**Building Your Base Map of Boston**

|  |
| --- |
| **Before we start: Saving your data and map** |

The base map that we create today will be perfect for your final project – you will be able to take this base map, add your crime data to it, and start analyzing. That means that you will probably want to be very careful as you build your base map, and you will want to carefully save it so that you can open it again later! Having to recreate it later would be a big waste of time, so make sure that you follow “best practices” as you download and store your data and when you save your map. Remember that saving your map as a .mxd file will not save the data with it, so if you move the .mxd file without the accompanying data, it will “break” and none of your layers will display correctly. I recommend saving your data and your .mxd file in the same folder on your USB storage drive and leaving them there – that’s the best way to make sure everything stays together.

|  |
| --- |
| **Downloading the necessary shapefiles** |

At this point, you might not know exactly which map features you will need for your final project, so let’s build a really comprehensive base map so you have lots of options later – you can always hide layers that you don’t need.

**Your map should have at least the following layers:**

|  |  |
| --- | --- |
| **Layer** | **Where to find it** |
| City of Boston Boundary | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |
| Boston Neighborhoods | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |
| Boston Police Districts | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |
| Boston 2010 **Census Tracts**, **Block Groups**, and **Blocks** | [MassGIS](https://docs.digital.mass.gov/dataset/massgis-data-datalayers-2010-us-census?_ga=2.174341730.1319140079.1549384499-1669285022.1546791647) (three layers in one download) |
| The ADDRFEAT (Address Range-Feature) streets layer\* | [Census TIGER (for Suffolk County)](https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2018&layergroup=Relationship+Files) |
| Open Space | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |
| Water | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |

**You may want to consider some of the following optional layers to deepen your analysis:**

|  |  |
| --- | --- |
| **Layer** | **Where to find it** |
| Boston Public Schools | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |
| Boston Police Stations | Analyze Boston ([data.boston.gov](https://data.boston.gov/group/geospatial)) |
| MBTA bus routes and stops | [MassGIS](https://docs.digital.mass.gov/dataset/massgis-data-mbta-bus-routes-and-stops) |
| MBTA parking lots | [MassGIS](https://docs.digital.mass.gov/dataset/massgis-data-layers-massdot?_ga=2.17912697.1319140079.1549384499-1669285022.1546791647) |
| Trains and MBTA Commuter Rail | [MassGIS](https://docs.digital.mass.gov/dataset/massgis-data-trains?_ga=2.178290532.1319140079.1549384499-1669285022.1546791647) |
| MBTA Rapid Transit (the T) | [MassGIS](https://docs.digital.mass.gov/dataset/massgis-data-mbta-rapid-transit?_ga=2.17912697.1319140079.1549384499-1669285022.1546791647) |

|  |
| --- |
| **Adding layers to your map** |

Once you have downloaded everything you want and saved the files to your working directory (a folder on your USB drive, hopefully!), you can start adding layers to your map. I recommend starting with the Boston city boundary, because you know this will be your biggest layer (so it goes on the ‘bottom’). The other benefit of adding this layer is that you can use it to “cut out” the pieces you need from map layers for larger areas.

For example, look at what happens when I add the Boston city boundary and then add the census tracts layer that I downloaded from MassGIS:



What happened?

Well, remember that we grabbed the census data package (census tracts, block groups, and blocks) from the MassGIS website, so the layer is for **all of Massachusetts**, not just Boston. We could leave it like this, I guess, and just put our Boston layers on top of it, but that’s pretty ugly. A better idea would be to get rid of all the of the stuff we don’t need. We can do this by selecting only the census tracts that are within the Boston city boundary and exporting the data for those tracts as its own shapefile. Then we add *that* shapefile to our map and get rid of this all-Massachusetts shapefile. Let’s do that now.

|  |
| --- |
| **“Select by location” to include only Boston features** |



Go up to the Selection menu and choose “Select by location…”. The selection method should say “select features from.” We want to select the features **from** the census tract layer – that’s the layer with the things we want. So for the “Target layer(s)” click “CENSUS2010TRACTS\_POLY”. Below that, you will see a dropdown menu that says “Source layer” – this is the layer that we want to use as a boundary. Make the source layer “City\_of\_Boston\_Boundary.”

The next option says “Spatial selection method for target layer feature(s):”. This is asking you how you want the program to select the features from the target layer. We only want the census tracts that are within the city boundary, so we’ll select “**have their centroid in the source layer feature**.” (Be careful here, there are lots of options – make sure you pick the right one.)

If you click “Apply” and look at your map (you might need to move the Select By Location window out of the way, but don’t close it yet), you should see the census tracts in Boston light up in a bright blue – this means they are selected. If you see census tracts outside of Boston selected, something probably went wrong, so go back and check your work.



If you are happy with the selection, click “OK.”

Now, leaving those census tracts selected, right-click on the CENSUS2010TRACTS\_POLY layer and click “Data” then “Export Data.” An Export Data window will open. What we are doing here is telling the program to take *only* the selected data and to save it as a new file. Next to Export it should say “Selected Features.” Use the same coordinate system as this layer’s source data. Then click the little folder icon next to the “Output feature class” box. Use the dropdown menu to find your working directory (your USB drive). Name your file “**Boston 2010 Census Tracts**” and change the “Save as type” option to **Shapefile**. Then click Save. When you return to the Export Data window, click OK. When it asks if you want to add the exported data to your map as a layer, click ‘Yes.’



You can do this procedure to “cut down” any shapefile to just the area that you want. For example, you may need to do this with the other MassGIS and Census TIGER shapefiles you will use to create your base map. We **only** want Boston features displayed on our final map.

|  |
| --- |
| **Leaving the nest!** |

Now that you know how to download files, add layers to your map, and use “select by location” to take only the map features that are within your chosen boundary (like the Boston city boundary), you are free to build your base map! Earlier in this document I listed the layers you should definitely have, then a lot of optional layers. Concentrate on getting the required layers downloaded and added to your map, and on adjusting the symbols (colors and outlines) and the order of layers so that your map looks nice. Then you can focus on adding the extra options.

|  |
| --- |
| **OPTIONAL ADVANCED STEP: ACS Data** |

Recently (like last week), the Boston Area Research Initiative (BARI) released data from the American Community Survey that can be linked to either census tracts or block groups (which you already have on your map, if you completed the steps above). The American Community Survey is a survey administered by the Census Bureau and collects a wide range of information (you can read more about it here: <https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018_ch01.pdf>).

You can download the ACS data files for block groups (recommended) or census tracts (larger, less fine detail) from Blackboard – I have posted them there to make it a bit easier for you.

If you download these layers, add them to your map, and open them (right-click and Open Attribute Table), you’ll see that they are not visible layers yet. They are just spreadsheets of information. But if you look at the ACS\_1317\_BLKGRP file, for example, you will see that the first column is BG\_ID\_10. This is a 10-digit unique identifier for each block group – like a unique label that only this block group has. Every block group in the US has one! It is based on a FIPS code, the Federal Information Processing codes.

The awesome thing about FIPS codes is that they are unique and so they can be used to join data from multiple sources. If you have a map layer of block groups with FIPS codes, and then a data set of information for block groups identified by FIPS codes… you can match them up! Let’s do that with our ACS file and our Boston block groups layer.

We’ve already identified that the ACS data has the variable **BG\_ID\_10** with the block group FIPS code. Let’s see which column in our block group layer has the FIPS code in it. Open the attribute table for the Boston block group layer you created.



In the table above, you can see that our **BG\_ID\_10** column from the ACS file best matches up with the **GEOID10** column in the block groups layer. HOWEVER, when I tried to go ahead and join these layers up, I realized that we have a problem – the **BG\_ID\_10** column is coded as a numeric variable (“double”, in ArcGIS language) and the **GEOID10** column is coded as a text variable (“string”). They won’t match. But we can fix that!

Open the attribute table for the Boston block group layer. In the top left corner, there is a button that says “Table Options” if you hover over it. Click the downward-facing arrow next to it and select ‘Add Field.’ We are going to add a new column to this table and then copy GEOID10 into it and make it a “double.”



Click ‘Add Field’ and a new window will open. Next to “Name”, type BG\_ID\_10. We might as well use the same variable name to make our match even easier! Change “Type” to Double. Change scale to 10. Click OK.

If you scroll over to the far right of your table, you will now see your new column added to the end. All of the values are blank. Now we need to copy the values from GEOID10. Right-click the name of the column and select ‘Field Calculator.’ We are going to tell ArcGIS to just copy all of the values from the GEOID10 column. All you need to do is double-click GEOID10 from the list, then click OK. You should now see all of the identification numbers in the column.



NOW we can join our layers together. You can close the attribute table. Right-click on the block group layer and select **Joins and Relates… > Join…**. We want to “Join attributes from a table.” First we have to choose which column in our block groups layer we want to link up to the ACS data. We just created it! Select BG\_ID\_10 from the drop-down menu.

Next, we need to choose which table we want to join it to. Select **ACS\_1317\_BLKGRP.csv**. We want to match up to the **BG\_ID\_10** column in this table, too, so select that for step 3. Then click OK.

That should be it! If you right-click your block group layer now and open the attribute table, you should be able to scroll all the way across and see all of the new columns of ACS data added on to the end of it. You did it! Now you can use this information to create different thematic maps or to add extra information to your final project analysis.

If you would like, you can also repeat the steps above but for the **census tracts** layer and the **ACS\_1317\_TRACT.csv** table.

Here’s a list of some of the variables available to you now…

|  |  |
| --- | --- |
| PopDen | Population per square mile |
| SexRatio | The ratio of female to male individuals |
| AgeU18 | Percentage of residents aged under 18 |
| Age1834 | Percentage of residents with ages between 18 and 34 |
| Age3564 | Percentage of residents with ages between 35 and 64 |
| AgeO65 | Percentage of residents age 65 and older |
| White, Black, Asian, Hispanic | Each column indicates the proportion of residents of that race or ethnicity. White is White Non-Hispanic, Black is Black Non-Hispanic, Asian is Asian Non-Hispanic. |
| EthHet | This is supposed to be a measure of ethnic heterogeneity, but it appears to be broken in this data set – I am contacting the authors to see if there was an error. **Do not use this variable at this time.** |
| TwoOrMore | Proportion of residents who identify as two or more races. |
| ForeignBorn | Proportion of residents born in a foreign country; **not available for block groups** (use with census tracts table and layer only). |
| MedHouseIncome | Median household income – **note: median household incomes over 250,000 are coded as 250,001.** |
| PubAssist | Proportion of the population receiving public assistance (type not specified). |
| GINI | Gini index, a measure of income inequality. Ranges from 0 (perfect equality) to 1 (perfect inequality). **Available for tracts only, not block groups.** |
| FamPovPer | Proportion of families with income falling below the poverty line. |
| UnempRate | Proportion of residents over age 16 who are unemployed. |
| FamHousePer | Proportion of households occupied by a family (children plus guardian(s)). |
| SamSexCoupPer | Proportion of households with unmarried couples for which the couple is same-sex. **Available for tracts only, not block groups.** |
| GrandHeadPer | Proportion of family households that are grandparent-headed. **Available for tracts only, not block groups.** |
| LessThanHS | Proportion of residents who have not completed high school or equivalent. |
| HSGrad | Proportion of residents who have completed high school or GED. |
| SomeCollorAssoc | Proportion of residents who have completed some years of college but not attained any college degree. |
| Bach | Proportion of residents who have completed a Bachelor’s Degree |
| Master | Proportion of residents who have completed a Master’s Degree |
| Prof | Proportion of residents who have completed a professional degree. |
| Doc | Proportion of residents who have completed a doctoral degree (PhD, MD) |
| ByCar, ByPubTrans, ByBike, ByWalk | Proportion of residents who commute (to work) by car, public transit, bike, or walking. |
| CommuteLess10, Commute1030, Commute3060, Commute6090, CommuteOver90 | Proportion of residents who commute less than 10 minutes, 10-29 minutes, 30-59 minutes, 60-89 minutes, or 90+ minutes. |
| VacantUnitsPer | Proportion of vacant units |
| RentersPer | Proportion of renter households |
| HomeOwnPer | Proportion of homeowner h ouseholds. |
| MedGrossRent | Median gross rent in dollars. **Note: median rents over $2000 are coded as $2001.** |
| MedYrBuilt | Median year of construction for buildings, in ten-year intervals. |
| MedHomeVal | Median value of homes. **Note: median home values over $1,000,000 are coded as $1,000,001.** |