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

DOCUMENTATION FOR CENSUS INDICATORS (2020 GEOGRAPHIES)

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 2016-2020, and cover census tracts and block groups for Census 2020 geographies. The Decennial Census indicators are based on the 2020 Decennial Census, which counts every person in the population, and cover census tracts, block groups, and blocks for Census 2020 geographies.

The database is available through the Boston Data Library's repository and map (i.e., Boston Area Research Map). It includes a series of files of the form *ACS_range_TRACT.csv* or *ACS_range_BLKGRP.csv* (e.g., *ACS_1620_BLKGRP.csv* for block group-level 2016-2020 estimates) and *DEC_CENSUS_2020_TRACT.csv*, *DEC_CENSUS_2020_BLKGRP.csv*, and *DEC_CENSUS_2020_BLK.csv* for 2020 Census indicators.

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 6

1.2.6. Transport to Work 7

1.2.7. Housing 7

1.2.8. Geographic Contextual Information 8

2. Appendix 9

2.1. Appendix A: Variable Names *9*

2.2. Appendix B: 2020 Decennial Census Availability *12*

2.3. Appendix C: Block Groups and Tracts Covering Multiple Municipalities *15*

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., 2016-2020). The indicators are provided across the Commonwealth of Massachusetts for 2020 census geographies at two levels: census block groups (*N* = 5,116) and census tracts (*N* = 1,620). Census block groups are nested within census tracts. Indicators are available for census blocks (N=107,278) for only the Decennial Census. Census blocks are nested within census block groups. The U.S. Census Bureau's 2020 Census Disclosure Avoidance System is designed to better protect respondents' information while preserving accuracy of the census. It adds statistical noise based on a mathematical algorithm that is embedded in the metrics curated here. More detailed information about this method can be found here, https://www.census.gov/library/fact-sheets/2021/protecting-the-confidentiality-of-the-2020-census-redistricting-data.html.

The database begins with 2016-2020 and goes through the most recent release, 2020-2024. For each set of estimates there is an independent .csv file of the naming convention *ACS_range_unit.csv* (e.g., *ACS_1620_BLKGRP.csv* for block group-level 2016-2020 estimates). Estimates beginning with 2016-2020 have also been visualized at the tract level on the Boston Area Research Map. Those looking to work with the data in a spatial form can download the .shp from Boston Area Research Map. Additionally, for analyses confined to Boston one can merge the .csv with BARI's Geographical Infrastructure for the City of Boston, using BLK_ID_20 (blocks), BG_ID_20 (block groups), or CT_ID_20 (tracts) as the ID variable.[1]

[1] ACS variables were downloaded and organized with assistance from the 'tidycensus' package in R (https://cran.r-project.org/web/packages/tidycensus/tidycensus.pdf), developed by Kyle Walker, Matt Herman, and Kris Eberwein.

1.2. Description of Variables

The database contains 52 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. These are accompanied by a series of variables providing geographic context. 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.
    - *Note:* Not available for census blocks.
- *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.
- *AgeO65* indicates the percentage of residents ages 65 and older.

1.2.2. Racial and Ethnic Composition

- *ForBorn* indicates the proportion of residents born in a foreign country.
    - *Note:* Not available for census block groups.
- *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.
- *AIAN* indicates the proportion of American Indian and Alaska Native Non-Hispanic residents.
- *NHOPI* indicates the proportion of Native Hawaiian and Other Pacific Islander Non-Hispanic residents.
- *Other* indicates the proportion of Some Other Race Non-Hispanic residents.
- *TwoOrMore* indicates the proportion of residents who identify as two or more races on the census.
- *EthHet* is a measure of ethnic heterogeneity.
    - Calculated using the Census's racial and ethnic diversity index (DI), or the chance that two people chosen at random will be from different racial and ethnic groups, also known as the Herfindahl Index[2]:

$$DI = 1 - (White^2 + Black^2 + Asian^2 + Hispanic^2 + AIAN^2 + NHOPI^2 + Other^2 + TwoOrMore^2)$$

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.
- *FamPovPer* indicates the proportion of families with income falling below the poverty line.

[2] Equation accessed from https://www2.census.gov/programs-surveys/decennial/2020/data/demographic-and-housing-characteristics-file/age-group-visualization/exploring-age-groups-technical-documentation.pdf

- *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.

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)).
- *FemHeadPer* indicates the proportion of family households that are female-headed.

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).
- *SomeColl* 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.

- *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.
- *ByAuto* 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.

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.

- *VacantUnitPer* 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.
    - Per the coding scheme used by the Census, median rents over $3,500 are coded as $3,501.
- *MedHomeVal* indicates the median value of homes in a given census geography.
    - Median home values over $2,000,000 are coded as $2,000,001.
- *MedYrBuiltRaw* indicates the median year of construction for buildings.
- *MedYrBuilt* indicates the median year of construction for buildings in ten-year intervals.
- *MedYrMovedIn* indicates the median year of move-in for all residents.
- *MedYrRentMovedIn* indicates the median year of move-in for residents who are identified as renters.
- *MedYrOwnMovedIn* indicates the median year of move-in for residents who are identified as homeowners.

1.2.8. Geographic Contextual Information

- *County* indicates which county a geography is located in.
- *Town* indicates which town (or city) a geography is located in according to the Census. Tracts or block groups split across multiple town boundaries are either designated as "Multiple Municipalities" if that geography is the only one that represents one of the towns (Appendix C). Or for cases where the towns are also represented in another tract or block group, it was assigned to the town that has the greatest percentage of the tract or block group population.

2. Appendix

2.1. Appendix A: Variable Names

The following table displays a translation of variable names from the Data Repository to the Boston Area Research Map.

Indicator	Data Repository Variable Name	Boston Area Research Map Variable Name
Total Population	TotalPop	Not included in map
Total Households	TotalHouseH	Not included in map
Total Housing Units	TotalHouseUnits	Not included in map
Pop Density	PopDen	Population Density
Sex Ratio (Female to Male)	SexRatio	Sex Ratio
Age 18 and under (%)	AgeU18	Age Under 18
Age 18-34 (%)	Age1834	Age 18-34
Age 35-64 (%)	Age3564	Age 35-64
Age 65 and over (%)	AgeO65	Age Over 65
Born in a foreign country	ForBorn	Not included in map
White (Non-Hispanic)	White	White
Black (Non-Hispanic)	Black	Black
Asian (Non-Hispanic)	Asian	Asian
Hispanic	Hispanic	Hispanic
American Indian and Alaska Native (Non-Hispanic)	AIAN	Not included in map
Native Hawaiian and Other Pacific Islander (Non-Hispanic)	NHOPI	Not included in map
Some Other Race (Non-Hispanic)	Other	Not included in map
Two or More Races	TwoOrMore	Two or More Races
Ethnic Heterogeneity	EthHet	Ethnic Heterogeneity

Indicator	Data Repository Variable Name	Boston Area Research Map Variable Name
Median Household Income	MedHouseIncome	Median House Income
Public Assistance Recipients	PubAssist	Public Assistance Recipients
GINI	GINI	Not included in map
Family Poverty (%)	FamPovPer	Not included in map
Unemployment Rate	UnempRate	Unemployment Rate
Family Households (%)	FamHousePer	Not included in map
Female-Headed Households (%)	FemHeadPer	Not included in map
Less than HS	LessThanHS	Highest Education: Less than High School
HS or Equivalent	HSGrad	Highest Education: High School
Some College	SomeColl	Highest Education: Some College
Bachelor's Degree	Bach	Highest Education: Bachelor's Degree
Master's Degree	Master	Highest Education: Master's Degree
Professional Degree	Prof	Highest Education: Professional Degree
Doctoral Degree	Doc	Highest Education: Doctorate Degree
Commute: <10 min	CommuteLess10	Commute Less Than 10 Mins
Commute: 10-30 min	Commute1030	Commute 10-30 Mins
Commute: 30-60 min	Commute3060	Commute 30-60 Mins
Commute: 60-90 min	Commute6090	Commute 60-90 Mins
Commute: >90 min	CommuteOver90	Commute Over 90 Mins

Indicator	Data Repository Variable Name	Boston Area Research Map Variable Name
By Auto	ByAuto	Commute By Auto
By Public Transit	ByPubTrans	Commute By Public Transportation
By Bike	ByBike	Commute By Bike
By Walk	ByWalk	Commute By Walking
Vacant Units (%)	VacantUnitPer	Not included in map
Renters (%)	RentersPer	Not included in map
Homeowners (%)	HomeOwnPer	Not included in map
Median Gross Rent	MedGrossRent	Not included in map
Median Home Value	MedHomeVal	Not included in map
Median Year Built (Raw)	MedYrBuiltRaw	Not included in map
Median Year Built (10-year intervals)	MedYrBuilt	Not included in map
Median Year Moved In – All Residents	MedYrMovedInraw	Not included in map
Median Year Moved In – Renters	MedYrRentMovedIn	Not included in map
Median Year Moved In – Homeowners	MedYrOwnMovedIn	Not included in map
County	County	Not included in map
Town	Town	Not included in map

2.2. Appendix B: 2020 Decennial Census Availability

The following table displays which indicators are available from the 2020 Decennial Census data as the ACS includes a broader range of measures.

Indicator	Data Repository Variable Name	Available in 2020 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
Born in a Foreign Country	ForBorn	
White (Non-Hispanic)	White	X
Black (Non-Hispanic)	Black	X
Asian (Non-Hispanic)	Asian	X
Hispanic	Hispanic	X
American Indian and Alaska Native (Non-Hispanic)	AIAN	X
Native Hawaiian and Other Pacific Islander (Non-Hispanic)	NHOPI	X
Some Other Race (Non-Hispanic)	Other	X
Two or More Races	TwoOrMore	X
Ethnic Heterogeneity	EthHet	X
Median Household Income	MedHouseIncome	

Indicator	Data Repository Variable Name	Available in 2020 Census Data
Public Assistance Recipients	PubAssist	
GINI	GINI	
Family Poverty (%)	FamPovPer	
Unemployment Rate	UnempRate	
Family Households (%)	FamHousePer	X
Female-Headed Households (%)	FemHeadPer	X
Less than HS	LessThanHS	
HS or Equivalent	HSGrad	
Some College	SomeColl	
Bachelor's Degree	Bach	
Master's 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	

Indicator	Data Repository Variable Name	Available in 2020 Census Data
Commute: 30-60 min	Commute3060	
Commute: 60-90 min	Commute6090	
Commute: >90 min	CommuteOver90	
Vacant Units (%)	VacantUnitPer	X
Renters (%)	RentersPer	X
Homeowners (%)	HomeOwnPer	X
Median Gross Rent	MedGrossRent	
Median Year Built	MedYrBuilt	
Median Home Value	MedHomeVal	
Median Year Moved in – All Residents	MedYrMovedln	
Median Year Moved in – Renters	MedYrRentMovedln	
Median Year Moved in – Homeowners	MedYrOwnMovedln	
County	County	X
Town	Town	X

2.3. Appendix C: Block Groups and Tracts Covering Multiple Municipalities

The following table displays which block groups and tracts were assigned as "Multiple Municipalities" and which towns those geographies overlap. Towns with only one occurrence are bolded.

Census Block Group ID (BG_ID_20)	**Towns**
250039343002	**EGREMONT, MOUNT WASHINGTON**
250110401001	CHARLEMONT, **MONROE, ROWE**
250110401002	COLRAIN, **HEATH**
250110401004	CHARLEMONT, **HAWLEY**
250110402001	BERNARDSTON, **LEYDEN**
250110404001	**ERVING**, WARWICK
250138130012	**BLANDFORD, TOLLAND**
250158227003	**MIDDLEFIELD, WORTHINGTON**
Census Tract Group ID (CT_ID_20)	**Towns**
25003931300	**FLORIDA, SAVOY**
25003931400	**PERU, WINDSOR**
25003932200	**BECKET, WASHINGTON**
25003933200	**MONTEREY, TYRINGHAM**
25003933400	**OTIS, SANDISFIELD**
25003934300	**ALFORD, EGREMONT, MOUNT WASHINGTON**
25003935100	**HANCOCK, NEW ASHFORD, RICHMOND**
25007200400	**AQUINNAH, CHILMARK, GOSNOLD, WEST TISBURY**
25011040100	**CHARLEMONT, COLRAIN, HAWLEY, HEATH, MONROE, ROWE**
25011040200	**BERNARDSTON, GILL, LEYDEN**
25011040400	**ERVING, WARWICK, WENDELL**
25011040600	**LEVERETT, NEW SALEM, SHUTESBURY**

Census Tract Group ID (CT_ID_20)	Towns
25011040800	**SUNDERLAND, WHATELY**
25011041501	**ASHFIELD, CONWAY**
25011041502	**BUCKLAND, SHELBURNE**
25013813001	**BLANDFORD, CHESTER, TOLLAND**
25013813002	**GRANVILLE, MONTGOMERY, RUSSELL**
25013813801	**HOLLAND, WALES**
25015820205	BELCHERTOWN, **PELHAM**
25015822607	**CHESTERFIELD, GOSHEN, WILLIAMSBURG**
25015822700	**CUMMINGTON, MIDDLEFIELD, PLAINFIELD, WORTHINGTON**
25027723100	**HARDWICK, NEW BRAINTREE**
25027746100	HOPEDALE, **MENDON**