

| Dataset 1: Daily Census Tract-Level PM2.5 Concentrations, 2011-2014 Provides the modeled predicted mean and associated standard deviation of PM2.5 concentrations in various locations and census tracts. (Source) (Dataset) | |
|---|--|
| Column Name | Description |
| year | Year of estimate |
| date | Day-month-year of estimate |
| statefips | State FIPS Code, uniquely identifies each US State |
| countyfips | County FIPS Code, uniquely identifies each US county |
| ctfips | Census Tract FIPS Code, uniquely identifies each US county subdivision |
| latitude | Latitude of census tract centroid |
| longitude | Longitude of census tract centroid |
| ds_pm_pred | Mean estimated 24-hour average PM2.5 concentration in $\mu\text{g}/\text{m}^3$ |
| ds_pm_stdd | Standard error of the estimated PM2.5 concentration |

| Dataset 2: Daily Census Tract-Level Ozone Concentrations, 2011-2014 Provides the modeled predicted mean and associated standard deviation of Ozone Levels in various locations and census tracts. (Source) (Dataset) | |
|---|--|
| Column Name | Description |
| year | Year of estimate |
| date | Day-month-year of estimate |
| statefips | State FIPS Code, uniquely identifies each US State |
| countyfips | County FIPS Code, uniquely identifies each US county |
| ctfips | Census Tract FIPS Code, uniquely identifies each US county subdivision |
| latitude | Latitude of census tract centroid |

| | |
|------------|--|
| longitude | Longitude of census tract centroid |
| ds_o3_pred | Mean estimated 8-hour average ozone concentration in parts per billion (ppb) within 3 meters of the surface of the earth |
| ds_o3_stdd | Standard error of the estimated ozone concentration |

| |
|--|
| <p>Dataset 3: American Community Survey Census Data Census data for 3142 United States counties spanning from 2000 to 2019 (Source)(Dataset)(Data Dictionary)</p> |
|--|

| <p>Dataset 4: U.S. Chronic Disease Indicators: Asthma Provides 124 indicators of state-based chronic disease data, filtered specifically for asthma (Source)(Dataset) *Relevant columns to project analysis were included below*</p> | |
|--|---|
| Column Name | Description |
| YearEnd | Year |
| LocationAbbr | Location Abbreviation (Postal) |
| Question | Category of Chronic Disease Data |
| DataValue | Value representing the question category for a given location, year |
| Stratification1 | Classification of each row, represented by race/ethnicity or gender |

| <p>Dataset 5: Mapping Inequalities Provides redlining data in the San Francisco Bay Area. This dataset was not used to address the research questions but rather helped us formulate our questions. (Source)(Dataset)</p> | |
|--|---|
| Column Name | Description |
| holc_id | ID given by the Home Owners' Loan Corporation (HOLC) for a given Neighborhood |
| holc_grade | Redlining grade given by HOLC |

| | |
|-----------------------|---|
| area_description_data | Each row contains a dictionary of descriptions for the neighborhoods in the dataset (city name, characterization of the city, etc.) |
| geometry | Coordinates used to create the map, representing a specific area |