

Boston Area Research Initiative ⦁ Northeastern & Harvard Universities  
E-Mail: bari@northeastern.edu ⦁ Web: [www.bostonarearesearchinitiative.net](http://www.bostonarearesearchinitiative.net)

**Documentation for Cross-sectional Property Assessment Database 2017**

**Overview**

This document describes the structure and organization of the City of Boston Assessing Department’s centralized database for property-specific data for all uniquely identifiable properties in the city (n = 170,910) for the year 2017. The data are released by the Assessing Department annually as part of the City of Boston’s open data initiative. Boston’s Assessing Department is responsible for determining ownership and physical characteristics for all properties in the city in order to ensure fair assessment of both taxable and non-taxable property in Boston. This dataset can be used to analyze valuations for all classifications and types of property.

The main dataset *(PADCross.Record.2017.csv)* is derived from the tax assessor's annual release through *data.boston.gov* but is curated by BARI to contain a handful of additional variables to facilitate informed analysis.

BARI has also created aggregate measures (i.e. ecometrics) that describe neighborhoods (in the form of census geographies). These variables are provided in a spreadsheet format (*PADCross.CT.2017.csv*) and in 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.

Table of Contents

[1. Summary of Assessing Department Data *(PADCross.Record.2017.csv)* 3](#_Toc483927581)

[1.1 Description of Variables 3](#_Toc483927582)

[1.2 Identifying Characteristics 3](#_Toc483927583)

[1.3 Property and Building Characteristics 5](#_Toc483927584)

[1.3 Geographical information 9](#_Toc483927585)

[2. Summary of Aggregate Ecometrics *(PADCross.CT.2017.csv* & *PADCross.CBG.2017.csv & corresponding shapefiles)* 11](#_Toc483927586)

[2.1 Description of Variables 11](#_Toc483927587)

[3. Appendix 12](#_Toc483927588)

[3.1 Appendix A: Codes for Land Use 12](#_Toc483927589)

[3.2 Appendix B: Property Type (PTYPE) Codes 13](#_Toc483927590)

1. Summary of Record-level Property Assessment Data *(PADCross.Record.2017.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 as a result of new construction, remodeling, and changes in ownership. The data contained within describes the property-specific address, ownership, type, structure, class, and valuation data. Upon annual review and re-assessment, property-specific data is updated and changes in calculated values are adjusted to reflect the most up to date status for each property. Property taxes (as indicated within this dataset as “GROSS\_TAX”) are also adjusted annually to reflect the annual taxation rates for residential and commercial properties.

Tax rates are determined by taking the portion of property tax levy for properties deemed “residential” and dividing that by the total assessed value of all residential properties in the City. Commercial tax rates are calculated in the same manner.[[1]](#footnote-1)

This dataset is a modified version of the original Tax Assessor’s Data Set, and includes only variables and measures that were derived from the original data. These variables and measures provide additional insight into the current physical makeup of properties in Boston.

This dataset was also used to conduct a measure of the historical life of structures that are residential in type, and contain certain characteristics that indicate the historical evolution of a structure's life as well as analysis of structure types' actual locations within the city as they relate to Census demographics

## 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. Building characteristics include information on the physical attributes of the building containing the property. Geographical information provides further detail on the location of the property and the other geographies that contain it.

## 1.2 Identifying Characteristics

* *parcel\_num* is the 10-digit property identification number, unique to each property. The first two digits indicate the Ward, digits 3 thru 7 are the parcel, and digits 8 thru 10 are the sub-parcel.
* *CM\_ID* is the 10-digit parcel number of the main condo building parcel. All condo units in each building are related to this number.
* *GIS\_ID* is another 10-digit property identification number. It is the unique identifier for the land the property is in (this is slightly different than *Land\_Parcel\_ID*, however, as the latter combines some land parcels into one).
* *ST\_NUM* is the street number of the property.
* *ST\_NAME* is the street name of the property.
* *ST\_NAME\_SUF* is the suffix of the street name. This variable contains two-character short-forms of each type of suffix (St, Av, BL, PL, etc…).
* *UNIT\_NUM* is the specific unit number within a multi-unit building.
* *ZIPCODE* is the zip code of the property.

## 1.3 Property and Building Characteristics

* *PTYPE* is the land usage Property Type of the property. Property Types are classified according to State Class Codes, which are three digit codes. PTYPES between 001 and 299 are Residential properties. PTYPES between 300 and 399 are Commercial properties. PTYPES between 400 and 465 are Industrial properties. PTYPES greater than 900 are Exempt properties, meaning they are fully (or partially) exempt from property taxes. See appendix B for a full list of PTYPES.
* *LU* is the Land Use type for the property. Codes for land use can be found in Appendix A.
* *OWN\_OCC* is a one-character code that indicates if an owner receives a residential exemption for the property. A "Y" indicates that the owner claims to live within the property (a.k.a. the property is "owner-occupied") and a "N" indicates the opposite.
* *OWNER* is the primary owner of the property, as of the date of the prior calendar year.
* *MAIL\_ADDRESSEE* is the name associated with the street mailing address of the owner, if it is not the owner.
* *MAIL\_ADDRESS* is the street mailing address of the owner, to which the property tax bill is mailed.
* *MAIL.CS* is the city and state to which the property tax bill is mailed.
* *MAIL\_ZIPCODE* is the zip code of the property where the tax bill is mailed.
* *AV\_LAND* is the assessed value of the land.
* *AV\_BLDG* is the total assessed value for the building on the property.
* *AV\_TOTAL* is the total assessed value for the property. It is a summation of the assessed values of the land and building.
* *GROSS\_TAX* is the amount billed to the owner as property excise tax. It is based on the total assessed value multiplied by the tax rate. Tax rates are adjusted each year for Residential and Commercial property types.
* *LAND\_SF* is the total size of the property in square feet. This is also known as the lot size.
* *YR\_BUILT* is the year in which the property was built. The original dataset held many properties whose year of construction was listed as zero. It was fixed by updating the YR\_BUILT variable, which now contains a "NA" value where it previously showed a "0".
* *YR\_REMOD* is the year in which the property was last remodeled. For some properties the year of its most recent remodel was listed as zero. It was fixed by updating the YR\_REMOD variable, which now contains a "NA" value where it previously showed a "0".
* *GROSS\_AREA* is the gross floor area for commercial properties.
* *NUM\_FLOORS* is the number of levels in the structure that is located on the property.
* *STRUCTURE\_CLASS* is the structural classification of commercial buildings. Classes include: *A* for Steel Structure, *B* for Reinforced Concrete, *C* for Brick & Concrete, *D* for Wood/Frame, *E* for Metal, and *R* for Residential.
* *R\_BLDG\_STYL* is the building style for residential properties. The styles are: BL for Bi-Level, BW for Bungalow, CL for Colonial, CN for Contemporary, CP for Cape, CV for Conventional, DK for Decker, DX for Duplex, L for Tri-Level, Oth for Other, RE for Row End, RM for Row Middle, RN for Ranch, RR for Raised Ranch, SL for Split Level, TF for Two-Family Stack, TD for Tudor, SD for Semidetached, and VT for Victorian.
* *R\_ROOF\_TYP* is the roof type for residential structures. Types include: F for Flat, G for Gable, H for Hip, L for Gambrel, M for Mansard, O for Other, and S for Shed.
* *R\_EXT\_FIN* is the exterior finish type for residential properties. Types include: A for Asbestos, B for Brick/Stone, C for Cement Board, F from Frame/Clapboard, G for Glass, K for Concrete, M for Vinyl, O for Other, P for Asphalt, S for Stucco, U for Aluminum, V for Brick/Stone Veneer, and W for Wood Shake.
* *R\_TOTAL\_RMS* is the total number of rooms in a residential structure.
* *R\_BDRMS* is the total number of bedrooms in a residential structure.
* *R\_FULL\_BTH* is the total number of Full Bathrooms in a residential structure. A full bath is also known as a four-piece bath, which includes a shower, a tub, a sink, and a toilet.
* *R\_HALF\_BTH* is the number of Half Bathrooms in a residential structure. A half bath typically includes a sink and a toilet, also known as a powder-room.
* *R\_BTH\_STYLE* is the bath style of the first bathroom in a residential structure. Types include: L for Luxury, M for Modern, N for No Remodeling, S for Semi-Modern.
* *R\_BTH\_STYLE2* is the bath style of the second bathroom in a residential structure. Types are above.
* *R\_BTH\_STYLE3* is the bath style of the third bathroom in a residential structure. Types are above.
* *R\_KITCH* is the total number of kitchens in a residential structure.
* *R\_KITCH\_STYLE* is the kitchen style of the first kitchen in a residential structure. Types include: L for Luxury, M for Modern, N for No Remodeling, S for Semi-Modern.
* *R\_KITCH\_STYLE2* is the kitchen style of the second kitchen in a residential structure. Types are above.
* *R\_KITCH\_STYLE3* is the kitchen style of the third kitchen in a residential structure. Types are above.
* *R\_HEAT\_TYP* is the type of heating in a residential structure. Types include: N for None, S for Space Heater, W for Hot Water, E for Electric, P for Heat Pump, F for Forced Air, and O for Other.
* *R\_AC* indicates if the residential structure has air conditioning. Types include: C for Central Air Conditioning, D for Ductless Air Conditioning, and N for None.
* *R\_FPLACE* is the total number of fireplaces in a residential structure.
* *R\_EXT\_CND* is the exterior condition of a residential structure. Types include: A for Average, E for Excellent, F for Fair, G for Good, and P for Poor.
* *R\_OVERALL\_CND* is the overall condition for a residential structure. Types are above.
* *R\_INT\_CND* is the interior condition of a residential structure. Types are above.
* *R\_INT\_FIN* is the interior finish of a residential structure. Types include: E for Elaborate, N for Normal, and S for Substandard.
* *R\_VIEW* is the view for a residential structure. Types include: A for Average, E for Excellent, F for Fair, G for Good, P for Poor, and S for Special.
* *S\_NUM\_BLDG* is the number of buildings in condo main.
* *S\_BLDG\_STYL* is the building style of of a Condo Main. Styles include: DK for Decker, DX for Duplex, CV for Conventional, FS for Free-Standing, HR for High Rise, LR for Low Rise, MD for Mid Rise, RE for Row End, RM for Row Middle, OT for Other, SD for Semi-Dethatched, and TH for Townhouse.
* *S\_UNIT\_RES* is the number of residential units in a condo main.
* *S\_UNIT\_COM* is the number of commercial units in a condo main.
* *S\_UNIT\_RC* is the number of mixed-use units in a condo main.
* *S\_EXT\_FIN* is the exterior finish type for condominium buildings. Types include: A for Asbestos, B for Brick & Stone, C for Cement Board, F for Frame/Clapboard, G for Glass, K for Concrete, M for Vinyl, O for Other, P for Asphalt, S for Stucco, U for Aluminum, V for Brick/Stone veneer, and W for Wood Shake.
* *S\_EXT\_CND* is the exterior condition of a condo main. Types include: A for Average, E for Excellent, F for Fair, G for Good, and P for Poor.
* *U\_BASE\_FLOOR* indicates the base floor number of a condo unit. This would indicate if a condo unit is greater than one floor.
* *U\_NUM\_PARK* is the number of parking spaces associated with a property.
* *U\_CORNER* indicates if a condo unit is located in the corner of a building. Y indicates Yes and N indicates No.
* *U\_ORIENT* indicates the orientation of the condo unit within the building. Types include: A for Rear Above, B for Rear Below, C for Face Courtyard, E for End, F for Front/Side, M for Middle, T for Through
* *U\_TOT\_ROOMS* is the total number of rooms in a condominium unit.
* *U\_BDRMS* is the total number of bedrooms in a condominium unit.
* *U\_FULL\_BTH* is the total number of Full Bathrooms in a condominium unit. A full bath is also known as a four-piece bath, which includes a shower, a tub, a sink, and a toilet.
* *­U\_HALF\_BTH is* the number of Half Bathrooms in a condominium unit. A half bath typically includes a sink and a toilet, also known as a powder-room.
* *U\_BTH\_STYLE* is the bath style of the first bathroom in a condo unit. Types include: L for Luxury, M for Modern, N for No Remodeling, and S for Semi-Modern.
* *U\_BTH\_STYLE2* is the bath style of the second bathroom in a condo unit. Types are above.
* *U\_BTH\_STYLE3* is the bath style of the third bathroom in a condo unit. Types are above.
* *U\_KITCH\_TYP* is the type of kitchen in a condominium unit. Types include: F for Full Eat In, N for None, O for One-Person, and P for Pull/Alcove
* *U\_KITCH\_STYLE* is the kitchen style in a condo unit. Types include: L for Luxury, M for Modern, N for No Remodeling, and S for Semi-Modern.
* *U\_HEAT\_TYP* is the type of heating in a condominium unit. Types include: N for None, S for Space Heater, W for Hot Water, E for Electric, P for Heat Pump, F for Forced Air, and O for Other.
* *U\_AC* indicates if a condominium unit has self-controlled air conditioning. Types include: C for Central Air Conditioning, D for Ductless Air Conditioning, and N for None.
* *U\_FPLACE* is the number of fireplaces in a condominium unit.
* *U\_INT\_FIN* is the interior finish of a condo unit. Types include: E for Elaborate, N for Normal, and S for Substandard.
* *U\_INT\_CND* is the interior condition of a condo unit. Types include: A for Average, E for Excellent, F for Fair, G for Good, and P for Poor.
* *U\_VIEW* is the view for a condo unit. Types include: A for Average, E for Excellent, F for Fair, G for Good, P for Poor, and S for Special.
* *LIVING\_AREA* is the total living area for residential properties.
* *AV\_BLDG\_PER\_SF* expresses the assessed value of a property’s building, divided by its gross floor area in square feet.
* *AV\_LAND\_PER\_SF* is the assessed value of a property’s land, divided by the total property area in square feet.
* *SIMPLIFIED\_LU* The Tax Assessor's dataset includes 17 different types of land use, including 6 different codes for residential use at varying densities (single floor houses, two-floor, etc). The many different classifications for similar uses can result in "artificially" high scores when used to calculate the diversity in land use distribution for a given area. SIMPLIFIED\_LU reflects land use according to a simplified classification, with eight core uses: Residential, Commercial, Condo, Mixed Residential/Commercial, Agricultural, Industrial, Tax Exempt, and Tax Exempt by the Boston Redevelopment Authority (applies to properties that are undergoing renovation projects).
* *EE\_SCORE* This is an aggregate variable that combines the HEAT\_SCORE, COOL\_SCORE and AGE\_SCORE in weighted sum [EE\_SCORE = AGE\_SCORE+0.75\*HEAT\_SCORE+0.75\*COOL\_SCORE]. It indicates the property specific composite energy efficiency index. This variable is only defined for properties with land usage R1, R2, or R3.
  + *Note: This variable takes into account three other variables which are otherwise not included in visualization data but necessary to calculate energy efficiency scores.*
  + *HEAT\_SCORE* represents the residential heating system types in form of energy efficiency score. Each residential heating system type is allocated a numeric score based on its energy efficiency performance.
  + *COOL\_SCORE* This variable represents residential cooling types in form of energy efficiency score. Each residential cooling system type is allocated a numeric score based on its energy efficiency performance.
  + *AGE\_SCORE* This variable represents residential unit energy efficiency based on the age of building. The scores were allocated on the assumption that older buildings are more energy inefficient.
* *BLDG\_AGE* is 2017 minus the year in which the building was most recently remodeled or the year in which it was first built if it was never remodeled.

**1.3 Geographical information**

* *X* is the longitude of the property.
  + This is derived from the City of Boston’s *Parcels 2017* shapefile.
* *Y* is the latitude of the property.
  + This is derived from the City of Boston’s *Parcels 2017* shapefile.
* *TLID* is the identifier for the segment of road containing the property.
  + This is found by subsetting the 2013 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* is the 2010 Census Block ID number.
  + This is found by spatially overlaying the longitude and latitude of the property onto the Census Blocks shapefile.
* *BG\_ID\_10* is the 2010 Census Group ID number.
* *CT\_ID\_10* is the 2010 Census Tract ID number.
* *NSA\_NAME* is the name of the Inspectional Service Department Neighborhood Statistical Area in which the building is located.
* *BRA\_PD* is the name of the Boston Redevelopment Authority Planning District in which the building is located.
* *Land\_Property\_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 2017.

# 2. Summary of Aggregate Measures *(e.g. PADCross.CT.2017.csv & corresponding shapefiles)*

Neighborhood-level datasets were created that describe aggregate features of neighborhood properties. Aggregate files are included for both the census tract level and block group level. 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 parentheses following the original variable names. Variable names for shapefiles are in parentheses.

## 2.1 Description of Variables

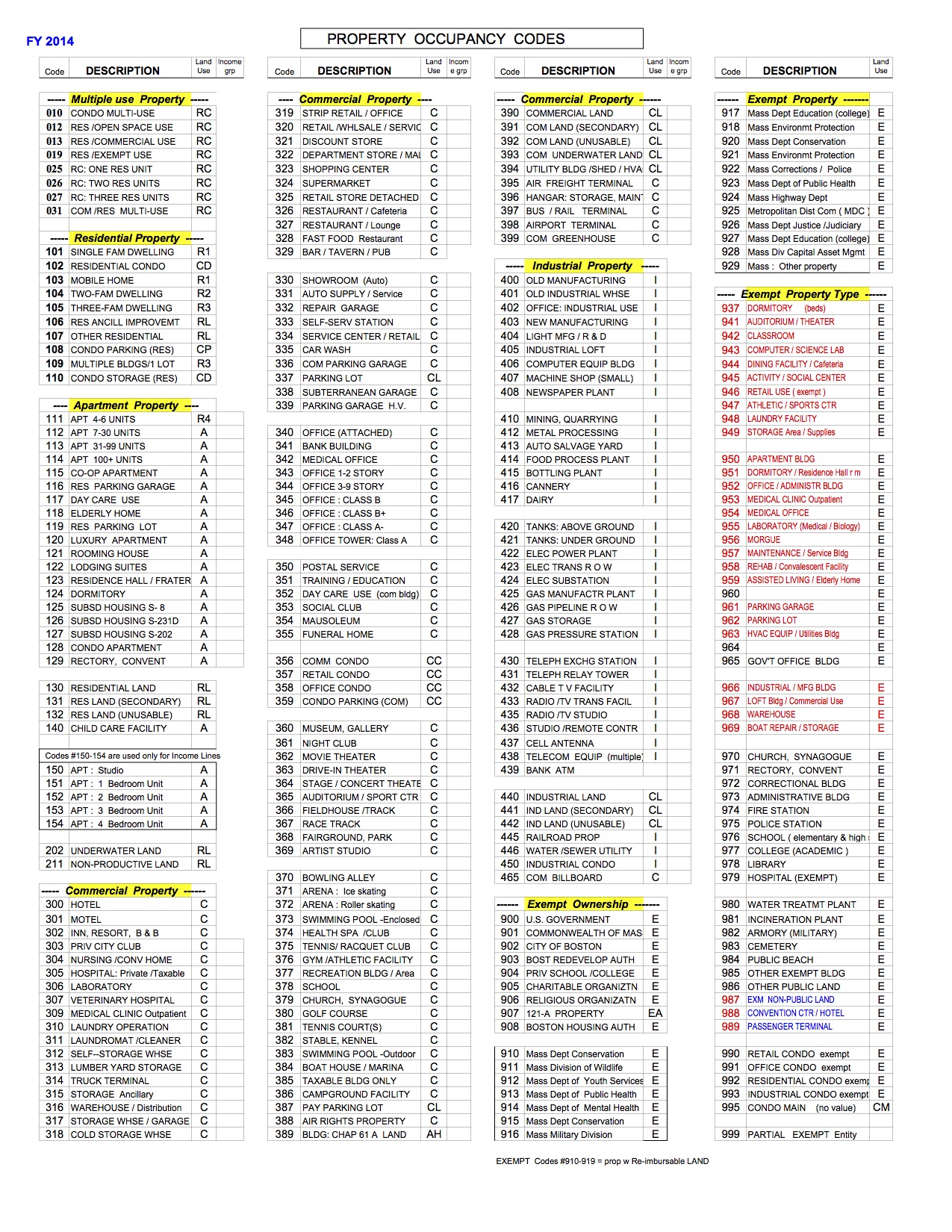
* *EE\_SCORE.res* *(EESR)* is the average energy efficiency index (*EE\_SCORE* above) for residential properties in the area. For more information on how the energy efficiency index is calculated, see the documentation above.
* *AV\_LAND\_PER\_SF.res (ALPSFR)* is the average assessed value of a property’s land, divided by the total property area in square feet (*TOTAL\_PER\_SF* above) for all residential properties in the area
* *AV\_LAND\_PER\_SF.nonres (ALPSFN)* is the average assessed value of a property’s land, divided by the total property area in square feet (*TOTAL\_PER\_SF* above) for all non-residential properties in the area
* *AV\_BLDG\_PER\_SF.res (ABPSFR)* is the assessed value of a property’s building, divided by its gross floor area in square feet (*AV\_BLDG\_PER\_SF* above) for all residential properties in the area.
* *AV\_BLDG\_PER\_SF.nonres (ABPSFN)* is the assessed value of a property’s building, divided by its gross floor area in square feet (*AV\_BLDG\_PER\_SF* above) for all non-residential properties in the area.
* *YR\_BUILT\_REMOD.res (YBRR)* is the mean value of the latest year remodeled or the year built for all residential properties in the area.
* *YR\_BUILT\_REMOD.nonres (YBRN)* is the mean value of the latest year remodeled or the year built for all non-residential properties in the area.
* *DEC\_BUILT\_REMOD.res (DBRR)* is the modal value of the latest decade it was remodeled in or the decade it was built in for all residential properties in the area.
* *DEC\_BUILT\_REMOD.nonres (DBRN)* is the modal value of the latest decade it was remodeled in or the decade it was built in for all non-residential properties in the area.

3. Appendix

## 3.1 Appendix A: Codes for Land Use



## 3.2 Appendix B: Property Type (PTYPE) Codes



1. Tax rate calculation information published by the City of Boston Assessing Department through the department’s website: http://www.cityofboston.gov/assessing/taxrates.asp [↑](#footnote-ref-1)