



Dan O'Brien • Research Director • Boston Area Research Initiative  
Radcliffe Institute for Advanced Study • Harvard University • 10 Garden Street, Cambridge, MA 02138  
Phone: 617-496-9975 • E-Mail: [daniel\\_obrien@radcliffe.harvard.edu](mailto:daniel_obrien@radcliffe.harvard.edu) • Web: [www.bostonarearesearchinitiative.net](http://www.bostonarearesearchinitiative.net)

## DOCUMENTATION FOR CENSUS INDICATORS

### Overview

This document summarizes a catalog of indicators describing the population and housing characteristics across the Commonwealth of Massachusetts, accessed from the U.S. Census Bureau's American Community Survey (ACS) and Decennial Census and curated by the Boston Area Research Initiative. The ACS indicators are based on five-year estimates that renew yearly, beginning with 2005-2009, and cover census tracts and block groups for Census 2010 geographies. The Decennial Census indicators are based on the 2010 Decennial Census, which counts every person in the population, and cover census tracts, block groups, and blocks for Census 2010 geographies.

The database is available through the Boston Data Library's repository and map (i.e., BostonMap). It includes a series of files of the form *ACS\_range\_TRACT.csv* or *ACS\_range\_BLKGRP.csv* (e.g., *ACS\_0812\_BLKGRP.csv*) and *DEC\_CENSUS\_2010\_TRACT.csv*, *DEC\_CENSUS\_2010\_BLKGRP.csv*, and *DEC\_CENSUS\_2010\_BLK.csv* for 2010 Census indicators.

The ACS 2005-2009 tract indicators are translated from 2000 census geographies to 2010 census geographies using the Longitudinal Tract Data Base (Logan, Xu, and Stults 2014). An earlier version of the 2008-2012 census data was uploaded to BostonMap with different variable names. An appendix document translates these earlier names to the conventions utilized for all other data. In addition, a separate appendix document details which indicators are available in the 2010 Decennial Census data because the Decennial Census includes only population and housing data.



## *Table of Contents*

1. Summary of Census Indicators	3
1.1. Description of Contents	3
1.2. Description of Variables	4
1.2.1. Sex, Age, and Density	4
1.2.2. Racial and Ethnic Composition	4
1.2.3. Economic	5
1.2.4. Family and Household	6
1.2.5. Highest Level of Education	7
1.2.6. Transport to Work	7
1.2.7. Housing	8
2. Appendices	9
2.1. Appendix A: Variables from 2008-2012 BostonMap Visualizations	10
2.2. Appendix B: 2010 Decennial Census Availability	13



## 1. Summary of Census Indicators

### 1.1. Description of Contents

The U.S. Census Bureau gathers information about population and housing characteristics across the country from a small percentage of the population each year, and from as much of the population as possible every ten years. This data set contains a series of indicators accessed or calculated from the decennial census and the American Community Survey (ACS; five-year estimates, e.g., 2008-2012). The indicators are provided across the Commonwealth of Massachusetts for 2010 census geographies at two levels: census block groups ( $N = 4985$ ) and census tracts ( $N = 1478$ ). Census block groups are nested within census tracts. Indicators are available for census blocks ( $N=157,508$ ) for only the Decennial Census. Census blocks are nested within census block groups.

The database begins with 2005-2009 as those are the first estimates available for 2010 census geographies, and go through the most recent release, 2011-2015. For each set of estimates there is an independent .csv file of the naming convention *ACS\_range\_unit.csv* (e.g., *ACS\_0812\_BLKGRP.csv*). Estimates beginning with 2008-2012 have also been visualized at the tract level on BostonMap. Those looking to work with the data in a spatial form can download the .shp from BostonMap. Additionally, for analyses confined to Boston one can merge the .csv with BARI's Geographical Infrastructure for the City of Boston, using BLK\_ID\_10 (blocks), BG\_ID\_10 (block groups), or CT\_ID\_10 (tracts) as the ID variable.

The ACS 2005-2009 tract indicators are translated from 2000 census tracts to 2010 census tracts using the Longitudinal Tract Data Base. For more information, see Logan, John R., Zengwang Xu, and Brian Stults. 2014. "Interpolating US Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database," *The Professional Geographer* 66(3): 412–420. ACS variables were downloaded and organized with assistance from the 'acs' package in R (<https://cran.r-project.org/web/packages/acs/acs.pdf>), developed by Ezra Glenn. 2010 Decennial Census variables were downloaded and organized with assistance from the 'UScensus2010' package in R (<https://cran.r-project.org/web/packages/UScensus2010/UScensus2010.pdf>), developed by Zack W. Almquist.



## 1.2. Description of Variables

The database contains 45 indicators organized into 7 main categories: 1) Sex, Age, and Density; 2) Racial and Ethnic Composition; 3) Economic measures; 4) Family and Household information; 5) Highest Level of Education completed; 6) Method of Transportation to Work; 7) Housing. All indicators are included for all five-year ranges at both geographic levels except where noted.

Not described under one of these categories but also included are: *TotalPop* or total population, *TotalHouseH* or total households, and *TotalHouseUnits* or total housing units.

### 1.2.1. Sex, Age, and Density

- *PopDen* indicates the population per square mile (population density) within the census geographical area
- *SexRatio* indicates the ratio of female to male individuals.
  - *Note:* Not calculated for census geographies with residents of only one sex.
- *AgeU18* indicates the percentage of residents ages under 18.
- *Age1834* indicates the percentage of residents with ages between 18 and 34.
- *Age3564* indicates the percentage of residents with ages between 35 and 64.
- *Age065* indicates the percentage of residents ages 65 and older.

### 1.2.2. Racial and Ethnic Composition

- *White* indicates the proportion of White Non-Hispanic residents.
- *Black* indicates the proportion of Black Non-Hispanic residents.
- *Asian* indicates the proportion of Asian Non-Hispanic residents.



- *Hispanic* indicates the proportion of Hispanic residents.
- *EthHet* is a measure of ethnic heterogeneity.
  - Calculated using a Herfindahl index,  $1 - \sum s_i^2$ , where  $s_i$  is the proportion of residents belonging to ethnicity  $i$ . Higher values indicate a more diverse school. This index represents the percentage of times two randomly selected people in this school would differ by race/ethnicity.
  - *Note:* Includes only White, Black, Asian, and Hispanic, with the exception of the map version of 2008-2012 estimates, which also incorporates individuals of two or more races.
- *TwoOrMore* indicates the proportion of residents who identify as two or more races on the census.
- *ForeignBorn* indicates the proportion of residents born in a foreign country.
  - *Note:* Not available for census block groups.

### 1.2.3. Economic

- *MedHouseIncome* indicates the median household income.
  - Median household incomes over \$250,000 are coded as \$250,001.
- *PubAssist* indicates the proportion of the population receiving public assistance.
- *GINI* is the Gini index, a commonly-used measure of income inequality.
  - Gini has a range from 0 = perfect equality to 1 = perfect inequality.
  - *Note:* Not available for census block groups.
  - *Note:* GINI estimates on mapped version for 2008-2012 differ slightly from those later accessed from the ACS. For rigorous analysis of this five-year estimate, please use the version in the data repository.
- *FamPovPer* indicates the proportion of families with income falling below the poverty line.



- *UnempRate* indicates the proportion of residents who are unemployed (unemployment rate).
  - *Note:* Available for census block groups, though not for the 2006-2010 and 2005-2009 estimates.
  - *Note:* Calculated based on the population greater than 16, which differs slightly in the map version of 2008-2012. For rigorous analysis of this five-year estimate, please use the version in the data repository.

#### 1.2.4. Family and Household

Indicators in this category that reflect proportions of the population are based on the number of households, or, when appropriate, the number of households that are families.

- *FamHousePer* indicates the proportion of households occupied by a family (children plus guardian(s)).
- *FamHeadPer* indicates the proportion of family households that are female-headed.
- *SameSexCoupPer* indicates the proportion of households with unmarried couples for which the couple is same-sex.
  - *Note:* Not available for census block groups.
  - *Note:* Calculation may have changed after the creation of the map version for 2008-2012. For rigorous analysis of this five-year estimate, please use the version in the data repository.
- *GrandHeadPer* indicates the proportion of family households that are grandparent-headed.
  - *Note:* Not available for census block groups.
  - *Note:* Calculation may have changed after the creation of the map version for 2008-2012. For rigorous analysis of this five-year estimate, please use the version in the data repository.



### *1.2.5. Highest Level of Education*

All indicators in this category reflect proportions of the population of 25 years or older that have achieved each as their highest level of education.

- *LessThanHS* indicates the proportion of residents who have not completed a high school degree or equivalent.
- *HSGrad* indicates the proportion of residents who have completed a high school education or equivalent (e.g. GED).
- *SomeCollorAssoc* indicates the proportion of residents who have completed some years of college but have not attained any college degree.
- *Bach* indicates the proportion of residents who have completed a Bachelor's Degree.
- *Master* indicates the proportion of residents who have completed a Master's Degree.
- *Prof* indicates the proportion of residents who have completed a professional degree.
- *Doc* indicates the proportion of residents who have completed a doctoral degree (e.g., PhD, MD).

### *1.2.6. Transport to Work*

Note that for all commute time calculations, the calculation of the denominator may have changed after the creation of the map version for 2008-2012 estimates. Though results are highly consistent, for rigorous analysis of this five-year estimate please use the version in the data repository.

- *ByCar* indicates the proportion of residents who commute to their place of employment by car.
- *ByPubTrans* indicates the proportion of residents who commute by public transit.



- *ByBike* indicates the proportion of residents who commute by bicycle.
- *ByWalk* indicates the proportion of residents who commute by walking.
- *CommuteLess10* indicates the proportion of residents who have a commute of less than 10 minutes to their place of employment.
- *Commute1030* indicates the proportion of residents who have commuting time of between 10 and 29 minutes.
- *Commute3060* indicates the proportion of residents who have commuting time of between 30 and 59 minutes.
- *Commute6090* indicates the proportion of residents who have commuting time of between 60 and 89 minutes.
- *CommuteOver90* indicates the proportion of residents who have commuting time of 90 minutes or more.

### 1.2.7. Housing

Housing variables that reference proportions are based on the number of housing units rather than the total population. Also note that for owner- and renter-occupied measures, the calculation of the denominator for the map version for 2008-2012 estimates was over all units; thereafter it has been of the proportion of occupied units (i.e., subtracting vacant units). Though results are highly consistent, for rigorous analysis of this five-year estimate, please use the version in the data repository.

- *VacantUnitsPer* indicates the proportion of vacant units.
- *RentersPer* indicates the proportion of renter households.
- *HomeOwnPer* indicates the rate of homeownership.
- *MedGrossRent* indicates the median gross rent in USD.
  - Median rents over \$2000 are coded as \$2001.





- *MedYrBuilt* indicates the median year of construction for buildings in ten-year intervals.
- *MedYrBuiltRaw* indicates the median year of construction for buildings in ten-year intervals.
  - *Note:* Not included in map.
- *MedHomeVal* indicates the median value of homes in a given census geography.
  - Median home values over \$1,000,000 are coded as \$1,000,001.



## 2. Appendix

### 2.1. Appendix A: Variable Name Change

The following table displays a translation of variable names as they have changed from the 2008-2012 census indicators uploaded to BostonMap to the variable names reported in the remainder of this document. 2008-2012 indicators accessed through the data repository have already been adjusted.

Indicator	Data Repository Variable Name	BostonMap Variable Name
Total Population	TotalPop	totalpop
Total Households	TotalHouseH	households
Total Housing Units	TotalHouseUnits	units
Pop Density	PopDen	popdens
Sex Ratio (Female to Male)	SexRatio	sexratio
Age 18 and under (percent)	AgeU18	punder18
Age 18-34 (percent)	Age1834	p18to34
Age 35-64 (percent)	Age3564	p35to64
Age 65 and over (percent)	AgeO65	p65older
White (Non-Hispanic)	White	pwhite
Black (Non-Hispanic)	Black	pblack
Asian (Non-Hispanic)	Asian	pasian
Hispanic	Hispanic	phispanic
Ethnic Heterogeneity	EthHet	HI



Two or More Races	TwoOrMore	ptwomore
Foreign Born	ForeignBorn	pforborn
Median Household Income	MedHouseIncome	medincome
Public Assistance Recipients	PubAssist	ppubassis
GINI	GINI	GINI
Family Poverty (%)	FamPovPer	pfampov
Unemployment Rate	UnempRate	punemploy
Family Households (%)	FamHousePer	pfamh
Female-Headed Households (%)	FemHeadPer	pfhh
Same Sex Couple Households (%)	SameSexCoupPer	psamesex
Grandparent-Headed Households (%)	GrandHeadPer	pgrandresp
Less than HS	LessThanHS	PlessHS
HS or Equivalent	HSGrad	pHS
Some College	SomeCollorAssoc	psomecol
Bachelor Degree	Bach	pBA
Masters Degree	Master	pMA
Professional Degree	Prof	pprofd
Doctoral Degree	Doc	pdoc
By Car	ByCar	pcar
By Public Transit	ByPubTrans	ppublic
By Bike	ByBike	pbike



By Walking	ByWalk	pwalk
Commute: <10 min	CommuteLess10	pless10
Commute: 10-30 min	Commute1030	ptenthir
Commute: 30-60 min	Commute3060	pthirsix
Commute: 60-90 min	Commute6090	psixnine
Commute: >90 min	CommuteOver90	pninemo
Vacant Units (%)	VacantUnitPer	pvac
Renters (%)	RentersPer	prentocc
Home Owners (%)	HomeOwnPer	pownocc
Median Gross Rent	MedGrossRent	grossrent
Median Year Built	MedYrBuilt	medyrblt
Median Home Value	MedHomeVal	ownmedval



## 2.2. Appendix B: 2010 Decennial Census Availability

The following table displays which indicators are available from the 2010 Decennial Census data.

Indicator	Data Repository Variable Name	Available in 2010 Census Data
Total Population	TotalPop	x
Total Households	TotalHouseH	x
Total Housing Units	TotalHouseUnits	x
Pop Density	PopDen	x
Sex Ratio (Female to Male)	SexRatio	x
Age 18 and under (percent)	AgeU18	x
Age 18-34 (percent)	Age1834	x
Age 35-64 (percent)	Age3564	x
Age 65 and over (percent)	AgeO65	x
White (Non-Hispanic)	White	x
Black (Non-Hispanic)	Black	x
Asian (Non-Hispanic)	Asian	x
Hispanic	Hispanic	x
Ethnic Heterogeneity	EthHet	x
Two or More Races	TwoOrMore	x
Foreign Born	ForeignBorn	
Median Household Income	MedHouseIncome	



Public Assistance Recipients	PubAssist	
GINI	GINI	
Family Poverty (%)	FamPovPer	
Unemployment Rate	UnempRate	
Family Households (%)	FamHousePer	x
Female-Headed Households (%)	FemHeadPer	x
Same Sex Couple Households (%)	SameSexCoupPer	x (not available by block group or block)
Grandparent-Headed Households (%)	GrandHeadPer	
Less than HS	LessThanHS	
HS or Equivalent	HSGrad	
Some College	SomeCollorAssoc	
Bachelor Degree	Bach	
Masters Degree	Master	
Professional Degree	Prof	
Doctoral Degree	Doc	
By Car	ByCar	
By Public Transit	ByPubTrans	
By Bike	ByBike	
By Walking	ByWalk	
Commute: <10 min	CommuteLess10	
Commute: 10-30 min	Commute1030	



Commute: 30-60 min	Commute3060	
Commute: 60-90 min	Commute6090	
Commute: >90 min	CommuteOver90	
Vacant Units (%)	VacantUnitPer	x
Renters (%)	RentersPer	x
Home Owners (%)	HomeOwnPer	x
Median Gross Rent	MedGrossRent	
Median Year Built	MedYrBuilt	
Median Home Value	MedHomeVal	