

DOCUMENTATION FOR THE CLIMATE READY BOSTON SOCIAL VULNERABILITY DATASET

Overview

This document describes the structure and contents of an updated Climate Ready Boston (CRB) Social Vulnerability dataset, which examines the distribution of socially vulnerable populations across 180 Census tracts and 23 neighborhoods in the City of Boston. Social vulnerability is defined as “the disproportionate susceptibility of some social groups to the impacts of hazards, including death, injury, loss, or disruption of livelihood.” Data are sourced principally from 2008–2012 American Community Survey (ACS) five-year estimates, with one exception (the *MedIllnes* variable) described below.

The analysis and definitions used in Climate Ready Boston (2016) are based on “A framework to understand the relationship between social factors that reduce resilience in cities: Application to the City of Boston,” published in the *International Journal of Disaster Risk Reduction* by Atyia Martin, Northeastern University, 2015.

Additional datasets are linked to the original CRB dataset to assist in different analyses. Specifically, the Boston Area Research Initiative (BARI) 311 Requests data between 2015-2019, that looks at the types of calls received by the City of Boston’s 311 system, is used to add a construct related to weather conditions to the Social Vulnerability dataset. In addition, variables in the Census tracts data of the BARI Geographical Infrastructure for Boston (2017) are utilized for some preliminary inferential statistics.

Table of Contents

1. Summary of Social Vulnerability Data (<i>CRB_SocialVulnerability_Updated.csv</i>).....	2
1.1 Description of Contents	2
1.2 Description of Variables	2
1.2.1 Initial Report Content.....	2
1.2.1.1 General Tract Information	2
1.2.1.2 Raw Population Counts.....	3
1.2.2 New Variables	4
1.2.2.1 Proportion and Density Measures	4
1.2.2.2 Risk Categorization.....	4
1.2.2.3 Updated Measures.....	5
1.2.2.4 Mobility	5
1.2.2.5 Snow Removal	5
1.2.2.6 Linked Data.....	5
1.2.3 Aggregate Measures	6
1.2.3.1 Neighborhood Level	6
1.2.3.2 City Level	6

1. SUMMARY OF SOCIAL VULNERABILITY DATA (*CRB_SocialVulnerability_Updated.csv*)

1.1 Description of Contents

The Social Vulnerability dataset contains the counts of socially vulnerable populations for the 180 2010 U.S. Census tracts that comprise the City of Boston. The data itself is constructed predominantly from the 2008-2013 American Community Survey five-year estimates from the Census Bureau. The updated dataset includes percentages of vulnerable populations by population and acreage and new metrics at the tract level, borrowed variables from linked data, and some original constructs related to mobility and snow conditions.

1.2 Description of Variables

The dataset has 59 variables grouped into multiple files, with three broad categories and 10 more specific sub-categories:

1. Initial Report Content, or information included in the original dataset file
 - i. General Tract Information describes basic information associated with each tract, or record, in the dataset
 - ii. Raw Population Counts describes the counts of selected demographic populations
2. New Variables, record-level variables added to the dataset file
 - i. Proportion and Density Measures describes the representation of selected demographic populations as percentages of total population or acreage of a tract
 - ii. Risk Categorization includes information to describe a given tract's significant vulnerability, based on demographic groups
 - iii. Updated Measures describes variables in original dataset with modified values
 - iv. Mobility includes the different variables used to measure the ease of movement within a tract
 - v. Snow Removal includes measures of snow-related 311 calls by tract
 - vi. Linked Data includes variables from additional datasets that are merged with our original CRB data and used for some inferential analyses
3. Aggregate Measures gives an overview of variables measured at aggregate levels
 - i. Neighborhood-level aggregates
 - ii. Citywide aggregates

The aggregate measures are taken at a neighborhood-level and city-level instead of at the Census tract level. The accompanying R syntax file describes the procedure to create these measures, which are not included in this record-level file, but in alternate aggregate files.

1.2.1 Initial Report Content

1.2.1.1 General Tract Information

- *FID* is a unique numeric identifier for each tract.
- *GEOID10* is the geographic identifier for each tract assigned in the 2010 U.S. Census.

- *Data note:* The value for each row begins with the Massachusetts state code, “25,” and the U.S. country code, “025.”
- *AREA_SQFT* is the tract area in square feet.
- *AREA Acres* is the tract area in acres.
- *HU100_RE* is the count of housing units in the tract.
- *Name* is the Boston neighborhood associated with the tract.
 - *Note:* 23 Possible Values – Allston, Back Bay, Bay Village, Brighton, Charlestown, Dorchester, East Boston, Fenway, Harbor Islands, Hyde Park, Jamaica Plain, Leather District, Longwood Medical Area, Mattapan, Mission Hill, North End, Roslindale, Roxbury, South Boston, South Boston Waterfront, South End, West End, West Roxbury
- *Shape__Area* is the geographic information system (GIS) code corresponding to the area and coordinates of the tract.
- *Shape__Length* is the GIS code corresponding to the length of the tract outline.

1.2.1.2 Raw Population Counts

The definitions of types of people in this section apply throughout the document. For example, “count of children” is understood to mean the number of children under the age of five in a given tract.

- *POP100_RE* is the tract population count.
- *TotDis* is the count of tract residents with disabilities (hearing, vision, cognitive, ambulatory, self-care, and/or independent living difficulties).
- *TotChild* is the count of tract residents who are children (under the age of five).
- *OlderAdult* is the count of tract residents who are older adults (over the age of 65).
- *Low_to_No* is the count of tract residents with low to no income (up to 149% of the federal poverty level).
- *LEP* is the count of tract residents with limited English proficiency.
- *POC2* is the count of tract residents who are people of color.
 - *Note:* Possible Values – Black, Native American, Asian, Island, Other, Multi, and Non-white Hispanics
- *MedIllnes* is the count of tract residents with medical illnesses.
 - *Data source:* This is a proxy measure based on EASI data accessed through Simply Map. Health data at the local level in Massachusetts is not available beyond zip codes. EASI modeled the health statistics for the U.S. population based upon age, sex, and race probabilities using U.S. Census Bureau data. The probabilities are modeled against the census and current-year and five-year forecasts. Medical illness is the sum of asthma in children, asthma in adults, heart disease, emphysema, bronchitis, cancer, diabetes, kidney disease, and liver disease. A limitation is that these numbers may be over-counted as the result of people potentially having more than one medical illness. Therefore, the analysis may have greater numbers of people

with medical illness within census tracts than actually present. Overall, the analysis was based on the relationship between social factors.

1.2.2 New Variables

1.2.2.1 Proportion and Density Measures

- *pct_TotDis* is the proportion of tract residents with disabilities, as a percentage of the total tract population count.
- *pct_TotChild* is the proportion of tract residents who are children, as a percentage of the total tract population count.
- *pct_OlderAdult* is the proportion of tract residents who are older adults, as a percentage of the total tract population count.
- *pct_LowtoNo* is the proportion of tract residents with low to no income, as a percentage of the total tract population count.
- *pct_LEP* is the proportion of tract residents with limited English proficiency, as a percentage of the total tract population count.
- *pct_POC* is the proportion of tract residents who are people of color, as a percentage of the total tract population count.
- *pct_MedIll* is the proportion of tract residents with medical illnesses, as a percentage of the total tract population count.
- *dens_TotDis* is the average count of tract residents with disabilities per acre.
- *dens_TotChild* is the average count of tract residents who are children per acre.
- *dens_OlderAdult* is the average count of tract residents who are older adults per acre.
- *dens_LowtoNo* is the average count of tract residents with low to no income per acre.
- *dens_LEP* is the average count of tract residents with limited English proficiency per acre.
- *dens_POC* is the average count of tract residents who are people of color per acre.
- *dens_MedIll* is the average count of tract residents with medical illnesses per acre.
- *house_dens* is the average count of housing units per acre in the tract.
- *res_dens* is the average count of tract residents per housing unit.
- *pop_dens* is the average count of tract residents per acre.

1.2.2.2 Risk Categorization

- *vul_sc* is the count of socially vulnerable groups in the tract considered significantly vulnerable, with “significantly vulnerable” defined as the proportion of tract residents in a socially vulnerable group being higher than the proportion of total (i.e., citywide) residents in that socially vulnerable group.
 - *Note:* Possible values range from 0 to 7
- *risk* is the level of social vulnerability for the tract, based on the value of *vul_sc*.

- *Note:* Possible values – low (*vul_sc* = 0–2), medium (*vul_sc* = 3–4), high (*vul_sc* = 5–7)

1.2.2.3 Updated Measures

- *round_MedIll* is the value of *MedIllnes* rounded down to the nearest integer to allow for clean comparison as a number of residents in a tract.
- *pct_round_MedIll* is the updated proportion of tract residents with medical illnesses, as a percentage of the total tract population count, based on *round_MedIll* instead of *MedIllnes*.

1.2.2.4 Mobility

- *mob_count* is the average mobility count of a given tract, defined as the mean of selected vulnerable population counts (*TotChild*, *OlderAdult*, *MedIllnes*, *TotDis*) identified as having potential limitations in movement, in a given row.
- *avg_mob* is the percentage of mobility in a given tract, as measured by the percentage of the average mobility count of the total tract population count. Higher percentages indicate less mobility.

1.2.2.5 Snow Removal

These measures are calculated in part by using the BARI 311 Requests data to find “Snow Removal” cases, in addition to the Social Vulnerability dataset.

- *snow_count* is the count of “Snow Removal” requests received by the 311 system in a given tract.
- *snow_prop* is the number of “Snow Removal” requests received by the 311 system proportional to the number of residents within a given tract.

1.2.2.6 Linked Data

The variables in this section are taken from the BARI Geographical Infrastructure data, and are specifically used to find differences in levels of our created mobility construct (*avg_mob*) between different types of tract groups.

- *Res* indicates whether the tract is generally a residential area.
 - *Note:* Possible values are 0 (no) and 1 (yes)
- *Type* describes the type of neighborhood for a given tract.
 - *Note:* Possible values – R (Residential), D (Downtown), I (Institutional, e.g., industrial, college campuses), and P (Park)

1.2.3 Aggregate Measures

The code to create the following variables is shown in the accompanying R syntax file but not added to the record-level file, as they are aggregate calculations based on the record-level file. These variables are included in the following aggregate files:

- i. *CRB_SocialVul_Neighborhood.csv* describes neighborhood-level measures of socially vulnerable populations and population densities.
- ii. *CRB_SocialVul_City.csv* describes city-wide measures of socially vulnerable populations.

1.2.3.1 Neighborhood Level

- *pct_n_TotDis* is the proportion of neighborhood residents with disabilities, shown as a percentage of the total neighborhood population count.
- *pct_n_TotChild* is the proportion of neighborhood residents who are children, shown as a percentage of the total neighborhood population count.
- *pct_n_OlderAdult* is the proportion of neighborhood residents who are older adults, shown as a percentage of the total neighborhood population count.
- *pct_n_LowtoNo* is the proportion of neighborhood residents with low to no income, shown as a percentage of the total neighborhood population count.
- *pct_n_LEP* is the proportion of neighborhood residents with limited English proficiency, shown as a percentage of the total neighborhood population count.
- *pct_n_POC* is the proportion of neighborhood residents who are people of color, shown as a percentage of the total neighborhood population count.
- *pct_n_MedIll* is the proportion of neighborhood residents with medical illnesses, shown as a percentage of the total neighborhood population count.
- *med_pop_n_dens* is the median value of the average count of housing units per acre, across tracts in a neighborhood.
- *med_house_n_dens* is the median value of the average count of tract residents per acre, across tracts in a neighborhood.

1.2.3.2 City Level

- *pct_b_TotDis* is the proportion of city residents with disabilities, shown as a percentage of the total city population count.
- *pct_b_TotChild* is the proportion of city residents who are children, shown as a percentage of the total city population count.
- *pct_b_OlderAdult* is the proportion of city residents who are older adults, shown as a percentage of the total city population count.
- *pct_b_LowtoNo* is the proportion of city residents with low to no income, shown as a percentage of the total city population count.
- *pct_b_LEP* is the proportion of city residents with limited English proficiency, shown as a percentage of the total city population count.

- *pct_b_POC* is the proportion of city residents who are people of color, shown as a percentage of the total city population count.
- *pct_b_MedIll* is the proportion of city residents with medical illnesses, shown as a percentage of the total city population count.