

California Household Travel Survey 2012: Tour mode choice data from the San Francisco Bay Area

This data was originally collected as part of the California Household Travel Survey (CHTS) in the year 2012. Individuals belonging to sampled households were asked to report their complete activity diary data over an observation period of one day, including which activities were conducted where, when, for how long, with whom and using what mode of travel. More information on the raw data can be found in [NuStats, LLC \(2013\)](#).

The data included here corresponds to individuals from the subset of households located in the [nine-county San Francisco Bay Area](#). The raw trip data was processed into home-based tours that can be used for the purpose of tour-based travel mode choice analysis. The resulting dataset includes 27,054 tours made by 17,717 individuals from 8,228 households.

For each tour, six possible travel mode alternatives are defined: private vehicle, private transit, walk to public transit, drive to public transit, bike, and walk. Private vehicle refers to cases where the individual used a motorized vehicle owned by themselves (or someone they know) as a driver or a passenger. Private transit includes the use of travel modes such as taxis, Uber, carshare, rental cars and private shuttles. Walk to public transit captures all cases in which an individual only used non-motorized travel modes to access public transit, and drive to public transit captures all cases in which a motorized travel mode was used to access public transit.

The level-of-service attributes, namely travel times and costs, for each of the six travel modes for each tour are determined using [network skims from the SF MTC for 2010](#), generated using version 3 of their travel demand model. We are unable to decompose travel time into its constituent elements, such as in-vehicle time and waiting time, as this information was unavailable at the time of processing. Travel costs are in 2000 US dollars.

The data folder contains four files: (1) `chtsTMC`: the processed data file; (2) `Tour generation`: an html version of the Python script used to process the raw data; (3) `Data analysis`: an html version of a Python script included as an example on how to use the data file for analysis; and (4) `caltrans_data_dictionary`: data dictionary of the raw CHTS files, containing additional information about each of the variables

A subset of this data was originally used by [Vij et al. \(2017\)](#) for understanding modal preference shifts in the San Francisco Bay Area over time. For more details, please refer to the original study. And if you have any questions, feel free to contact Akshay.Vij@unisa.edu.au.

References

Nustats, LLC, 2013. 2010–2012 California Household Travel Survey Final Report.

Vij, A., Gorripathy, S., & Walker, J. L. (2017). From trend spotting to trend'splaining: Understanding modal preference shifts in the San Francisco Bay Area. *Transportation Research Part A: Policy and Practice*, 95, 238–258.