Boston Area Research Initiative ● Northeastern & Harvard Universities
E-Mail: bari@northeastern.edu ● Web: www.bostonarearesearchinitiative.net

Documentation for Longitudinal Property Assessment Database 2025

Overview

This document describes the structure and organization of the City of Boston Assessing Department's centralized database describing all uniquely identifiable properties in the city (n = 187,149) for the years 2001-2025. Boston's Assessing Department is responsible for determining accurate values for all properties in the city. To this end, the Department maintains property ownership and value information to ensure fair assessment of both taxable and non-taxable property in Boston. The data is released by the Assessing Department annually as part of the City of Boston's open data initiative. The Boston Area Research Initiative accesses, cleans, and merges these data at the property level across years to create this longitudinal database. This database can be used to analyze trends in valuations and types of property over time.

The main dataset (*PADLong.Record.YEAR.csv*) contains a series of variables generated annually by the city as well as a set of variables calculated by BARI to track changes in assessed value and land-use over time.

We also offer aggregated datasets that track change in assessment and use over time at the tract and block group levels for the 2010 and 2020 Census Geographies, such as the *PADLong.CT10.YEAR.csv* file. This file contains aggregate measures, also known as ecometrics, that describe neighborhoods at the 2010 census tract level. These variables are available in a spreadsheet format (*.csv*) and as mappable shapefiles (*.shp*).

In past years, BARI has released these datasets using the name "Tax Assessor's Database" rather than "Property Assessment Database." We have changed names in order to match the names used by the City of Boston. Finally, beginning with the 2024 release (previous release) the year in each column represents the fiscal year rather than the year the data was collected. This change was made to better reflect the names used by the City of Boston. The city did not provide any data for the 2014 fiscal year.

Table of Contents

1. Summary of Record-level Property Assessment Data (PADLong.Record.YEAR.csv) 3
1.1. Description of Variables 3
1.1.1. Identifying Characteristics 3
1.1.2. Property and Building Characteristics 4
1.1.3. Geographical information 5
2. Summary of Aggregate Measures (e.g., PADLong.CT.YEAR.csv and corresponding shapefiles) 6
2.1. Description of Variables 6
3. Appendix 7
3.1. Appendix A: Codes for Land Use 7

1. Summary of Record-level Property Assessment Data *(PADLong.Record.YEAR.csv)*

The City of Boston's Assessing Department is responsible for determining accurate values for all properties in the city. To this end the Department maintains property ownership and value information to ensure fair assessment of both taxable and non-taxable property in Boston. Assessing records are compiled and reviewed annually to reflect changes to properties arising from new construction, remodeling, and changes in ownership. Upon annual review and re-assessment, data is updated and changes in calculated values are adjusted to reflect the most up to date status for each property. BARI downloads these data and links them longitudinally at the property level. The resulting data capture valuation, land use, and owner-occupancy across years. More detail on properties can be accessed through the Cross-Sectional Property Assessment Data, which are released annually in wide form (i.e., containing many variables for a single year).

1.1 Description of Variables

Property assessment variables are split into three categories: identifying characteristics, property and building characteristics, and geographical information. Identifying characteristics include variables regarding the basic identity and attributes of the address. Property and building characteristics include information on the physical attributes, zoning, and valuations of the property. Geographical information provides further detail on the location of the property and the other geographies that contain it. In certain character names, (YYYY) indicates that the variable is repeated for each fiscal year between 2001 and 2025.

Additional variables have been added with specific regard to interest and relevance towards analyzing assessment data in the city of Boston before, during, and after the "Great Recession" economic crash of 2008.

1.1.1. Identifying Characteristics

- *PID*, the 10-digit property identification number, unique to each property. The first two digits indicate the Ward, digits 3 through 7 are the parcel, and digits 8 through 10 are the sub-parcel.

• *CM_ID*, the 10-digit parcel number of the main condo building property. All condo units in each building are related to this number.

• *ST_NUM,* the street number of the property.

• *ST_NAME,* the street name and the street suffix of the property.

• *ZIPCODE* is the zip code of the property.

1.1.2 Property and building characteristics

• *FY(YYYY).LU,* the Land Use type for the property in a given year. Codes for land use can be found in Appendix A. For more information about land use codes, visit the Boston Assessing Department website.

• *FY(YYYY).AV,* the total assessed value for the property in a given year. It is a summation of the assessed values of the land and building. Values of $0 are set to "NA".

• *FY(YYYY).RESEX,* a one character code that indicates if an owner received a residential exemption for the property in a given year. A "Y" indicates that the owner claims to live within the property (i.e., the property is "owner-occupied"), having applied for the associated tax exemption. "N" indicates that the owner has not applied for the owner-occupancy exemption.

• *DiffAV(YYYY),* the difference in value between *YYYY* and the previous year

• *PercChangeAV(YYYY)*, the percent change in value between *YYYY* and the previous year as a percentage of the previous year.

• *LU(YYYY)FourCat* groups the 17 original land use codes into four broad categories.

  o *Note*: These categories are Res (consisting of original codes R1, R2, R3, R4, A, CD, CP, CM, and RL), Comm (original codes C, CC, RC, and CL), Ind (consisting of original codes I and AH), and Exem (original codes E and EA).

• *GrowthDiffAV* reflects the difference in value over the growth period before the Great Recession (defined here as 2001-2007).

• *GrowthPercChangeAV* reflects the percent change in value over the growth period before the Great Recession (defined here as 2001-2007).

• *CrashDiffAV* represents the difference in value over the crash period around the Great Recession (defined here at 2007-2011).

• *CrashPercChangeAV*, the difference in the percent change in value over the crash period around the Great Recession (defined here as 2007-2011).

• *RecoveryDiffAV* represents the difference in value over the recovery period around the Great Recession (defined here at 2011-2017).

• *RecoveryPercChangeAV*, the difference in the percent change in over the recovery period after the Great Recession (defined here as 2011-2017).

1.1.3 Geographical Information

• *X* is the longitude of the property.

  o This is derived from the City of Boston's *Parcels 2024*[1] shapefile.

• *Y* is the latitude of the property.

  o This is derived from the City of Boston's *Parcels 2024* shapefile.

• *GIS_ID* is another 10-digit property identification number. It is the unique identifier for the plot the property is in.

• *Land_Parcel_ID* is the unique ID of the land parcel containing the property. For more information on this ID and the geography to which it corresponds, see BARI's Geographical Infrastructure for the current year.

• *TLID*, the identifier for the segment of road containing the property.

  o This is found by subsetting the 2013 (for years before 2023) or 2023 (for years 2023 or after) TIGER lines street segments to only those that match the street name of the property, and then finding the one that is geographically closest to the property.

• *Blk_ID_10,* the 2010 Census Block ID number.

  o This is found by spatially overlaying the longitude and latitude of the property onto the Census Blocks shapefile.

• *BG_ID_10,* the 2010 Census Block Group ID number.

• *CT_ID_10,* the 2010 Census Tract ID number.

• *Blk_ID_20,* the 2020 Census Block ID number.

  o This is found by spatially overlaying the longitude and latitude of the property onto the Census Blocks shapefile.

• *BG_ID_20,* the 2020 Census Block Group ID number.

• *CT_ID_20,* the 2020 Census Tract ID number.

[1] We use 2024 data, as that is the most recent parcel data provided by the city at the time of release.

2. Summary of Aggregate Measures *(e.g., PADLong.CT10.YEAR.csv and corresponding shapefiles)*

The variables below describe recorded levels of change in assessed value and land use in a region. Aggregations are made at the census tract or block group level for the 2010 and 2020 Census Geographies. CT10, CT20, CBG10, or CBG20 in the filename indicates whether data is aggregated to 2010 Census tracts (CT10), 2020 Census tracts (CT20), 2010 Census block groups (CBG10), or 2020 Census block groups (CBG20). Variable suffixes (YYYY) indicate the year of measurement. Aggregate measures are provided in both standard format (.csv) and as mappable shape files (.shp). Truncated variable names for the latter format are included in brackets following the original variable names.

2.1 Description of Variables

- *CT_ID_10, CT_ID_20, BG_ID_10,* or *BG_ID_20,* the ID number for a specific Census geography that data is aggregated to (see section 1.1.3 for details).
- *SumDiffAV(YYYY) [SDAV(YYYY)]* the sum of the difference in value between *YYYY* and the previous year for all parcels within the tract.
- *PercChangeAV(YYYY) [PCAV(YYYY)]* the median percent change in value between *YYYY* and the previous year as a percentage of the previous year for all parcels within the tract.
- *NewCondo(YYYY) [NwC(YYYY)]* reflects the number of new condo parcels between previous year and (*YYYY*) for a given census tract.
- *GrowthPercChangeAV [GrwPCAV]* reflects the median percentage change in assessed value from 2001 – 2007.
- *CrashPercChangeAV [CrsPCAV]* reflects the median percentage change in assessed value from 2007-2011.
- *RecoveryPercChangeAV [RcvPCAV]* reflects the median percentage change in assessed value from 2011 – Current Year.

3. Appendix

3.1 Appendix A: Codes for Land Use

USE CODE	DESCRIPTION
A	Residential 7 or more units
AH	Agricultural/Horticultural
C	Commercial
CC	Commercial condominium
CD	Residential condominium unit
CL	Commercial land
CM	Condominium main (physical structure housing all related condo units with no assessed value)
CP	Condo parking
E	Tax-exempt
EA	Tax-exempt (121A)
I	Industrial
R1	Residential 1-family
R2	Residential 2-family
R3	Residential 3-family
R4	Residential 4 or more family
RC	Mixed use (res. and comm.)
RL	Residential land