Data Sources & Methods

The California Tobacco Health Assessment Tool (CTHAT) was developed by the Stanford Prevention Research Center and GreenInfo Network, with funding from the Tobacco-Related Disease Research Program grant #22RT-0142, PI: Lisa Henriksen, PhD. Last updated April 2024.

Sources

Licensed Tobacco Retailers: California Department of Tax and Fee Administration (CDTFA) www.cdtfa.ca.gov/. The map includes 29,481 retailers with valid addresses from a list of tobacco retail licenses obtained in 2023.

Adult Smoking Rates: California Health Interview Survey (CHIS). These statistics are point estimates and 95% confidence intervals for adult (18 yrs and older) current (past 30 day) smokers. Estimates are reported by county and statewide. Counties that are small in population size are computed by combining data from multiple counties, but are still reported per county. Grouped Counties:

- Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
- Colusa, Glenn, Tehama
- Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity
- Napa, Sonoma
- San Benito, Santa Cruz
- Santa Barbara, Ventura
- The remaining 35 counties estimates are county specific.

School Boundaries, Centroids, University Lands, Community Colleges: California School Campus Database 2021

Local Policies: California Tobacco Control Program's Policy Evaluation Tracking System, in partnership with Americans for Nonsmokers' Rights Foundation, Policy data as of January, 2024 Source: American Nonsmokers' Rights Foundation

Social Vulnerability Index: Agency for Toxic Substances and Disease Registry, Social Vulnerability Index (SVI) 2020 SVI indicates the relative vulnerability of every U.S. Census tract. SVI ranks the tracts on 16 social factors, including unemployment, racial and ethnic minority status, and disability, and further groups them into four related themes. CTHAT displays the overall ranking for each Census tract.

Demographics: American Community Survey 2018-2022 5-year estimates
The following fields are used:
- Population: B01001_001E
- Under 5: B01001_003E + B01001_027E
- Age 5 to 17: (B01001_004E to B01001_006E) + (B01001_028E to B01001_030E)
- Age 18 to 19: B01001_007E + B01001_031E

- Under 21 (Age 0 to 20): (B01001_003E to B01001_008E) + (B01001_027E to B01001_032E)
- White, not Hispanic or Latino: B03002_003E
- Hispanic/Latino, of any race: B03002_012E
- Asian/Pacific Islander, not Hispanic or Latino: B03002_006E + B03002_007E
- Black, not Hispanic or Latino: B03002_004E
- All Other Races, not Hispanic or Latino [American Indian, Some other race, Two or more races]: B03002_005E + B03002_008E + B03002_009E
- Households: B19001_001E
- Median Household Income (block group, tract, city/place, county, ZCTA): B19013_001E
- Median Household Income (unincorporated and half mile areas): B19001_001E to B19001_017E
- SNAP Benefits: B22010_001E, B22010_002E, B22010_005E

After the removal of block groups that have no land (water areas) and no population, demographic attributes are displayed by quantiles rounded to one decimal place.

School Districts (Unified, Secondary, Elementary): TIGER 2022

City, Designated Place, Zip Code Tabulation Area, Tract, Block Group: TIGER 2022

Parks and Open Space: California Protected Areas Database www.calands.org (CPAD 2022b), December 2022

DATA	SOURCE	DATE
Licensed Tobacco Retailers	CA Department of Tax and Fee Administration	12/2023
Adult Smoking Rates CA	Health Interview Survey	2016-2018
Schools	CA School Campus Database & CA Department of Education	2021
Tobacco Retail Licensing	California Policy Evaluation Tracking System (PETS)	01/2024
Flavor Sales Restrictions	California Policy Evaluation Tracking System (PETS)	01/2024
Smokefree Multi-Unit Housing (MUH)	California Policy Evaluation Tracking System (PETS)	01/2024
Cigarette Min Price and/or Pack Size	California Policy Evaluation Tracking System (PETS)	01/2024
Cigar Min Price and/or Pack Size	California Policy Evaluation Tracking System (PETS)	01/2024

Demographics	American Community Survey	2018-2022
City, County	TIGER	2022
School Districts	TIGER	2022
Block Group, Tract, ZCTA	TIGER	2022
Parks and Open Space	CA Protected Areas Database	12/2022

Methods and Processing

Tobacco Retailers:

CDTFA locations were geocoded using ArcGIS World Geocoding Service. Matched records = 29,481 Unmatched - 0. All stores with a match confidence less than 90% were reviewed for accuracy and improved as needed. The final match rate was 100%.

Vape shops are included only if they had a state tobacco retail license at the time the source list was obtained.

Retailers are each assigned a County, ZCTA, City/Designated Place, and Unincorporated County for analysis purposes. For retailer counts, retailers are only reported in the geography which contains the point location.

Demographics:

When viewing the map (outside of a search), demographic data is shown at a tract level up to zoom level 12 (about 1:150,000). At zoom level 13 (about 1:70,000) demographics are shown at a block group level. When demographic details are displayed as a result of clicking the map, the same selection of tracts and block groups is used. The change in geography provides a balance of site performance and detailed attribute reporting.

Demographics for counties, cities, designated places, and ZCTAs were extracted from the Census API available from the Census Bureau.

Demographics for school selections and half-mile buffers are estimated based on the census block groups intersecting the displayed area. For statistics other than median household income, this is a weighted proportion based on the population of each block group and the area which intersects the block group. The income figure is the median of block groups with their centroid within the buffer. These data should be interpreted with caution.

Legend breaks for demographics are quartiles. If a user performs a search, the system allows for the selection of statewide or county quartiles. When county quartiles are used, the system:

- Searches for the county, city, zip code, school, address, or latitude and longitude
- Looks up the county the searched area is in*
- Filters demographic data to only those block groups in the relevant county
- Looks up the county quartiles of the selected demographic variable
- Zooms to the selected location, draw census block groups with county quartiles

*Some zip code, address, and latitude/longitude searches will return a location that spans more than one county. In these instances, demographics are only displayed for one county. It is recommended that users consider reverting to statewide quartiles in these cases.

ZIP Codes and ZIP Code Tabulation Areas (ZCTAs):

ZIP Codes are mail routes and do not have an associated polygon geometry. ZCTAs are a generalized representation of service areas. More information and resources are available at: www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html ZIP Codes are a trademark of the US Postal Service, ZIP Code Tabulation Areas are a trademark of the United States Census Bureau.

The ZIP Code of each retailer from CDTFA was compared to the TIGER ZCTA layer. Some retailers had a different ZCTA than reported by CDTFA. This may be due to the difference between ZCTA and ZIP Codes or differences in the accuracy of the ZCTA layer from TIGER. All retailers were assigned to the ZCTA they fall within for spatial consistency. Retailers not within a ZCTA were assigned the ZIP Code given by CDTFA.

Unincorporated areas:

Unincorporated areas were created by subtracting the incorporated cities (MTFCC = G4110) from each respective county. Designated places are included in unincorporated geographies. Statistics are not reported from the Census Bureau for unincorporated areas, the estimates on CTHAT are the county totals minus incorporated city figures. Caution should be used when referencing the household income for these areas, as they are estimated using the available information.

Public Schools:

Includes tribal education facilities listed by the Bureau of Indian Education, December 2020 bie.edu/Schools/

Schools offering mixed or combined levels of elementary, middle and high are included in the highest level.

- Elementary: Elementary
- Middle: Elementary-Intermediate/Middle/Junior High, Intermediate/Middle/Junior High
- High: High School, Elementary-High Combination, Middle/High Combination

The CSCD data set was developed from the California Department of Education public school list as of November 2020 cde.ca.gov/ds/si/ds/pubschls.asp. Extensive research was done to improve and expand the geographic representation of schools, including: corrections to erroneously located schools, development

of school property as a boundary, corrected attributes, and inclusion of all known locations when schools operate at multiple locations. Not all school types are included, users can review the dataset user manual for more details.

Schools are each assigned a County, ZCTA, City/Designated Place, and Unincorporated County for analysis purposes. If the 500/1000 foot buffer of the school spans multiple geographies, it is only assigned to the geography which contains the school campus. For example, there are schools near the Los Angeles/Orange County border where the buffer extends into the neighboring county. If the school is located in Los Angeles County the entire buffer is assigned to Los Angeles County, and any retailers in that buffer are counted.

Retailer/School Proximity Methods:

Starting in 2019-2020 the data on school and retailer proximity in the CTHAT application is generated through a series of scripts. Below is an outline of the manual steps one can follow to achieve the same results.

Each retailer is reviewed to see if it is within 500/1000 feet of any school (Euclidean distance)

Each school is reviewed to see if it has retailers within 500/1000 feet of its campus (per CSCD boundary)

Use the "Select by location" tool, to determine the retailers within 500/1000 feet of a school, with the following settings:

- Input Features: Tobacco Retailers (point)
- Selecting Features: School Boundary (polygon)
- Spatial selection method: are within a distance of the input layer feature - 500/1000 feet

The selected features can then be summarized for the reported geographies: California, City/Census Designated Place, Unincorporated, ZCTA

To determine the schools with a tobacco retailer within 500/1000 feet, use the settings:

- Input Features: School Boundary (polygon)
- Selecting Features: Tobacco Retailers (point)
- Spatial selection method: are within a distance of the input layer feature - 500/1000 feet

The selected features can then be summarized for the reported geographies: California, City/Census Designated Place, Unincorporated, ZCTA

Each school and retailer is assigned only one: County, City/Census Designated place, Unincorporated, ZCTA. If a school buffer crosses a geographic boundary, it is only counted in the geography which contains the school boundary. If a school boundary itself crosses a geographic boundary, the centroid is used.

Median Household Income:

Median Household Income is only available at predefined TIGER geometries (block group, tract, city/designated place, county, ZCTA).

Unincorporated figures are estimated by using available income range data (B19001_001E to B19001_017E). By determining the midpoint household ([total households]/2) one can define the income range the median household value is between (for example the range of $40,000 to $49,999). To narrow the range to a value, equal distribution of households within the range is assumed. Detailed methods are described by the California Department of Finance in their publication: [Re-calculating Medians and their Margin of Errors for Aggregated ACS Data (January 2011)](#)

The value reported for a half mile buffer is the median of all block groups with their centroid inside the buffer area.

Parks and Open Space:

Park and open space land are from the California Protected Areas Database (CPAD), 2022b release.