Final project proposal

Kimberly Nicholls
1/11/2026

Problem statement

Is Bay Area traffic getting worse?

The total population of the Bay Area has increased by 13% from 2000 to 2019 according to census data (2019 ACS B01003_001 and 2000 decennial P001001). How did that impact the time people spent commuting to work? How is commute time associated with urbanization? This may include factors such as distance from urban cores, population density, etc.

This project will use data from the 2000 decennial census and the 2019 ACS to create two maps, one for each year. Each map will show the mean commute time by census block, as well as urbanization status from the Census Bureau's urban-rural classification.

Study Area

Each map will include the nine Bay Area counties: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, Sonoma.

Because census tract boundaries may change over time, the data for 2000 and 2019 will be displayed on separate maps.

Unit of Analysis (UA)

The maps will display data at the census tract level. Census tracts are designed to have approximately equal population, which makes them useful in several ways.

Mapping census tracts makes it easy to visually identify population density (smaller tract areas are more dense than larger), and enables quickly comparing dense areas with less dense areas.

Census tracts can illustrate variability within larger geographic areas such as cities. For example, the city limits of San Jose covers a large and heterogeneous area: downtown San Jose may look different than an outlying neighborhood such as Almaden.

Many other datasets are available at the census tract level that can provide useful context. This project will include urban-rural classification data, available via tigris.

While data is reported at the block group level, this resolution is too noisy to be visually useful. In addition, the smaller sample size comes with a larger margin of error in each bucket. Comparing the 2019 data for Santa Clara county shows a relative margin of error of 93.5% at the block group level, versus 52.6% at the census tract level.

Mapping at the county level is too coarse: Bay Area counties are large in area and heterogeneous in land use. Mapping at the county level could potentially hide interesting features. For example, residents of Santa Clara County include rural areas in the Santa Cruz Mountains, urban areas in downtown San Jose, and suburban areas such as Sunnyvale and Palo Alto. Residents of these areas may have very different commute time experiences that would be invisible on a county-scale map.

Measurement Scales & Variables

2019 map

ACS dataset `acs5` with quantitative values:

- B08303_001 total travel time to work (discrete)
- B08303_001 - B08303_013 travel time to work by time range: "Less than 5 minutes", "5 to 9 minutes", "10 to 14 minutes", etc. (discrete)
- The R script will calculate a weighted average travel time for each tract (discrete)

2000 map

Decennial census dataset `sf3` with quantitative values:

- P031002 total commuters (“Did not work at home”) (discrete)
- P031003 - P031014 travel time to work for workers 16 years and over by time range: "Less than 5 minutes" "5 to 9 minutes", "10 to 14 minutes", etc. (discrete) These time ranges match the values in the 2019 data.
- The R script will calculate a weighted average travel for each tract (discrete)