



# NEIGHBORHOOD EQUITY AND SERVICE REQUESTS

Analyzing the Relationship Between 311 Requests and Household Income in San Francisco

Prepared by

Felix Zhang

[felixzhang04@berkeley.edu](mailto:felixzhang04@berkeley.edu)

3038372755

Leslie Veliz

[lveliz955@berkeley.edu](mailto:lveliz955@berkeley.edu)

3037973499

Manuel Martinez Garcia

[manpazito@berkeley.edu](mailto:manpazito@berkeley.edu)

3038324551

## SEPTEMBER 26, 2025

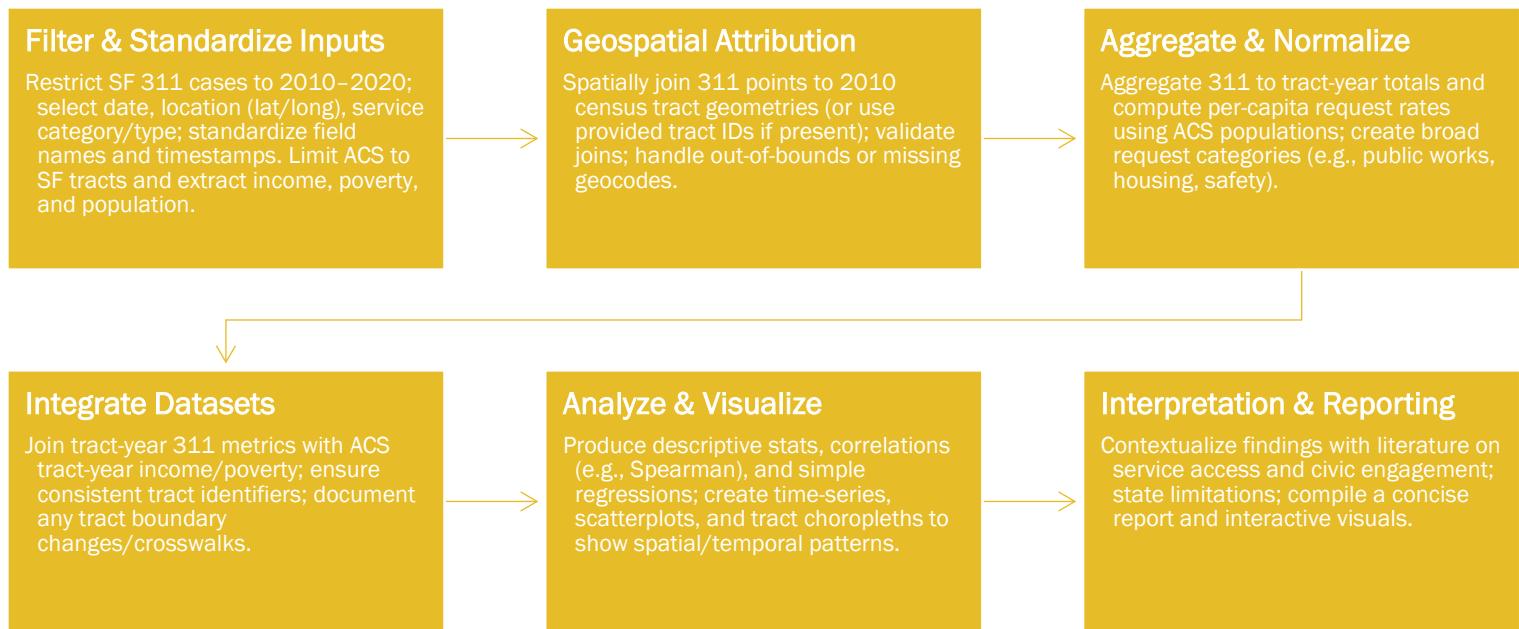


## Research Statement

Our project asks: *What is the relationship between 311 requests and household income at a per capita basis within San Francisco between 2010 and 2020?*

We focus on San Francisco at the census tract level, using data over ten years (2010-2020). The intended outcome is to identify whether neighborhoods of differing income levels show distinct patterns in the frequency and type of 311 requests. By bounding the analysis geographically to San Francisco, we aim to create a quantifiable comparison that can highlight disparities in access to city services.

## Step-by-Step Approach



## Data Description

### Dataset #1: San Francisco 311 Cases

- Coverage: All recorded 311 requests within San Francisco, 2008–2025 (our study uses 2010–2020).
- Granularity: Individual service requests (row = request).
- Type: Administrative/transactional city records (event-based).
- Format: CSV (with geocoded fields).
- Source: DataSF Open Data Portal | <https://data.sfgov.org>

### Dataset #2: American Community Survey (ACS) 5-Year Estimates

- Coverage: U.S. Census tracts; 2010–2020 (SF tracts subset).
- Granularity: Tract-level aggregates (median household income, poverty rate, population).
- Type: Survey-based socioeconomic statistics (rolling 5-year).
- Format: CSV (via API extracts or prepared tables).
- Source: U.S. Census Bureau | <https://www.census.gov/programs-surveys/acs>

## Methodology

We plan to address our question by cleaning and combining two major datasets: the San Francisco 311 cases and the American Community Survey economic data. For the 311 cases, we will extract the specific variables and years relevant to our analysis, then group them into census tract/year combinations. This will allow us to track the number and types of 311 requests within each census tract per year. We will also create new columns to represent these statistics on a per-capita basis, ensuring comparisons are consistent across different parts of the city. Similarly, we will clean the American Community Survey data to include only the census tracts in San Francisco for the years under study.

With both datasets prepared, we will overlay variables to examine relationships between 311 requests and economic indicators such as request type, request volume, median household income, and poverty rate across time.

## Timeline & Milestones

Conduct background research on 311 requests and historical inequities in San Francisco (Felix & Manny)	• Sept 29 – Oct 5
Clean 311 and ACS data to relevant scope (Manny & Leslie)	• Oct 6 – Oct 12
Combine datasets into tract-level metrics; begin early analysis (Leslie & Felix)	• Oct 13 – Oct 19
Midterm Submission: cleaned dataset, one working interactive map, and a 2–3 page report of data preparation and early analysis (All)	• Deadline: Oct 22
Examine variable combinations; create draft visuals across time (Felix & Manny)	• Oct 23 – Nov 9
Develop polished interactive visuals and integrate contextual information (Manny & Leslie)	• Nov 10 – Dec 7
Final Submission: complete deliverables (maps, analysis, final report) (All)	• Deadline: Dec 10

## Potential Challenges

Census tract boundaries sometimes change over time, which can make it harder to compare results across years. To address this, we will limit our study to years between the 2010 and 2020 censuses. Alternatively, results can be analyzed at the council district level as council boundaries are systematically redrawn after each census to reflect population shifts.

The ACS 5-year estimates average data across several years, so they are less precise for tracking year-to-year changes; we will be cautious when interpreting these trends. 311 data also reflect who reports issues, which can be shaped by awareness, language, or internet access. For that reason, we will treat them as measures of reporting behavior, not absolute need, and discuss the equity implications. Finally, because the 311 dataset is very large, we will use efficient tools (most likely DuckDB or pandas) and pre-aggregate data to keep processing manageable.