



DOCUMENTATION FOR BOSTON AREA RESEARCH INITIATIVE'S BUILDING PERMITS DATA

Overview

This document describes the structure and organization of the Building Permit Applications in Boston database (*BuildPermits.Records.2017.csv*), which contains information on 354,682 building permit applications collected between September 26, 2006 and January 22, 2018. This data was gathered and released by the Inspectional Service Department (ISD) of the City of Boston, where information pertaining to applications for building permits is stored in the city. The original dataset contains 21 variables by which these applications can be described; some of these variables are identifying characteristics of building permits, and the others are geographic information about the permits. This document also describes modifications to the original data made by the Boston Area Research Initiative (BARI), including new variables to facilitate their study, as well as a series of aggregate measures (i.e., ecometrics) that use the permits to describe patterns of investment and growth across neighborhoods (approximated with census geographies). These variables are provided in a spreadsheet format (e.g. *Permits.Ecometrics.CT.2017.csv*) and in mappable shapefiles (.shp).

Table of Contents

1. Building Permit Applications (e.g. <i>Permits.Records.Geocoded.20180408.csv</i>)	2
1.1. Description of Contents.....	2
1.2. Summary of Variables.....	2
1.2.1. <i>Building Permit Application Characteristics</i>	2
1.2.2. <i>Geographic Information</i>	3
1.2.3 <i>Types of Work</i>	5
2. Neighborhood Ecometrics (e.g., <i>Permits.Ecometrics.CT.2017.csv</i>)	5
2.1. Description of Contents.....	5
2.2. Summary of Variables.....	6



1. Building Permit Applications (e.g. Permits.Records.Geocoded.20180408.csv)

1.1. Description of Contents

The City of Boston requires that property owners apply for permits before performing substantial work on any structures or constructing new structures. Each permit references a particular type of work to be conducted (e.g., gas, plumbing, excavation, demolition), for which reason a single project could require multiple permits. Once approved, permits are public record. The BARI building permits database contains 354,682 building permit applications collected between September 26, 2006 and January 22, 2018.

1.2. Summary of Variables

The building permits include variables that can be grouped into three categories: application characteristics, all of which are original to the data; geographic information, which includes variables introduced by BARI to make the data more easily compatible with other data sources; and types of work, categorized by BARI based on the types of permits granted.

1.2.1. Building Permit Application Characteristics

- *PermitNumber* is the unique identifier assigned to the building that can be linked to Inspectional Service Department of the City of Boston.
- *WORKTYPE* is a one-word description of the type of work of the building permit.
- *PermitTypeDescr* describes the type of building permit in more detail than “WORKTYPE.”
- *DESCRIPTION* explains what “WORKTYPE” stands for and explains the specific nature of the work to be completed.
- *NOTES* are case specific details about each permit; provides more details than the “DESCRIPTION” field.
- *APPLICANT* is the person or organization applying for the permit.
- *DECLARED_VALUATION* is the fair estimation of market value for the proposed building permit action.

- *TOTAL_FEES* is the cost of the building permit.
- *ISSUED_DATE* is the date the permit was issued.
- *EXPIRATION_DATE* is the date the permit expires.
- *STATUS* indicates the status of the permit. A permit can be either open or expired.
- *OWNER* indicates the owner of the building.
- *OCCUPANCY* refers to the type of people the building houses. For example, the occupancy can be “Commercial,” “Residential,” “Mixed,” “1-2 Family,” “6 Units,” etc.
- *sq_feet* is the square footage.
- *ADDRESS* is the physical address of the building.
- *CITY* indicates the city or neighborhood in which the building is located.
- *STATE* indicates the state in which the building is located.
- *ZIP* indicates the zip code in which the building is located.

1.2.2. Geographic Information

Geolocation information for parcels is provided in this dataset as additional identifying data that can improve analysis of Boston’s properties. The latitude and longitude is provided in the City of Boston’s original dataset. BARI uses these coordinates and the parcel ID to link the data to BARI’s Geographical Infrastructure,¹ which organizes the city’s locations at multiple levels of spatial resolution.

U.S. Census information is aggregated for each unique parcel as a way of grouping parcels using a recognized set of attributes. The census block, census block group, and census tract identifiers for each parcel are set by the government in order to better understand the demographic makeup of cities, towns, states, etc. This information is appended to the building permits to better enable their coordination with data generated by the 2010 census and all subsequent American Community Surveys, as well as any other data set similarly mapped with the Geographical Infrastructure.

¹ https://dataverse.harvard.edu/dataverse/geographical_infrastructure_2017



- *Location* is the latitude, longitude coordinates of the building. It can also be read as X, Y.
- *Property_ID* is the unique identifier given to the building property to track the building in the document. This was derived by the Inspectional Service Department of the City of Boston.
- *parcel_num* is the 10-digit parcel identification number, unique to each parcel. The first two digits indicate the Ward, digits 3 thru 7 are the parcel, and digits 8 thru 10 are the sub-parcel.
- *X* indicates the x-coordinate or latitude location of the building.
- *Y* indicates the y-coordinate or longitude location of the building.
- *Land_Parcel_ID* is the 10-digit parcel identification number, unique to each parcel. The first two digits indicate the Ward, digits 3 thru 7 are the parcel, and digits 8 thru 10 are the sub-parcel. (Note: it is identical to the *parcel_num* variable, but is used to verify BARI's geocoding process).
- *TLID* is the unique identifier for the road segment containing the building.
- *Blk_ID_10* is the unique identifier for the 2010 U.S. Census Block in which the building is located.
- *BG_ID_10* is the unique identifier for the 2010 U.S. Census Block Group (2010-present) in which the building is located.
- *CT_ID_10* is the unique identifier for the 2010-present U.S. Census Tract in which the building is located.
- *NSA_NAME* is the name of the Boston Planning and Development Agency neighborhood statistical unit in which the region is located.
- *BRA_PD* is the name of the Boston Planning and Development Agency planning district in which the region is located.



1.2.3 Types of Work

Using the *Description* and *PermitTypeDescr* variables, BARI constructed new dummy-coded variables (i.e., “0” = no, “1” = yes, relative to the name of the variable) to match types of building permits to the types of work that might be done there.

- *newcon* refers to permits indicating “new construction” where there was not previously a structure. This includes *Erect/New Construction* and similar descriptions of work types.
- *addition* refers to permits indicating extensive change (or an addition) to a structure that is not New Construction, Demolition or Renovation. This includes *Amendment to a Long Form* and *Long Form/Alteration Permit* as descriptions of permit types.
- *reno* refers to permits indicating a small change to the building or a required improvement (renovation) for a building code. This includes many types of work not described in the above categories.
- *PermitDuration* is the length of time between the permit’s issue date and expiration date in days. The variable was created by subtracting the permit’s expiration date from the permit issue date, and converting all negative values to NA since negative values are the result of input errors on the permit application.

2. Neighborhood Ecometrics (e.g., *Permits.Ecometrics.CT.2017.csv*)

2.1. Description of Contents

This section describes the econometric variables created by BARI. These econometrics describe patterns of neighborhood growth and investment through the volume and profile of building permits using U.S. Census Block Groups (e.g. *Permits.Ecometrics.CBG.2015.csv*), U.S. Census Tracts (e.g. *Permits.Ecometrics.CT.2015.csv*), and land parcels (e.g. *Permits.Ecometrics.LP.2015.csv*). Each variable is included both in an annual file and a composite longitudinal file. In the longitudinal files, variables have suffixes indicating the year of measurement.

Aggregations for the econometrics are first made at the parcel level, in which work at every parcel is categorized as being new construction, addition, or renovation. This categorization is mutually exclusive based on the types of permits that were pulled for a

parcel. The categorization is cascading in that any parcel with a permit indicating new construction cannot be an addition or renovation, and any features permits indicating an addition cannot be a renovation. This is because more ambitious projects still require permits that would be included in less intensive projects. These categorizations at the parcel level are then used to calculate the census tract (*CT*; ID variable = *CT_ID_10*) and block group measures (*CBG*; ID variable = *BG_ID_10*).

2.2. Summary of Variables

Because shapefiles require shortened variable names, these alternate names are included in parentheses.

- *numLandParcels (NP)* total number of parcels in the specific geographic region (block groups or tracts), per the City of Boston's Master Address List.
- *NEWCON_count (NC_c)* is the count of parcels in a region with permits indicating new construction where there was not previously a structure. This includes *Erect/New Construction* permit types.
- *NEWCON_PP (NC_PP)* the proportion of parcels in a region with permits indicating construction where there was not previously a structure. This includes *Erect/New Construction* permit types.
- *NECON_DV_SUM (NC_DV_SUM)* is the total change in valuation in a region due to New Construction.
- *NEWCON_DV_MED (NC_DV_MED)* is the median change in valuation in a region due to New Construction.
- *ADDDEMO_count (AD_c)* is the count of parcels in a region with permits indicating extensive change to a structure that is not New Construction or Renovation. This includes *Amendment to a Long Form, Long Form/Alteration Permit, and Demolition* as Permit Types.
- *ADDDEMO_PP (AD_PP)* is the proportion of parcels in a region with permits indicating extensive change to a structure that is not New Construction or Demolition. This includes *Amendment to a Long Form, Long Form/Alteration Permit, and Demolition* as Permit Types.



- *ADDDEMO_DV_SUM (AD_DV_SUM)* is the total change in valuation in a region due to Additions and Demolitions.
- *ADDDEMO_DV_MED (AD_DV_MED)* is the median change in valuation in a region due to Additions and Demolitions.
- *RENO_count (R_c)* is the count of parcels in a region with permits indicating a small change to the building or a required improvement (renovation) for a building code that is not a New Construction, Demolition, or Addition. This includes many permit types.
- *RENO_PP (R_PP)* is the proportion of parcels in a region with permits indicating a small change to the building or a required improvement (renovation) for a building code that is not a New Construction, Demolition, or Addition. This includes many permit types.
- *RENO_DV_SUM (R_DV_SUM)* is the total change in valuation in a region due to Renovations.
- *RENO_DV_MED (R_DV_MED)* is the median change in valuation in a region due to Renovations.
- *ALTSTRUCT_count (AS_c)* is the count of parcels in a region qualifying as either an Addition or Renovation (alterations to existing structures).
- *ALTSTRUCT_PP (AS_PP)* is the proportion of parcels in a region qualifying as either an Addition or Renovation.
- *ALTSTRUCT_DV_SUM (AS_DV_SUM)* is the total change in valuation in a region due to Additions or Renovations.
- *ALTSTRUCT_DV_MED (AS_DV_MED)* is the median change in valuation in a region due to Additions and Renovations.