Caltrans Transportation Equity Index (EQI) Documentation

February 2023

PUBLIC DISCUSSION DRAFT

Contents

Figures

Introduction

The California Department of Transportation (Caltrans) acknowledges that communities of color and under-served communities have experienced fewer benefits and a greater share of the burdens associated with California's transportation system. These disparities largely reflect a history of transportation decision-making, policy, processes, planning, design, and construction that has quite literally put-up barriers, divided communities, and amplified racial inequities, particularly in Black and Brown neighborhoods.[1]

To operationalize Caltrans' commitments to equity, the department is developing the Caltrans Transportation Equity Index (EQI). The EQI is a screening and evaluation tool that utilizes multiple transportation-specific and socioeconomic indicators to identify transportation-based priority populations at the Census block level.

Many tools exist to evaluate the impact of the built environment. Still, these tools typically consider a wide range of factors that are not explicitly focused on burdens caused or exacerbated by the transportation system.

Caltrans aims to bridge this gap by creating an index to inform how the Department can best address and mitigate inequities exacerbated by the transportation system.

Broadly speaking, the EQI will be used for the following purposes:

1. Identify transportation-specific priority population areas for applicable funding programs (e.g., Reconnecting Communities: Highways to Boulevard Pilot Program).
2. Provide guidance to improve the analysis of project impacts and identify opportunities to advance equitable outcomes during project planning, development, and design.

Additionally, the EQI is designed in a manner to support partner agencies and other entities who may voluntarily use the EQI to analyze impacts and evaluate the effectiveness of various transportation projects and solutions.

[1] Caltrans Equity Statement December 2020

EQI Concept

The EQI contributes to the advancement of spatial analysis tools/methods by using Census blocks instead of larger geographies (such as tracts), enabling a more granular level of analysis. Given the nature of the EQI's transportation indicators, this level of granularity is necessary, as an indicator can have significant variance within a tract-scale area, such as a neighborhood, that the EQI is designed to capture.

Furthermore, the EQI only includes variables with spatial significance, meaning that their distribution across the state is largely determined by the spatial nature of the state's transportation system. Central to the EQI's concept is the identification of transportation-based priority populations and the targeting resources to said populations. This concept requires spatially significant data that does not have a uniform distribution throughout a geographic area. If the top 20% of Census blocks were screened using an indicator with a uniform distribution, approximately 80% of the population impacted by the indicator would be excluded. While such indicators may still be highly important when considering transportation equity, they require different approaches and are less useful in a spatial index such as the EQI. The use of thresholds with spatially significant variables confirms that geographies are screened for inclusion that contains a much higher share of burdened populations. This—however—does not preclude the use of non-spatial equity data in other Caltrans processes and decision rules.

Lastly, the EQI relies on both publicly available and internally developed datasets, including:

- Race/Ethnicity and Household Income data from the U.S. Census Bureau Decennial Census and American Community Survey (ACS) 5-year estimates
- Traffic proximity/volume data from Caltrans and the United States Department of Transportation (USDOT)
- Crash data from the California Highway Patrol and UC Berkeley Safe Transportation Research and Education Center (SafeTREC)
- Access to destinations data from Caltrans tools and analysis

EQI Screens

Three distinct screens were developed for the EQI, each intended to identify and address distinct problems. All three are coupled with the Demographic Overlay. Below is a summary of the EQI's three screens, which are discussed in greater detail in this document's data sources and methodology section.

- **Traffic Exposure Screen.** The Traffic Exposure Screen identifies low-income and majority non-white Census blocks that bear a negative traffic burden as measured by proximity, volume, and vehicle type for all interstates, highways, principal arterials, and minor arterials in the state — or safety burden as measured through a statewide crash exposure calculation. The screen identifies Census blocks for inclusion with traffic proximity and volume at or above the 80th percentile or blocks with a crash exposure score at or above the 80th percentile. Traffic exposure is used as a proxy for multiple environmental burdens, including diesel particulate matter, diesel exhaust, noise, and traffic safety impacts on communities.
- **Access to Destinations Screen.** The Access to Destinations Screen identifies low-income and majority non-white Census blocks with poor relative multimodal access to destinations. For this screen, access to destinations is operationalized as the ratio of transit and walking access to destinations compared to auto access to destinations, with blocks having a score equal to or less than 0.12 being screened for inclusion as having poor relative multimodal access to destinations. The analysis is run separately for both jobs and non-work destinations and the block screened for inclusion if the previously discussed threshold is met for either destination type. Poor multimodal access to destinations leads to higher transportation costs and fewer reachable destinations via non-auto modes.[2] Caltrans is also in the process of operationalizing bicycle access to destinations metric which will be considered for inclusion in the EQI upon further development.
- **Priority Populations Screen.** This screen includes both the Traffic Exposure and Access to Destinations Screens and identifies the priority populations of the state that are the most burdened by traffic exposure but also benefit the least from the multimodal transportation network (as measured by access to destinations).

[2] https://htaindex.cnt.org/about/HTMethods_2016.pdf

EQI Data Sources and Methodology

The Caltrans Transportation Equity Index (EQI) includes three components:

1. Demographic indicators focused on household income and race/ethnicity measures.
2. Traffic indicators, as measured by traffic proximity and volume and crash exposure.
3. Access to destinations indicators, measured as the ratio of weighted transit/walk access to destinations to weighted automobile access to destinations.

2020 Census blocks are the geographic unit of measurement used in the EQI. All indicators are either measured at the block level or interpolated to the block level from the block group level.

Demographic Indicators

All EQI screening scenarios include demographic indicators of low-income status and race/ethnicity. If either of the household income level or race/ethnicity criteria are met (as outlined below), the block is screened for inclusion for further analysis with the transportation-specific indicators (traffic exposure and access to destinations indicators).

Income Data

To determine a block's status as either 'low-income' or 'not low-income,' two measures were used in alignment with AB 1550. A Census block group was designated as a 'low-income' community if either 1) its median household income[3] was at or below 80% of the statewide median household income, OR 2) its median household income was at or below the 2022 county low-income limit[4] established by the California Department of Housing and Community Development. If either criterion was met, the block group was identified as a low-income community and screened for inclusion for the Demographic

[3] Median household income was retrieved from the US Census Bureau API using 2020 ACS 5-year estimates table B19013_001.

[4] For the county low-income limit, the average household size was retrieved for each block group using the 2020 ACS 5-year estimates B25010_001 table and used to find the appropriate household-size-adjusted low-income limit.

Overlay. Block group income data was joined to Census block data using the GEOID20 ID in both datasets.

Race/Ethnicity Data

2020 ACS data[5] and 2020 Decennial Census data was used to determine whether a block's population was greater than 63.4902% non-white[6]. The non-white percentage was determined using the following formula:

1 – (Total Not Hispanic or Latino Population of One Race White Alone / Total Population)

If the resulting non-white percentage was greater than or equal to 63.4902%, the block was screened for inclusion for the Demographic Overlay.

It is important to note that race and ethnicity data is available for Decennial Census years on the block level. While block group-level ACS data was used to allow for annual updates, Census blocks falling outside of a screened block group that were greater than or equal to 63.4902% non-white in 2020 are also screened for inclusion.

Traffic Exposure Indicators

The EQI utilizes two sets of traffic exposure indicators: 1) traffic proximity and volume from the highway system and arterial roads in the state and 2) Census block-level crash exposure based on crash history.

Traffic Proximity and Volume Data

1. **Input Data.** Highway Performance Monitoring System (HPMS) data was used and accessed via the Bureau of Transportation Statistics for this analysis. This spatial dataset contains all roads in the United States and car and truck Average Annual Daily Traffic (AADT) for highways and arterials.
2. **Traffic Exposure Python Script.** Using the data described in the previous step as the input, a Python script[7] is used to calculate block-level exposure. The script performs the following steps:
    a. A buffer is created around each road segment at a specified distance.
    b. That buffer is intersected with Census blocks.[8]

[5] American Community Survey (ACS) Table B03002: Hispanic or Latino Origin by Race
For non-Decennial Census years, ACS 5-year estimates will be used to calculate non-white percentage, but data is only available at the block group level.

[6] For the race and ethnicity indicator, a threshold of 63.4902% is used since it is the percentage of the state's 2020 population that is non-white.
https://data.census.gov/table?q=Hispanic+or+Latino+by+race&g=0400000US06&tid=ACSDT5Y2020.B03002

[7] https://github.com/hhmckay/Caltrans-Equity-Index-EQI-

[8] 2020 Census blocks were used in this version of the EQI.

c. For each block, the maximum AADT value for a given route is selected (in some cases, there is overlap between buffers for the same route, so the maximum AADT value is used to provide a more conservative estimate).
d. For blocks with AADT exposure from multiple routes, the separate AADT exposure values are summed to capture the additive impact of proximity to multiple facilities.
e. Steps a-d are repeated for all specified radii.
f. A CSV file with the cumulative AADT exposure for each Census block at each specified radius is exported.

3. **Post-Processing in R.** After the Traffic Exposure data is created in ArcGIS/Python, a simple R script is used to create Traffic Exposure bands and apply decay weights[9], finalizing the Traffic Exposure data for inclusion in the EQI. First, the script creates individual traffic exposure bands for a given radius by subtracting the cumulative AADT of all narrower radii from a given radii. Once these unique bands are created, each value is decay weighted, and the resulting decay-weighted values are summed to create a final block-level Traffic Exposure value. For this version of the EQI, ten radii were used [500 meters, 450 meters, 400 meters, 350 meters, 300 meters, 250 meters, 200 meters, 150 meters, 100 meters, and 50 meters] and the corresponding inverse-distance decay weights were: [.1, .11, .125, .14, .166, .2, .25, .25, .33, .5, and 1]. Using this approach, traffic exposure occurring closer to the linear source was weighted heavier than traffic exposure occurring further away. Figure 1 shows the relationship between distance and decay weights used for the EQI traffic proximity and a volume indicator. Lastly, a percentile rank value was calculated for each Traffic Exposure score, with any block scoring above 0.8 being screened for inclusion.

[9] Decay weighting is a quantitative method of giving some elements in a set more weight or influence than others, based on certain characteristics. In the EQI, decay weights are representative of spatial distance, discounting elements occurring further from the point being measured.

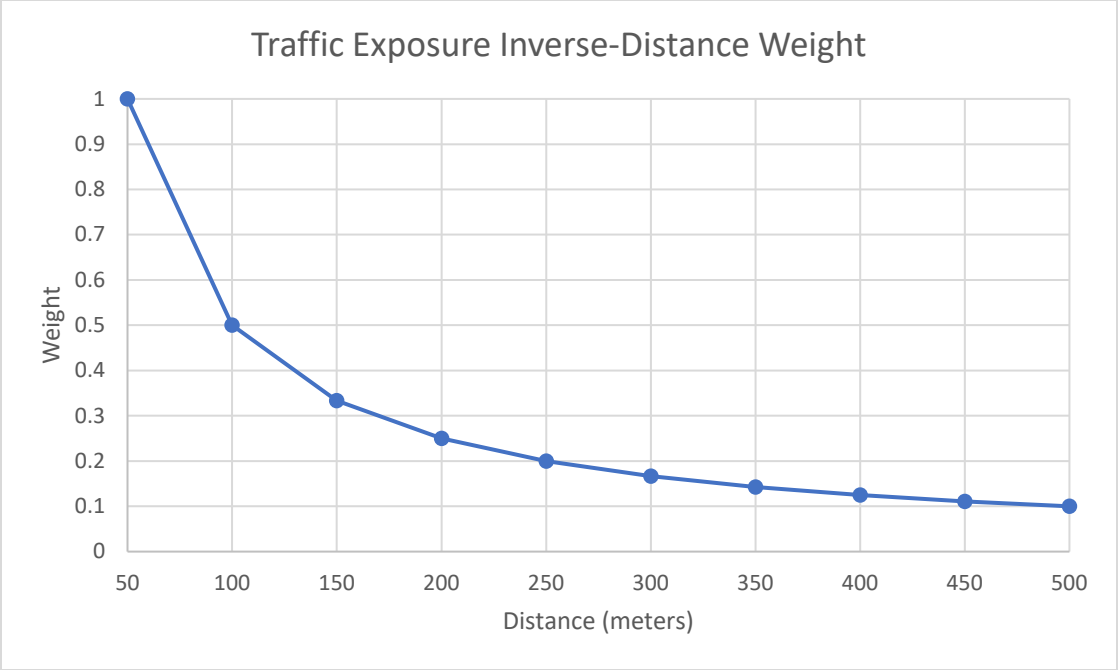

Figure 1. Traffic Proximity and Volume Weight

Crash Exposure Data

1. **Input Data.** The EQI's crash exposure indicator uses data from the Transportation Injury Mapping System (TIMS) developed by UC Berkeley Safe Transportation Research and Education Center (SafeTREC).[10] This data is derived from the Statewide Integrated Traffic Records System (SWITRS), maintained by the California Highway Patrol. For the EQI, a five-year window of the most-recent non-provisional data is used.
2. **Data Cleaning and Weighting.** The EQI crash exposure indicator only considers crashes resulting in injuries or fatalities, so property damage-only crashes are filtered out. Next, each crash was weighted by the highest level of injury in the crash. Weighting factors were derived from the Cal-Benefit-Cost model[11] and are consistent with the costs used in other Caltrans benefit-cost analyses. The following weighting factors were used:
    - Injury (minor): 1
    - Injury (moderate): 1.96
    - Injury (severe): 7.19
    - Fatality: 157.97

    The EQI crash exposure indicator excludes crashes that occur on closed-access highways, as these crashes are less spatially relevant to their

[10] https://tims.berkeley.edu/
[11] https://dot.ca.gov/programs/transportation-planning/division-of-transportation-planning/data-analytics-services/transportation-economics

surrounding communities than crashes occurring on local roads or main street sections of the highway network. State Highway System (SHS) bicycle access status was used to determine highway access status, with crashes being removed from sections of the SHS with prohibited bicycle access.[12] Crashes occurring on highway ramps were kept in the dataset since ramp crashes have a greater impact on local traffic and safety. Figure 2 visually depicts which crashes were removed from or kept within the dataset for the City of Sacramento as an example.

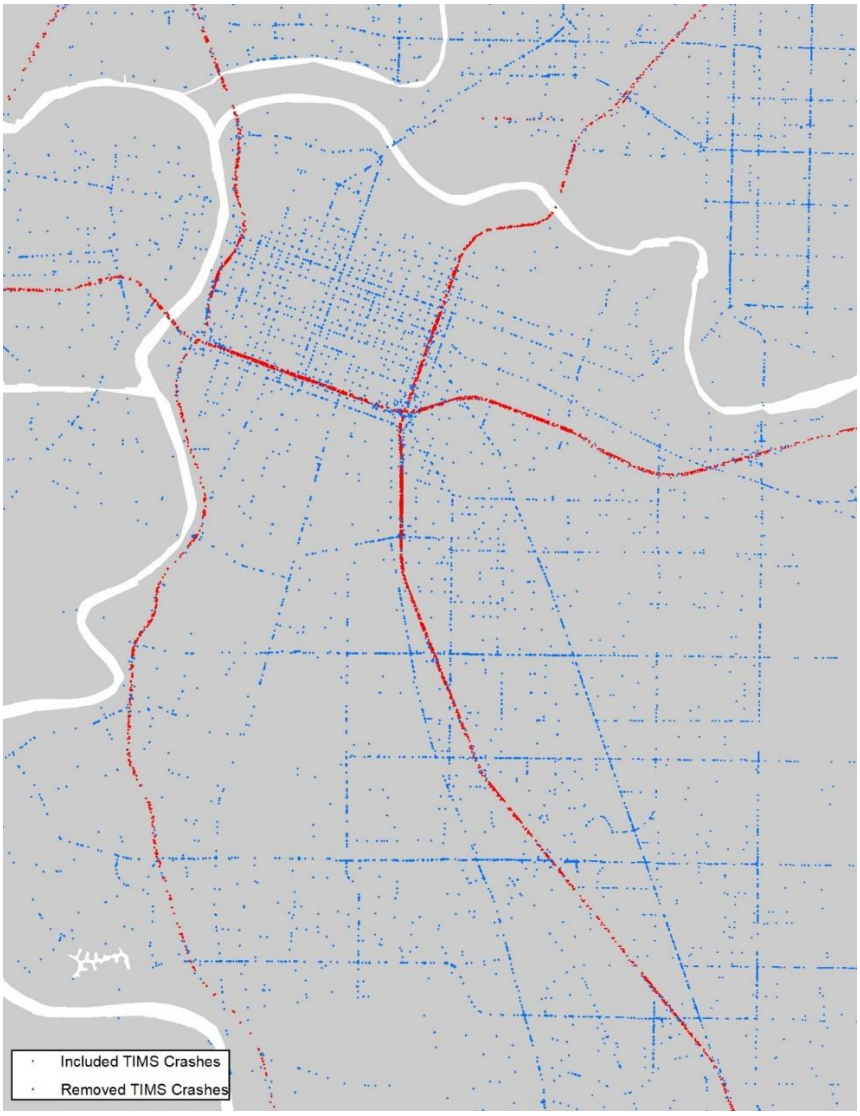

Figure 2. TIMS crash data in Sacramento, CA

[12] https://sv03tmcpo.ct.dot.ca.gov/portal/apps/webappviewer/index.html?id=49cfd2cfa06b4e078d131df264fee437

3. **Data Aggregation.** Once crash data was processed and weighted, crashes were mapped and aggregated to Census blocks. A 250-foot buffer was applied around each Census block, and all crashes within that block (and surrounding buffer) were included in the block score. Since many block boundaries run along roads, crashes occurring along those roads should be associated with all immediately surrounding blocks, not just the block on the side of the road where the crash occurred. A 250-foot buffer was used since it is the approximate diagonal distance across a large intersection and is a standard buffer distance for intersection safety screening.
   The weighted crash values for all crashes occurring in the block and surrounding 250-foot buffer were summed and divided by the area of the Census block (in square miles) to calculate a density score for every Census block in the state. Lastly, a percentile rank was calculated for every Census block with a land area greater than zero and a demographic overlay score. Blocks at or above the 80th percentile were screened for inclusion.

Access to Destinations Indicators

The EQI measures multimodal access to destinations on the Census block level using internal Caltrans tools and analysis methods. The access to destinations score used in the EQI is the ratio of decay-weighted multimodal access to destinations (transit and walking) to decay-weighted auto access to destinations for each Census block. Using this methodology, geographies with comparatively better multimodal access to destinations will have higher ratios. The analysis was run separately for both jobs and non-work destinations. If the threshold was met for either analysis, the block was screened for inclusion. The following steps were used to develop an access to destinations score various input data sources and internal Caltrans tools:

1. **Access to Destinations Calculations.** Caltrans uses the Conveyal accessibility platform[13] to perform access to destinations calculations. The platform uses a routing engine to calculate the number of opportunities that can be reached from a given origin using a given mode, or combination of modes. These reachable opportunities are then decay-weighted based on the time it takes to reach them. Figure 3 shows the exponential decay curves used in the analysis. The sum of decay-weighted cumulative reachable opportunities from a given origin is that origin's accessibility.

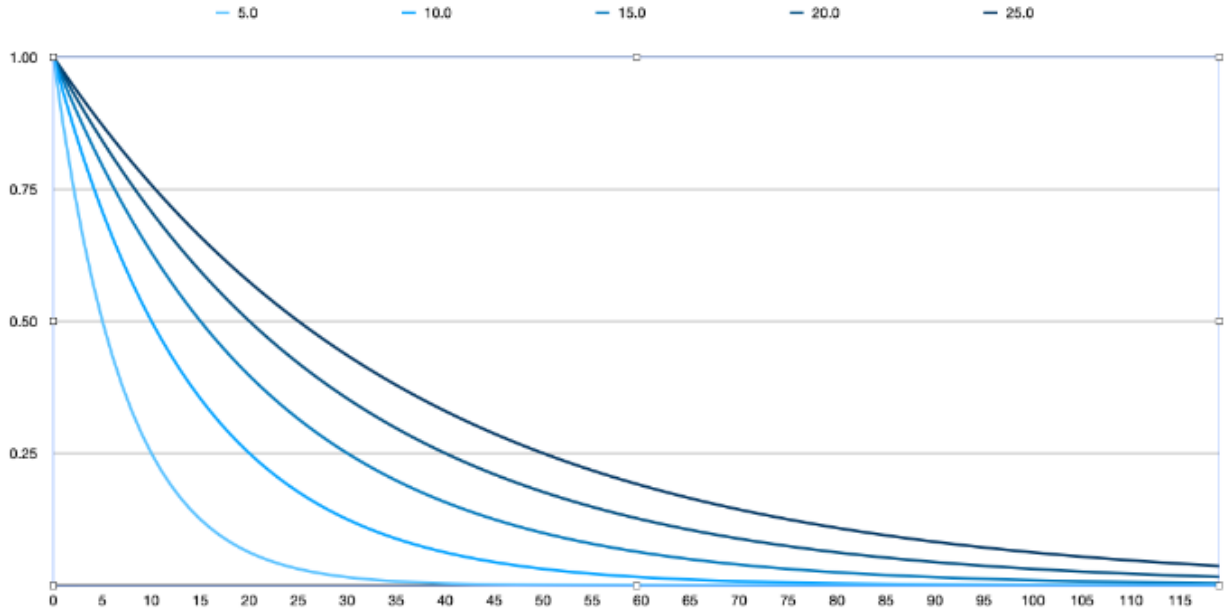

Figure 3. Access to Destinations Exponential Decay Function (Source: Conveyal)

[13] https://conveyal.com/

These calculations are repeated across the state for a series of origin points making up a grid. This process is repeated for both auto and non-auto (transit and walking) modes and for different destination types (work and non-work).

From Conveyal, accessibility scores can be downloaded as raster files, with each grid cell representing the accessibility of a given point. Raster files were then used to interpolate accessibility scores for each Census block in the state, using a methodology detailed in Appendix 3.

2. **Multimodal Access to Destinations Ratio.** Once weighted multimodal and auto access to destinations scores have been calculated for each Census block in the state (or study area), a ratio can be determined by dividing the weighted auto access to destinations score by the weighted multimodal access to destinations score. On a simple level, this score represents a geography's relative multimodal access compared to its auto access.[14]

[14] In most areas so the state, auto access to destinations is quite good. If auto access to destinations is low, this is generally due to a lack of destinations, not a lack of auto infrastructure. That is to say, auto access to destinations is a good proxy for what can be considered good access to destinations in a given region. Using this approach, a dense walkable downtown core in a small rural town in the north state would have a similar access to destinations score to that of downtown San Francisco or Los Angeles, despite a large difference in the actual number of jobs that can be reached in the two place types. The metric, in this case, is relative.

EQI Screening Thresholds

Once the three indicator datasets were created, they were merged into one dataset using the shared block 'GEOID20' variable. Before screens were applied, all Census blocks with a land area of 0 were removed from the dataset. Using this combined dataset, screening thresholds were applied to create the EQI's three distinct screens.

Demographic Overlay

The Demographic Overlay contains a screen for both income and race/ethnicity.[15] If the criteria for either-indicators were met, the block was screened for inclusion and was further analyzed with the traffic exposure and access to destinations indicators.

Traffic Exposure Indicator

The Traffic Exposure indicator includes a measurement for both traffic proximity and volume as well as crash exposure. The Traffic Exposure indicator identifies blocks at or above the 80th percentile[16] of traffic proximity and volume and/or the statewide crash exposure calculation.

Access to Destinations Indicator

The Access to Destinations indicator provides a ratio of the multimodal access to destinations (transit and walking) to automobile access to destinations. Census blocks with an Access to Destinations score less than or equal to 0.2 for either work or non-work destinations are screened as having poor relative multimodal access to destinations.

[15] These screens are detailed on pages 7-8 of this document.

[16] This percentile is only measured for blocks that have AADT exposure and or crash exposure (by being within 500 meters of the SHS and NHS or within 250 ft of a crash location).

Screening Scenarios

Three distinct screening scenarios are envisioned for the EQI, each intended for different uses. Each of the three screening scenarios includes the demographic overlay by default.

Traffic Exposure Screen

The Traffic Exposure screen shows which areas of the state are most impacted by traffic and the negative externalities it produces (noise, pollution, etc.), as well as crash exposure. Figure 4 shows a map of the traffic exposure screen as currently operationalized.

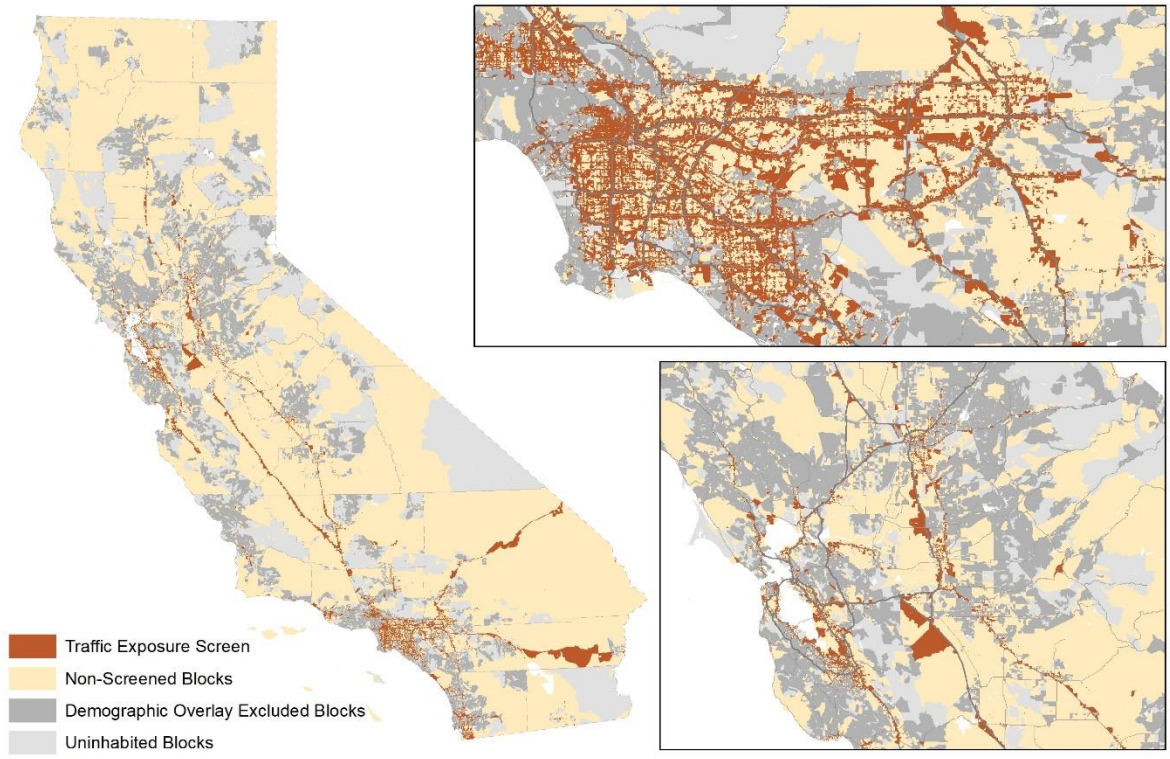

Figure 4. Traffic Exposure Screen

Access to Destinations Screen

The Access to Destinations screen shows which areas of the state have poor multimodal access to destinations and may disproportionately suffer from the negative impacts of having poor access, including but not limited to spending large amounts of money on owning and maintaining a vehicle. Figure 5 shows a map of the Access to Destinations screen.

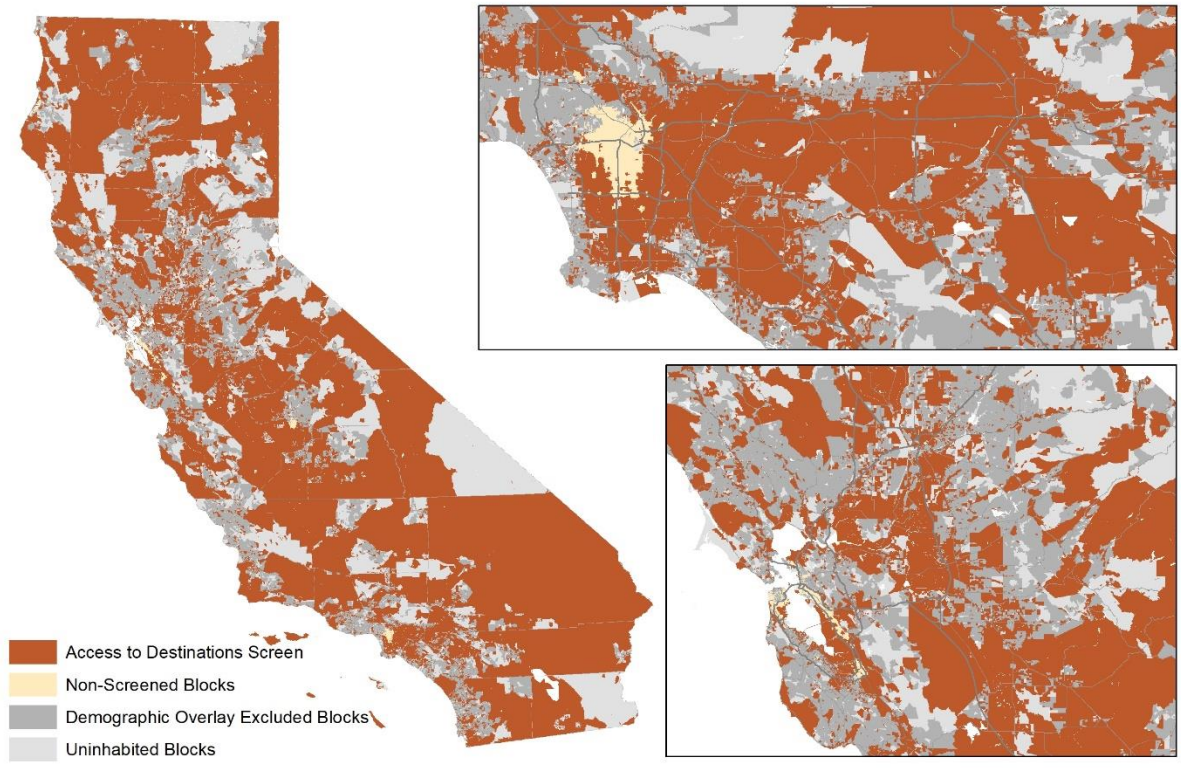

Figure 5. Access to Destinations Screen

Priority Populations Screen

The Priority Populations screen includes all previous screens and identifies the areas of the state that are the most impacted by traffic (traffic proximity and volume and crash exposure) and that benefit the least from the multimodal transportation network (multimodal access to destinations). Figure 6 shows a map of the Priority Populations screen.

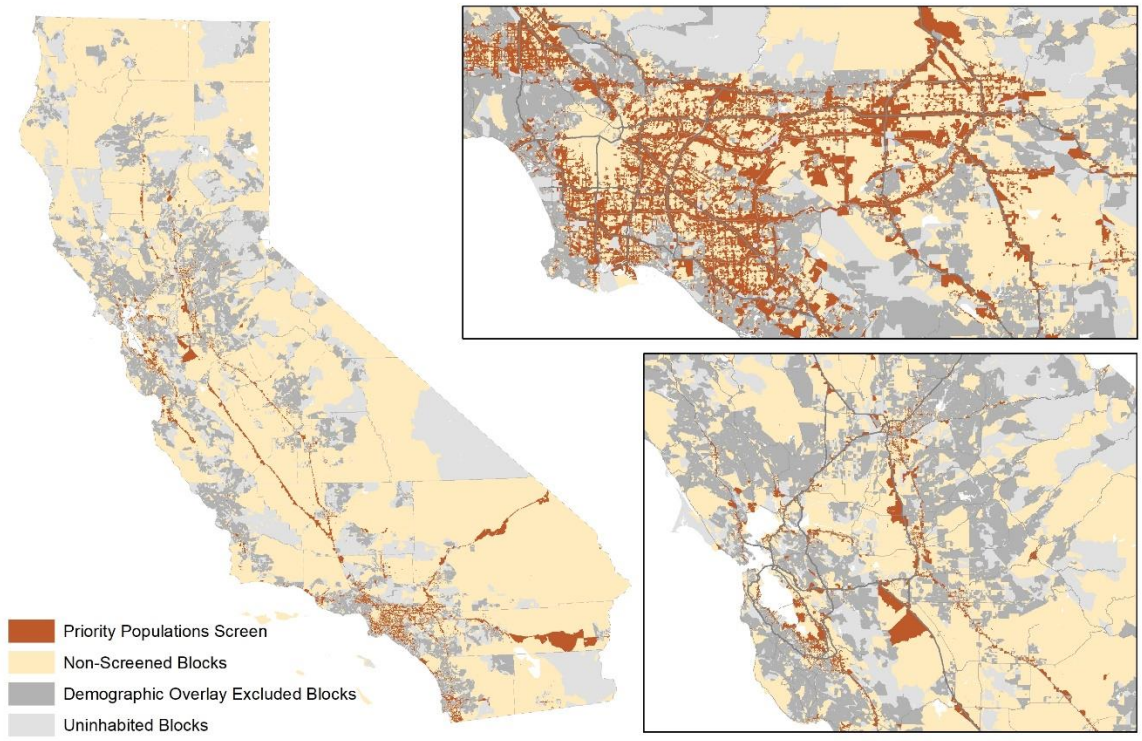

Figure 6. Priority Populations Screen

Lastly, figure 7 shows the relationship between the EQI's indicators, demographic data, and three screening scenarios.

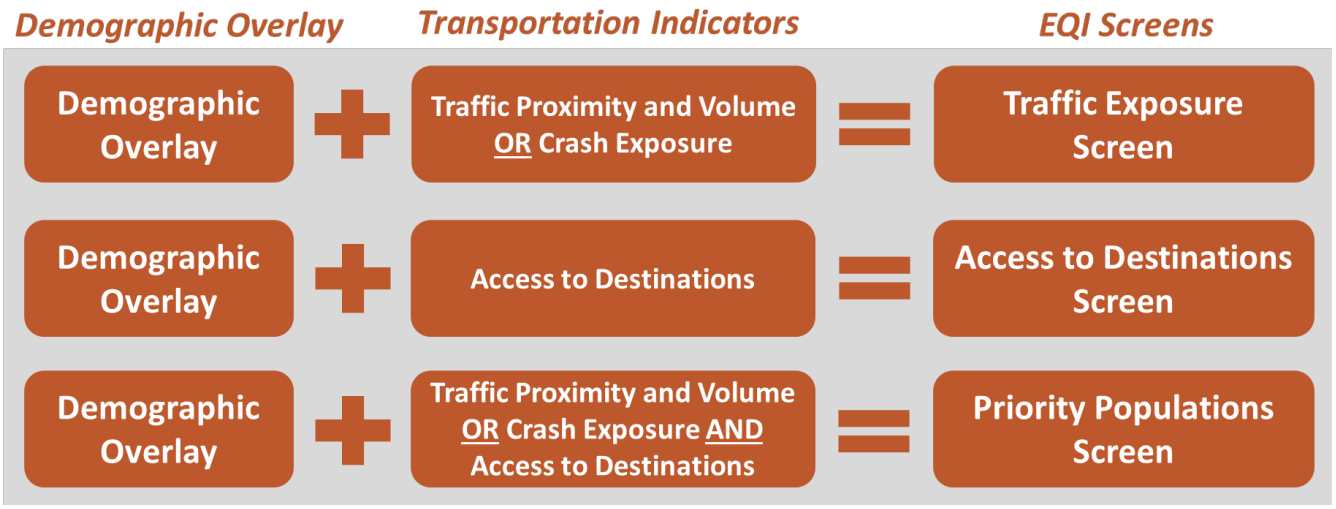

Figure 7. EQI Indicators and Screens

Appendix 1: EQI Indicator Maps

Traffic Proximity and Volume

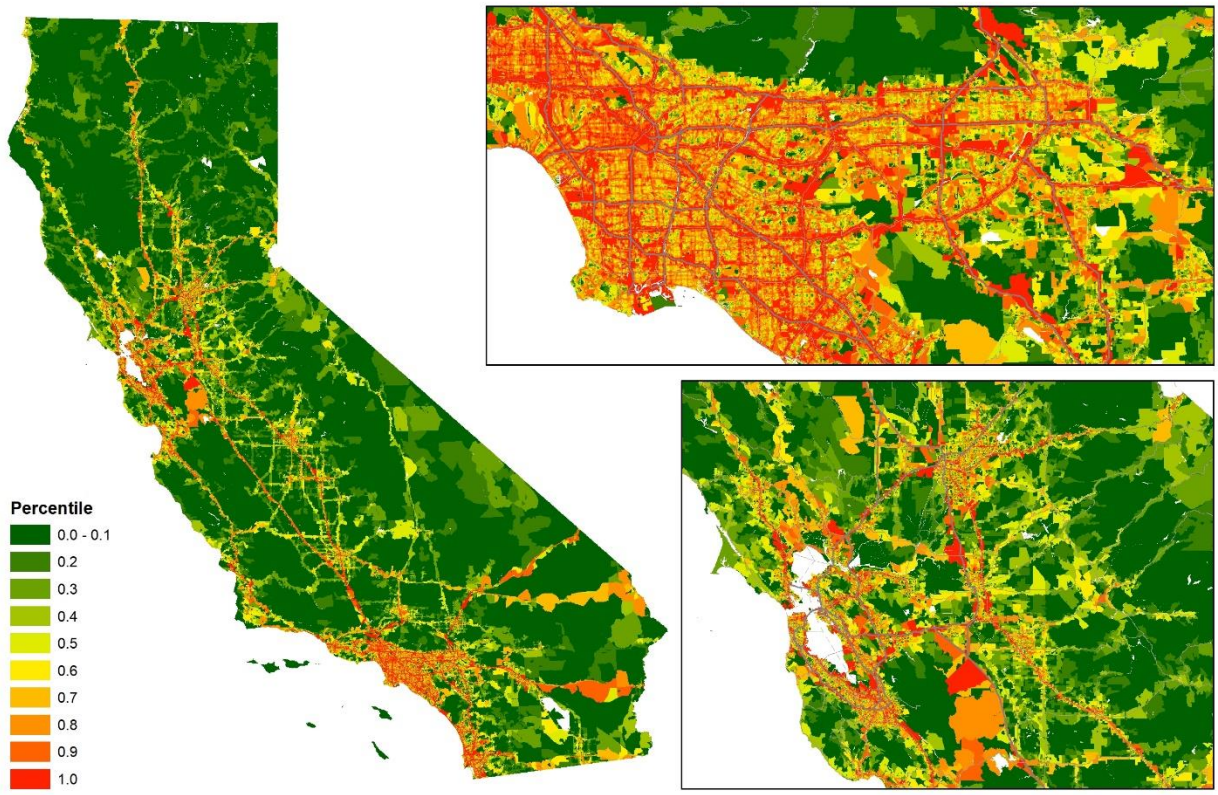

Figure A1-1. EQI Traffic Proximity and Volume Indicator

Crash Exposure

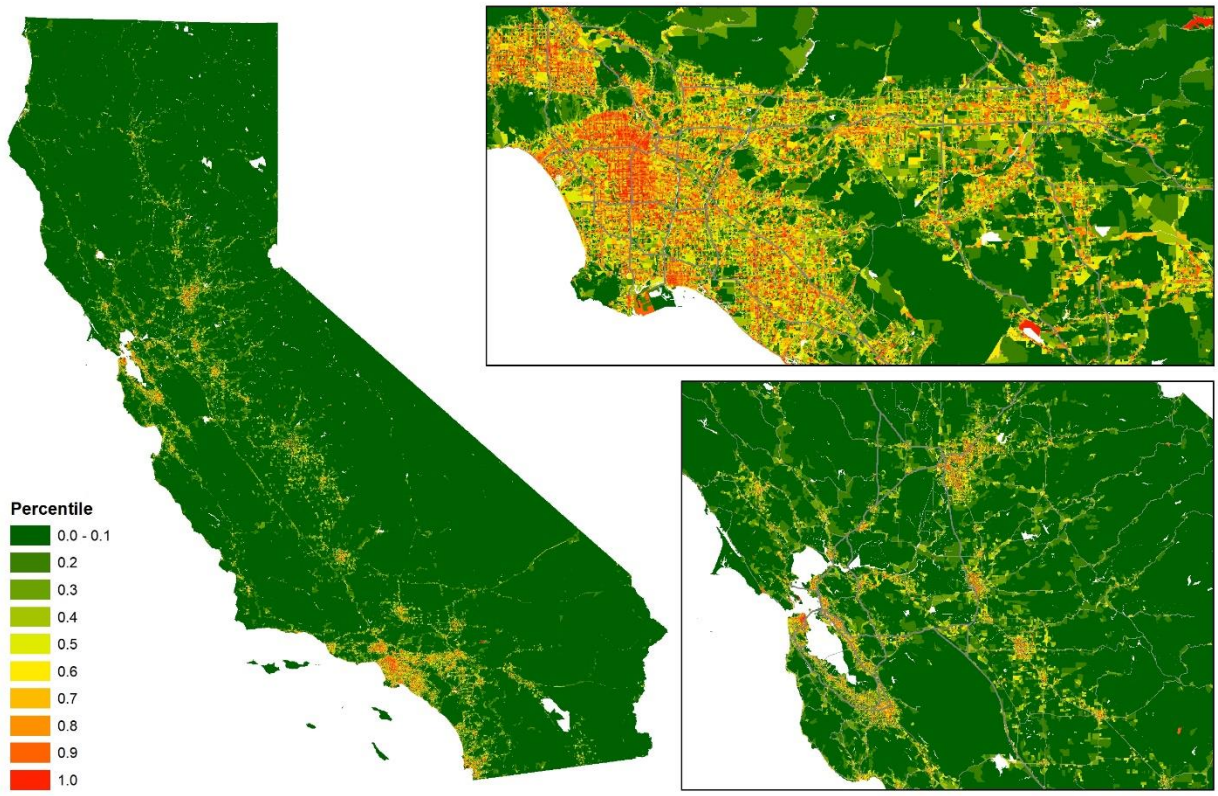

Figure A1-2. EQI Crash Exposure Indicator

Work Access to Destinations

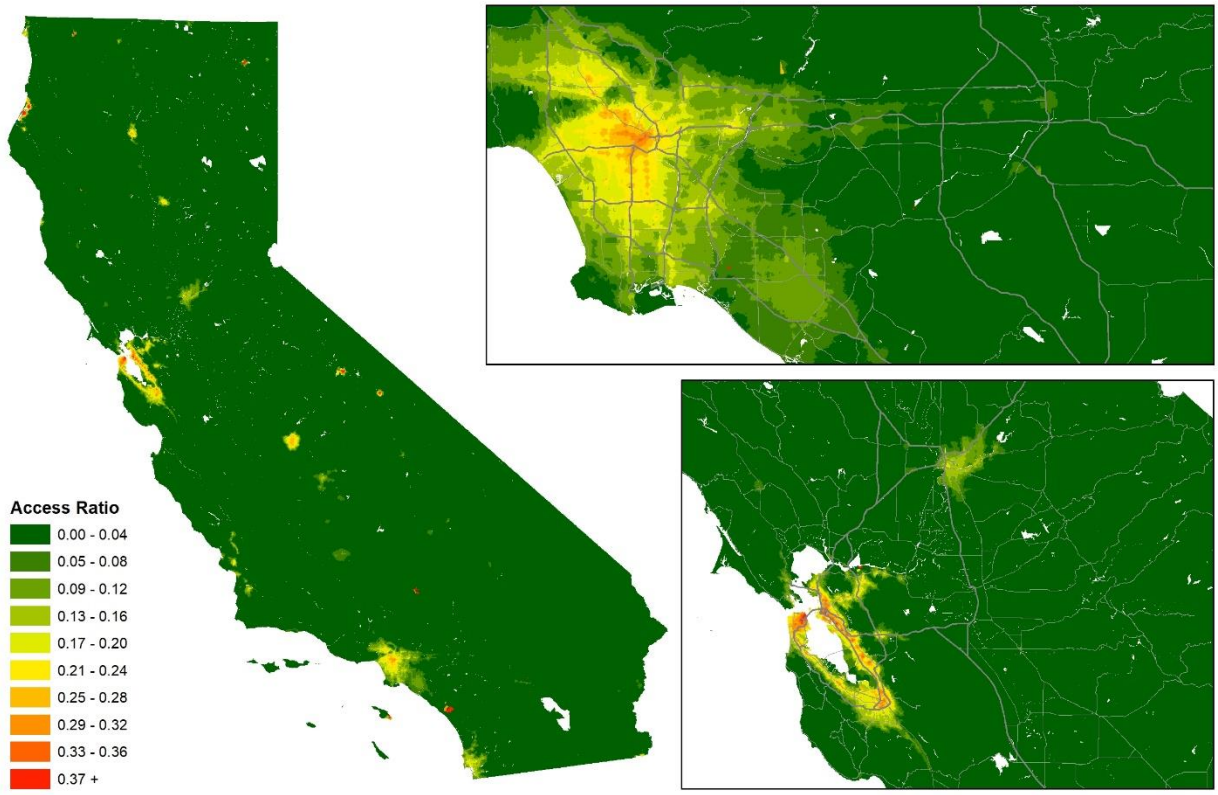

Figure A1-3. EQI Work Access to Destinations Indicator

Non-Work Access to Destinations

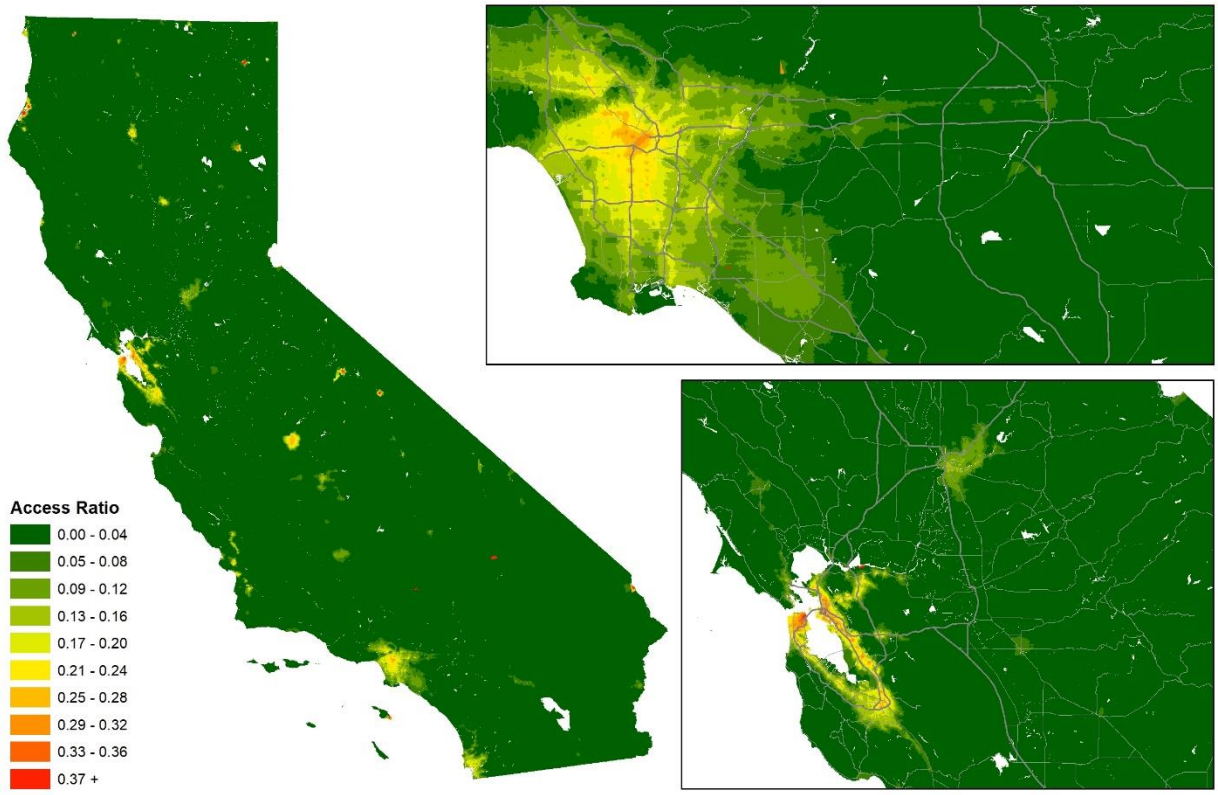

Figure A1-4. EQI Non-Work Access to Destinations Indicator

Appendix 2: Non-Work Access to Destinations Categories

Category	Destination Type	Count	Status	Source
Medical Service	Healthcare	57,823	Core	HERE
Grocery Store/Convenience Store	Grocery	24,390	Core	SNAP
School	Education	15,543	Core	HERE
Park/Recreation Area	Recreation	8,928	Core	HERE
Pharmacy	Healthcare	5,874	Core	HERE
Post Office	Public Service	1,848	Core	HERE
Government Office	Public Service	1,784	Core	HERE
Library	Public Service	1,145	Core	HERE
Hospital	Healthcare	700	Core	HERE
Higher Education	Education	532	Core	HERE
City Hall	Public Service	429	Core	HERE
Court House	Public Service	244	Core	HERE
Civic/Community Centre	Public Service	181	Core	HERE
Restaurant	Food & Drink	79,208	Other	HERE
Specialty Store	Shopping	27,464	Other	HERE
Place of Worship	Other	18,522	Other	HERE
Clothing Store	Shopping	15,664	Other	HERE
ATM	Bank	12,632	Other	HERE
Petrol/Gasoline Station	Shopping	9,531	Other	HERE
Bank	Bank	8,510	Other	HERE
Coffee Shop	Food & Drink	7,496	Other	HERE
Consumer Electronics Store	Shopping	6,577	Other	HERE
Home Improvement & Hardware Store	Shopping	6,108	Other	HERE
Home Specialty Store	Shopping	5,114	Other	HERE
Sporting Goods Store	Shopping	3,971	Other	HERE
Office Supply & Services Store	Shopping	3,788	Other	HERE
Department Store	Shopping	3,388	Other	HERE
Shopping	Shopping	1,341	Other	HERE
Bookstore	Shopping	584	Other	HERE

Appendix 3: Raster to Census Block Interpolation Methodology

Access to Destinations calculations are performed and aggregated using a statewide grid made up of cells measuring 216 by 216 meters. For this data to be used in the EQI, it must be interpolated to Census block geometries. Figure X shows data for non-auto non-work access to destinations as both raw raster output and interpolated Census block data.

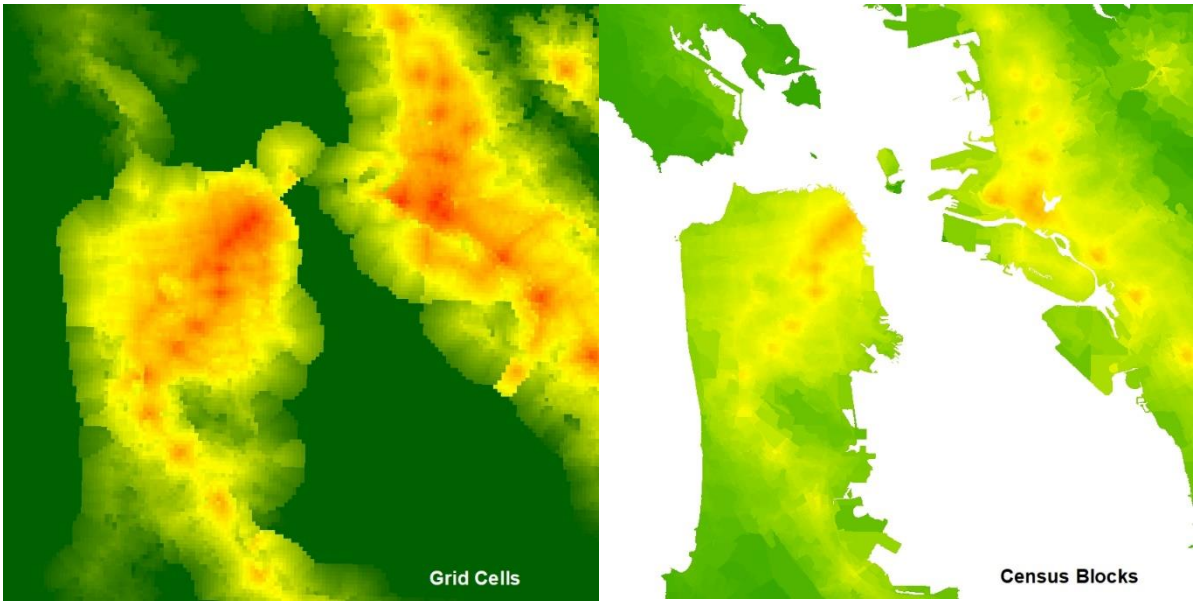

Figure A3-1. Raster to Census block data conversion maps.

The following steps are used to interpolate access to destinations data from raw raster output to Census block geometries:

1. Resample raster cells to a finer-grained resolution. This analysis resamples raster cells to 5 by 5 meters, from 216 by 216 meters. This is necessary to ensure that there are enough raster cells within each Census block to calculate summary statistics. However, resampling raster cells to a finer-grain resolution significantly increases computation time.
2. The ArcGIS Zonal Statistics as Table tool is used to calculate the mean access to destinations score of all raster cells that fall within each Census block. Figure X. shows the ArcGIS work flow used to process raster inputs.
3. This table is attached to the 2020 Census blocks dataset.
4. The process is repeated for four different types of access to destinations (auto work, auto non-work, multimodal work, and multimodal non-work).

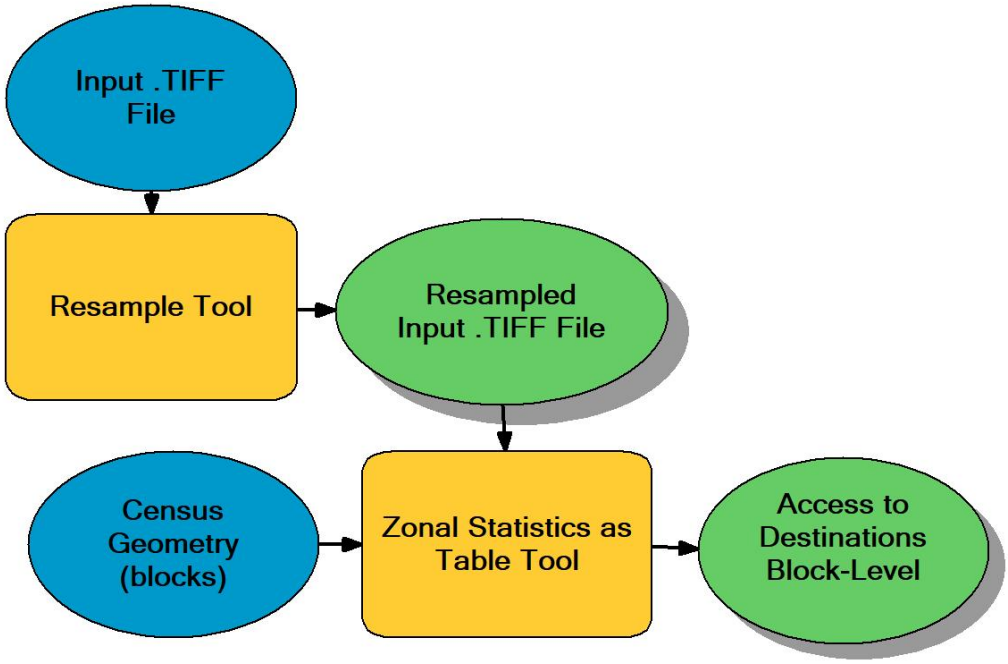

Figure A3-2. Raster to Census block data conversion workflow.