**Creating Choropleths (or Shaded Grid Maps)**

**by Joining Map Layers**

|  |
| --- |
| **Introduction** |

**Before we begin:** Please use the files you downloaded in the “Dr Stone’s Map Layers and Data” folder to assemble a map of Boston. Use the City Boundary, Open Space, and ADDRFEAT layers. *Remember to put your map layers in the correct order and change the symbology for each layer to something professional. I recommend an outlined, no-fill option for the City Boundary, some shade of green for parks, and black thin lines for roads.*

**Make sure your crime data is displayed on your map:** You should have a saved layer of point features for your 2017 and 2018 crime data by this point. Make sure that these layers are saved as shapefiles, not as “Events.” If you have not saved these layers as shapefiles, follow the steps below:

1. Drag your Excel file of crime data to your Table of Contents.
2. Right-click it and select “Display XY data…”
3. Set the **X Field** to *Longitude* or *Long*, and the **Y Field** to *Latitude* or *Lat*.
4. Click OK, and your crime data will be displayed on your map in a point layer called something like “2018 Boston YourCrimeType.csv Events”
5. Right-click this layer, select **Export Data…** and save this layer as a shapefile in the same place that you are saving all other layers for this map.
6. Add this new shapefile to your map when prompted.
7. Right-click the old “YourCrimeType.csv Events” layer and select “Remove” to take it off your map. Leave the shapefile layer on.
8. Make sure you do these steps for your 2017 and 2018 crime data files.

Your table of contents should look something like this, but with whichever crime type you have chosen:

For today’s exercise, it will be more useful to us if we **merge** our 2017 and 2018 data into a single layer.

To do this, open ArcToolbox  (the button is located just below your Windows and Help menus at the top of the screen). In ArcToolbox, click the plus sign next to **Data Management Tools** to expand that menu, then click the plus sign next to **General** to expand that menu. You should see an option that says **Merge.** Double-click it.

The **Merge** window will open up. The first drop-down menu says *Input Datasets*. This is where we tell ArcGIS which layers or tables on our map that we want to combine. Using this drop-down menu (click the downward-pointing arrow), select your 2017 and 2018 crime layers. You should see both of them pop into the window below this selection option.

Below that, you will see the *Output Dataset* option. This is asking you where you want to save the resulting data layer. Click the yellow folder icon, navigate to your working directory (wherever all of your other layers are saved), and save your file as something descriptive – I called mine “2018 and 2018 Street Robberies.” The **type** should be set as *Feature classes*. Click save. You will return to the Merge window, where you can click OK.

The program will think for a moment (depending on the size of the files you are merging), and then your new layer should appear on the map. There is an easy way to make sure this worked correctly:

1. Open the attribute table for your old 2017 crime data layer and look at the bottom of the window. It should say something like “0 out of X Selected” where X is the number of records in the file. Write this number down.
2. Open the attribute table for your old 2018 crime data later and get the same information.
3. Add these two numbers together.
4. Now open the attribute table for your new combined layer and see how many records are in there. It should equal the sum of the 2017 and 2018 records. If it doesn’t, something went wrong!

Once you have added your new combined 2017 and 2018 crime data layer to your map, you can uncheck the boxes for your single 2017 and 2018 layers to hide them – we only need the combined layer to be showing, and that is what we will work with for the rest of this exercise.

|  |
| --- |
| **What else do we need for a choropleth?** |

The whole point of a choropleth map is to color areas of the map (polygons) according to some quality or quantity present there. For example, in the Week 4 lecture slides, we saw an example of a choropleth map created by shading Mexican municipalities (cities and towns) according to their homicide rates.



This is similar to what we will be doing today. We will use our map layers to shade census tracts or block groups (types of polygons) according to the *count* and *rate* of your specific crime type. This is an important distinction: the *count* is how many crimes happened within that polygon, and the *rate* is how many crimes happened there *per* some measure of population. The rate is often much more informative when it comes to comparing crime across different areas, because – typically – areas with more people in them will just have more crime than areas with very few people! But we will look at both versions of our map to see the difference.

So what else do we need on our maps? We have our base map and now we have our combined 2017 and 2018 crime data. We need some sort of polygon layer that we can shade according to how many crimes happened there. Let’s start with the **census tracts** layer.

Drag and drop the **Boston Census Tracts 2010** layer from your Catalog into your Table of Contents. I recommend placing it just above your **City Boundary** layer, but below the **Open Space** layer (otherwise it will cover up the smaller details).

As you might be able to see (try unchecking the boxes for your ADDRFEAT and Open Space layers just for a moment), the **Census Tracts** layer is made up of irregularly-shaped polygons. Census tracts are small geographic areas defined by the U.S. Census Bureau for the purpose of measuring population change over time. That’s good for us, because it means that when we downloaded this census tract layer, it came with information about the number of people living in each tract! Let’s look at that now.

Right-click your **Census Tract** layer and open the attribute table. If you scroll across, you can see some columns like *AREA\_SQFT, POP100\_RE*, and *HU100\_RE.*

|  |  |
| --- | --- |
| **AREA\_SQFT** | The area of the census tract, measured in square feet. |
| **POP100\_RE** | The 100% population count for the tract. |
| **HU100\_RE** | The 100% count of housing units for the tract. |

What we *don’t* have – yet – is a count of how many crimes happened in each census tract. We can see them on our map and we could technically zoom in and count them by hand, but that would be very slow, boring, and probably prone to error (for example, some crimes might happen at the same location and the dots are stacked on top of each other, making it appear like a single dot!).

***So the first thing we will do is tell ArcGIS to count up all the dots in each census tract and then create a new column in the attribute table with the “count” of crimes for each tract.***

|  |
| --- |
| **Creating a “count” variable for tracts through *Joins*** |

We can create this “count” variable by *joining* our layer of crime points to our census tract layer. To do this, right-click the Boston Census Tracts 2010 layer and select “Joins and Relates” and then “Join…”. The **Join Data** window will open.

What do we want to join to this layer? We want to join the crime points based on where they are located – within individual census tracts. Select “*Join data from another layer based on spatial location*.”

The next option asks us to choose the layer we want to join to the census tracts layer. That would be our crime data points, so choose your combined 2017/2018 points layer.

Note that next to 2., ArcGIS can tell that we are joining “points to polygons” – crime points to our census tract polygons. It then gives us two different options for how to do this join. We want the first option, which says “Each polygon will be given a summary of the numeric attributes of the points that fall inside it, and **a count field showing how many points fall inside it**.” That’s all we need! There is an option that asks “How do you want the attributes to be summarized?” but we don’t need that right now – all we want is the count field showing how many points fall inside the census tract. So leave those boxes unchecked.

Finally, down at 3., ArcGIS asks us how to save the resulting layer. Save your layer in the same place as all of your other data, and give it a descriptive name – I called mine “Census Tracts with Crime Count.” Click OK and ArcGIS will create the new map layer for you.

To check that it worked, right-click your new layer and open the attribute table. You should see a new column at the far end of the table that is called “Count” and there should be a number in each cell.

Your new layer will be added on top of everything else on your map and, as you can see, will probably cover everything up. That's fine, because we are about to mess with the symbology anyway!

|  |
| --- |
| **Shading the map by the “Count” variable** |

Now we have what we need to make a choropleth: a layer of polygons with some sort of quantity that we can use to determine what color each polygon should be. We can now tell ArcGIS how we want it to shade our polygons.

Right-click your “census tracts with crime count” layer and choose “Properties,” then go to the **Symbology** tab.



On the left side of the window, you can see the different options for symbolizing this layer. When you first open the window it will be on “Features” and will just be set on one solid color. Instead, select **Quantities.** This allows us to set different colors for different polygons depending on some *quantity* (some numeric value) in the attribute table for this layer.

Next, under Fields, use the drop-down menu to set the Value to **Count\_**. This is the name of the column in the attribute table that lists how many crimes happened in each census tract.

For now, leave Normalization on **none.**

In the bottom half of the window, you can see the color options. Normally ArcGIS defaults to this weird yellow-brown scheme which can be pretty ugly, so I changed my **Color Ramp** to a single hue that increases in intensity, just like we learned in the Week 5 lecture: ***“The same hue in increasing darkness or saturation can be used with orderable categories or numerical data. Darker or more saturated means higher, greater, stronger.”*** As you can see here, the higher the count range, the darker the color. You can also see which count ranges are associated with each color. The lowest category is 0-4, and the highest category is 46-140. These values can all be changed, but we will leave them for now. Click OK and see below for what my map looked like:



This map represents *just* the number of crimes that happened in each census tract, *not* controlling for how many people live in that area. As a result, although it does show where crime is happening, it may also just be showing us where people live or spend their time, because that is where they are most likely to commit crimes or be the victim of crimes! If we want to “control” for population density, we can *normalize* our count variable by the population of each tract. Let’s do that next.

Right-click your “census tract and crime count” layer again and open the symbology tab. It should look the same as where you left it. All you need to change now is where it says **Normalization**, change it from “none” to “POPRE\_100” – this was our variable measuring the population in each tract.



One thing that you can see now is that the range of values has changed. We still have 5 classes (5 different color distinctions), but the number associated with them is very different. Our lowest class used to be 0-4 crimes, but now it is 0 to 0.0018. *For example, the first census tract in the attribute table, Census Tract 104.05, has a population of 5522 people and a count of 8 street robberies.* ***8 divided by 5522 is 0.0014****, so on my color ramp, this census tract will be colored in the lightest shade of red.*

Click OK and take a look at how your map has changed. Does this “normalized” map look different to your original count-only map? Have the darkest areas moved? **HINT: You might want to check the “Open Space” layer relative to your census tracts.**



|  |
| --- |
| **Finishing up** |

This assignment is a little bit different to the others, so pay careful attention to the submission directions.

**First**, go back to your **count only** symbolization – so open up your symbology for your map layer again and just set the Normalization to *none* and click OK.

Switch to **Layout** view. Uncheck the box next to your 2017 and 2018 crime layer to hide them (they are probably covered up by the shaded layer anyway!).

Go to Insert > Legend and insert a legend on your map. Use the black arrow tool to *carefully* move your legend into place. You may also want to use the white hand tool to position your map neatly, making sure you are zoomed in so that the map fits the space neatly.

Then go to Insert > Title and insert a title at the top of your map. I called this first map “*Choropleth of 2017-2018 Street Robberies by Count*.” This title ends up being a little bit long to fit at the top of my map, so I double-clicked the title to open the **Properties** and click **Change Symbol…** to change the font size. I changed mine from 21 to 18.

When you are satisfied with your map, export it as a JPEG.

**NEXT,** change the symbology of your shaded layer back to the **Normalization: POP100\_RE option.** You don’t need to switch out of layout view or anything – your map and legend will automatically update.

You do need to change the title, though, so double-click the title and in the Text box, type your new title. I called this map “*Choropleth of 2017-2018 Street Robberies per Person*.” I also needed to **Change Symbol…** to make this font size smaller so that it would fit. When you are satisfied with your map, export it as a JPEG.

Now you should have two JPEG maps: one shaded by count, one shaded by count normalized by population.

Go to Blackboard and start a new submission to the ArcGIS #4 dropbox. Attach both maps, and *in the submission text box*, please write a paragraph explaining the differences you see between your two maps and why you think there is a difference. I also encourage you to identify the highest-rate census tracts on your map and see if you can look them up on Google Maps to see what is in that area – **do a bit of sleuthing to see what you can learn about where your crime types are most concentrated and why that might be.**

 