. This technical methodology document describes the TC Explorer as released in January 2025.

The TC Explorer consists of *indicator* variables—measures of specific demographic or physical traits of a community—as well as *disadvantage score* variables—percentile-ranked measures of transportation disadvantage in terms of an *overall disadvantage* and *component scores* and *subcomponent scores* that measure different components of disadvantage. All of these variables are calculated for and provided in terms of 2020 *Census tracts*. In addition, a number of *display layers* that may be of use in preparing Statewide Transportation Improvement Programs (STIPs), *Metropolitan Planning Organization* (MPO) Transportation Improvement Programs (TIPs), Notice of Funding Opportunity (NOFO) applications, and other transportation planning documents are provided.

The disadvantage score variables form a hierarchy, beginning with thirteen subcomponent scores, calculated from indicator variables, which quantify specific subcomponents of transportation disadvantage. These subcomponent scores are then used to calculate three component scores—*transportation insecurity*, *place-based burden*, and *population-based vulnerability*—which are in turn used to calculate the overall disadvantage scores.

The relationships between indicators, subcomponent, and component scores are shown Table 1 on the next page. Details on the calculation of the indicator scores can be found in Section 2 on page 6, while details on the calculation of disadvantage scores based on them can be found in Section 3 on page 21.

In general, communities—identified in terms of Census tracts—are identified as experiencing *disadvantage* if their overall disadvantage score is at or above 65th percentile nationally, and are identified as experiencing a component or subcomponent of disadvantage if their score for that component or subcomponent is above 65th percentile nationally. However, all Census tracts in the US territories—where most of the data to calculate indicators is not available—are considered to experience overall disadvantage and its subcomponents.

The TC Explorer is intended to allow the evaluation of *project areas* made up of multiple Census tracts as well as individual tracts. Disadvantage scores for project areas made up of multiple tracts or portions of tracts are calculated as population-weighted averages (based on 2020 Decennial Census populations) of the disadvantage scores for the tracts (or portions of tracts) in the project area. A project areas is identified as disadvantaged if either the majority of its population is in disadvantaged tracts and/or the majority of tracts it is located in are disadvantaged.

Table 1: Relationship Between Component Scores, Subcomponent Scores, and Indicators

Component Scores	Subcomponent Scores	Indicators
Transportation Insecurity	Destination Access Vulnerability	Pedestrian Access Score
		Cyclist Access Score
		Motorist Access Score
	Vehicle Access Vulnerability	Children, Elderly, and Disabled
		Vehicles Per Adult
		Households Without Vehicles
	Transportation Cost Burden	Transportation Cost Burden
	Traffic Fatality Burden	Traffic Fatalities—Buffered
Placed-Based Burden	Extreme Weather Hazard	Extreme Heat
		Extreme Precipitation
		Freeze-Thaw Cycles
		Drought
		Impervious Surface
		Wildfire Risk
		Flood Inundation
	Infrastructure Proximity	Railroad Proximity (½- & 1- mile)
		Freeway Proximity (½- & 1- mile)
		High-Traffic Road Proximity (½- & 1- mile)
		Major Airport Proximity (½- & 1- mile)
		Port Proximity (½- & 1- mile)
	Air Pollution Burden	Diesel Particulates Concentration
		Nitrogen Dioxide Concentration
		Air Toxics Cancer Risk
		Air Toxics Respiratory Risk
		Air Toxics Concentration
	Surface Pollution Burden	Hazardous Waste Biennial Reporter Proximity
		Toxic Release Inventory Site Proximity
		Risk Management Plan Site Proximity
		Hazardous Waste Site Proximity
		Leaking Underground Storage Tanks
		Active Mine Proximity
Population-Based Vulnerability	Communication Vulnerability	Population With Limited English
		Households Without Internet Access
	Employment Vulnerability	Population With Limited Education
		Population Not Currently Employed
	Income Vulnerability	Population Below 200% Poverty Line
		Population Without Health Insurance
		Relative Household Income
	Housing Vulnerability	Households Renting Housing
		Households With Cost-Burdened or Inadequate Housing
	Health Vulnerability	Asthma Prevalence
		High Blood Pressure Prevalence
		Cancer Prevalence
		Diabetes Prevalence
		Poor Mental Health Prevalence

The TC Explorer online tool consists of five pages: the home page, which contains a basic introduction to the tool, the TC Explorer – National Results page, the TC Explorer – State Results page, the Transportation Insecurity Analysis Tool (TIAT) page, and the Data and Methodology Download page.

The TC Explorer – National Results page displays disadvantage scores calculated via percentile-ranking over all tracts within the *United States*, while the TC Explorer – State Results pages displays disadvantage scores calculated via percentile ranking of tracts within a given state. Both pages also contain a "Raw Data" tab which displays indicators as opposed to disadvantage scores; since these are not percentile-ranked, the values are the same on both pages.[1] In addition, the maps in the two tabs show the same set of display layers. The sources for these display layers are discussed in Section 6.6 on page 57.

The updated Transportation Insecurity Analysis Tool (TIAT) linked to on the fourth tab of the TC Explorer online tool is a separate tool, developed by the Bureau of Transportation Statistics (BTS) that provides information on transportation costs and cost burdens as well as other forms of transportation insecurity. It is described in detail in a separate technical methodology document.

The Data and Methodology Download tab of the TC Explorer online tool provides links to this technical methodology document, the TC Explorer data dictionary, the TIAT technical methodology document, user guides for the TC Explorer and TIAT, as well as a number of downloadable datasets, collectively referred to herein as the TC Explorer Technical Data Download files.

The TC Explorer Technical Data Download files are listed on the next page:

[1] When using the indicators reported for project areas in the TC Explorer Online Tool, it is important to note that the "Average Median Household Income" indicator is not a simple median household income value. Due to technical limitations, it is the average of the values of median household income for all tracts in the project area.

TC Explorer Technical Data Download Files

- **TC_Tracts_GDB_2025_01_03.zip** – A geodatabase of Census tracts with indicator and disadvantage scores. A data dictionary is also included.
- **TC_Tracts_CSV_2025_01_03.zip** – A CSV file of Census tracts with indicator and disadvantage scores. A data dictionary is also included. Includes the same data as the GDB version.
- **TC_Counties_2025_01_03.zip** – A geodatabase of county equivalents. A data dictionary is also included.
- **TC_Centroids_and_Neighbors_2025_01_03.zip** – Geodatabases of Census tract population and land area centroids, along with CSV files of the fifteen nearest neighbors for each Census tract (Section 5.3 on page 40).
- **TC_Display_Disadvantage_2025_01_03.zip** – A geodatabase containing layers representing the Opportunity Zone and Area of Persistent Poverty (Section 6.6.1 on page 57) display layers in the web tool.
- **TC_Display_Geographies_2025_01_03.zip** – A geodatabase containing layers representing *state*, *county*, locality, *Metropolitan Statistical Area*, *Metropolitan Planning Organization*, *Urban Area*, and Census Bureau Tribal Land Area (Section 6.6.2 on page 57) display layers in the web tool.[2]
- **TC_Display_Infrastructure_2025_01_03.zip** – A geodatabase containing the transportation infrastructure, traffic fatalities, and alternative fueling station (Section 6.6.3 on page 58) display layers in the web tool.
- **TC_Display_NonAttainment_2025_01_03.zip** – A geodatabase containing the EPA Non-Attainment Area (Section 6.6.4 on page 60) display layers in the web tool.
- **TC_Travelsheds_(Pedestrian)_2025_01_03.zip** – Pedestrian travelsheds for Census tracts constructed for use in calculating the destination access scores (Section 4.1 on page 26).
- **TC_Travelsheds_(Cyclist)_2025_01_03.zip** – Cyclist travelsheds for Census tracts constructed for use in calculating the destination access scores (Section 4.1 on page 26).
- **TC_Travelsheds_(Motorist)_2025_01_03.zip** – Motorist travelsheds for Census tracts constructed for use in calculating the destination access scores (Section 4.1 on page 26).
- **TIAT_Cost_Burden_Tracts_2019_v01_2024_12_31.zip** – Transportation Insecurity Analysis Tool (TIAT) 2019 data.
- **TIAT_Cost_Burden_Tracts_2021_v01_2024_12_31.zip** – Transportation Insecurity Analysis Tool (TIAT) 2021 data.

[2] While Bureau of Indian Affairs Land Area Representation layers are available in the web tool, we are unable to make them available for download at this time.

2. Calculating the Indicators

A total of forty-five *indicators* are used in the calculation of disadvantage scores in the TC Explorer, while an additional four (housing cost burden, traffic fatalities—non-buffered, 2.5-micron particulates concentration, and median household income) are displayed in the raw data tabs of the online tool but not used in calculating disadvantage scores.[3] These indicators can be divided into three categories: *transportation insecurity* indicators (Section 2.2 on page 7), *place-based burden indicators* (Section 2.3 on page 11), and *population-based vulnerability* indicators (Section 2.4 on page 16), depending on which disadvantage component score they are related to.

Below, these variables are listed—along with ACS population and housing counts (Section 2.1 on page 6) used in the processing of the indicator data—followed by the variable names used for them in the TC Technical Data Download file. With the exception of the ACS population and household counts indicators, each indicator is represented by three variables in the technical data download, suffixed with (`_R`) for the *raw indicator*, (`_N`) for the *normalized indicator*, and (`_A`) for a text field indicating the availability of the indicator.

2.1 ACS Population and Household Counts

Population and household count data variables from the 2019-2023 American Community Survey (ACS) / Puerto Rico Survey (PRS) 5-year Estimates produced by the Census Bureau are included in the technical data download, and the total population variable is presented in the online tool. Details on the ACS / PRS 5-year Estimates can be found in Section 6.2 on page 46.

Although these variables lack availability fields in the technical data download, they are available for all *Census tracts* in the fifty states, the District of Columbia, and Puerto Rico. (They are not available for tracts in the *insular areas*.)

Total Population (ACS_POP)

Total population of the Census tract, ACS variable `B01003_001`.

Household Population (ACS_POPHH)

Population of the Census tract living in households, ACS variable `B25008_001`.

Group Quarters Population (ACS_POPGQ)

Population of the Census tract living in group quarters (i.e. all residents not living in households), calculated as: `B01003_001 - B25008_001`.

Total Households (ACS_HH)

Total number of households in the Census tract, ACS variable `B25002_002`. (Note that this will necessarily be less than or equal to the number of housing units, with the difference being the number of unoccupied housing units.)

---

[3] Three of these, housing cost burden, traffic fatalities—non-buffered, and median household income, are supplied to add context to indicators (transportation cost burden, traffic fatalities—buffered, and relative household income) that are used in the model. The fourth, 2.5-micron particulates concentration, is provided at the request of the Federal Railroad Administration (FRA) but is not included in the disadvantage scores because it is unavailable for Alaska and Hawaii.

Population Density (ACS_DENSE)

Population density of the Census tract, ACS variable B01003_001 divided by the land area reported in the *TIGER/Line shapefile* for the tract, ALAND20.

2.2 Transportation Insecurity Indicators

Transportation Insecurity occurs when people are unable to get to where they need to go to meet the needs of their daily life regularly, reliably, affordably, and safely. The *indicators* used to measure transportation insecurity are discussed below. Detailed information on the data sources can be found in Section 6.3 on page 47.

2.2.1 Destination Access Vulnerability Indicators

Destination Access Vulnerability measures access to essential destinations, such as public transit, medical facilities, education, groceries, and jobs. Higher scores reflect project areas where residents have less access to essential destinations within 30 minutes of driving, walking, or biking.

The destination access vulnerability indicators are unique in that, unlike the other indicators in the TC Explorer, they were developed specifically for this tool. Details on how they were calculated can be found in the appendix at Section 4 on page 26. The indicators measure the accessibility of common destinations from the population centroids (or land area centroids for *low-population tracts*) of Census tracts based on the numbers of these destinations reachable within a half hour of travel by three different modes of transportation: walking, cycling, and driving. *Travelsheds* were generated using ESRI ArcGIS Pro using ESRI's proprietary routable network and ten subindicators were measured for each indicator:

- **Area –** the area within the travelshed
- **Education Facilities –** the number of colleges and K-12 schools within the travelshed
- **Groceries –** the number of grocery stores within the travelshed
- **Libraries –** the number of public libraries within the travelshed
- **Medical Facilities –** the number of hospitals, outpatient care facilities, and pharmacies within the travelshed
- **Parks –** the number of parks within the travelshed
- **Post Offices –** the number of post offices within the travelshed
- **Transit –** the number of transit trips serving the travelshed in a typical service week
- **Population –** the population living within the travelshed
- **Jobs –** the number of jobs within the travelshed

The destination access vulnerability indicators were then calculated by normalizing the subindicators over tracts in the *United States* and then averaging them to create overall indicators for each mode of transportation.

Pedestrian Access Score (ACC_PED)

The pedestrian access score indicator is based on a 30-minute walking trip, operationalized as ½ mile on roads and paths where pedestrians are permitted. Note that this indicator does not take into account safety features such as sidewalks and signalized intersections due to the lack of data.

Cyclist Access Score (ACC_CYC)

The cyclist access score indicator is based on a 30-minute cycling trip, operationalized as 5 miles on roads and paths where pedestrians are permitted. Note that this indicator does not take into account the effect of terrain or safety features such as protected bike lanes due to the lack of data.

Motorist Access Score (ACC_MTR)

The motorist access score indicator is based on a 30-minute driving trip away from the tract centroid, starting at 8:30am on a typical Wednesday.

Contrary to the usual case, higher values of these indicators indicate less disadvantage. As a result, the normalized versions of these indicators have been subtracted from 1 to produce normalized indicators with 1 indicating maximum disadvantage and 0 indicating minimum disadvantage.

2.2.2 Vehicle Access Vulnerability Indicators

Vehicle Access Vulnerability measures whether residents in the project area have access to a vehicle and/or have the ability to drive. Higher scores reflect project areas where households are less likely able to drive to essential destinations.

The indicators used to measure vehicle access vulnerability are all demographic measures based on the 2019-2023 American Community Survey (ACS) / Puerto Rico Survey (PRS) 5-year estimates. Details on the ACS / PRS can be found in Section 6.2 on page 46.

Because these indicators were calculated based on demographic data, values for *Census tracts* with too-few residents or households for the data to be meaningful were replaced with nearest-neighbor interpolated values as though they had missing values, as described in Section 3.1 on page 21. The criterion for replacement is specified for each indicator.

Children, Elderly, and Disabled (VEH_ABL)

The children, elderly, and disabled indicator measures the share of the population that is under age 18, age 65 or older, or disabled.[4] The indicator is calculated with the expression:

```
((DP05_0001 - DP05_0021) + DP05_0024 + DP02_0076)/DP05_0001
```

This value is replaced with a nearest-neighbor interpolated value for tracts with ACS_POP < 100.

[4] As per ACS subject definitions: a person is disabled if they "are deaf or have serious difficulty hearing"; "are blind or have serious difficulty seeing even when wearing glasses"; "have serious difficulty walking or climbing stairs"; "have difficulty dressing or bathing"; and/or "have difficulty doing errands alone such as visiting a doctor's office or shopping." Disability status is not reported for active-duty military service members and residents of custodial *group quarters*, and these populations are not included in the denominator.

Vehicles Per Adult (`VEH_VPA`)

The vehicles per adult indicator measures the ratio of total vehicles available to [*households*] to the total number of adults, age 18 or older living in households.[5] The indicator is calculated with the expression:

```
B25046_001 / B09021_001
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_HH < 100`.

Contrary to the usual case, higher values of this indicator indicate less disadvantage. As a result, the normalized version of this indicator has been subtracted from `1` to produce a normalized indicator with `1` indicating maximum disadvantage and `0` indicating minimum disadvantage.

Households Without Vehicles (`VEH_NVH`)

The households without vehicles indicator measures the share of *households* with no vehicles available.[6] The indicator is calculated with the expression:

```
DP04_0058 / DP04_0057
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_HH < 100`.

2.2.3 Transportation Cost Burden Indicators

Transportation Cost Burden measures the share of income that households in the project area spend on daily transportation. Higher scores reflect project areas where households spend a higher percentage of their income on transportation.

Along with the transportation cost burden indicator, which is the sole indicator used to calculate the transportation cost burden subcomponent score, housing cost burden is displayed in the raw data tab of the online tool. Both indicators are sourced from the Transportation Insecurity Access Tool (TIAT) profile 1 (all households) datasets for 2021 (Section 6.3 on page 47). It is important to keep in mind that, while the housing costs and household income reported in this dataset are direct observations taken from the American Community Survey, the transportation costs are estimates based on models of daily travel behavior and expenses, and are likely an underestimate of transportation costs in *isolated Alaska tracts*, where air travel is necessary to carry out essential trips of daily life that would occur by driving, cycling, walking, or public transportation in other parts of the *United States*.

[5] As per ACS subject definitions, “vehicles available” refers to “passenger cars, vans, and pickup or panel trucks of one-ton (2,000 pounds) capacity or less kept at home and available for the use of household members. Vehicles rented or leased for one month or more, company vehicles, and police and government vehicles are included if kept at home and used for non-business purposes. Motorcycles or other recreational vehicles are excluded. Dismantled or immobile vehicles are excluded. Vehicles kept at home but used only for business purposes also are excluded.”

[6] As per ACS subject definitions, “vehicles available” refers to “passenger cars, vans, and pickup or panel trucks of one-ton (2,000 pounds) capacity or less kept at home and available for the use of household members. Vehicles rented or leased for one month or more, company vehicles, and police and government vehicles are included if kept at home and used for non-business purposes. Motorcycles or other recreational vehicles are excluded. Dismantled or immobile vehicles are excluded. Vehicles kept at home but used only for business purposes also are excluded.”

Transportation Cost Burden (`CST_TCB`)

The transportation cost burden indicator measures the ratio of estimated average [*household*] transportation cost to median household income. The indicator is taken from the `transportation_cost_burden` variable in profile 1 (all households) of the 2021 TIAT dataset. This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_HH < 100`.

Housing Cost Burden (`CST_HCB`) – Displayed Only on the Raw Data

The housing cost burden indicator is not included in the disadvantage score calculations but is reflected in the raw data tab of the online tool. This indicator measures the ratio of average *household* housing cost to median household income. The indicator is the ratio of the `housing_cost` and `median_income` variables in profile 1 (all households) of the 2021 TIAT dataset. This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_HH < 100`.

2.2.4 Traffic Fatality Burden Indicators

Traffic Fatality Burden measures traffic fatalities (both motorist non-motorist) using the National Highway Transportation Safety Administration's (NHTSA) Fatality Analysis Reporting System (FARS) data for 2018-2022. Higher scores reflect project areas with higher number of traffic fatalities.

The traffic fatality burden component score is calculated from a single indicator, "traffic fatalities—buffered", which is a measure of the number of traffic fatalities in and within 250 ft of a Census tract's border using National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS) datasets over the five-year period 2018-2022.

To enable users to calculate the total traffic fatalities in a group of Census tracts[7] a "traffic fatalities—not buffered" variable is included on the "Raw Data" tabs of the National and State Result pages of the online tool. Both variables are available in the Technical Data Download files.

Traffic Fatalities—Buffered (`SFT_BFT`)

The traffic fatalities—buffered indicator measures the total number of traffic fatalities that occurred within a Census tract and a 250-foot buffer around the tract borders between 2018 and 2022, as reported in FARS datasets (Section 6.3 on page 47). The 250-foot buffer, based on an approach taken from the California Department of Transportation Caltrans Transportation Equity Index, is intended to approximate the width of a wide arterial road or intersection. This is an important consideration for traffic fatality counts at the Census tract level because arterial roads are commonly chosen as Census tract boundaries and are also frequently the location of traffic fatalities. Using this buffer avoids the arbitrary assignment of fatalities on roads forming tract boundaries to one or the other tract and instead results in fatalities on tract boundaries being assigned to both tracts. This does, however, mean that the traffic fatalities—buffered indicator cannot be summed across multiple tracts.

Traffic Fatalities—Non-Buffered (`SFT_PTT`) – Displayed Only on the Raw Data

The traffic fatalities—not buffered indicator is equivalent to the traffic fatalities—buffered indicator except that the 250-foot buffer is not used, meaning that fatalities occurring near Census tract borders are assigned to only one tract. This indicator can be summed across tracts to produce traffic fatality counts for larger areas.

[7] The traffic fatalities—buffered values cannot be summed across tracts without double-counting fatalities.

2.3 Place-Based Burden Indicators

Place-Based Burden is the disadvantage inherent in a location and experienced by all residents of the location. These indicators are important because they provide transportation decision makers the information needed to develop transportation plans and make funding decisions that ensure a community's transportation infrastructure is safe, resilient, and minimizes negative health and economic impacts. The *indicators* used to measure place-based burden are discussed below. Detailed information on the data sources can be found in Section 6.4 on page 49.

2.3.1 Extreme Weather Hazard Indicators

Extreme Weather Hazard measures the predicted change in extreme weather hazards or variability in the project area by 2050 which may have impacts on transportation system performance, safety, and reliability. These impacts in turn have major implications to supply chains, emergency response and the longevity of transportation investments. Higher scores reflect project areas that are likely to see increased extreme weather impacts to transportation infrastructure.

Four of the extreme weather hazard indicators—extreme heat, extreme precipitation, freeze-thaw cycles, and drought—are taken from data tabulated at the Census tract level in the National Oceanographic and Atmospheric Administration (NOAA) Climate Mapping for Resilience & Adaptation (CMRA) dataset. The impervious surface indicator is extracted from the United States Geologic Survey (USGS)-led Multi-Resolution Land Characteristics (MLRC) Consortium 2021 National Land Cover Database (NLCD) and the National Oceanographic and Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP) High-Resolution Land Cover rasters. The wildfire risk indicator is extracted from the United States Forest Service (USFS) Spatial Datasets of Probabilistic Wildfire Risk Components for the United States (270m), 3rd Edition rasters. And the flood inundation indicator is extracted from the World Resources Institute (WRI) Aqueduct Flood Risk Rasters. Details on these datasets can be found in Section 6.4.1 on page 50.

Extreme Heat (`WTH_XHT`)

The extreme heat indicator measures the estimated annual number of days with high temperatures of at least 90°F by 2050 based on the RCP 8.5 climate model. The indicator is sourced from the CMRA variable:

```
RCP.8.5.Mid.century...Mean...Annual.number.of.days.with.a.maximum.temp
erature.greater.than.90degF
```

Extreme Precipitation (`WTH_XPR`)

The extreme precipitation indicator measures the estimated annual number of days with total precipitation equivalent to at least 2 inches of rain by 2050 based on the RCP 8.5 climate model. The indicator is sourced from the CMRA variable:

```
RCP.8.5.Mid.century...Mean...Annual.number.of.days.with.total.precipit
ation.greater.than.2.inches
```

Freeze-Thaw Cycles (`WTH_CLD`)

The freeze-thaw cycle indicator measures the estimated annual number of days with a minimum temperature below and maximum temperature above 32°F by 2050 based on the RCP 8.5 climate model. The indicator, intended to approximate the annual number of freeze-thaw cycles, is sourced from the CMRA variables:

```
RCP.8.5.Mid.century...Mean...Annual.number.of.frost.days..days.with.a.
minimum.temperature.less.than.32degF. -
RCP.8.5.Mid.century...Mean...Annual.number.of.icing.days..days.with.a.
maximum.temperature.less.than.32degF.
```

Drought (`WTH_DRO`)

The drought indicator measures the estimated increase in the annual number of days with total precipitation less than one one-hundredth of an inch by 2050 relative to the historical average based on the RCP 8.5 climate model. The indicator is sourced from the CMRA variables:

```
RCP.8.5.Mid.century...Mean...Annual.number.of.days.with.total.precipit
ation.less.than.0.01.inches -
Historical...Mean...Annual.number.of.days.with.total.precipitation.les
s.than.0.01.inches
```

Impervious Surface (`WTH_IMP`)

The impervious surface indicator measures the share of land area in the Census tract covered by impervious surface. The indicator is sourced from NLCD rasters for the contiguous *United States* and C-CAP rasters for Alaska, Hawaii, and the *US territories*.

Wildfire Risk (`WTH_WFR`)

The wildfire risk indicator measures the average annual wildfire burn probability for land in the Census tract. The indicator is sourced from the USFS Wildfire Risk Components for the United States rasters.

Flood Inundation (`WTH_FLD`)

The flood inundation indicator measures the share of land area in the Census tract predicted to be inundated by simultaneous 100-year riverine and 100-year coastal flooding events defined in terms of flooding probability in 2050 based on the RCP 8.5 climate model. The indicator is sourced from the non-partisan World Resources Institute (WRI) Aqueduct Flood Risk rasters. This is the only non-Federal governmental data source used in the TC Explorer. It is included because of the significant impact flooding has on transportation infrastructure and because, currently, there is not a government source for this data that includes riverine as well as coastal flooding.

2.3.2 Infrastructure Proximity Indicators

Infrastructure Proximity measures how close the project area is to freeways, high-volume roads, railways, airports, and ports, which may align with higher rates air and noise pollution, as well as divided communities. Higher scores reflect project areas close to one or more forms of transportation infrastructure.

The infrastructure proximity indicators were calculated as the shares of Census tract population (or land area for *low-population tracts*) within half-mile and one-mile buffers around transportation infrastructure facilities identified in US Department of Transportation and US Army Corps of Engineers datasets described in Section 6.4.2 on page 51. The ten variables are paired, with half-mile and one-mile buffers used for each of five types of infrastructure: railroads, freeways and expressways, high-traffic roads, major airports, and port facilities.

One-mile and half-mile buffers around the infrastructure facilities were calculated in the appropriate *UTM zones*. Population and land area intersections were calculated using the buffer-calculation geometries described in Section 5.4 on page 41.

Railroad Proximity (`INF_RRH` and `INF_RRF`)

The railroad proximity indicators measure the fraction of the tract's population within half-mile and one-mile buffers around rail lines.

Freeway Proximity (`INF_FWH` and `INF_FWF`)

The freeway proximity indicators measure the fraction of the tract's population within half-mile and one-mile buffers around freeways and expressway.

High-Traffic Road Proximity (`INF_TRH` and `INF_TRF`)

The high-traffic road proximity indicators measure the fraction of the tract's population within half-mile and one-mile buffers around roads with annual average daily traffic (AADT) of at least 50,000 trips.

Major Airport Proximity (`INF_ARH` and `INF_ARF`)

The major airport proximity indicators measure the fraction of the tract's population within half-mile and one-mile buffers around major airports, defined as all military airports and all other airports that in 2021 and/or 2022 saw at least 10,000 passenger enplanements or 100 million pounds of landed weight of cargo-only aircraft.

Port Proximity (`INF_PTH` and `INF_PTF`)

The port proximity indicators measure the fraction of the tract's population within half-mile and one-mile buffers around docks used for freight handling and shipping purposes. Docks not in use or used for museum or exhibit ships, military and other government operations, casino/showboats, or research and training purposes were excluded.

2.3.3 Air Pollution Burden Indicators

Air Pollution Burden measures the potential exposure of the project area to air pollutants and the adverse environmental conditions caused by air pollution. This measure does not incorporate carbon dioxide or greenhouse gases as a form of air pollution, but focuses on criteria air pollutants and on hazardous air pollutants (also known as air toxics), which are known to cause negative health outcomes. Higher scores reflect project areas with higher rates of air pollution.

There are five air pollution indicators—diesel particulates concentration, nitrogen dioxide concentration, air toxics cancer risk, air toxics respiratory risk, and air toxics concentration—used in the calculation of the air pollution burden subcomponent score. In addition, three other measures are discussed below. In addition, 2.5-micron particulates concentration is displayed in the raw data tab of the online tool and available in the Technical Data Download but was not included in the air pollution burden subcomponent scores because it is not available for Alaska and Hawaii.

All air pollution burden indicators are taken directly from the Environmental Protection Agency (EPA) Environmental Justice Screening and Mapping Tool (EJScreen) versions 2.2 (released 2023) and 2.3 (released 2024). Further details on this data source can be found in Section 6.4.3 on page 54.

Diesel Particulates Concentration (`AIR_DSL`)

The diesel particulates concentration indicator measures the concentration of diesel particulate matter in outdoors air in units of μg/m³. The source variable is `DSLPM` from EJScreen v2.3 for the contiguous *United States* and Hawaii and EJScreen v2.2 for Alaska.

Nitrogen Dioxide Concentration (`AIR_NO2`)

The nitrogen dioxide concentration indicator measures the annual concentration of nitrogen dioxide in outdoor air (as calculated from satellite imagery) in units of parts per billion (ppb). The source variable is `NO2` from EJScreen v2.3.

Air Toxics Cancer Risk (`AIR_CNR`)

The air toxics cancer risk indicator measures the estimated lifetime inhalation cancer risk from carcinogens in ambient outdoor air in units of persons per million residents. The source variable is `CANCER` from EJScreen v2.2.

Air Toxics Respiratory Risk (`AIR_RSP`)

The air toxics respiratory risk indicator is a non-cancer respiratory hazard index from toxic pollutants in ambient outdoor air in arbitrary units. The source variable is `RESP` from EJScreen v2.2.

Air Toxics Concentration (`AIR_RLS`)

The air toxics concentration indicator measures the toxicity-weighted concentrations of toxic chemicals in outdoor air modeled from toxic release inventory data in arbitrary units. The source variable is `RSEI_AIR` from EJScreen v2.3.

2.5-Micron Particulates Concentration (`AIR_P25`) – Displayed Only on the Raw Data

The 2.5-micron particulates concentration indicator measures the concentration of particulate matter with diameters of 2.5 microns or smaller in outdoor air in units of μg/m³. The source variable is `PM25` from EJScreen v2.3. This indicator is not available for Alaska and Hawaii and so is not included in the calculation of component scores but is available as a raw data value for use in decision making and grant application narratives.

2.3.4 Surface Pollution Burden Indicators

Surface Pollution Burden measures the potential exposure of the project area to land and surface water pollutants and the adverse environmental conditions cause by surface pollution. Higher scores reflect project areas with higher rates of surface pollution.

With the exception of the leaking underground storage tanks *indicator*—taken from the Environmental Protection Agency (EPA) Environmental Justice Screening and Mapping Tool (EJScreen) version 2.3 (released 2024)—all *surface pollution* indicators were calculated as the shares of Census tract population (or land area for *low-population tracts*) within one mile of a potential pollution point source. The point sources for the active mine proximity indicator are taken from the Department of Labor Mine Data Retrieval System (MDRS) and the point sources for the other four indicators—hazardous waste biennial reporter proximity, toxic release inventory site proximity, risk management plan site proximity and hazardous waste site proximity—were taken from the EPA Facility Registry Service (FRS). Further details on these data sources can be found in Section 6.4.4 on page 54.

One-mile buffers around the point sources were calculated in the appropriate *UTM zones*. Population and land area intersections were calculated using the buffer calculation geometries described in Section 5.4 on page 41.

Hazardous Waste Biennial Reporter Proximity (`SUR_BNL`)

The hazardous waste biennial reporter proximity indicator measures the fraction of the tract's population within one mile of a hazardous waste Biennial Reporter, listed as "`HAZARDOUS WASTE BIENNIAL REPORTER`" in the `INTEREST_TYPE` field of the EPA Facility Registry Service dataset.

Toxic Release Inventory Site Proximity (`SUR_TRI`)

The toxic release inventory site proximity indicator measures the fraction of the tract's population within one mile of a Toxic Release Inventory reporter site, listed as "`TRI REPORTER`" in the `INTEREST_TYPE` field of the EPA Facility Registry Service dataset.

Risk Management Plan Site Proximity (`SUR_RMP`)

The risk management plan site proximity indicator measures the fraction of the tract's population within one mile of a hazardous waste Risk Management Plan reporter site, listed as "`RMP REPORTER`" in the `INTEREST_TYPE` field of the EPA Facility Registry Service dataset.

Hazardous Waste Site Proximity (`SUR_TSG`)

The hazardous waste site proximity indicator measures the fraction of the tract's population within one mile of a hazardous waste large quantity generator or transportation, storage, or disposal facility, listed as "`TSD`" or "`LQG`" in the `INTEREST_TYPE` field of the EPA Facility Registry Service dataset.

Leaking Underground Storage Tanks (`SUR_UST`)

The leaking underground storage tanks indicator measures the estimated concentration of leakage from linking underground fuel storage tanks in arbitrary units. The source variable is `UST` from EJScreen v2.3.

Active Mine Proximity (SUR_MIN)

The active mine proximity indicator measures the fraction of the tract's population within one mile of an active mine (other than stone quarries and sand or gravel pits) listed in the Department of Labor Mine Data Retrieval System.

2.4 Population-Based Vulnerability Indicators

Population-Based Vulnerability is the disadvantage experienced by a population due to demographic and socioeconomic traits that make them particularly vulnerable. The *indicators* used to measure population-based vulnerability are all demographic measures based on the 2019-2023 American Community Survey (ACS) / Puerto Rico Survey (PRS) 5-year estimates and the Centers for Disease Control and Prevention (CDC) Population Level Analysis and Community Estimates (PLACES) dataset. Details on the ACS / PRS can be found in Section 6.2 on page 46 and details on PLACES can be found in Section 6.5 on page 55.

Because these indicators were calculated based on demographic data, values for *Census tracts* with too-few residents or households for the data to be meaningful were replaced with nearest-neighbor interpolated values as though they had missing values, as described in Section 3.1 on page 21. The criterion for replacement is specified for each indicator.

2.4.1 Communication Vulnerability Indicators

Communication Vulnerability measures whether community members can easily and reliably receive information and alerts about the transportation system. This information is important for planners developing emergency response and evacuation plans as well as information on infrastructures closures. Higher scores reflect project areas whose households have limitations in their ability to receive digital communications or understand English.

Population With Limited English (COM_LEP)

The population with limited English indicator measures the share of the population age 5 or older that speaks English less than "very well". The indicator is calculated from ACS / PRS data with the expression:

`DP02_0115 / DP02_0112`

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Households Without Internet Access (COM_INT)

The households without internet access indicator measures the share of *households* without internet access at home. The indicator is calculated from ACS / PRS data with the expression:

`S2801_C01_019 / S2801_C01_001`

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_HH < 100` as the indicator excludes *group quarters* residents.

2.4.2 Employment Vulnerability Indicators

Employment Vulnerability measures lack of employment and education. Higher scores reflect project areas whose residents have lower rates of employment and high school graduation.

Population With Limited Education (EMP_EDU)

The population with limited education indicator measures the share of the population age 25 or older without a high school diploma or equivalent (e.g. an GED). The indicator is calculated from ACS / PRS data with the expression:

```
(DP02_0059 – DP02_0067) / DP02_0059
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Population Not Currently Employed (EMP_WKF)

The population not currently employed indicator measures the share of the population age 16 or older that is unemployed or not in the labor force (i.e. the share of the population that is neither employed nor serving in the military). The indicator is from ACS / PRS data calculated with the expression:

```
(DP03_0005 + DP03_0007) / DP03_0001
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

2.4.3 Income Vulnerability Indicators

Income Vulnerability measures poverty and income inequality in the project area. Higher scores reflect project areas with higher poverty rates and median household incomes below the regional average.

Population Below 200% Poverty Line (INC_POV)

The population below 200% poverty line indicator measures the fraction of the population living below 200% of the Federal poverty line. The indicator is calculated from ACS / PRS data with the expression:

```
S1701_C01_042 / S1701_C01_001
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Population Without Health Insurance (INC_INS)

The population without health insurance indicator measures the fraction of the population without health insurance. It assumes that all active-duty military service members and residents of custodial *group quarters* (prisons and long-term healthcare facilities) have health insurance, as these populations are not included in the ACS / PRS estimates of the population without health insurance. The indicator is calculated from ACS / PRS data with the expression:

```
DP03_0099 / DP05_0001
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Median Household Income (INC_MED) – Displayed Only on the Raw Data

The median household income[8] indicator is not included in the disadvantage score calculations but is reflected in the raw data tab of the online tool. This indicator measures the median *household* income (ACS / PRS variable DP03_0062) for each Census tract. This value is replaced with a nearest-neighbor interpolated value for tracts with ACS_HH < 100 as the indicator excludes *group quarters* residents.

Contrary to the usual case, higher values of this indicator indicate less disadvantage. As a result, the normalized version of the indicator has been subtracted from 1 to produce normalized indicators with 1 indicating maximum disadvantage and 0 indicating minimum disadvantage.

Relative Household Income (INC_MSA)

The relative household income indicator measures the ratio of the median household income indicator to the overall median household income of the *region* (*metropolitan statistical area* or non-metropolitan portion of a state). Regional median household income is approximated using bucketed household income data tabulated at the county level from the ACS / PRS variable B19001 using the approximation that incomes within each bucket are distributed evenly across the income range of the bucket.

Contrary to the usual case, higher values of this indicator indicate less disadvantage. As a result, the normalized version of the indicator has been subtracted from 1 to produce normalized indicators with 1 indicating maximum disadvantage and 0 indicating minimum disadvantage.

2.4.4 Housing Vulnerability Indicators

Housing Vulnerability measures housing condition and the share of household income spent on housing. Higher scores reflect project areas where households spend a higher percentage of income on housing or there is a prevalence of household overcrowding and lack of indoor plumbing and kitchens.

Households Renting Housing (HOU_RNT)

The households renting housing indicator measures the share of *households* renting their primary *housing units*. As per ACS subject definitions, any household that does not include the owner or a co-owner of the housing unit is considered to be "renting," even if no rent is paid (e.g. resident managers, caretakers, housing provided for free by the owner to friends or relatives). However, a housing unit owned outright without mortgage by a household member on leased land is considered to be owner-occupied. The indicator is calculated from ACS / PRS data with the expression:

DP04_0047 / DP04_0045

This value is replaced with a nearest-neighbor interpolated value for tracts with ACS_HH < 100 as the indicator excludes *group quarters* residents.

[8] As per ACS subject definitions: "'Total income' is the sum of the amounts reported separately for wage or salary income; net self-employment income; interest, dividends, or net rental or royalty income or income from estates and trusts; Social Security or Railroad Retirement income; Supplemental Security Income (SSI); public assistance or welfare payments; retirement, survivor, or disability pensions; and all other income."

Households With Cost-Burdened or Inadequate Housing (HOU_CND)

The households with cost-burdened or inadequate housing measures the share of *households* occupying *housing units* satisfying one or more of the following conditions: lacking complete plumbing facilities,[9] lacking complete kitchen facilities,[10] with more than one occupant per room, selected monthly owner costs[11] as a percentage of household income greater than 30%, and gross rent[12] as a percentage of household income greater than 30%. The indicator is calculated from ACS / PRS data with the expression:

```
(B25123_003 + B25123_004 + B25123_005 + B25123_006 + B25123_009 +
B25123_010 + B25123_011 + B25123_012) / B25123_001
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_HH < 100` as the indicator excludes *group quarters* residents.

2.4.5 Health Vulnerability Indicators

Health Vulnerability measures the increased prevalence of health conditions that may result from exposure to pollutants, poor walkability, car dependency, and long commute times. Higher scores reflect project areas with higher rates of asthma, cancer, high blood pressure, diabetes, and poor mental health.

The five health vulnerability indicators are sourced from the CDC PLACES dataset and represent the prevalences of health conditions. The source variables are divided by 100 to convert them from percentages to fractions of the population, to correspond with the population shares derived from ACS / PRS data.

Asthma Prevalence (HEA_AST)

The asthma prevalence indicator measures the share of the population age 18 or older that has current asthma. The indicator is calculated with the expression:

```
CASTHMA / 100
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

---

[9] As per ACS subject definitions: "Complete plumbing facilities include: hot and cold running water and a bathtub or shower. Both facilities must be located inside the house, apartment, or mobile home, but not necessarily in the same room." Hot water is not required for "complete plumbing facilities" in Puerto Rico.

[10] As per ACS subject definitions: "A unit has complete kitchen facilities when it has all three of the following facilities: a sink with a faucet, a stove or range, and a refrigerator. All kitchen facilities must be located in the house, apartment, or mobile home, but they need not be in the same room. A housing unit having only a microwave or portable heating equipment such as a hot plate or camping stove should not be considered as having complete kitchen facilities. An icebox is not considered to be a refrigerator."

[11] As per ACS subject definitions: "Selected monthly owner costs are the sum of payments for mortgages, deeds of trust, contracts to purchase, or similar debts on the property (including payments for the first mortgage, second mortgages, home equity loans, and other junior mortgages); real estate taxes; fire, hazard, and flood insurance on the property; utilities (electricity, gas, and water and sewer); and fuels (oil, coal, kerosene, wood, etc.). It also includes, where appropriate, the monthly condominium fee for condominiums and mobile home costs (personal property taxes, site rent, registration fees, and license fees)."

[12] As per ACS subject definitions: "Gross rent is the contract rent plus the estimated average monthly cost of utilities (electricity, gas, and water and sewer) and fuels (oil, coal, kerosene, wood, etc.) if these are paid by the renter (or paid for the renter by someone else)."

High Blood Pressure Prevalence (HEA_HBP)
The high blood pressure prevalence indicator measures the share of the population age 18 or older that has ever been diagnosed with high blood pressure, excluding high blood pressure only during pregnancy and "borderline hypertension." The indicator is calculated with the expression:

```
BPHIGH / 100
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Cancer Prevalence (HEA_CNR)
The cancer prevalence indicator measures the share of the population age 18 or older that has ever been diagnosed with cancer, other than non-melanoma skin cancer. The indicator is calculated with the expression:

```
CANCER / 100
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Diabetes Prevalence (HEA_DIA)
The diabetes prevalence indicator measures the share of the population age 18 or older that has ever been diagnosed with diabetes, excluding diabetes only during pregnancy. The indicator is calculated with the expression:

```
DIABETES / 100
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

Poor Mental Health Prevalence (HEA_PMH)
The asthma prevalence indicator measures the share of the population age 18 or older that reports that their mental health was "not good" for at least 14 days out of the last thirty. The indicator is calculated with the expression:

```
MHLTH / 100
```

This value is replaced with a nearest-neighbor interpolated value for tracts with `ACS_POP < 100`.

3. Preparing Data and Calculating Disadvantage Scores

As outlined in Section 1 on page 2, the *disadvantage scores* in the TC Explorer are calculated in stages from *indicator* values. First, *nearest-neighbor interpolation* is used to supply missing indicator values (Section 3.1) Indicator values are then normalized on a 0-to-1 scale (Section 3.2 on page 22). National-ranked and state-ranked disadvantage scores are calculated are calculated (Section 3.3 on page 23) and, finally, *disadvantaged status* is assigned based on these scores (Section 3.4 on page 24).

3.1 Interpolation of Missing Indicator Values

In general, each indicator value was measured directly for each Census tract. Indicator values are left as `NA` for territories where no data is available. [13] However, in some cases an indicator is missing for a small number of isolated tracts. Furthermore, as discussed in Section 2 on page 6, measures that depend on demographic counts are not meaningful or reliable when the underlying population—residents or households—being sampled is too small.[14] In these cases, in which a direct measurement of indicator values is unavailable for individual tracts, nearest-neighbor interpolation is used to estimate the missing value.

For each indicator (except the ACS Population and Housing Counts, Section 2.1 on page 6, for which no interpolation was performed), a data availability field (suffixed `_A` in the Technical Data Download tracts files) is assigned for each variable as follows:

- "`AVAILABLE`" – for tracts where the data source from which the indicator is calculated are available; these values are taken directly from the data sources listed in Section 2.
- "`INTERPOLATED`" – for tracts where the data source from which the indicator is calculated is unavailable (or the tract has fewer than 100 residents or households per 2022 ACS / PRS data for variables where such a cut-off is listed in Section 2). For these tracts, nearest-neighbor interpolation is used to produce values of the indicator.
- "`UNAVAILABLE`" – for all *remote island tracts* and for tracts in Puerto Rico and the insular areas where data to calculate the indicator is generally unavailable for Puerto Rico and/or the insular areas. Also used for the 2.5-micron particulates concentration measures in Alaska and Hawaii, where this measure is not available.

For each tract-variable combination with a value of "`INTERPOLATED`", the interpolated value is calculated as the average of the values for the three nearest tracts from the nearest neighbor lists (Section 5.3 on page 40) with a value of "`AVAILABLE`". Note that, because only "`AVAILABLE`" tracts are used for this purpose, interpolated values are never used in the interpolation of values for other tracts. Indicators for *transportation insecurity* and *population-based vulnerability* are nearest-neighbor interpolated using population centroid nearest neighbors because these indicators represent traits of the

---

[13] With the exception of the 2.5-micron particulates concentration measure—not available for Alaska and Hawaii and not used in calculating disadvantage scores—all indicators are available for the fifty states and District of Columbia. All indicators except for the cost burden, health vulnerability, and some of the extreme weather hazard indicators are available for Puerto Rico. However, most indicators are not available for the *insular areas*.

[14] This is especially true since the demographic data sources the indicators are based on are all produced through sampling, meaning that substantial sampling error is possible. However, even if an exact count was available, a very small population would make measures of the share of the population with a given trait questionable and likely to be unstable if a small number of people moved into or out of the tract.

populations within tracts, while indicators for *place-based burden* are nearest-neighbor interpolated using land area centroid nearest neighbors because these indicators represent traits of the tracts themselves.

3.2 Normalization of Indicators

The indicators calculated as discussed in Section 2 on page 6 and interpolated as discussed in the previous section are *raw indicators*—direct measures of some demographic or physical property of Census tracts. These indicators, suffixed "_R" in the Technical Data Download tracts file, have a variety of units—shares of Census tract population, shares of households, parts per billion, and so on—depending on what is being measured, which makes them useful for assessing specific community conditions.

This variation in units of measurement makes raw indicators unsuitable for averaging together for the calculation of *subcomponent scores*. For this reason, DOT converted raw indicators with units to *normalized indicators*, suffixed "_N" in the Technical Data Download tracts file, which are averaged in the calculation of subcomponent scores. (The ACS Population and Household Counts indicators, described in Section 2.1 on page 6 are not normalized as they are not needed in the calculation of subcomponent scores.) Indicators are normalized to a 0-1 scale with 0 indicating the least and 1 the most disadvantage using a modified version of a traditional minimum-maximum normalization intended to ensure that both outlier and non-outlier data is represented fairly in the normalized indicators.

Simply defining 1 as the highest value the indicator takes in any tract, and 0 as the lowest value it takes in any tract—as in a traditional minimum-maximum normalization procedure—causes issues with outliers, as a small number of tracts with values far outside the range found in other tracts can compress all other variation in the indicator into a small subset of the 0-1 range. This compression of nearly all variation between tracts into a very small part of the range is undesirable, because it results in the indicator having little effect in disadvantage calculations other than to distinguish those tracts with extreme values from the other tracts: the indicator essentially becomes binary. Furthermore, extreme outliers of this sort are likely to be due to measurement errors in the data sources or circumstances where the measured values are not a good representative of the actual situation.[15]

To ameliorate this issue, normalized subindicator values for the TC Explorer are calculated as follows:

- Values of the indicator for tracts in the *fifty states and District of Columbia*, excluding *remote island tracts*, are percentile ranked and the value of the 1st percentile tract is set equal to a normalized value of 0 while the value of the 99th percentile tract is set equal to a normalized value of 1.
- All tracts (both in the fifty states and District of Columbia and in the *territories*) with values below the 1st percentile or above the 99th percentile are given normalized values of 0 and 1 respectively.
- Tracts (both in the fifty states and District of Columbia and in the *territories*) with values between the 1st percentile and 99th percentile are given normalized values equal to the difference between their

[15] For example, the campuses of elite universities where most or all students are full-time and do not have outside jobs often show extremely high poverty rates, because the students—although they are largely being supported financially by family and will likely obtain high-income jobs after graduating—do not currently have income.

value and the 1st percentile value divided by the difference between the 99th percentile and 1st percentile values.

This approach produces minimum-maximum normalized results based on the range of values covered by the vast majority of tracts. In cases where extreme outliers are not present—for example, most of the indicators that measure the share of tract population within the buffer area—it is very similar to or identical to the values produced by a traditional minimum-maximum normalization. However, when extreme outliers are present, it produces a broader measure of the variation between most tracts, while maintaining the position of the outliers at the edges of the distribution.

The final step in the normalization process concerns four transportation insecurity indicators—pedestrian access score, cyclist access score, motorist access score, and vehicles per adult—and two population-based vulnerability indicators—median household income and relative household income—which, unlike all other indicators in the TC Explorer, have higher values for tracts that are more advantaged (have more access to destinations, more vehicles, and higher incomes).

For these six indicators, the raw values remain unchanged, but the normalized values have been subtracted from `1` to produce normalized indicators with `1` indicating maximum disadvantage and `0` indicating minimum disadvantage.

3.3 Calculation of Disadvantage Scores

Once the *normalized indicators* are calculated, they are used to calculate the thirteen *subcomponent scores*. The normalized indicators that correspond to each subcomponent, as listed in Section 1 on page 2 and discussed at length in Section 2 on page 6, are averaged to produce raw subcomponent scores, suffixed "`_R`" in the Technical Data Download tracts files. These scores potentially can range between 0 and 1, but in practice usually have a smaller range, since it is rare for any tract to have values of either 0 or 1 for all normalized indicators.

The national-ranked subcomponent scores (suffixed "`_P`" in the Technical Data Download tracts file) are calculated for *Census tracts* in the *fifty states and District of Columbia* by percentile ranking all such tracts (excluding *remote island tracts*) and reporting the values on a 0-to-1 range, with 1 the most disadvantaged (highest subcomponent score) and 0 the least disadvantaged (lowest subcomponent score).

After national-ranked subcomponent scores have been calculated based on indicators, the *transportation cost burden* and *traffic fatality burden* scores for *isolated Alaska tracts* are set to 1 (most disadvantaged) in recognition of the fact that the transportation cost burden and traffic fatality burden indicators do not fully reflect the costs and safety issues inherent in transportation in these extremely isolated areas.

Although Census tracts in Puerto Rico are not included in the calculation of overall national scores, comparable national-ranked subcomponent scores are provided for Puerto Rico for those subcomponents for which all necessary indicators are available[16] to aid in developing grant application narratives and to provide an understanding of the tracts relative to disadvantage in the fifty states and

[16] Puerto Rico subcomponent scores are available for all subcomponents except transportation cost burden, extreme weather hazard, and health vulnerability.

District of Columbia. These national-ranked scores for Puerto Rico are determined by setting the ranked value equal to the ranked value of the tract in the fifty states and District of Columbia with a raw subcomponent value closest to that of the Puerto Rico tract.

The approach for calculating state-ranked subcomponent scores (suffixed "_SP" in the Technical Data Download tracts file) is the same as that used for national-ranked subcomponent scores, except that raw scores are percentile ranked solely among tracts in a given state (or tracts in Puerto Rico, for subcomponents available in Puerto Rico).

With percentile-ranked subcomponent scores calculated, the next step is the calculation of the three *component scores*. Raw component scores, suffixed "_R" "_SR" and in the Technical Data Download tracts files for nationally-ranked and state-ranked scores respectively, are calculated by averaging national-ranked and state-ranked subcomponent scores, suffixed "_P" and "_SP" in the Technical Data Download tracts file, for the subcomponents that correspond to a particular subcomponent score. National-ranked (suffixed "_P" in the Technical Data Download tracts file) and state-ranked (suffixed "_SP" in the Technical Data Download tracts file) component scores are then calculated from these raw values in the same manner as subcomponent scores were calculated. (Note, however, that no component scores are available for Puerto Rico, due to the unavailability of data.)

Finally, the *overall disadvantage score* is calculated from the three component disadvantage scores in the same manner as the component scores are calculated from the subcomponent scores: the raw score, suffixed "_R" for nationally-ranked and "_SR" for state-ranked component scores, is the average of the national-ranked component scores, and the national-ranked overall disadvantage score suffixed "_P" and state-ranked overall disadvantage score suffixed "_SP" are calculated from it.

As noted in Section 1 on page 2, disadvantage scores for project areas made up of multiple tracts or portions of tracts are calculated as population-weighted averages (based on 2020 Decennial Census populations) of the disadvantage scores for the tracts in the project area. If only part of a tract is included in a project area, the population of the area of the tract within the project area should be used as a weight.

3.4 Identification of Disadvantaged Communities

Once the national-ranked (suffixed "_P" in the Technical Data Download tracts file) and state-ranked (suffixed "_SP" in the Technical Data Download tracts file) *disadvantage scores* have been calculated, they are used to identify *disadvantaged communities*. Disadvantaged communities, in terms of *overall disadvantage* and the *component scores* and *subcomponent scores*, are identified in the fields suffixed "_B" (for national-ranked disadvantage) and "_SB" (for state-ranked disadvantage) in the Technical Data Download tracts file. These fields have values of "1" to indicate disadvantaged communities and "0" for other *Census tracts*, including all *remote island tracts*.

In general, and excluding remote island tracts, all tracts in the *US territories* (including Puerto Rico) are coded as disadvantaged communities, while tracts in the *fifty states and the District of Columbia* are coded as disadvantaged communities for overall disadvantage and components or subcomponents of disadvantage if the corresponding disadvantage score is at or above the 65th percentile (0.65 in the Technical Data Download tracts file). To make this more visually obvious, such scores are shown in red in the TC Explorer Online Tool.

Project areas consisting of multiple tracts or portions of multiple tracts have disadvantage scores calculated as population-weighted averages of the scores for the tracts in question. However, they are coded as disadvantaged if the majority of their population is in disadvantaged tracts or if the majority of the tracts that they intersect with are disadvantaged.

4. Appendix: Calculating Access to Destinations

The three *destination access vulnerability indicators* discussed in Section 2.2.1 on page 7 are pedestrian, cyclist, and motorist access scores, calculated as the averages of ten normalized subindicators measuring the availability of different destination types within thirty-minute pedestrian, cyclist, and motorist travelsheds of the population centroids (or land area centroids, for *low-population tracts*) of each *Census tract*. This appendix details the methodology for the generation of the travelsheds (Section 4.1 below) and the calculation of the availability of destinations within the travelsheds (Section 4.2 on page 29).

4.1 Travelshed Generation

Three sets of travelsheds were produced for the population centroids of each Census tract, one each for driving, cycling, and walking. All three sets of travelsheds were produced for thirty-minute travel times: this was operationalized as one mile for walking and five miles for cycling, and as 30 minutes driving outbound in morning rush-hour traffic for driving. The travelsheds were generated in ESRI ArcGIS Pro Desktop using ESRI's proprietary routing network and the Service Area Analyst tool.

4.1.1 Pedestrian and Cyclist Travelsheds

The thirty-minute pedestrian travelsheds were computed using a travel distance of 1 mile minus 200 meters, plus a polygon trim distance of 200 meters, while the thirty-minute cyclist travelsheds used a travel distance of 5 mile minus 200 meters, plus a polygon trim distance of 200 meters. That is, travel along the routing network was limited to 1 or 5 miles minus 200 meters (0.876 or 4.876 miles). A 200-meter buffer around the traversed portion of the routing network was added to account for the length of driveways and parking lots in suburban and rural areas, for a total of 1 mile of travel.

The pedestrian and cyclist travelsheds were generated excluding ferries, limited access roads, roads where walking is prohibited, and roads marked "unsuitable for pedestrians" in the ESRI routing network. Routes with gates were avoided where possible, and routes "preferred for pedestrians" were preferred where possible.[17] Pedestrian settings were used for cyclist travelsheds because the ESRI routing network does not offer separate settings for cyclists.

It was decided to exclude ferry travel from the pedestrian and cyclist travelsheds on the basis that they are essentially a form of public transportation—which was not otherwise included in these travelsheds—and would act to substantially increase potential travel distance. Furthermore, since ferries are often infrequent, including destinations only reachable by ferry would overestimate the destinations that could be reached within thirty minutes.

[17] The travel mode settings for the routing network were:

- **Avoid Ferries – Prohibited** (ferry travel was not allowed)
- **Avoid Gates – Avoid** (routes without gates were preferred)
- **Avoid Limited Access Roads – Prohibited** (limited-access roads were excluded)
- **Avoid Roads Unsuitable for Pedestrians – Prohibited** (roads "unsuitable for pedestrians" were excluded)
- **Preferred for Pedestrians – Prefer** (roads "preferred for pedestrians" were preferred)
- **Walking – Prohibited** (roads where walking is prohibited were excluded)

For most tracts, this approach produced reasonable travelsheds. However, in fewer than 100 cases, no travelshed or implausibly-small travelsheds were produced. These cases were evaluated individually as discussed below.

Adjusting Starting Locations for Pedestrian and Cyclist Travelsheds

In general, travelsheds were calculated for the population centroids of Census tracts (or land area centroids for low-population tracts). However, in some cases, it was necessary to adjust starting location to produce a viable travelshed.

An initial review of the pedestrian travelsheds produced found that nine centroids produced no pedestrian and cyclist travelsheds at all. All of these centroids were in extremely rural Alaska, where the tract centroid was not within 20 km (the maximum search distance) of any road. For each of these tracts, the starting location was moved from the centroid of the tract to the municipal building or community center of the largest community in the tract:

- Tract 1 in Aleutians East Borough, AK (FIPS code `02013000100`)
  Moved starting location to Akutan, AK, the largest community.
- Tract 1 in Aleutians West Census Area, AK (FIPS code `02016000100`)
  Moved starting location to Unalaska, AK, the largest community.
- Tract 1 in Copper River Census Area, AK (FIPS code `02066000100`)
  Moved starting location to Glennallen, AK, the largest community.
- Tract 1 in Dillingham Census Area, AK (FIPS code `02070000100`)
  Moved starting location to Togiak, AK, the largest community.
- Tract 1 in Kusilvak Census Area, AK (FIPS code `02158000100`)
  Moved starting location to Hooper Bay, AK, the largest community.
- Tract 1 in Lake and Peninsula Borough, AK (FIPS code `02164000100`)
  Moved starting location to Newhalen, AK, the largest community.
- Tract 2 in North Slope Borough, AK (FIPS code `02185000200`)
  Moved starting location to Wainwright, AK, the largest community.
- Tract 1 in Yukon-Koyukuk Census Area, AK (FIPS code `02290000100`)
  Moved starting location to Fort Yukon, AK, the largest community.
- Tract 3 in Yukon-Koyukuk Census Area, AK (FIPS code `02290000300`)
  Moved starting location to Galena, AK, the largest community.

In addition to the nine tracts where no travelshed was generated, 86 tracts were generated where the cyclist travelshed was smaller than 4.876 miles. These were examined individually and, in four cases, the travelsheds were found to be likely accurate, due to small, isolated communities with only freeway or similar access, where it is likely impossible to safely and legally cycle or walk in or out of the community. In the remaining cases, however, the small travelsheds appeared to be artifacts due to the location on the routing network closest to the tract Centroid being on a small, disconnected section of parking lot or walkway. In these cases, the starting location was shifted a small distance (less than 100 m) to the nearest connected segment of public road.

4.1.2 Motorist Travelsheds

The thirty-minute driving travelsheds were computed using a travel time model, using ESRI's routing network's estimated travel speeds at 8:30am on a typical Wednesday for travel away from the tract centroids to approximate morning rush-hour traffic. As with the pedestrian and cyclist travelsheds, a 200-meter buffer (polygon trim distance) was added. Given the much larger size of motorist travelsheds, however, this buffer adds a negligible distance to the potential distance travelled, and so it was not necessary to subtract it from the travel distance.

The motorist travelshed were generated excluding carpool-only lanes and roads, tolled express lands, roads closed for construction, and roads where driving an automobile is prohibited. In addition, ferries, roads with gates, private roads, unpaved roads, and roads where through traffic is prohibited were avoided where possible.[18] Ferries were included (though discouraged) in generating the motorist travelsheds because the use of travel time estimates with a specific start time eliminates the issue of unknown wait times for ferries for pedestrians and cyclists, and because vehicle ferries differ from other modes of public transportation in that they can be included as part of an automobile trip rather than only as an add-on at the start or end of a trip.

Adjusting Starting Locations for Motorist Travelsheds

In cases where the starting locations for cyclist and pedestrian travelsheds had to be shifted to generate valid travelsheds, discussed in the previous section, these shifted starting locations were also used for motorist travelsheds. Despite these shifts, four tracts that had generated pedestrian and cyclist travelsheds generated no motorist travelsheds and an additional fourteen motorist travelsheds were significantly smaller than expected.

Two of the four tracts that produced no motorist travelsheds are located in extremely rural areas of Alaska where the nearest community to the tract centroid was only accessible by air with no road connections, and where the community is small enough that driving is unlikely to be a substantial means of transportation within the community. The other two tracts were located on islands and it appears that no public roads are present within the tracts. The four tracts in question are:

- Tract 1 in Bethel Census Area, AK (FIPS code `02050000100`)
  The community nearest the centroid is Tuntutuliak, AK, a very small Inuit community that is only accessible by air and which is small enough—less than a mile across—that driving is unlikely to be a significant means of transportation within the community.
- Tract 1 in Northwest Arctic Borough, AK (FIPS code `02188000100`)
  The community nearest the centroid is Noorvik, AK, a very small Inuit community that is only

[18] The travel mode settings for the routing network were:

- **Avoid Carpool Roads – Prohibited** (carpool-only lanes and roads were excluded)
- **Avoid Express Lanes – Prohibited** (tolled express lanes were excluded)
- **Avoid Ferries – Avoid** (routes avoiding ferries were preferred)[18]
- **Avoid Gates – Avoid** (routes avoiding roads with gates were preferred)
- **Avoid Private Roads – Avoid** (routes avoiding private roads were preferred)
- **Avoid Unpaved Roads – Avoid (high)** (paved roads were strongly preferred)
- **Driving an Automobile – Prohibited** (roads were automobiles are prohibited were excluded)
- **Roads under Construction Prohibited – Prohibited** (roads closed due to construction were excluded)
- **Through Traffic Prohibited – Avoid (high)** (routes avoiding roads where through traffic is prohibited were strongly preferred)

accessible by air and which is small enough—less than a mile across—that driving is unlikely to be a significant means of transportation within the community.

- Tract 412 in Kauaʻi County, HI (FIPS code 15007041200)
  Consists entirely of the privately-owned island of Ni'ihau.
- Tract 9518 in Manu'a District, AS (FIPS code 60020951800)
  It was unclear that any public roads exist within the tract.

For the fourteen tracts that produced especially small motorist travelsheds, ten represented small islands with no bridges and either no ferry connections or ferry connections that would take more than thirty minutes to traverse, and the travelsheds were deemed accurate. For the other four, the small travelsheds appeared to be artifacts due to the point on the ERSI routing network nearest the centroid being on a disconnected section of parking lot (or, in one case, an airport taxiway). In these four cases, the starting location was shifted a short (less than 100 meters) distance to the nearest segment of connected public road.

4.2 Calculating Subindicators

For each of the three destination access vulnerability indicators—the pedestrian access score, the cyclist access score, and the motorist access score—discussed in Section 2.2.1 on page 7, ten subindicators were calculated to measure the accessibility of different destination types within the travelsheds for each transportation mode. The raw subindicators (suffixed with "_R" in the Technical Data Download tracts file), in the form of measures of destinations of each type within the travelshed, were converted to normalized subindicators (suffixed with "_N" in the Technical Data Download tracts file) as follows:

- Values of the indicator for tracts in the *fifty states and District of Columbia*, excluding *remote island tracts*, are percentile ranked and the value of the 1st percentile tract is set equal to a normalized value of 0 while the value of the 99th percentile tract is set equal to a normalized value of 1.
- All tracts (both in the fifty states and District of Columbia and in the *territories*) with values below the 1st percentile or above the 99th percentile are given normalized values of 0 and 1 respectively.
- Tracts (both in the fifty states and District of Columbia and in the *territories*) with values between the 1st percentile and 99th percentile are given normalized values equal to the difference between their value and the 1st percentile value divided by the difference between the 99th percentile and 1st percentile values.

This is the same normalization process used for indicators, as discussed in Section 3.2 on page 21. Unlike with the calculation of indicators, however, no interpolation is needed before normalization can occur, since the subindicators are not missing for any individual tracts. (Not all subindicators, however, are available for Puerto Rico and the *insular areas*, although travelsheds were calculated for all non-remote island tracts in the *territories*.)

It is important to note that the normalized subindicator scores—and thus the raw indicator scores, which are produced by averaging them—are one a scale of 0 (least access) to 1 (most access), so higher scores indicate better access, and thus less disadvantage. The details of the ten sets of subindicators are described below. Each of the ten measures is included only once, because each subindicator is calculated equivalently for pedestrian, cyclist, and motorist access, with the only difference being the travelshed used.

Area Subindicator

The area subindicator is calculated as the area of the travelshed, found with the `st_area` function from the `sf` R package in the appropriate *UTM zone* for the tract.

Education Facilities Subindicator

The educational facilities subindicator is calculated as the total number of public and private primary, secondary, and post-secondary educational institutions within the travelshed. Three distinct data sources were used for these facility types, all from the Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data – Open data repository:

- Post-secondary educational facilities were sourced from the Colleges and Universities Campuses dataset, with coverage as of the end of the 2021-2022 academic year. As this dataset provides polygon geometries for campuses, they were converted to point geometries by calculating the area centroids of the polygons, for compatibility with the other destinations data.
- Public primary and secondary educational facilities were sourced from the Public Schools dataset, with coverage as of the end of the 2022-2023 academic year. This data source includes all public elementary and secondary education facilities included in the National Center for Education Statistics (NCES) Common Core of Data (CCD).
- Private primary and secondary educational facilities were sourced from the Private Schools dataset, with coverage as of the end of the 2021-2022 academic year. This data source includes all private elementary and secondary education facilities included in the National Center for Education Statistics (NCES) Private School Survey (PSS).

Groceries Subindicator

The groceries subindicator is calculated as the total number of grocery stores within the travelshed. Grocery store locations were sourced from the United States Department of Agriculture (USDA) Supplemental Nutrition Assistance Program (SNAP) historical retailer locations dataset. The SNAP retailers data used reflects stores that had not closed (`End.Date == " "`) as of the end of calendar 2023. Only retailers listed as "`Large Grocery Store`", "`Supermarket`", or "`Super Store`" in the `Store.Type` field were included in the calculation of the subindicator to avoid the inclusion of small retailers that are less likely to sell a broad selection of groceries, including fresh meat and vegetables.

Libraries Subindicator

The libraries subindicator is calculated as the total number of public libraries within the travelshed. Library locations were sourced from the Institute of Museum and Library Services (IMLS) Public Libraries Survey (PLS) for fiscal year 2021. Using the `C_OUT_T` field, only central libraries ("`CE`") and branches ("`BR`") were included, excluding bookmobiles and books-by-mail services, since the listed locations for the latter represent their home bases or administrative locations, but not necessarily the areas served.

Medical Facilities Subindicator

The medical facilities subindicator is calculated as the total number of hospitals, outpatient and ambulatory medical facilities, and pharmacies located within the travelshed. Three distinct data sources were used for these facility types:

- The Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data – Open (HIFLD-Open) Hospitals dataset, downloaded on 19 April 2024, was used for hospital locations. This dataset includes the locations of hospital facilities acquired from state departments and Federal sources, as tabulated by DHS.
- A commercial dataset made available through licensing by the Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data – Secure (HIFLD-Secure) data repository—the SafeGraph Retail Trade commercial dataset from May 2024—was used to source pharmacy locations, identified as those with a value of "`Pharmacies and Drug Stores`" in the `sub_category` field.
- A commercial dataset made available through licensing by the Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data – Secure (HIFLD-Secure) data repository—the SafeGraph Outpatient Care and Labs commercial dataset from May 2024—was used to source outpatient and ambulatory medical facility locations, identified as those with a value of "`Medical and Diagnostic Laboratories`" or "`Outpatient Care Centers`" in the `top_category` field, as well as those with "`Ambulatory Health Care Services`" in the `top_category` field and "`All Other Miscellaneous Ambulatory Health Care Services`" in the `sub_category` field.

In addition, because these datasets contain numerous duplicate listings (such as the names of different doctors at the same facility treated as separate points), medical facilities located within 0.001 degrees of each other in latitude and longitude were dropped as duplicates. Between 30°N and 45°N latitude, this corresponds to a distance of roughly 100 meters.

Parks Subindicator

The parks subindicator is calculated as the total number of parks within the travelshed. Because there is no national-in-scope public dataset reporting all park locations in the US, a commercial dataset made available through licensing by the Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data – Secure (HIFLD-Secure) data repository was used. Park locations were taken from the HERE Points of Interest commercial dataset from January 2024, using those locations listed as "parks" and "wildlife refuges".

Post Offices Subindicator

The post offices subindicator is calculated as the total number of post offices within the travelshed. Because the United States Postal Service (USPS) only makes post office locations available for a fee, a commercial dataset made available through licensing by the Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data – Secure (HIFLD-Secure) data repository was used. Post office locations were taken from the "HERE Points of Interest" commercial dataset from January 2024, using those locations listed as "post offices".

Transit Subindicator

The transit subindicator is calculated as the total number of public transit trips serving stops in the travelshed in a typical service week. That is, each trip with stops in the travelshed was counted only once, even if there were stops at multiple locations in the travelshed. A route served by a number of trips in a typical service week would count the number of times trips, on a route, stopped in the service area.

Duplicate stops within a single travelshed are excluded to better represent the amount of transit service reachable in the travelshed. Including duplicate stops would weight service with frequent stops—such as local buses—more heavily than services that provide faster and longer-distance service—such as rapid transit. In addition, inconsistencies in the ways different transit agencies reported stops in the underlying data mean that stop counts are not necessarily comparable between different agencies.

Transit trips and stops for this subindicator were sourced from the General Transit Feed Specification (GTFS) feeds supplied to the Bureau of Transportation Statistics (BTS) by National Transit Database (NTD) reporting agencies service weeks in June 2024. These GTFS feeds are used by BTS to produce the National Transit Map (NTM), but the raw feeds used to generate the transit subindicator are not currently publicly available.

Population Subindicator

The population subindicator is calculated as the total 2020 Decennial Census population in the travelshed. This is estimated using the population *block group* geometries created as discussed in Section 5.4 on page 41 and the `st_intersection` function from the `sf` R package to determine what share of each block group's area is located within the travelshed.

Jobs Subindicator

The jobs subindicator is calculated as the total jobs (reported at the *Census block* level) in the travelshed. This is estimated using job block group geometries produced by a method equivalent to that used to create population block group geometries as discussed in Section 5.4 on page 41; the difference is that here, block group geometries are assembled out of those blocks containing jobs rather than those blocks containing population.

The jobs data used is taken from the Census Bureau's Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics dataset, LODES v8.1. The full technical document for the LODES v8.1 dataset is posted on the Census website. The LODES v8.1 data is tabulated for 2020 Census blocks, and the most recent data available was used for each state. This means 2016 data for Alaska, 2018 data for Arkansas and Mississippi, and 2021 data for the rest of the *United States*. LODES data is not available for the *territories*. The data was downloaded using the `lehdr` R package and was downloaded for workplace area characteristics (`lodes_type = c("wac")`), all jobs (`job_type = c("JT00")`), and total number of jobs in all industries (`segment = c("S000")`).

5. Appendix: Preparing the Geographies

The basic geographic building blocks of the TC Explorer are the 2023 release of the 2020 *Census tracts* established by the Census Bureau. *TIGER/Line shapefiles* for 2023 Census geographies, including Census tracts and *counties* were downloaded as described in Section 6.1 on page 42. *Water tracts*—Census tracts consisting entirely or nearly entirely of water—were dropped from the tract dataset.[19] These county and tract geometries were then assigned to *metropolitan and non-metropolitan regions* and planar coordinate reference systems. Next, tracts were classified into types based on the availability of data for calculating *indicators* and identifying *disadvantaged communities* and were assigned to *urban areas* and identified as *low-population tracts*. Finally, tract centroids and nearest neighbors were identified and modified geometries for use in buffer calculations and raster extractions in the calculation of indicators were produced.

5.1 Classifying Counties

While Census tracts are the primary geographies for the TC Explorer, two classifications were performed on counties and county equivalents and then carried over to the tracts within those counties. First, counties were assigned to *regions* for use in the calculation of the relative household income *indicator* (Section 2.4.3 on page 17). Second, counties were assigned to *Universal Transverse Mercator (UTM) zones* to provide planar coordinate reference systems for spatial calculations.

5.1.1 Assigning Regions to Counties

The calculation of the relative household income indicator requires the median household incomes of Census tracts to be compared to the median household income for the region in which the tract is located. For tracts in 2023 *Metropolitan Statistical Areas (MSAs)*, these regions correspond to the MSA in which the tract is located. For tracts not located in MSAs, the portion of each state not part of any MSA is treated as a region. In the TC Explorer technical data download, the `METRO_FIPS` and `METRO_NAME` variables contain the FIPS codes and names of MSA regions and an assigned code of the form "`NM-`" followed by the two-digit state FIPS code for non-MSA regions. The `METRO_TYPE` field lists each region as either "Metropolitan" or "`Non-Metropolitan`" depending on whether it is an MSA. These fields were first assigned to counties and then carried over to the Census tracts that are part of these counties.

[19] All Census tracts with *FIPS codes* of the form xxxxx99xxxxxx—that is Census tracts with tract number of the form 99xx.xx—were identified as water tracts. In addition, Tract 9801.00 in Hillsborough County, FL (FIPS code 12057980100) is treated as a water tract in the TC Explorer because it contains minimal land area and is excluded from Census Bureau cartographic boundary files.

5.1.2 Assigning Coordinate Reference Systems to Counties

All spatial operations in the processing of data for the TC Explorer were performed in planar *Universal Transverse Mercator* (UTM) coordinate reference systems based on the North American Datum of 1983 (NAD83). UTM zones were identified for each *county* based on the midpoint between the easternmost and westernmost points in the Census Bureau TIGER/Line geometry (Section 6.1 on page 42) for the county and were used for all spatial operations involving the county and tracts in the county.[20] EPSG codes are given in the `CNTY_EPSG` and `TRACT_EPSG` fields for counties and tracts in the TC Explorer technical data download The EPSG codes of the coordinate reference systems used for each UTM zone are given in table 2 below.

Table 2: UTM Zones and EPSG Codes Used

UTM Zone	EPSG Code	Longitudes	States and Territories Covered
1N	26901	174°W – 180°W	Alaska
3N	26903	162°W – 168°W	Alaska
4N	26904	156°W – 162°W	Alaska and Hawaii
5N	26905	150°W – 156°W	Alaska and Hawaii
6N	26906	144°W – 150°W	Alaska
7N	26907	138°W – 144°W	Alaska
8N	26908	132°W – 138°W	Alaska
9N	26909	126°W – 132°W	Alaska
10N	26910	120°W – 126°W	Contiguous US
11N	26911	114°W – 120°W	Contiguous US
12N	26912	108°W – 114°W	Contiguous US
13N	26913	102°W – 108°W	Contiguous US
14N	26914	96°W – 102°W	Contiguous US
15N	26915	90°W – 96°W	Contiguous US
16N	26916	84°W – 90°W	Contiguous US
17N	26917	78°W – 84°W	Contiguous US
18N	26918	72°W – 78°W	Contiguous US
19N	26919	66°W – 72°W	Contiguous US
20N	26920	60°W – 66°W	Puerto Rico and US Virgin Islands
55N	8693	144°E – 150°E	Guam and Northern Mariana Islands
2S	2195	168°W – 174°W	American Samoa

[20] Honolulu County, HI was treated as a special case: since all of its population is located on the island of Oahu but the county also contains a remote island tract made up of Northwestern Hawaiian Islands, an archipelago that extends over 1,000 miles west of the island of Oahu. Instead of using the bounds of the TIGER/Line geometry, the coordinates of the island of Oahu were approximated as 21.5°N, 158°W.

5.2 Classifying Census Tracts

The TC Explorer uses *Census tracts* as its main building blocks for measuring *indicators* and calculating *disadvantage scores*. However, fundamental differences in tracts – what data is available for them, what indicators are meaningful for them, and whether they fundamentally represent residential areas or not—require tracts to be treated differently for some operations. In the following sections, the processes for identifying tracts where not all indicators are available—*remote island tracts* as well as Puerto Rico and the *insular areas*—tracts where *disadvantaged community status* or *disadvantaged community scores* are calculated differently—the *territories* and *isolated Alaska tracts*—and tracts with low populations that necessitate a different approach to spatial processing—*low-population tracts*—are discussed. In addition, the assignment of *urban areas* and urban area size to Census tracts is discussed.

5.2.1 Puerto Rico, Insular Areas, and Remote Island Tracts (`TRACT_LCTN`)

Ideally, all indicators would be calculated for all Census tracts but, in practice, data for calculating many of the indicators is unavailable for the *US territories*, or is available for Puerto Rico but not for the *insular areas* (American Samoa, Guam, the Northern Mariana Islands, and the US Virgin Islands). The `TRACT_LCTN` field for tracts in the TC Explorer Technical Data Download indicates tracts in the *fifty states and District of Columbia* (for which all indicators should be available) with the value "`US`", those in Puerto Rico with the value "`PR`", and those in the insular areas with the value "`IA`". The `CNTY_LCTN` field for counties uses the same codes.

In addition, twelve Census tracts represent islands with no permanent population—meaning that demographic indicators are unavailable—that are sufficiently far from other tracts that indicator values for them cannot reasonably be *nearest-neighbor interpolated* from adjacent tracts. These tracts are identified as *remote island tracts* and no indicator values are supplied for them; they are identified with the value "`RI`" in the `TRACT_LCTN` field, regardless of whether they are located in the fifty states and District of Columbia, Puerto Rico, or the insular areas.

Remote island tracts are excluded from the count of Census tracts in the fifty states and District of Columbia in the percentile rankings used to calculate disadvantage scores. Likewise, three *county-equivalents* in the territories—Rose Island and Swains Island in American Samoa and the Northern Islands Municipality in the Northern Mariana Islands—each consist of a single remote island tract, and so are excluded as *remote island counties* and assigned as "`RI`" in the `CNTY_LCTN` field in the county dataset. A list of the remote island tracts and descriptions of the areas they represent is given below.

List of Remote Island Tracts

- **Tract 9804.01 in San Francisco County, CA (FIPS 06075980401)** – Consists of the Farallon Islands National Wildlife Refuge, located in the Pacific Ocean approximately 30 miles from the mouth of San Francisco Bay.
- **Tract 9801 in Santa Barabara County, CA (FIPS 06083980100)** – Consists of the portion of Channel Islands National Park and the Santa Cruz Island Reserve in Santa Barabara County, CA. The four islands in the tract—San Miguel, Santa Cruz, Santa Rosa, and Santa Barbara—are in the Pacific Ocean 20 miles or more from shore.
- **Tract 9800 in Ventura County, CA (FIPS 06111980000)** – Consists of San Nicholas Island in the California Channel Islands, which is owned by the US Navy and consists of Naval Outlying Landing Field San Nicolas Island.
- **Tract 9802 in Hillsborough County, FL (FIPS 12057980200)** – Consists of Egmont Key National Wildlife Refuge and State Park, at the mouth of Tampa Bay approximately two miles from other barrier islands in Hillsborough and Manatee Counties, FL.
- **Tract 9801 in Monroe County, FL (FIPS 12087980100)** – Consists of Dry Tortugas National Park, located in the Gulf of Mexico approximately 70 miles west of Key West.
- **Tract 9812 in Honolulu County, HI (FIPS 15003981200)** -- Consists of the uninhabited Northwest Hawaiian Islands, which make up Papahānaumokuākea Marine National Monument. These islands are located in the Pacific Ocean over 100 miles northwest of the inhabited portion of Hawaii. Midway Atoll, which is not part of the state of Hawaii, is not part of this tract and, as a United States Minor Outlying Island, is not included in any Census tract.
- **Tract 9800 in Maui County, HI (FIPS 15009980000)** – Consists of the uninhabited islands of Kahoʻolawe (Kahoʻolawe Island Reserve) and Molokini (Molokini Seabird Sanctuary) in Maui County, Hawaii. Kahoʻolawe is located in the Pacific Ocean approximately 7 miles southwest of the island of Maui while Molokini is located in the Pacific Ocean approximately 2 miles southwest of the island of Maui.
- **Tract 9801 in Keweenaw County, MI (FIPS 26083980100)** – Consists of Isle Royale National Park, located in Lake Superior approximately 15 miles from the nearest shoreline, which is part of the western portion of the Canadian Province of Ontario.
- **Tract 9519 in Rose Island, AS (FIPS 60030951900)** – Consists of Rose Atoll National Wildlife Refuge and Rose Atoll Marine National Monument, approximately 170 miles east of Tutuila, the main island of American Samoa. Also a remote island county.
- **Tract 9520 in Swains Island, AS (FIPS 60040952000)** – Consists of the uninhabited island of Swains Island, administered as part of American Samoa, and located approximately 110 miles south of the nearest other land, the island of Fakaofo in the New Zealand territory of Tokelau. Also a remote island county.
- **Tract 9501.02 in Tinian Municipality, MP (FIPS 69120950102)** – Consists of the uninhabited island of Aguijan, located approximately 5 miles from the island of Tinian.
- **Tract 9501 in the Northern Islands Municipality, MP (FIPS 69085950100)** – Consists of the uninhabited Northern Islands Municipality of the Northern Mariana Islands, all at least 50 miles north of the northernmost permanently-inhabited island. Also a remote island county.

5.2.2 US Territories and Isolated Alaska Tracts (**TRACT_DAC**)

In general, Census tracts are identified as experiencing overall disadvantage or a component or subcomponent of disadvantage if the associated disadvantage score has a value above the 65th percentile. However, all tracts in the *territories* are automatically assigned disadvantaged status regardless of their scores, and disadvantage scores and disadvantage status are not calculated at all for *remote island tracts*. While the 65th percentile rule for identifying disadvantaged tracts does apply to *isolated Alaska tracts*, discussed below, the *transportation cost burden* and *traffic fatality burden* subcomponent scores for these tracts are automatically set at 100th percentile.

The `TRACT_DAC` field in the TC Explorer Technical Data Download for Census tracts has the values "`Typical`" for tracts where disadvantage scores and status are calculated normally, "`Isolated Alaska`" for isolated Alaska tracts, "`US Territory`" for tracts in the territories, and "`Not Calculated`" for remote island tracts.

Isolated Alaska Tracts

The indicators included in the TC Explorer were chosen to effectively measure the components and subcomponents of disadvantage across the United States. However, the wide variety of social and physical conditions across the country mean that no indicator will be equally useful everywhere.

A particularly extreme example of this issue is found in remote areas of Alaska where road travel to even small urban clusters is impossible much or all of the year. In these areas, residents need to rely on aviation—often using small aircraft and landing fields with limited improvements—for purposes where road travel would be used elsewhere in the country, including medical care (both routine and emergency), access to government offices, resupply of groceries and other consumables, and so on.

The indicators used to measure transportation insecurity do not provide useful information in these areas: the traffic fatality burden subcomponent is based on a FARS reports of roadway accidents, but does not include the safety risk associated with frequent flights in small aircraft, while the transportation cost burden subcomponent excludes air travel, which plays a uniquely important role in these areas.

To provide a more accurate understanding of the transportation disadvantage experienced by residents of these remote areas in Alaska, we defined as "isolated Alaska tracts" those *Census tracts* in Alaska where the majority of the population lacked year-round road access to one of the six Census-defined urban areas or urban clusters in Alaska with at least 10,000 residents as of the 2020 Decennial Census: Anchorage, Anchorage Northeast, Fairbanks, Wasilla/Knik-Fairview/North Lakes, Juneau, and Ketchikan.[21] Road access was identified using a road network geometry provided by the Alaska Department of Transportation and excluding roads labeled as "priority level 5" / "No Winter Maintenance." These tracts—listed below–were deemed to experience transportation cost burden and traffic fatality burden at the 100th percentile level.

---

[21] Four of these six urban areas or clusters—Anchorage, Anchorage Northeast, Fairbanks, and Wasilla/Knik-Fairview/North Lakes are connected to each other—and to Kenai and Soldotna, which together total over 10,000 residents though neither has 10,000 residents on its own—by year-round roads. Juneau and Ketchikan, in Southeastern Alaska, lack year-round road connections to each other and the rest of the state.

List of Isolated Alaska Tracts and Boroughs/Census Areas

The isolated Alaska tracts identified by the methodology discussed above encompass 41 (23%) of the 177 Census tracts in Alaska and 118,585 (16%) of 733,391 the residents of Alaska reported by the 2020 Decennial Census. All Census tracts in the following *county-equivalents* (boroughs and Census areas) in Alaska are designated as isolated Alaska tracts:

- Aleutians East Borough, AK (FIPS 02013)
- Aleutians West Census Area, AK (FIPS 02016)
- Bethel Census Area, AK (FIPS 02050)
- Bristol Bay Borough, AK (FIPS 02060)
- Dillingham Census Area, AK (FIPS 02070)
- Haines Borough, AK (FIPS 02100)
- Hoonah-Angoon Census Area, AK (FIPS 02105)
- Kodiak Island Borough, AK (FIPS 02150)
- Kusilvak Census Area, AK (FIPS 02158)
- Lake and Peninsula Borough, AK (FIPS 02164)
- Nome Census Area, AK (FIPS 02180)
- North Slope Borough, AK (FIPS 02185)
- Northwest Arctic Borough, AK (FIPS 02188)
- Petersburg Borough, AK (FIPS 02195)
- Prince of Wales-Hyder Census Area, AK (FIPS 02198)
- Sitka City and Borough, AK (FIPS 02220)
- Skagway Municipality, AK (FIPS 02230)
- Wrangell City and Borough, AK (FIPS 02275)
- Yakutat City and Borough, AK (FIPS 02282)

In addition, six tracts in other county-equivalents in Alaska are designated as isolated Alaska tracts:

- Census Tract 2 in Chugach Census Area, AK (FIPS 02063000200)
- Census Tract 1 in Kenai Peninsula Borough, AK (FIPS 02122000100)
- Census Tract 12 in Kenai Peninsula Borough, AK (FIPS 02122001200)
- Census Tract 1 in Yukon-Koyukuk Census Area, AK (FIPS 02290000100)
- Census Tract 3 in Yukon-Koyukuk Census Area, AK (FIPS 02290000300)
- Census Tract 4 in Yukon-Koyukuk Census Area, AK (FIPS 02290000400)

5.2.3 Identifying Low-Population Census Tracts

In defining *Census tracts*, the Census Bureau generally targets a population size between 1,200 and 8,000 residents, with an optimum size of 4,000 residents. These tract sizes are consistent regardless of population density: in rural areas, tracts cover larger areas to encompass similar populations. However, some Census tracts fall substantially outside this range: in particular, tracts designated to contain non-residential land uses—for example parkland, wilderness areas, industrial and office parks, shopping malls, and non-residential portions of colleges and military bases—often have minimal or no population.

Tracts with minimal population pose a challenge for spatial calculations that involve the population distribution of tracts: calculation of population centers of tracts (Section 5.3 on page 40), identification of what urban area the majority of a Census tract's population falls within (Section 5.2.4 below), and calculations of what share of a tract's population falls within the buffer around another feature (used in calculation of some *indicators* as discussed in Section 2 on page 6) For tracts with no residents, these measures are undefined. For tracts with very small but non-zero populations, they cannot be reliably calculated because the Census Bureau's implementation of *differential privacy* in the 2020 Decennial Census results means that the population counts and spatial distribution of population in tracts with very low populations are essentially noise. Furthermore, since Census tracts with populations substantially below the 1,200-resident lower limit targeted for ordinary tract populations are generally drawn to contain non-residential land uses, including employment clusters, population distribution may not accurately reflect the portions of the tract where the most people spend the most time.

To avoid these issues, tracts with 2020 Decennial Census populations of fewer than 100 residents have been identified as *low-population tracts* and their land area centroids and land area shares within buffers are substituted for population centroids and population shares within buffers when these measures are used for typical tracts. The `TRACT_INTP` field in the TC Explorer Technical Data Download tract file codes tracts with at least 100 residents as "`Typical`", those with fewer than 100 residents as "`Low-Population`", and *remote island tracts* as "No Data".

5.2.4 Assigning Tracts to Urbanized Areas

Assigning Census tracts to Census *urban areas* requires accounting for the fact that urban areas are defined in terms of *Census blocks*, and so a single Census tract can contain portions in multiple urban areas and in no urban area. Tracts are first identified as urban or rural depending on whether the majority of their population (or land area for *low-population tracts*) is located in an urban area of at least 50,000 residents. Urban tracts were then assigned to the urban area with the largest share of their population (or land area for low-population tracts) and assigned a size based on the population of that urban area ("small urban" under 200,000 residents and "large urban" if at least 200,000 residents). In the TC Explorer Technical Data Download, the `URBAN_FIPS` field contains the FIPS code of the urban area (empty for rural tracts), the `URBAN_NAME` field contains the name of the urban area (empty for rural tracts), the `URBAN_SIZE` field has values of "`Rural`", "`Small Urban`", "`Large Urban`", or "`Remote Island`", and the `URBAN_TMA` field has values of "`TMA`" or "`Non-TMA`" depending on whether the tract is in an urban area designated as a *transportation management area (TMA)*.

5.3 Identifying Census Tract Centroids and Nearest Neighbors

The centroids of *Census tracts* are used in the TC Explorer for two purposes: the calculation of *travelsheds* for use in the calculation of *destination access vulnerability* *indicators* and the identification of nearest neighbors for use in the *nearest-neighbor interpolation* procedure used for tracts missing raw values of indicators (Section 3.1 on page 21). In fact, two sets of Census tract centroids are needed: population centroids for travelshed calculations and interpolation of *transportation insecurity* and *population-based vulnerability* indicators land area centroids for interpolation of *place-based burden* indicators.

Both sets of centroids are calculated using the *2023 TIGER/Line shapefiles* for *Census blocks*. Since no information is available about the internal distribution of land area and population within blocks, the centroids of each block was calculated using the `st_centroid` function from the `sf` R package in the appropriate *UTM zone*. The coordinates for all of the blocks in each tract were then extracted and population-weighted and area-weighted centroids of the set of blocks for each tract were calculated using land area and 2020 Decennial Census populations of tracts as weights. Finally, since population centroids cannot be reliably calculated for *low-population tracts*, the land area centroids of these tracts were substituted as population centroids for these tracts as well.

Nearest-neighbors for Census tracts were calculated twice: once with land area centroids and once with population centroids. To simplify the identification of nearest neighbors, the `poly2nb` function from the `spdep` R package was used to identify the states bordering each state. (The territories, Alaska, and Hawaii were considered to have no neighbors.) Distances to all tract centroids from each tract centroid was calculated with the `st_distance` function in the R `sf` package[22] and an ordered list of the nearest 16 (including self) were recorded. These lists can be found in the TC Explorer Technical Data Download.

[22] Contrary to usual practice in spatial calculations for the TC Explorer, nearest neighbors were identified in the North American Datum 83 (NAD83) coordinate system, not planar coordinate systems. This was done for reasons of computational tractability and because exact distances were not needed, only a list of nearest tracts.

5.4 Preparing Geometries for Use in Spatial Calculations

A number of the *place-based burden indicators* (Section 2.3 on page 11) are either derived from raster data, which needs to be tabulated over the land area portions of *Census tracts* or are based on the share of a tract's population that falls within a half-mile or one-mile buffer of another feature. To aid in these calculations, two special variant geometry sets were developed.

For land area raster extraction, land-only Census tract geometries were created by performing a spatial union on all *Census blocks* within a tract with non-zero land area. While the resulting tracts may have non-zero water area, since they include blocks that are entirely land and blocks that contain both land and water, they remove the vast majority of the water area contained in Census tracts, which can be the majority of the total area of the tracts for tracts on coastlines. In addition to their use in raster extraction, they were used to calculate the share of land area within buffers for *low-population tracts*, for which calculations of population within buffers was not viable.

For population buffer calculations, however, an additional step was used, because population density is not uniform within Census tracts. First, all blocks with no population were dropped. A spatial union was then performed on remaining blocks to aggregate them into *Census block groups* which were treated as having uniform population density. Since there are roughly three times as many Census block groups as Census tracts, this approach provides a better approximation of the distribution of population within Census tracts.

6. Appendix: Data Sources

A number of data sources were used in the development of TC Explorer. This appendix lists, by type of disadvantage measured, all datasets used in the production of TC Explorer, along with details on data vintage, format, and processing.

6.1 Census Geography Datasets

The basic geographic building blocks of the TC Explorer are the 2023 release of the 2020 *Census tracts* established by the Census Bureau. These Census tract geometries, along with supplemental Census geometries and data at the *Census block*, *block group*, *county equivalent*, *metropolitan statistical area* (MSA), *urban area* (UA), and *state* or *territory* levels, were downloaded in their 2023 vintages as TIGER/Line shapefiles as discussed below, and crosswalk files were used to handle variations in *FIPS codes* and the importation of data tabulated for 2010 Census tracts.

2023 TIGER/Line Shapefiles

Geometry objects for Census geographies were downloaded as 2023 *TIGER/Line shapefiles* using the R package `tigris` in both high-resolution (the default) and *cartographic boundary file* (1:500,000 resolution) forms. The high-resolution geometries were used for all spatial data processing, while the cartographic boundary file geometries were used for the display geometries in the online tool and technical data download and, for some geographies, as a source of human-readable names. In addition, land area, 2020 Decennial Census population and *housing unit* counts were taken from the Census block TIGER/Line shapefiles and summed to produce values for the other Census geographies as needed.[23] 2023 TIGER/Line shapefiles were downloaded for the following geographies:

- State or Territory
- County Equivalent
- Census Tract
- Block Group
- Census Block (only high-resolution geometries available)
- 2020 Census Urbanized Area (only high-resolution geometries available)
- 2023 Metropolitan Statistical Areas

To maintain the consistency of the state-county-tract-block group-block FIPS code hierarchy, the FIPS codes of Census blocks in Connecticut were updated to match the FIPS codes of the councils of government used as county equivalents in Connecticut from 2022 forward, as established using the 2020 to 2022 Connecticut Census Tract Crosswalk discussed below.[24]

---

[23] It is worth noting that 2020 Decennial Census housing unit counts are exact, while population counts are not, due to the implementation of *differential privacy* by the Census Bureau.

[24] While Census tract and block group FIPS codes are changed when necessary to correspond with updated county equivalent FIPS codes, Census block FIPS codes are invariant between Decennial Census years. Changing Census block FIPS codes to match the values they "would have had" if the 2022-onward county definitions were in force in 2020 allowed computational simplicity.

Census Bureau 2020 Census Tract to 2010 Census Tract Relationship File

The Census Bureau 2020 Census Tract to 2010 Census Tract relationship file, which provides land area intersections between 2010 Census tracts (using 2010 FIPS codes) and 2020 Census tracts (using 2020 FIPS codes) was used for the area-weighted interpolation transfer of CMRA data from 2010 to 2020 Census tracts, as discussed in Section 6.4.1 on page 50.

A complication with using this dataset is that, while Census tracts are generally updated only in Decennial Census years, the FIPS codes that identify Census tracts change more often, usually due to changes in county boundaries.[25] While the datasets used in the TC Explorer that are tabulated for 2010 Census tracts are coded with 2019 Census tract FIPS codes, the Census Bureau 2020 Census tract to 2010 Census tract relationship file (and the NHGIS 2010-2020 Crosswalk discussed below) identifies Census tracts using 2010 Census tract FIPS codes. For block groups where the FIPS code changed between 2010 and 2019, the FIPS codes in the crosswalk were updated individually using an R script; a list of FIPS codes that were changed can be found on page 44.

NHGIS 2010-2020 Crosswalk

Because the Census Bureau 2020 Census tract to 2010 Census tract relationship file discussed in the previous section only provides land area intersections between 2010 and 2020 tracts, it cannot be used for the population-weighted interpolation transfer of demographic data, such as the 2022 PLACES data. Instead, the 2010 Block Groups to 2020 Census Tracts Crosswalk[26] published by the IPUMS National Historic Geographic Information System project of the University of Minnesota, was used for the population-weighted interpolation transfer of 2022 PLACES data from 2010 to 2020 Census tracts, as discussed in Section 6.5 on page 56. This crosswalk estimates the population intersection between 2010 and 2020 Census geographies using a methodology based on population counts at the block level.

As with the Census Bureau 2020 Census Tract to 2020 Census Tract Relationship File discussed above, the difference between 2010 and 2019 FIPS codes poses a complication for using this crosswalk. While the PLACES dataset identifies Census tracts using 2019 FIPS codes, this crosswalk identifies tracts using 2010 FIPS codes. For block groups where the FIPS code changed between 2010 and 2019, the FIPS codes in the crosswalk were updated individually using an R script; a list of FIPS codes that were changed can be found in on page 44.

---

[25] Census tract FIPS codes have the FIPS code of the county they are located in as their first five digits; if the county or county name changes, its FIPS code and the FIPS codes of all tracts in the county changes as well. In addition, several Census tracts in California, New York, and Arizona changed FIPS codes between 2010 and 2012 due to corrections made by the Census Bureau to the initial assignment of tracts and FIPS codes.

[26] Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. 2021. http://doi.org/10.18128/D050.V16.0

List of 2010 Census Tracts With Different FIPS Codes in 2010 and 2019

A number of 2010 *Census tracts* had changes to their *FIPS codes* between their initial definition in 2010 and their final edition in 2019. The following corrections were made to the FIPS codes of 2010 Census tracts in the Census Bureau and NHGIS 2010-2020 crosswalks used to convert data tabulated for 2010 Census tracts to 2020 Census tracts.

- **Renaming of Wade-Hampton Census Area, AK to Kusilvak Census Area, AK**
  - `02270xxxxxx` changed to `02158xxxxxx`
- **Correction of erroneous FIPS codes for seven tracts in Pima County, AZ**
  - `04019002701` changed to `04019002704`
  - `04019002903` changed to `04019002906`
  - `04019410501` changed to `04019004118`
  - `04019410502` changed to `04019004121`
  - `04019410503` changed to `04019004125`
  - `04019470400` changed to `04019005200`
  - `04019470500` changed to `04019005300`
- **Correction of erroneous FIPS code for one tract in Los Angeles County, CA**
  - `06037930401` changed to `06037137000`
- **Correction of erroneous FIPS codes for nine tracts in Madison County, NY**
  - `36053940101` changed to `36053030101`
  - `36053940102` changed to `36053030102`
  - `36053940103` changed to `36053030103`
  - `36053940200` changed to `36053030200`
  - `36053940300` changed to `36053030300`
  - `36053940401` changed to `36053030401`
  - `36053940403` changed to `36053030403`
  - `36053940600` changed to `36053030600`
  - `36053940700` changed to `36053030402`
- **Correction of erroneous FIPS codes for three tracts in Oneida County, NY**
  - `36065940000` changed to `36065024800`
  - `36065940100` changed to `36065024700`
  - `36065940200` changed to `36065024900`
- **Deletion an erroneous all-water tract in Richmond County, NY**
  - `36085008900` deleted
- **Renaming of Shannon County, SD to Oglala Lakota, SD**
  - `46113xxxxxx` changed to `46102xxxxxx`
- **Retrocession of Bedford City, VA into Bedford County, VA**
  - `51515xxxxxx` changed to `51019xxxxxx`

2020 to 2022 Connecticut Census Tract Crosswalk

As documented in Federal Register Notice 87 FR 34235, starting with the 2022 data year, the Census Bureau switched from using the historical counties of Connecticut—abolished as governmental and administrative entities in 1960—to using the state's current councils of government as county equivalents. This resulted in changes to the FIPS codes of every Census tract in Connecticut, since the first five digits of a Census tract FIPS code are the FIPS code of the county in which it is located. Since the TC Explorer is based on 2022 Tiger/Line geometries (Section 6.1 on 42), 2022 FIPS codes are used for Census tracts. However, 2020 Census tract FIPS codes are also assigned to tracts to facilitate the joining of data tabulated for 2020 tracts using the original 2020 FIPS codes.

Because the Census Bureau did not provide a crosswalk between the 2020 and 2022 FIPS codes for Census tracts, the 2020 to 2022 Connecticut Census Tract Crosswalk published by the CTData Collaborative, a non-profit organization that provides data for non-profit and government organizations in the state of Connecticut, was used. This crosswalk provides 2020 and 2022 FIPS codes for every Census tract in Connecticut, except for the *water tracts* in Long Island Sound, which changed their boundaries in 2022 and so cannot be exactly converted.

6.2 American Community Survey / Puerto Rico Survey

The Census Bureau's American Community Survey (ACS) 5-year estimates and parallel Puerto Rico Survey (PRS) 5-year estimates, both produced annually, are the largest and most detailed *Census tract*-level datasets covering the *United States* and Puerto Rico produced by the Federal government. (Equivalent data is not available for the *insular areas*.) The following discussion is based on the ACS, but the details of the PRS are similar.

While both ACS 1-year and 5-year estimates are published by the Census Bureau, 1-year estimates are only published for geographies with populations of at least 65,000 residents, and so are not available for Census tracts. The 5-year estimates published each year are based on surveys conducted in the five previous years but intended to be used as estimates of the characteristics of each geography in the final year. That is, the 2019-2023 ACS 5-year estimates are based on survey data from 2019 through 2023 but are intended to be used as estimates of the actual characteristics of the geography in 2023.

The ACS is based on two samples: one of *households* and one of *group quarters* (both non-custodial, e.g. college dorms and military barracks, and custodial, e.g. prisons and long-term medical care facilities). The ACS households sample has, in recent years, involved roughly 2 million interviews a year, and thus 10 million interviews over the 5-year period used to create 5-year estimates. Since the United States population is approximately 330 million, this means that roughly 1 in 33 households are represented in the 5-year estimates, although this varies substantially as some areas are sampled more heavily (especially low-population municipalities, school districts, counties, and so on) and others less heavily. It is important to keep this in mind in understanding the uncertainty of tract-level indicators based on the ACS, or on other datasets that use ACS data as an input.

All data from the ACS and PRS used in the TC Explorer was downloaded through the Census Bureau APIs using the `tidycensus` package in R. Except where otherwise noted, ACS/PRS data is from the 2019-2023 ACS/PRS 5-year estimates tabulated for 2020 Census tracts. However, there are two exceptions to this:

- The relative household income indicator (page 18) requires the calculation of median household income for *regions* not found in the 2019-2023 ACS 5-year estimates: the *Metropolitan Statistical Areas (MSAs)* defined in 2023 and the portions of states not in any MSA. Values for these were estimated based on ACS/PRS bucketed household income data tabulated at the *county* level.
- The *population-weighted interpolation* of demographic data between 2010 block groups and 2020 tracts discussed in Section 6.5 on page 56 required the use of 2015-2019 ACS 5-year estimates tabulated at the block group level.

Descriptions of ACS variables used can be found on the Census Bureau website by type: general variables (B/C), subject variables (S), and demographic profile variables (DP). Note that ACS variables with names beginning "`DP02`" have names beginning "`DP02PR`" for the PRS; all other ACS and PRS variables used in the TC Explorer have the same names.

6.3 Sources for Transportation Insecurity Indicators

The data sources used for population-based vulnerability indicators are listed in Table 3 and discussed in depth below.

Table 3: Data Sources for Transportation Insecurity Indicators

Subcomponent	Indicators	Data Source
Destination Access Vulnerability	Pedestrian Access Score	See Section 4 on page 26 for details.
	Cyclist Access Score	See Section 4 on page 26 for details.
	Motorist Access Score	See Section 4 on page 26 for details.
Vehicle Access Vulnerability	Children, Elderly, and Disabled	ACS 2019-2023 5-year estimates
	Vehicles Per Adult	ACS 2019-2023 5-year estimates
	Households Without Vehicles	ACS 2019-2023 5-year estimates
Transportation Cost Burden	Transportation Cost Burden	TIAT 2021 estimates
	Housing Cost Burden	ACS 2017-2021 5-year estimates
Traffic Fatality Burden	Traffic Fatalities—Buffered	NHTSA FARS 2018-2022 data
	Traffic Fatalities—Not Buffered	NHTSA FARS 2018-2022 data

Three of the four sets of *transportation insecurity indicators* are each sourced from a single dataset, the exception being the calculation of the *destination access vulnerability* indicators (Section 2.2.1 on page 7), which are in the form of pedestrian access, cyclist access, and motorist access scores calculated based on a number of subindicators described in detail in Section 4 on page 26. The calculation of the access scores used as destination access vulnerability measures is described in detail in Section 4 on page 26, including the details of the data sources used.

The *vehicle access vulnerability* indicators (Section 2.2.2 on page 8) are sourced from the 2019-2023 American Community Survey (ACS) / Puerto Rico Survey (PRS) 5-year estimates.

The *traffic fatality burden* indicator (Section 2.2.4 on page 10) is sourced from the National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS) datasets, released annually, which contain data on and coordinates for every traffic fatality on public roads in the United States and Puerto Rico. Finalized FARS datasets for 2018, 2019, 2020, and 2021 and the provisional FARS dataset for 2022 were used in the calculation of the indicator.

Transportation Cost Burden Estimates from the Transportation Insecurity Analysis Tool (TIAT)

The *transportation cost burden* indicators (Section 2.2.3 on page 9) are sourced from the profile 1 (representing all households) data for 2021 from the BTS Transportation Insecurity Analysis Tool (TIAT) prototype. The housing costs and median household income data in the TIAT come from 2017-2021 ACS 5-year estimates at the Census tract level, while the transportation costs are modeled based on data from the ACS, the Federal Highway Administration (FHWA) National Household Travel Surveys (NHTS) from 2017 and 2022, and regional/statewide household travel surveys conducted by the (Minneapolis-St. Paul) Metropolitan Council, the Puget Sound Regional Council, the Ohio Department of Transportation, the Utah Department of Transportation, the Spokane Regional Transportation Commission, and the Community Planning Association of Southwest Idaho.

The household transportation costs model used in the TIAT data calculates total transportation cost as the sum of four components:

- **Automotive Ownership Cost –** An estimate of depreciation costs, finance charges, vehicle insurance, property tax on vehicles, and registration fees, based on the rate of automobile ownership for households in the Census tract.
- **Automotive Operating Cost –** An estimate of fuel spending and maintenance/repair costs, based on the average vehicle miles travelled (VMT) for households in the Census tract.
- **Regional Public Transit Cost –** A cost estimate based on vehicle miles travelled and cost per mile of travel.
- **Taxi & Ride-Hailing Service Costs –** A cost estimate based on vehicle miles travelled and cost per mile of travel.

More details on the TIAT model can be found in the TIAT Technical Documentation.

6.4 Sources for Place-Based Burden Indicators

The data sources used for *place-based burden* indicators are listed in Table 4 and discussed in depth separately in their own subsections:

- The sources for the *extreme weather hazard* indicators are detailed in Section 6.4.1 on page 50.
- The sources for the *infrastructure proximity* indicators are detailed in Section 6.4.2 on page 51.
- The sources for the *air pollution burden* indicators are detailed in Section 6.4.3 on page 54.
- The sources for the *surface pollution burden* indicators are detailed in Section 6.4.4 on page 54.

Table 4: Data Sources for Place-Based Burden Indicators

<table>
<tr><th>Subcomponent</th><th>Indicators</th><th>Data Source</th></tr>
<tr><td rowspan="7">Extreme Weather Hazard</td><td>Extreme Heat</td><td>2023 NOAA CMRA (estimates for 2050)</td></tr>
<tr><td>Extreme Precipitation</td><td>2023 NOAA CMRA (estimates for 2050)</td></tr>
<tr><td>Freeze-Thaw Cycles</td><td>2023 NOAA CMRA (estimates for 2050)</td></tr>
<tr><td>Drought</td><td>2023 NOAA CMRA (estimates for 2050)</td></tr>
<tr><td>Impervious Surface</td><td>USGS NLCD (2021) for contiguous US
NOAA C-CAP (2020-2023) for AK/HI/territories</td></tr>
<tr><td>Wildfire Risk</td><td>USFS Probabilistic Wildfire Risk (2022 data)</td></tr>
<tr><td>Flood Inundation</td><td>WRI Aqueduct v2 (estimates for 2050)</td></tr>
<tr><td rowspan="5">Infrastructure Proximity</td><td>Railroad Proximity
(½- & 1- mile)</td><td>BTS/FRA North American Rail Network (2024)</td></tr>
<tr><td>Freeway Proximity
(½- & 1- mile)</td><td>BTS/FHWA Highway Performance Monitoring System (2020)</td></tr>
<tr><td>High-Traffic Road Proximity
(½- & 1- mile)</td><td>BTS/FHWA Highway Performance Monitoring System (2020)</td></tr>
<tr><td>Major Airport Proximity
(½- & 1- mile)</td><td>BTS/FAA Aviation Facilities (2024)</td></tr>
<tr><td>Port Proximity
(½- & 1- mile)</td><td>US Army Corps of Engineers
Master Docks File (2024)</td></tr>
<tr><td rowspan="6">Air Pollution Burden</td><td>Diesel Particulates Concentration</td><td>2024 EPA EJScreen v2.3 (except AK)
2023 EPA EJScreen v2.2 (AK)</td></tr>
<tr><td>Nitrogen Dioxide Concentration</td><td>2024 EPA EJScreen v2.3</td></tr>
<tr><td>Air Toxics Cancer Risk</td><td>2023 EPA EJScreen v2.2</td></tr>
<tr><td>Air Toxics Respiratory Risk</td><td>2023 EPA EJScreen v2.2</td></tr>
<tr><td>Air Toxics Concentration</td><td>2024 EPA EJScreen v2.3</td></tr>
<tr><td>2.5-Micron Particulates Concentration</td><td>2024 EPA EJScreen v2.3</td></tr>
<tr><td rowspan="6">Surface Pollution Burden</td><td>Hazardous Waste Biennial Reporter Proximity</td><td>EPA Facilities Registry System (2024)</td></tr>
<tr><td>Toxic Release Inventory Site Proximity</td><td>EPA Facilities Registry System (2024)</td></tr>
<tr><td>Risk Management Plan Site Proximity</td><td>EPA Facilities Registry System (2024)</td></tr>
<tr><td>Hazardous Waste Site Proximity</td><td>EPA Facilities Registry System (2024)</td></tr>
<tr><td>Leaking Underground Storage Tanks</td><td>2024 EPA EJScreen v2.3</td></tr>
<tr><td>Active Mine Proximity</td><td>Depart. of Labor MDRS (2024)</td></tr>
</table>

6.4.1 Extreme Weather Hazard Datasets

The extreme weather hazard indicators (Section 2.3.1 on page 11) were sourced from five different datasets. Four of the indicators—extreme heat (`WTH_XHT`), extreme precipitation (`WTH_XPR`), Freeze-Thaw Cycles (`WTH_CLD`), and drought (`WTH_DRO`)—were taken from the National Oceanographic and Atmospheric Administration (NOAA) Climate Mapping for Resilience & Adaptation (CMRA) dataset, an annually-updated set of tract-level predictions of extreme weather hazard conditions. The other three indicators—impervious surface (`WTH_IMP`), wildfire risk (`WTH_WFR`), and flood inundation (`WTH_FLD`)—were derived from raster datasets produced by NOAA, the US Geologic Survey (USGS), the US Forest Service (USFS), and a non-profit organization, the World Resources Institute (WRI).

NOAA Climate Mapping for Resilience & Adaptation (CMRA)

The NOAA Climate Mapping for Resilience & Adaptation (CMRA) dataset, updated annually, provides tract-level estimates of extreme weather hazard outcomes based on several choices of climate models. Four of the extreme weather hazard indicators—extreme heat (`WTH_XHT`), extreme precipitation (`WTH_XPR`), Freeze-Thaw Cycles (`WTH_CLD`), and drought (`WTH_DRO`)—are sourced from the mid-century predictions based on the RCP 8.5 model from the 2023 version (version 2) of the CMRA data.

For two of these variables, it was necessary to interpolate for missing data. Certain tracts in the Anchorage, AK area had values of `NA` for the extreme heat indicator; as they were surrounded by tracts with values of `0`, the `NA`s were replaced with `0`s. Similarly, certain tracts in the Florida Keys had values of `NA` for the freeze-thaw cycles indicator; as they were surrounded by tracts with values of `0`, the `NA`s were replaced with `0`s.

As an additional complication, the CMRA dataset is tabulated for 2010 Census tracts. Since the extreme weather hazard indicators represent physical characteristics of tracts, they were converted using *area-weighted interpolation* using the Census Bureau 2020 Census Tract to 2010 Census Tract Relationship File (Section 6.1 on page 43). The four CMRA indicators, which are intensive with respect to area were joined to the tract relationship file and values for 2020 tracts were calculated as weighted means of the values for 2010 tracts, weighted by the `AREALAND_PART` field from the crosswalk, indicating the land area intersections between the 2010 and 2020 tracts.

USFS Spatial Datasets of Probabilistic Wildfire Risk Components

The United States Forest Service (USFS) Spatial Datasets of Probabilistic Wildfire Risk Components for the United States (270m), 3rd Edition[27] is a set of 270-m pixel rasters of the United States providing data on present-day wildfire burn probabilities. Average burn probability values (from the BP rasters) were averaged within the land portion of each Census tract to estimate the tract's average wildfire burn probability for the wildfire risk (`WTH_WFR`) indicator.

World Resources Institute (WRI) Aqueduct Flood Risk Raster

Flooding—both coastal and riverine—is an important contributor to extreme weather hazard because it can damage or destroy transportation infrastructure, including bridges and roads, often at the times they are most needed for emergency evacuations. As no Federal government dataset provided estimates of

[27] Dillon, Gregory K.; Scott, Joe H.; Jaffe, Melissa R.; Olszewski, Julia H.; Vogler, Kevin C.; Finney, Mark A.; Short, Karen C.; Riley, Karin L.; Grenfell, Isaac C.; Jolly, W. Matthew; Brittain, Stuart. 2023. Spatial datasets of probabilistic wildfire risk components for the United States (270m). 3rd Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2016-0034-3

flood inundation based on climate models that took into account both coastal and riverine flooding, the Aqueduct Flood Risk Rasters version 2 (October 2020) from a well-respected, non-partisan non-profit organization, the World Resources Institute (WRI), was used to quantify flooding.

The Aqueduct flood risk rasters represent estimated flood inundation depths in 30 arc-second pixels. The mid-century predictions for 100-year floods based on the RCP 8.5 climate model were used for the flooding (`WTH_FLD`) indicator in TC Explorer. To represent the potential flooding from simultaneous 100-year coastal and riverine flooding[28], `NA` values were replaced with `0`s in both the coastal and the riverine flooding rasters and the rasters were summed to produce estimates of total inundation depth. The share of the land area of each Census tract with an inundation depth of at least 0.1 m was then calculated and identified as the fraction of the tract exposed to inundation in a 100-year flood event.

Impervious Surface Rasters

The impervious surface indicator (`WTH_IMP`) is sourced from two different sets of impervious surface rasters. Data for the contiguous United States is sourced from the United States Geologic Survey (USGS)-led Multi-Resolution Land Characteristics (MLRC) Consortium 2021 National Land Cover Database (NLCD) while data for Alaska, Hawaii, and the US territories is sourced from the National Oceanographic and Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP) High-Resolution Land Cover rasters.

Data for the contiguous United States from the 2021 NLCD is provided as rasters files—one for land cover type and one for percent impervious surface—with 30-m pixels. Areas of open water (raster value `11`) were identified using the land cover raster and excluded from the area considered. The average percent impervious surface from the impervious surface raster was then found for the land area of each Census tract (as identified from TIGER/Line geometries), excluding open water pixels.

The C-CAP data was used for Alaska, Hawaii, and the territories is of several vintages: 2020 for Alaska and the US Virgin Islands, 2021 for Hawaii, 2022 for American Samoa, the Northern Marianas Islands, and Puerto Rico, and 2023 for Guam. This data is provided as impervious surface and water area rasters with 1-m pixels. For computational simplicity and comparability with the NLCD data, the impervious surface rasters were aggregated to 30-m pixels by averaging the impervious surface shares of each 1-m pixel within a 30-m pixel. The water area rasters were also aggregated to 30-m pixels and converted to Boolean water or not water values for comparability as well: since the C-CAP pixels represent share of pixel area that is water, these values were averaged, and the 30-m pixels were identified as water or non-water based on whether they were at least 50% water. The average percent impervious surface from the impervious surface raster was then found for the land area of each Census tract (as identified from TIGER/Line geometries), excluding open water pixels.

6.4.2 Infrastructure Proximity Datasets

The infrastructure proximity indicators (Section 2.3.2 on page 13) were produced using a number of transportation infrastructure datasets published by the Bureau of Transportation Statistics (BTS), the Federal Aviation Administration (FAA), and the US Army Corps of Engineers (ACE), as detailed below. The ½-mile and 1-mile buffered versions of each of these datasets used in the calculation of the indicators

[28] Coastal and riverine flooding are not independent events, since hurricanes and other large storms coming ashore often produce both a storm surge (leading to coastal flooding) and heavy rainfall (leading to riverine flooding) simultaneously.

were produced by dividing the datasets into sections for each *UTM zone* and performing the buffering in the appropriate UTM coordinate reference system before recombining the buffered geometries.

BTS / FRA North American Rail Network (NARN)

The BTS / Federal Railroad Administration (FRA) North American Rail Network (NARN) Lines dataset consists of all rail lines in the United States, Mexico, and Canada. The dataset is updated as needed, and data downloaded on 17th September 2024 was used in TC Explorer. This data was used in calculating the railroad proximity indicators.

To exclude former or unused rail lines and rail ferry routes, rail lines with the following values of the `NET` field were dropped from the NARN dataset:

- "`A`" – indicates abandoned rail lines.
- "`F`" – indicates rail ferry service; this is primarily the Alaska Railroad's ferry connection to the main US rail network.
- "`R`" – indicates abandoned rail lines that have been physically removed.
- "`T`" – indicates a trail on a former or railbanked right-of-way.
- "`X`" – indicates rail lines that are out-of-service.

In addition, rail lines with the value "`R`" in the `PASSNGR` field, indicating rapid transit lines, were dropped. We excluded these lines because they represent rapid transit lines not part of the national rail network and only a relatively small fraction of rapid transit systems are present in the NARN database, so consistency required removing the ones that were present.

BTS / FHWA Highway Performance Monitoring System (HPMS)

The BTS / Federal Highway Administration (FHWA) Highway Performance Monitoring System (HPMS) dataset contains public roads in the United States, along with traffic data in the form of annual average daily traffic (AADT) values. Documentation for the dataset can be found in the FHWA HPMS field manual. Two versions of the US road network were created from this dataset for use in TC Explorer: a freeways-and-expressways network used for the freeway proximity indicators a high-traffic roads network used for the high-traffic road proximity indicators.

Both road networks were initially simplified by dropping non-mainline roads—those with values of "`4`" (ramp), "`5`" (non-mainline), or "`7`" (planned/unbuilt) in the `FACILITY_TYPE` field—as well as minor collector and local roads—those with values of "`6`" (minor collectors) or "`7`" (local) in the `F_SYSTEM` field. The freeways-and-expressways network was then produced by excluding all roads except those with values of "`1`" (Interstate highways) or "`2`" (other freeways and expressways) in the `F_SYSTEM` field, while the high-traffic roads network was produced by excluding those with values greater than or equal to 50,000 in the `AADT` field.

ACE Master Docks File

The ACE Master Docks File downloaded on 30 January 2024 was used to identify the locations of all dock facilities for calculation of the port proximity indicators. As the master docks file contains other navigational features besides docks, it was simplified by the removal of all features that were not either listed as "`Dock`" in the `FACILTY_TYPE` field or listed with the `FACILTY_TYPE` field blank but a value in the `PURPOSE` or `COMMODITIES` fields. In addition, docks used by the US military or Federal, state, or local government, as well as docks used for vessels on display, passenger vessels, research vessels, and ferries were excluded as follows:

- Docks operated by the US military (including the Coast Guard) and other Federal, state, and local government entities, as well as museums and universities were excluded based on values in the `OPERATORS` field.
- Docks not in use or used for museum or exhibit ships, passenger service, casino/showboats, or government, research, or training purposes were excluded based on values in the `PURPOSE` field.
- Docks with a value of "`Ferried Autos, Passengers, Railway Cars`" in the `COMMODITIES` field were excluded as ferries.
- Docks with no values in the `OPERATORS`, `OWNERS`, and `PORT_NAME` fields or with a date listed in the `SERVICE_TERMINATION` field were excluded as likely abandoned or not in use.

BTS / FAA Aviation Facilities

The BTS / Federal Aviation Administration (FAA) Aviation Facilities dataset, updated monthly, contains all airport and other aviation facility locations in the United States and territories. A copy of the dataset downloaded on 17 September 2024 was used for TC Explorer. However, because this dataset includes all aviation facilities, from major hub airports to rarely-used private airfields, it does not provide a clear way of distinguishing the impact of proximity to an aviation facility. Data from the 2023-2027 National Plan for Integrated Air Service and FAA data on 2021 and 2022 passenger enplanements and cargo service were used as well.

The proximity to aviation facilities indicator was based on proximity to major airports, defined as:

- Airports that, in 2021 and/or 2022 saw at least 10,000 passenger enplanements (the minimum to be classified as a primary commercial service airport by the FAA) or 100 million pounds of landed weight of cargo-only aircraft (a backstop to capture a small number of airports with very substantial cargo service but no or minimal passenger service).
- Military airports, identified by a value of "`CG`", "`MA`", "`MN`", or "`MR`" in the `OWNERSHIP_` field and a value of "`A`" (aircraft) in the `SITE_TYPE_` field. All military airports were treated as major due to the lack of reliable data to identify size or usage levels for these airports.

6.4.3 Air Pollution Burden Datasets

The air pollution burden indicators (Section 2.3.3 on page 14) are sourced from the Environmental Protection Agency (EPA) Environmental Justice Screening and Mapping Tool (EJScreen) dataset, updated annually, which provides tract-level estimates of air and surface pollution.

EJScreen v2.3, released in August 2024, was used for the air toxics concentration (`AIR_RLS`), nitrogen dioxide concentration (`AIR_NO2`), and 2.5-micron particulates concentration (`AIR_P25`) indicators, and for the diesel particulates indicator (`AIR_DSL`) except in Alaska. EJScreen v2.2, released in September 2023 was used for data not available from EJScreen v2.3: for the air toxics cancer risk (`AIR_CNR`) and air toxics respiratory risk (`AIR_RSP`) indicators and for the diesel particulates indicator (`AIR_DSL`) in Alaska, as this data was not available in EJScreen v2.3. Both datasets were tabulated using 2020 Census tracts.

6.4.4 Surface Pollution Burden Datasets

The surface pollution burden indicators (Section 2.3.4 on page 15) are drawn from several sources. The leaking underground storage tanks indicator (`SUR_UST`) was taken from EJScreen v2.3, discussed in the previous section. The other five surface pollution indicators were calculated from the portions of Census tracts within 1-mile buffers of pollution source point objects taken from the Environmental Protection Agency (EPA) Facility Registry Service (FRS) and the Department of Labor Mine Data Retrieval System (MDRS), discussed below. The 1-mile buffered versions of each of these datasets used in the calculation of the indicators were produced by dividing the datasets into sections for each *UTM zone* and performing the buffering in the appropriate UTM coordinate reference system before recombining the buffered geometries.

EPA Facilities Registry System (FRS)

The EPA Facility Registry Service (FRS) geospatial data downloads are updated weekly and contain locations and descriptions of a variety of facility types that handle, store, transport, generate, or dispose of hazardous waste. The file geodatabase version of the data was downloaded on 17 September 2024 and the `FACILITY_INTERESTS` layer was used. Facilities types were identified using the `INTEREST_TYPE` field and limited to those where the ACTIVE_STATUS field had values of "`ACTIVE`", "`OPEN`", or "`Y`".

Department of Labor Mine Data Retrieval System (MDRS)

The Department of Labor Mine Data Retrieval System (MDRS) maintains a regularly-updated database of descriptions and locations of all mines in the United States and territories. Mines with `CURRENT_MINE_STATUS` values of "`Abandoned`" or "`Abandoned and Sealed`" were excluded, and the dataset was limited to mines where either the `PRIMARY_CANVASS` or the `SECONDARY_CANVASS` field had a value of "`Coal`", "`Metal`", or "`Nonmetal`" to exclude stone quarries and sand and gravel pits.

6.5 Sources for Population-Based Vulnerability Indicators

The data sources used for population-based vulnerability indicators are listed in Table 5 and discussed in depth below.

Table 5: Data Sources for Population-Based Vulnerability Indicators

Subcomponent	Indicators	Data Source
Communication Vulnerability	Population With Limited English	ACS 2019-2023 5-year estimates
	Households Without Internet Access	ACS 2019-2023 5-year estimates
Employment Vulnerability	Population With Limited Education	ACS 2019-2023 5-year estimates
	Population Not Currently Employed	ACS 2019-2023 5-year estimates
Income Vulnerability	Population Below 200% Poverty Line	ACS 2019-2023 5-year estimates
	Population Without Health Insurance	ACS 2019-2023 5-year estimates
	Median Household Income	ACS 2019-2023 5-year estimates
	Relative Household Income	ACS 2019-2023 5-year estimates
Housing Vulnerability	Households Renting Housing	ACS 2019-2023 5-year estimates
	Households With Cost-Burdened or Inadequate Housing	ACS 2019-2023 5-year estimates
Health Vulnerability	Asthma Prevalence	CDC PLACES 2024 (2021 data)
	High Blood Pressure Prevalence	CDC PLACES 2024 (2022 data) CDC PLACES 2022 (2019 data) for FL
	Cancer Prevalence	CDC PLACES 2024 (2022 data)
	Diabetes Prevalence	CDC PLACES 2024 (2022 data)
	Poor Mental Health Prevalence	CDC PLACES 2024 (2022 data)

Four of the five sets of *population-based vulnerability indicators*—*communication vulnerability*, *employment vulnerability*, *income vulnerability*, and *housing vulnerability*—are sourced from the 2019-2023 American Community Survey (ACS) / Puerto Rico Survey (PRS) 5-year estimates (Section 6.2 on page 46). With the exception of the median household incomes for *regions*—used in the calculation of the relative household income indicator (Section 2.4.3 on page 18)—which were calculated from ACS/PRS data tabulated at the *county* level, all ACS/PRS data was tabulated for 2020 *Census tracts*.

The five *health vulnerability* indicators (Section 2.4.5 on page 19) are sourced from the Centers for Disease Control and Prevention (CDC) Population Level Analysis and Community Estimates (PLACES) dataset, released annually, which provides small-area estimates of the prevalence of various health conditions at the Census tract level. The health vulnerability indicators are taken from the PLACES 2024 dataset, with the exception of high blood pressure prevalence (`HEA_HBP`) data for Florida, which is absent from PLACES 2024. High blood pressure prevalence data for Florida is instead taken from PLACES 2022.

The high blood pressure prevalence, cancer prevalence, diabetes prevalence, and poor mental health prevalence data from PLACES 2024 represents prevalence estimates for 2022, while the asthma prevalence data from PLACES 2024 represents prevalence estimates for 2021. The PLACES 2022 high blood pressure data used for Florida represents prevalence estimates for 2019.

The PLACES 2024 dataset is tabulated for 2020 Census tracts, while the PLACES 2022 dataset is tabulated for 2010 Census tracts and was converted to 2020 Census tracts by *population-weighted interpolation* as discussed below.

Transfer of PLACES 2022 data from 2010 to 2020 Census tracts

The 2022 version of the PLACES dataset used for the high blood pressure prevalence indicator for Florida were tabulated for 2010 Census tracts and required conversion to 2020 Census tracts. Since this indicator represents demographic data, it was converted using population-weighted interpolation based on 2019 populations using the NHGIS 2010 Block Groups to 2020 Census Tracts Crosswalk (Section 6.1 on page 43).

The prevalence of high blood pressure indicator tabulated by 2010 tracts was joined to the National Historical Geographic Information System (NHGIS) 2010 *block group* to 2020 Census tract crosswalk (see page 43), assigning the same value to all 2010 block groups in a single tract. The 2019 populations of the block groups were estimated using 2015-2019 ACS 5-year estimates of total population (`B01001_001`).

To calculate the 2019 population of each 2020 tract, the 2019 block group populations were multiplied by the `wt_pop` population weight field from the crosswalk (representing the share of 2010 block group's population intersecting with each 2020 tract) and summed by 2020 tract. The process was repeated for the 2019 populations of the block groups with high blood pressure—calculated as the product of the 2019 population and the prevalence of high blood pressure—to calculate the population of each tract with high blood pressure. The 2020 tract high blood pressure prevalence values were then calculated as the number of residents of the 2020 tract with high blood pressure divided by the total number of residents of the 2020 tract.

6.6 Sources for Display Layers

Along with the four disadvantage layers showing *Census tracts* that qualify as *disadvantaged communities* in terms of *overall disadvantage*, *transportation insecurity*, *place-based burden*, and *population-based vulnerability*, a large number of display layers are provided in the TC Explorer. The data sources for these layers are described below.

6.6.1 Opportunity Zones and Persistent Poverty Tracts and Counties

The Opportunity Zones and Areas of Persistent Poverty display layers were sourced from the official definitions of these types of disadvantage areas.

The Opportunity Zones display layer was downloaded from the Department of the Treasury Community Development Financial Institutions Fund (CDFI) Opportunity Zones page on 8 November 2024. Note: this display layer consists of 2010 *Census tracts*.

The Areas of Persistent Poverty display layers for both *counties* and 2020 Census tracts that qualify as areas of persistent poverty were sourced from the Rebuilding American Infrastructure with Sustainability and Equity (RAISE) grant program areas of persistent poverty dataset for 2025-2026. Tracts listed as "`Not Identified`" in the dataset were excluded.

6.6.2 Tribal Land

Two different representations of tribal land areas are included as display layers: one sourced from the Bureau of Indian Affairs (BIA) and one from the Census Bureau.

Bureau of Indian Affairs (BIA) Alaskan Native Villages and Native American Land Layers

The Bureau of Indian Affairs tribal land layers were sourced from BIA data sources. The BIA "Land Area Representation" and "Supplemental Land Area Representation" datasets were used to create the BIA Native American Land layer, while the BIA Alaskan Native Villages dataset was used as-is as the BIA Alaskan Native Villages layer.

Census Bureau Tribal Land Layer

The Census Bureau Tribal Land Layer was sourced from the 2023 American Indian/Alaska Native Areas and Hawaiian Home Lands *TIGER/Line shapefile*, downloaded with the `tigris` package in R. The dataset was limited to Federally-recognized communities by selecting only geographies with an "`F`" value in the `AIANNHR` field. Land area types were assigned as follows:

- American Indian Reservation and/or Trust Land – `CLASSFP` value of "`D2`", "`D3`", "`D5`", or "`D8`"
- American Indian Tribal Statistical Area – `CLASSFP` value of "`D6`"
- Alaskan Native Village Statistical Area – `CLASSFP` value of "`E1`"
- Hawaiian Home Land – `CLASSFP` value of "`F1`"

In addition, joint-use areas had to be assigned by hand because the category includes both statistical areas and reservations/trust lands. The joint-use areas with `GEOID` values of "`5915R`", "`5950R`", "`5970R`", or "`5955R`" were assigned as American Indian Tribal Statistical Area and those with `GEOID` values of "`4910R`", "`4930R`", or "`4940R`" were assigned as American Indian Reservation and/or Trust Land.

6.6.3 Transportation Layers

The traffic fatalities, dock facilities, rail network, freeways and expressways, roads with over 50,000 AADT, and airports layers are based on datasets downloaded for the *traffic fatalities burden* (Section 2.2.4 on page 10) and *infrastructure proximity* (Section 2.3.2 on page 13) *indicators*. As discussed below, minor adjustments were made to the processing of these layers to allow for the display of additional information.

In addition, the public transportation, intercity bus, alternative fueling stations, and national bridge inventory display layers are also included in the TC Explorer display layers although they were not used to calculate indicator scores. As discussed in Section 4.2 on page 29, an internal Bureau of Transportation Statistics (BTS) GTFS dataset was used for measuring public transportation access for the *destination access* indicators. The national bridge inventory data, which is supplied as a data layer to aid users in characterizing the condition of transportation infrastructure in project areas, was not used in an indicator because of the large share of Census tracts nationally that do not contain any bridges.

Traffic Fatalities Layer

The 2018-2022 FARS fatality point data (Section 6.3 on page 47) was used for the traffic fatalities display layer with one addition: motor vehicle occupant and other road user fatalities were distinguished. Fatalities with values of "`Driver of a Motor Vehicle In-Transport`", "`Occupant of a Motor Vehicle Not In- Transport`", "`Passenger of a Motor Vehicle In-Transport`", or "`Unknown Occupant Type in a Motor Vehicle In- Transport`" in the `PER_TYPNAME` field were coded as motor vehicle occupants; all others were coded as other road users. The distinction between motor vehicle occupants and other road users was not included in the *traffic fatality burden* indicator because the share of Census tracts with other road user fatalities was too low to support a statistically significant indicator.

Dock Facilities, Rail Network, Freeways and Expressways, and Roads with over 50,000 AADT Layers

The dock facilities, rail network, freeways and expressways, and roads with over 50,000 AADT display layers were largely sourced from the infrastructure proximity datasets (Section 6.4.2 on page 51), however, there were some minor modifications. Rail segments with null values of in the `PASSNGR` field were coded as freight-only, while those with other values were coded as having passenger service. (No distinction between the two was included in the infrastructure proximity indicator) In addition, to reduce file size and avoid overloading the web application, resolution of the rail and roadway network display layers was decreased using the `st_simplify` function in the `sf` R package with a tolerance of 50 meters.

Airports Layer

The airports display layer is a modified version of the airports layer discussed in Section 6.4.2 on page 51, differing in that, along with major public airports and military airports, "other public airports"—defined as those with a value of "PU" in the FACILITY_U field and a value of "A" (aircraft), "C" (seaplane), "G" (glider), or "U" (ultralight) in the SITE_TYPE_ field—that did not qualify as major public airports or military airports are also included. These airports were included because, although they see substantially less traffic than major public airports, they still provide important transportation resources for communities that planners may wish to consider. In addition, airports that qualified as both major public airports (by passenger or cargo traffic) and military airports (by ownership) are displayed as major public airports.[29]

Alternative Fueling Stations Layer

The Department of Energy Alternative Fuels Data Center dataset of alternative fueling stations for road vehicles was downloaded on 5 October 2024. This dataset contains locations and information for alternative fueling stations across the United States and was used to generate the alternative fueling stations display layer, with the fuel_type_code field used to identify fueling station type.

Public Transportation and Intercity Bus Layers

The public transportation display layers were sourced from the September 2024 update of the Bureau of Transportation Statistics (BTS) / Federal Transit Administration (FTA) National Transit Map Stops and National Transit Map Routes layers. Transit agency names were sourced from the Agency_Name, City, and State fields in the September 2024 update of the BTS/FTA National Transit Map Agencies layer. Service types were identified from the route_type_text field as follows:

- Bus – blank, "Bus", or "Trolleybus"
- Rail – "Rail", "Subway, Metro", "Monorail", "Tram, Streetcar, Light rail", "Cable car", or "Gondola, Suspended cable car"
- Ferry – "Ferry"

The intercity bus display layers were sourced from the August 2024 update of the BTS / FTA Intercity Bus Atlas Stops and Intercity Bus Atlas Routes layers.

To reduce file size and avoid overloading the web application, resolution of both the public transit routes and intercity bus routes display layers was decreased using the st_simplify function in the sf R package with a tolerance of 50 meters.

[29] There are a number of joint-use airports that are owned by the US military but also have a public-use portion and serve as commercial airports. While these did not need to be distinguished for the major airport proximity indicator, they are displayed as public airports to better indicate their potential utility in transportation planning.

National Bridge Inventory Layer

The BTS / FHWA National Bridge Inventory dataset, updated annually, documents the location and condition of bridges and culverts on public roads and publicly-accessible bridges on Federal and Tribal lands. Details on the coding of bridges can be found in the Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges. The 29 August 2024 update was used in the TC Explorer. This dataset was used to generate the national bridge inventory display layer, with the following fields used for bridge condition information:

- Overall Condition – `BRIDGE_CONDITION`
- Deck Rating: – `DECK_COND_058`
- Superstructure Rating: – `SUPERSTRUCTURE_COND_059`
- Substructure Rating – `SUBSTRUCTURE_COND_060`
- Channel Rating – `CHANNEL_COND_061`
- Culvert Rating – `CULVERT_COND_062`

6.6.4 EPA Non-Attainment Area Layers

Environmental Protection Agency (EPA) air quality Non-Attainment Area display layers corresponding to the 28 September 2024 Green Book revision were downloaded from the EPA Green Book GIS Download site. Including this information is helpful to *states* and *Metropolitan Planning Organizations* as they develop their Statewide Transportation Improvement Programs (STIPs) and Transportation Improvement Programs (TIPs), which are required to conform with the Clean Air Act (42 U.S.C. 7506).

One non-attainment area, the Canton, OH lead (2008 standard) area, was not included in these shapefiles and was hand-drawn based on the text description in the Green Book. Maintenance and non-attainment areas were distinguished using the `naastatus` field, with values of "`M`" corresponding to maintenance areas and "`N`" to non-attainment areas. In addition, values of the `class` field for the New Haven, CT 10-micron particulate matter and Sheboygan, WI ozone (2008 standard) had to be added by hand, and the distinct values of the `class` for the state portions of the Cincinnati, OH/Cincinnati, KY and Louisville, KY / Louisville, IN ozone (2015 standard) had to be input by hand.

7. Glossary

Air Pollution Burden

The air pollution burden subcomponent of *place-based burden* measures the potential exposure of the project area to air pollutants and the adverse environmental conditions caused by air pollution. This measure does not incorporate carbon dioxide or greenhouse gases as a form of air pollution, but focuses on criteria air pollutants and on hazardous air pollutants (also known as air toxics), which are known to cause negative health outcomes. Higher scores reflect project areas with higher rates of air pollution. The *indicators* used to calculate air pollution burden are discussed in Section 2.3.3 on page 14.

Annual Average Daily Traffic (AADT)

Annual average daily traffic is a metric used in transportation planning and engineering, consisting of the total number of vehicles that pass over a roadway or roadway segment in a year divided by 365. The Federal Highway Administration (FHWA) guide for traffic monitoring provides further details on how this data is collected and tabulated.

Area-Weighted Interpolation

Area-weighted interpolation is the process of generating estimated values of a variable for a set of geographies based on the area of their intersections with a set of geographies for which the variable is known. Area-weighted interpolation (using land area values provided from a crosswalk file) was used to transfer extreme weather hazard data (the NOAA CMRA extreme weather hazard *indicators*) from 2010 to 2020 *Census tracts* for use in TC Explorer.

Cartographic Boundary File

Cartographic boundary files are lower-resolution (1:500,000) versions of *TIGER/Line shapefiles* published by the Census Bureau for most Census geographies. In addition to being lower resolution, these shapefiles generally exclude *Census tracts* and *Census block groups* that are entirely water, making larger geographies such as counties and states conform more closely to their land areas. The 2023 editions of cartographic boundary files are used for data presentation in the TC Explorer but are not used for data analysis. *Water tracts* are defined as those excluded from the 2023 tract-level cartographic boundary files.

Census Block (or Block)

Census blocks are the smallest geographic units identified by the Census Bureau, which all larger Census geographies are defined in terms of. Blocks are defined in Decennial Census years (i.e. 2010 and 2020) and remain the same, with the same *FIPS codes*, between Decennial Census years. The FIPS code of a block is 15 digits long and is part of the state-county-tract-block group-block hierarchy: 2 digits (state) + 3 digits (county) + 6 digits (tract) + 1 digit (block group) + 3 digits (block).

Census Block Group (or Block Group)

Census block groups are the smallest geographic units for which the Census Bureau reports ACS data. They are generally defined to contain between 600 and 30,00 residents and are proper subsets of *Census tracts*. Block groups are defined in Decennial Census years (i.e. 2010 and 2020) and remain the same between Decennial Census years. However, their *FIPS codes* change between Decennial Census years if the *county* they are part of changes or changes its FIPS code. The FIPS code of a group block is 12 digits long and is part of the state-county-tract-block group-block hierarchy: 2 digits (state) + 3 digits (county) + 6 digits (tract) + 1 digit (block group).

Census Tract (or Tract)

Census tracts are the Census Bureau's primary geography for reporting ACS data for areas smaller than counties. They are proper subsets of *counties* and made up of *Census block groups* and are generally defined to contain between 1,200 and 8,000 residents, though special-purpose tracts defined to represent non-residential areas such as parkland, wilderness areas, industrial and office parks, shopping malls, and non-residential portions of colleges and military bases may have minimal or no population. Tracts are defined in Decennial Census years (i.e. 2010 and 2020) and remain the same between Decennial Census years. However, their *FIPS codes* change between Decennial Census years if the county they are part of changes or changes its FIPS code. The FIPS code of a tract is 11 digits long and is part of the state-county-tract-block group-block hierarchy:
2 digits (state) + 3 digits (county) + 6 digits (tract).

The TC Explorer uses the Census tracts defined for the 2020 Census with the FIPS codes in use as of the 2022 release of TIGER/Line geometries by the Census Bureau. These FIPS codes differ from the Census tract FIPS codes initially assigned to tracts for the 2020 Census only in Connecticut, where they reflect the replacement of Connecticut's historical counties with Councils of Government as the state's county-equivalents.

Communication Vulnerability

The communication vulnerability subcomponent of *population-based vulnerability* measures whether community members can easily and reliably receive information and alerts about the transportation system. Higher scores reflect project areas whose households have limitations in their ability to receive digital communications or understand English. The *indicators* used to calculate communication vulnerability are discussed in Section 2.4.1 on page 16.

Community

As used in this document, "community" is synonymous with "*Census tract*."

Component Score

In the TC Explorer, a "component score" is one of the three *disadvantage scores* (*transportation insecurity*, *place-based burden*, and *population-based vulnerability*) that the *overall disadvantage score* is calculated from.

County Equivalent (or County)

In this document, "county" refers to Census Bureau-designated "county equivalents" in use as of the 2022 release of TIGER/Line geometries. Thus, counties include counties (in all states except Alaska, Connecticut, and Louisiana); independent cities (in Maryland, Missouri, Nevada, and Virginia); Alaska boroughs and Census areas; Connecticut councils of government; Louisiana parishes; Puerto Rico municipios; American Samoa districts and islands; municipalities in the Northern Mariana Islands; and islands in the US Virgin Islands. Guam and the District of Columbia each consist of a single county-equivalent.

Destination Access Vulnerability

The destination access vulnerability subcomponent of *transportation insecurity* measures access to essential destinations, such as public transit, medical facilities, education, groceries, and jobs. Higher scores reflect project areas where residents have less access to essential destinations within 30 minutes of driving, walking, or biking. The *indicators* used to calculate destination access vulnerability are discussed in Section 2.2.1 on page 7.

Differential Privacy

For the 2020 Decennial Census, the Census Bureau implemented a system known as differential privacy as part of their disclosure avoidance system for ensuring that personally-identifiable information could not be backed out of published Decennial Census data. This system does not affect the counts of *housing units*, but adds uncertainty or "noise" to the values of all other data, including population counts, reported for the 2020 Decennial Census. Specifically, published population counts for all Census geographies may be slightly decreased or increased over their true measured values. For geographies with large populations, these variations are small and may fall within the margin of error, but for geographies with no or very small populations, they may result in a significant overestimate of the population, especially since geographies with no population can only have their population increased, not decreased. Because of the differential privacy system, the populations of *Census tracts* with very low populations cannot be known reliably, nor can the spatial distribution of their population among *Census blocks*. This uncertainty is part of the reason that *low-population tracts* are handled specially in the TC Explorer as discussed in Section 5.2.3 on page 39.

Disadvantage Score

In the TC Explorer, a "disadvantage score" is a value between 0 (least disadvantaged) and 100 (most disadvantaged) calculated for a *Census tract* or *project area*. A disadvantage score for a Census tract—either the *overall disadvantage score* or one of the *component scores* or *subcomponent scores* it is made up of is calculated as a percentile ranking of the Census tract among all tracts in the *United States*. Disadvantage scores for project areas are calculated as the population-weighted average (using 2020 Decennial Census populations) of the Census tracts making up the project area. For tracts in Puerto Rico, those disadvantage subcomponent scores that are available are defined as the percentile rank of the tract in the United States with the closest value of the average of the normalized indicators to the Puerto Rico tracts.

DOT considers a census tract/project area as "disadvantaged" if its overall disadvantage score is in the 65th percentile or greater nationally, or if it is in the *US territories*.

Disadvantaged Community

As used in this document, a "disadvantaged community" is a Census tract in the *United States* that experiences *overall disadvantage* or one of its component or subcomponents at the 65th percentile level or higher relative to all tracts in the United States, or any tract in the *US territories*. Federally-recognized Native American/Alaskan Native tribal governments are also considered disadvantaged communities independently of their locations. The identification of disadvantaged communities is discussed in Section 3.4 on page 24.

Display Layer

The display layers in the TC Explorer are map layers provided for potential utility in preparing Notice of Funding Opportunity (NOFO) applications. The sources for these layers are listed in Section 6.6 on page 57.

Employment Vulnerability

The employment vulnerability subcomponent of *population-based vulnerability* measures lack of employment and education. Higher scores reflect project areas whose residents have lower rates of employment and high school graduation. The *indicators* used to calculate employment vulnerability are discussed in Section 2.4.2 on page 17.

EPSG Code

An EPSG code is an integer between `1024` and `32767` that identifies the specifications of a coordinate reference system in the *EPSG Geodetic Parameter Dataset* maintained by the Geomatics Committee of the International Association of Oil & Gas Producers (IOGP).

Extreme Weather Hazard

The extreme weather hazard subcomponent of *place-based burden* measures the predicted change in extreme weather or variability in the project area by 2050 which may have impacts on transportation system performance, safety, and reliability. These impacts in turn have major implications to supply chains, emergency response and the longevity of transportation investments. Higher scores reflect project areas that are likely to see increased extreme weather impacts to transportation infrastructure. The *indicators* used to calculate extreme weather hazard are discussed in Section 2.3.1 on page 11.

FIPS Code

In this document, "FIPS code" refers to the full, nationally-unique geographic identifier assigned to a geography by the Census Bureau, i.e. five-digit *county* FIPS codes and eleven-digit *Census tract* FIPS codes.

Group Quarters

Group quarters are one of two distinct types of living arrangements defined in Census Bureau data, the other being *households*. Group quarters consist of managed housing arrangements that may be custodial (prisons, long-term healthcare facilities) or non-custodial (some college dormitories, military barracks, shelters for the unhoused, or religious/monastic communities). A full list of group quarters types for the 2020 Decennial Census can be found on the Census Bureau website. Notably, a number of *indicators* are defined solely for residents of households and so have to be estimated by *nearest-neighbor interpolation* for Census tracts where most residents live in group quarters.

Census Bureau definition: *A Group Quarters (GQ) is a place where people live or stay in a group living arrangement that is owned or managed by an entity or organization providing housing and/or services for the residents. These services may include custodial or medical care, as well as other types of assistance, and residency is commonly restricted to those receiving these services. This is not a typical household-type living arrangement. People living in GQs usually are not related to each other. GQs include such places as college residence halls, residential treatment centers, skilled nursing facilities, group homes, military barracks, correctional facilities, workers' dormitories, and facilities for people experiencing homelessness.*

Health Vulnerability

The health vulnerability subcomponent of *population-based vulnerability* measures the increased prevalence of health conditions that may result from exposure to pollutants, poor walkability, car dependency, and long commute times. Higher scores reflect project areas with higher rates of asthma, cancer, high blood pressure, diabetes, and poor mental health. The *indicators* used to calculate health vulnerability are discussed in Section 2.4.5 on page 19.

Household

A household is ne of two living distinct types of arrangements defined in Census Bureau data, the other being *group quarters*. Households consist of groups of residents—whether related or not—living together in a *housing unit*. Housing units encompass typical owned or rented housing, including some college dormitories.

Census Bureau definition: *A household includes all the people who occupy a housing unit. (People not living in households are classified as living in group quarters.) A housing unit is a house, an apartment, a mobile home, a group of rooms, or a single room that is occupied (or if vacant, is intended for occupancy) as separate living quarters. Separate living quarters are those in which the occupants live separately from any other people in the building and which have direct access from the outside of the building or through a common hall. The occupants may be a single family, one person living alone, two or more families living together, or any other group of related or unrelated people who share living arrangements.*

Housing Unit

In Census Bureau data, a housing unit is a set of living quarters, whether owned or rented, inhabited by (or vacant but potentially inhabitable by) a *household*.

Census Bureau definition: *A housing unit may be a house, an apartment, a mobile home, a group of rooms or a single room that is occupied (or, if vacant, intended for occupancy) as separate living quarters. Separate living quarters are those in which the occupants live separately from any other individuals in the building and which have direct access from outside the building or through a common hall.*

Housing Vulnerability

The housing vulnerability subcomponent of *population-based vulnerability* measures housing condition and the share of household income spent on housing. Higher scores reflect project areas where households spend a higher percentage of income on housing or there is a prevalence of household overcrowding and lack of indoor plumbing and kitchens. The *indicators* used to calculate housing vulnerability are discussed in Section 2.4.4 on page 18.

Income Vulnerability

The income vulnerability subcomponent of *population-based vulnerability* measures poverty and income inequality in the project area. Higher scores reflect project areas with higher poverty rates and median household incomes below the regional average. The *indicators* used to calculate income vulnerability are discussed in Section 2.4.3 on page 17.

Indicator

In the TC Explorer, an indicator is a measure of a specific physical, environmental, or demographic property of a *Census tract* or *project area* that is used directly to quantify the tract or area or to calculate *disadvantage scores*. Indicators have two forms: *raw indicators* (discussed in Section 2 on page 6) and *normalized indicators* (discussed in Section 3.2 on page 22).

Infrastructure Proximity

The infrastructure proximity subcomponent of *place-based burden* measures how close the project area is to freeways, high-volume roads, railways, airports, and ports, which may align with higher rates air and noise pollution, as well as divided communities. Higher scores reflect project areas close to one or more forms of transportation infrastructure. The *indicators* used to calculate infrastructure proximity are discussed in Section 2.3.2 on page 13.

Insular Area

In this document, "insular areas" refers to the four US territories of American Samoa, Guam, the Northern Mariana Islands, and the US Virgin Islands that are not included in the American Community Survey/Puerto Rico Survey by the Census Bureau. Very limited *indicator* data is available for these territories, and none of the component and subcomponent scores are calculated.

Isolated Alaska Tract (IAT)

In the TC Explorer, isolated Alaska tracts are defined as those *Census tracts* in Alaska where the majority of the population lacks year-round road access to one of the six Census-defined *urban areas* or urban clusters in Alaska with at least 10,000 residents as of the 2020 Decennial Census: Anchorage, Anchorage Northeast, Fairbanks, Wasilla/Knik-Fairview/North Lakes, Juneau, and Ketchikan. Road access was identified using a road network geometry provided by the Alaska Department of Transportation and excluding roads labeled as “priority level 5” / “No Winter Maintenance.” Further explanation and a list of these tracts can be found in Section 5.2.2 on page 37.

The *transportation cost burden* and *traffic fatality burden* *subcomponent scores* are set at the 100th percentile level regardless of the values of the underlying indicator values.

Low-Population Tract (LPT)

Census tracts that were reported as containing fewer than 100 residents in the 2020 Decennial Census are designated “low-population tracts” in the TC Explorer. Spatial calculations of indicators for these tracts are performed using land area instead of population, because their low populations, combined with the *differential privacy* noise included in the 2020 Decennial Census, makes calculations involving population distribution unreliable. In addition, these tracts, which have populations much lower than the usual 1,200 residents lower bound for tract populations, generally represent land with non-residential purposes, making residential population distribution a less useful way to characterize where people are most often located in the tracts. For more information, see Section 5.2.3 on page 39.

Metropolitan Planning Organization (MPO)

A Metropolitan Planning Organization (MPO) is the policy board of an organization created and designated to carry out the metropolitan transportation planning process. All *urban areas* with populations over 50,000 residents are required to be represented by MPOs.

Metropolitan Statistical Area (MSA)

Metropolitan statistical areas (MSAs) are regions defined as groups of counties by the Office of Management and Budget (OMB) to represent metropolitan areas centered on core urban areas with populations of at least 50,000 and containing outlying areas connected to the core urban area by commuting patterns. The TC Explorer uses the MSAs defined by the OMB on 21 July 2023 in OMB Bulletin 23-01, which are based on the results of the 2020 Decennial Census. These definitions use the 2022-onward standard of treating councils of government as county-equivalents for Connecticut.

Nearest-Neighbor Interpolation

Nearest-neighbor interpolation is the process of generating estimated values of a variable for geographies for which it is unavailable from its values in proximate geographies, based on Tobler's First Law of Geography. Since the calculation of *disadvantage scores* requires *indicators* to have values for all *Census tracts* in the *United States*, nearest-neighbor interpolation is used to supply estimated values for missing *raw indicators*. Missing values are estimated as the average of the values of the same raw indicator for the three nearest Census tracts for which the value is not missing, with nearest neighbors defined by distance between Census tract population centroids (calculated for 2020 Decennial Census data) for *transportation insecurity* and *population-based vulnerability* indicators and land area centroids for *place-based burden* indicators. The use of nearest-neighbor interpolation to supply missing values for indicators is discussed in Section 3.1 on page 21.

Normalized Indicator

In the TC Explorer, normalized *indicators* are versions of the *raw indicators* that have normalized to a scale of 0 (least disadvantaged) to 1 (most disadvantage) for use in the calculation of *disadvantage scores*. The calculation of normalized indicators is discussed in Section 3.2 on page 22.

Overall Disadvantage

The overall *disadvantage score* is a measure of a Census tract or project area's overall *transportation disadvantage*. DOT considers a census tract/project area as "disadvantaged" if its overall disadvantage score is in the 65th percentile or greater nationally, or if it is in the *US territories*. Overall disadvantage is also broken down into three component scores: *transportation insecurity*, *place-based burden*, and *population-based vulnerability*.

Place-Based Burden

The place-based burden component of overall disadvantage measures disadvantage inherent in a location and experienced by all residents of the location. The subcomponents and indicators in this component are important because they provide transportation decision makers the information needed to develop transportation plans and make funding decisions that ensure a community's transportation infrastructure is resilient and minimizes negative health and economic impacts. This component score is calculated based on four subcomponents: *extreme weather hazard*, *infrastructure proximity*, *air pollution burden*, and *surface pollution burden*. The *indicators* used to calculate place-based burden are discussed in Section 2.3 on page 11.

Population-Based Vulnerability

The population-based vulnerability component of overall disadvantage measures disadvantage experienced by a population due to demographic and socioeconomic traits that make them particularly vulnerable. This component score is calculated based on five subcomponents: *communication vulnerability*, *employment vulnerability*, *income vulnerability*, *housing vulnerability*, and *health vulnerability*. The *indicators* used to calculate population-based vulnerability are discussed in Section 2.4 on page 16.

Population-Weighted Interpolation

Population-weighted interpolation is the process of generating estimated values of a variable for a set of geographies based on the populations of their intersections with a set of geographies for which the variable is known. Population-weighted interpolation (using population values from a crosswalk file) was used to transfer demographic data (the 2022 CDC PLACES high blood pressure prevalence data for Florida) from 2010 to 2020 *Census tracts* for use in TC Explorer.

Project Area

A project area is the area served by and/or affects by a transportation project. In the TC Explorer, a project area must be defined in terms of *Census tracts*.

Raw Indicator

In TC Explorer, raw *indicators* are direct measures of some demographic or physical property of Census tracts. These indicators, suffixed "`_R`" in the Technical Data Download tracts files, have a variety of units—shares of Census tract population, shares of households, parts per billion, and so on—depending on what is being measured, which makes them useful for assessing specific community conditions. However, this diversity of units means they cannot be used for calculating *subcomponent scores* directly. They are converted to *normalized indicators*, which are used for that. The calculation of raw indicators is discussed in Section 2 on page 6.

Region

Except where otherwise indicated by context, "region" refers to the *Metropolitan Statistical Area (MSA)* a Census tract is located in for tracts in MSAs and to the non-MSA portion of the *state* or *territory* a Census tract is located in for tracts not in MSAs. The assignment of regions is discussed in Section 5.1.1 on page 33.

Remote Island County

Three *counties*—Rose Island (FIPS 60030) and Swains Island (FIPS 60040) in American Samoa and the Northern Islands Municipality (FIPS 69085) in the Northern Mariana Islands—consist solely of *remote island tracts* and so, like remote island tracts, are excluded from the TC Explorer. More details can be found in Section 5.2.1 on page 35.

Remote Island Tract

In addition to *water tracts*, twelve *Census tracts* that consist of one or more islands with no permanent population (and sometimes no access to the general public) that are sufficiently far from other tracts that *indicator* values cannot reasonably be *nearest-neighbor interpolated* from nearby tracts are designated as "remote island" tracts and excluded from the TC Explorer. More details and a list of these tracts can be found in Section 5.2.1 on page 35.

State

In this document, "state" refers to any of the fifty states and the District of Columbia unless context explicitly indicates otherwise; the District of Columbia is treated equivalently to the fifty states in all calculations. However, "state" generally excludes the *US territories* defined as state-equivalent entities by the Census Bureau: Puerto Rico, American Samoa, Guam, the Northern Mariana Islands, and the US Virgin Islands. When these are to be included, the text will explicitly reference "state or Puerto Rico" (if *insular areas* are excluded) or "state or territory" (if insular areas are included).

Subcomponent Score

In the TC Explorer, a "subcomponent score" is one of the *disadvantage scores* that the three *component scores* (*transportation insecurity*, *place-based burden*, and *population-based vulnerability*) are calculated from. The calculation of subcomponent scores is discussed in Section 3.3 on page 23.

Surface Pollution Burden

The surface pollution burden subcomponent of *place-based burden* measures the potential exposure of the project area to land and surface water pollutants and the adverse environmental conditions cause by surface pollution. Higher scores reflect project areas with higher rates of surface pollution. The *indicators* used to calculate surface pollution burden are discussed in Section 2.3.4 on page 15.

Territory (or United States Territory)

In this document, "Territories" or "US Territories" refers to Puerto Rico and the four *insular areas* of American Samoa, Guam, the Northern Mariana Islands, and the US Virgin Islands. Normalization of *indicators* and percentile ranking of *disadvantage scores* in the TC Explorer excludes Census tracts in the territories. All Census tracts in the territories are considered to experience overall disadvantage and the components and subcomponents of disadvantage. In general, disadvantage scores are not calculated for tracts in the territories due to the absence of indicator data, but some subcomponent scores are calculated for tracts in Puerto Rico, where sufficient data is available for reliable calculations. For the purposes of the TC Explorer, this excludes the *United States Minor Outlying Islands*.

TIGER/Line Shapefile

TIGER/Line shapefiles are official descriptions of Census geographies as GIS geometry objects published by the Census Bureau. With the exception of the geometries for *Census blocks*, which contain Decennial Census population and housing unit counts, they do not generally contain demographic data, but they do contain information on the relationship between the geography represented and other geographies, as well as land area and other descriptive values. Technical documentation and errata and user notes are available online. Except where otherwise noted, all analysis for the TC Explorer uses the 2023 editions of TIGER/Line shapefiles. The downloading of these shapefiles is discussed in Section 6.1 on page 42.

Traffic Fatality Burden

The traffic fatality burden subcomponent of *transportation insecurity* measures traffic fatalities (both motorist non-motorist) using the National Highway Transportation Safety Administration's (NHTSA) Fatality Analysis Reporting System (FARS) data for 2018–2022. Higher scores reflect project areas with higher number of traffic fatalities. The indicators used to calculate traffic fatality burden are discussed in Section 2.2.4 on page 10. Note that the percentile-ranked traffic fatality burden subcomponent scores for *isolated Alaska tracts* are automatically set to the 100th percentile level.

Transportation Cost Burden

The transportation cost burden subcomponent of *transportation insecurity* measures the share of income that households in the project area spend on daily transportation. Higher scores reflect project areas where households spend a higher percentage of their income on transportation. The *indicators* used to calculate transportation cost burden are discussed in Section 2.2.3 on page 9. Note that the percentile-ranked transportation cost burden subcomponent scores for *isolated Alaska tracts* are automatically set to the 100^{th} percentile level.

Transportation Disadvantage

See *overall disadvantage*.

Transportation Insecurity

The transportation insecurity component of overall disadvantage measures disadvantage experienced when people are unable to get to where they need to go to meet the needs of their daily life regularly, reliably, affordably, and safely. This component score is calculated based on four subcomponents: *destination access vulnerability*, *vehicle access vulnerability*, *transportation cost burden*, and *traffic fatality burden*. The *indicators* used to calculate transportation insecurity are discussed in Section 2.2 on page 7.

Transportation Management Area (TMA)

A transportation management area (TMA) is an *urban area* with a population of over 200,000 residents in the Decennial Census for which it was defined, or an urban area with a lower population designated as a TMA by the Secretary of Transportation on request of the governor of the state in which it is located and the *metropolitan planning organization* or affected local officials. In addition, the Lake Tahoe MPO region in California and Nevada is defined as a TMA with an urban area population of 145,000 residents in California and 65,000 residents in Nevada as per 23 U.S.C. 134(r). Transportation management areas are subject to special transportation planning and programming requirements (Federal Transit Administration / Federal Highway Administration).

Travelshed

A travelshed is the area that can be reached in a specified amount of time—or by travelling a specified distance—by a given means of transportation. The access to destinations measures in the TC Explorer are based on thirty-minute travelsheds for three means of transportation: walking (defined as ½ mile walking along roads and paths where pedestrians are permitted); cycling (defined as 5 miles cycling along roads and paths where cyclists are permitted); and driving (defined as 30 minutes driving time at morning rush hour. The generation of travelsheds is discussed in Section 4.1 on page 26.

United States (US)

The TC Explorer follows the Census Bureau's definition and uses "United States" to refer to the fifty states and the District of Columbia, excluding the *territories* and *United States Minor Outlying Islands*.

United States Minor Outlying Islands

The United States Minor Outlying Islands are a set of small islands under US jurisdiction which are not part of the five US territories with permanent, year-round inhabitants. The Census Bureau does not define Census tracts for them, and they are not included in the TC Explorer.

Census Bureau definition: *The U.S. Minor Outlying Islands refers to certain small islands that are U.S Territories under U.S. jurisdiction in the Caribbean Sea and Pacific Ocean: Baker Island, Howland Island, Jarvis Island, Johnston Atoll, Kingman Reef, Midway Islands, Navassa Island, Palmyra Atoll, and Wake Island. These areas usually are not part of standard data products because they generally do not include population year-round.*

United States and Territories

The United States and Territories are the full area included in the TC Explorer. The *United States* (defined as the fifty states and the District of Columbia) and the *territories* (Puerto Rico and the four *insular areas* of American Samoa, Guam, the Northern Mariana Islands, and the US Virgin Islands). For the purposes of the TC Explorer, this excludes the *United States Minor Outlying Islands*.

Universal Transverse Mercator (UTM)

Universal Transverse Mercator (UTM) is a system for projecting the Earth's surface as planar coordinates. The globe is divided into standard zones (UTM zones) 6° of longitude wide for each hemisphere, in which latitude/longitude coordinates can be converted to planar coordinates defined in meters. All spatial calculations for the TC Explorer were performed by transforming features to the appropriate UTM zone based on the North American Datum of 1983 (NAD83). The assignment of UTM zones is discussed in Section 5.1.2 on page 34

Urban Area (UA)

In this document, "urban area" refers exclusively to 2020 Census urban areas that were reported to have populations of at least 50,000 residents as of the 2020 Decennial Census. This corresponds to the definition of "urban area" used by the Census Bureau prior to 2020 and excludes those areas with populations of less than 50,000, which were previously designated as "urban clusters." The assignment of *Census tracts* to urban areas is necessarily an approximation; the details are discussed in Section 5.2.4 on page 39.

Vehicle Access Vulnerability

The vehicle access vulnerability subcomponent of *transportation insecurity* measures whether residents in the project area have access to a vehicle and/or have the ability to drive. Higher scores reflect project areas where households are less likely able to drive to essential destinations. The indicators used to calculate *indicators* access vulnerability are discussed in Section 2.2.2 on page 8.

Water Tract

The TC Explorer excludes all *Census tracts* with FIPS codes of the form xxxxx99xxxxxx—that is Census tracts with tract number of the form 99xx.xx—as "water tracts." These FIPS codes are assigned by the Census bureau as tracts containing only water, and usually consists of offshore water areas, though some inland lakes (Lake Tahoe, for example) are also water tracts. In addition, Tract 9801.00 in Hillsborough County, FL (FIPS code 12057980100) is treated as a water tract in the TC Explorer because it contains minimal land area and is excluded from Census Bureau *cartographic boundary files*. Projects that pass through water tracts should be evaluated based on the *indicators* and *disadvantage scores* of the adjacent non-water Census tracts that they serve.