**Introduction to ArcGIS Exercise**

In this exercise you will use ArcGIS software to create several maps from scratch. Objectives of the exercise are:

* To gain a basic level of familiarity with ArcGIS software.
* To understand how to display and query GIS shapefile data.
* To make choropleth, dot density, and proportional symbol maps.
* To perform a table join, spatial join, map latitude / longitude coordinates, perform an aggregation, and create density maps.

**1) Downloading the workshop data**

In a web browser, go to <http://gis.harvard.edu>. Scroll down and click on the **Political Geography GIS Workshop** link on that page. Click the **Download Lab 1 Materials** link. Unzip the **Intro\_to\_GIS.zip** file onto your Desktop.

**2) Open ArcMap, add data, explore the tools**

Open ArcMap software by clicking **Start > ArcGIS > ArcMap 10.5.1**. In the dialog window, choose **A Blank map**. We want to open a blank map to start.

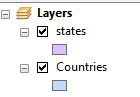
Click **File** > **Save** **As**, and browse to your **Desktop\Intro\_to\_GIS** folder. Name the map **Intro\_Map**, and click **OK**. This creates an ArcMap document (file type .mxd) and sets your Home location.

Click **File** > **Add Data > Add Data**. In the Add Data dialog window, click the drop down arrow after the **Look in:** prompt and choose the **Home – Desktop\Intro\_to\_GIS** location (see below). You may have to scroll up to see this option, it is the top one.

Double click the **Shapefile** folder, click the **Countries** shapefile, hold the Ctrl key and click the **States** and **Ocean\_background** shapefiles, and click **Add**. The map will now display these layers on the map, and list them on the Table Of Contents on the left. The source of these shapefiles is Esri Data and Maps 2016, a huge GIS data package that comes with ArcGIS.

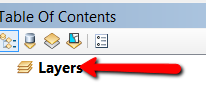
Experiment with the tools on the Tool toolbar to see what they do to the map:



Change the symbology of the Countries or States layers by clicking on the symbol square in the Table of Contents

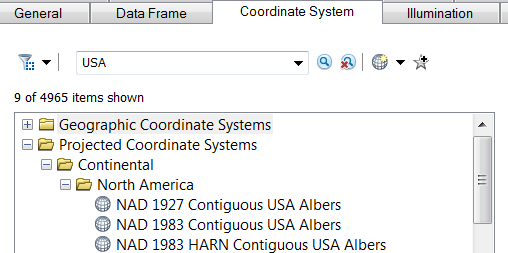
**Right click** the **Countries** layer and choose **Properties**, and **Open Attribute Table**. This displays all of the information linked to each country shape on the map.

**3)** **Change the map projection**

Map projections display the 3 dimensional earth in 2d on a flat plane. It is a standard practice to use a map projection that is meant for the area you are making the map for. For example, if making a map of the USA, choose a USA Projected Coordinate system. To change your map projection, in your Table of Contents, double click where it says Layers (see red arrow on screenshot at right). This will open **Layer Properties**.

Choose the **Coordinate System** tab. Under Projected Coordinate Systems, choose a coordinate system, and click OK to see how the change affects your map.

After experimenting with a few different coordinate systems, set your map to a USA coordinate system. We’ll be making maps of the USA (lower 48 states) next, so the best practice is to use a coordinate system designed for this area. In the Search area, type in “USA” and hit Enter. See below.

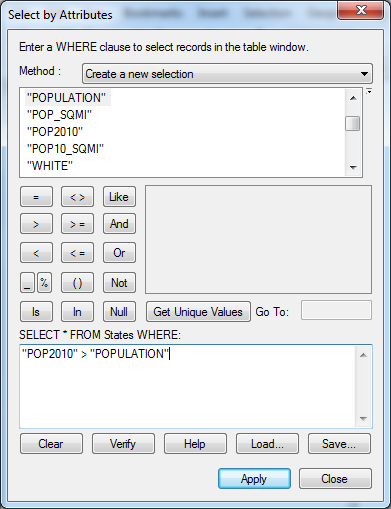


Click the **Projected Coordinate Systems > Continental > North America** folders. Choose one of the **USA** coordinate systems, and click **OK**. Your map will now be set to that coordinate system.

**4) Perform an attribute query on the attributes of the U.S. States shapefile.**

**Right click** the **States** layer and choose **Open Attribute Table**. This is mostly census information that Esri includes with this data package. See the field descriptions below.

|  |  |
| --- | --- |
| STATE\_NAME | Name of the State |
| STATE\_FIPS | The 2 character Federal Information Processing Standard for the state |
| SUB\_REGION | Census sub region the state is on |
| STATE\_ABBR | 2 character state abbreviation |
| POP2015 | Total population, 2015 |
| POP15\_SQMI | Population density, people per square mile, 2015 |
| POP2010 | Total population, 2010 |
| POP10\_SQMI | Population density, people per square mile, 2010 |
| WHITE | Total number of people identifying themselves as White, 2010 |
| BLACK | Total number of people identifying themselves as Black 2010 |
| AMERI\_ES | Total number of people identifying themselves as American Eskimo, 2010 |
| ASIAN | Total number of people identifying themselves as Asian, 2010 |
| HAWN\_PI | Total number of people identifying themselves as Hawaiian / Pacific Islander, 2010 |
| OTHER | Total number of people identifying themselves as Other, 2010 |
| HISPANIC | Total number of people identifying themselves as Hispanic, 2010 |
| MALES | Total males, 2010 |
| FEMALES | Total females, 2010 |
| MED\_AGE | Median age for the state population, 2010 |
| MED\_AGE\_M | Median age for the state population, 2010 |
| MED\_AGE\_F | Median age for the state population, 2010 |
| SQMI | The state’s area in square miles |

Let’s use these attributes table to answer a question: Did any states experience a population decrease between 2010 and 2015?

To make this query, click the **Table Options** button: and choose **Select by Attribute**. In the select by attribute dialog window, double click “**POP2010**”, then the greater than sign **“>”** and then double click “**POPULATION”**. Your select statement should look like the screenshot on the right. Click **Apply** to run the query. What state(s) experienced a population decrease? Experiment by making similar selections with other attributes.

**5) Experiment drawing different thematic map types.**

Clear your selected states from step 4 by clicking the Clear Selection button: 

**Double click** the **States** layer and choose **Properties**, and the **Symbology** tab. Here you can symbolize the map based on attributes in the table.

Each map layer is symbolized with a single symbol by default. To symbolize based on a numeric attribute, click **Quantities** on the left, and then for the **Field Value** choose **POP15\_SQMI** and click **Apply**. Now your map will be colored according to population density.

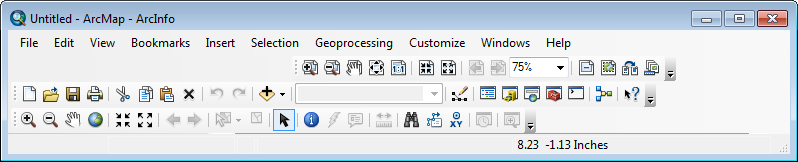
Under the Quantities Symbology list, experiment with making maps using the different types. Here are some cartographic guidelines:

* Graduated colors (choropleth maps) are best for data that is already normalized by population or area, such as population density or median age.
* Graduated and proportional symbols are best for showing total count data, like total population.
* Dot density and Chart maps are best for showing multiple attributes, like several racial groups.

**6) Make a choropleth map in PDF format of median age with a Title, Legend, Scale bar, and North Arrow.**

**Double click** the **States** layer and choose **Properties**, and the **Symbology** tab. for the **Field Value** choose MED\_AGE. Click Apply, and the map will now turn into a choropleth map showing median age values for each state.

Click **View > Layout View**. This changes the map to a page layout, and operates similar to graphical editing software. Use the Layout toolbar (see below) to navigate in the page layout space.



Click **Insert > Legend**, and click Next 4 times and then Finish. The legend is added to your map. It is highly customizable; just double click it to change content, fonts, background. It is also dynamically linked to the map table of contents, so if you change the name of a layer in the TOC, the legend reflects the name change. Optional: To individually change the elements in the legend, it can be converted to graphics. Do this by right clicking the legend and choosing Convert to Graphics. Then right click and choose Ungroup (repeat Ungroup until nothing is grouped.

Click **Insert > Scalebar**, and choose a scale bar style to insert.

Click **Insert > North Arrow**, and choose a north arrow.

Click **Insert > Title**, and type in a title for your map.

Click **Insert > Text** and type in “Map produced by <your name>. Sources: ESRI Data and Maps 2016”.

To undo mistakes you might make, click **Edit** > **Undo**. This can be a very useful thing to know!

All of these elements can be altered by double clicking on the element to open up the Properties box. Each element can be moved or resized to make the map as legible and meaningful as possible. The Draw toolbar arrow is used to select elements (see red box on Draw toolbar screenshot below). If your Draw toolbar is not displayed, click **Customize** > **Toolbars** > **Draw** to add this toolbar.



**Click File > Export Map** to export the map to PDF, PNG, or many other image formats that are easily transferable.

**7) Saving your Map**

Click **File** > **Map Document Properties**. In the resulting Map Document Properties window, check the “Store relative pathnames to data sources” box, and click OK (see screenshot below).



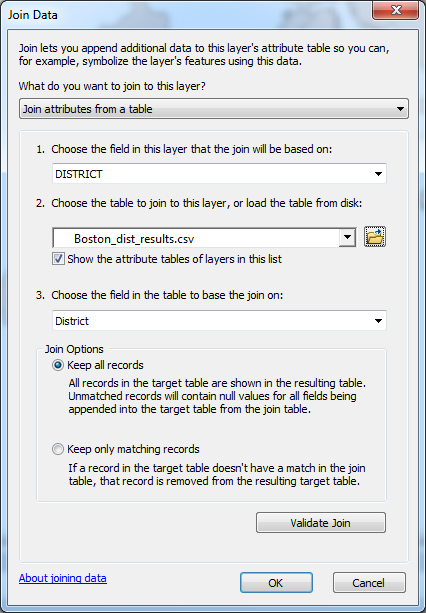
Doing this indicates to ArcMap that the data in the map is stored in the same folders as the map document.

Next, click **File** > **Save**. This creates an .mxd (ArcMap document) file, which saves the symbology you have chosen, the layers you’ve added, and the reference paths from the layers to the datasets in the Intro\_to\_GIS folder. This enables you to copy this folder, load it on to any computer running ArcGIS 10.5, and open this map as it appears here.

\*\*Optional – Add Alaska and Hawaii insets to your map using instructions in Appendix A below\*\*

**8) Map Boston election results from a table.**

Click **File** > **New** to open a new blank map. ***What coordinate system is the map in?*** Add the **Boston\_city\_council\_districts** shapefile from the **\Intro\_to\_GIS\Shapefile** folder by clicking **File > Add Data > Add Data**. ***What coordinate system is the map in now?*** The map’s coordinate system should match the layer’s coordinate system. Open the attribute table for the council districts layer, and note that it contains the District number and Council member. Save this map into your Intro\_to\_GIS folder, calling it Boston.

In Windows File Explorer, browse to the **\Intro\_to\_GIS\Tables** folder and double click the **Boston\_dist\_results.csv** table to open it in Excel. This table of data contains the percent of vote for the winner, loser, and Margin of Victory for a 2011 Boston election. Often one will encounter tables of data that refer to geographic places like this. To map this table of data to the corresponding geographic boundaries, performing a Table Join based on a common field between the tables is necessary. The common field between the districts shapefile and this table is the District number.

In ArcMap, **right click** the **Boston\_city\_council\_districts** layer and choose **Joins and Relates > Join**. In the Join Data window, specify “**Join attributes from a table**”, and choose DISTRICT for #1 For #2 click the browse button, browse to the \Intro\_to\_GIS\Tables folder and double click the Boston\_dist\_results.csv file. For #3, choose the District field. Your screen should look like the one on the right:

Click **OK** to activate the join. Now you should see the Winning\_Perc, Runner\_Perc, and MOV fields added to the Boston\_city\_council\_districts shapefile. These fields are now available to symbolize your map with.

**9) Map and label the winner’s percent.**

**Double click** the Boston\_city\_council\_districts layer, and click on the **Symbology** tab. Click **Quantities** and choose **Winning\_Perc** as the **Value** field. Choose a color scheme and click **Apply**. Now the districts will be shaded according to winner’s percent of the vote. This is informative, but displaying the actual percentages on the map would give the map reader even more information. Click the **Labels** tab in the Layer Properties window. At the top of the window, check the box next to “**Label features in this layer**”. For the Label Field choose **Winning\_Perc**. Change the font size to 12, and click **OK**. Now the labels will appear on the map. However, expressed as decimals they are not very readable. To format them as percentages **right click** the Boston\_city\_council\_districts layer and **Open Attribute Table**. Then **right click** the **Winning\_Perc** field and choose **Properties**. Click the **Numeric…** button. Click Percentage on the left, and choose “The number represents a fraction. Adjust it to show a percentage” Then click **Numeric Options**, and change the **Number of decimal places** to **2**. Click **OK** and **OK**. Now the labels will include the % symbol, and will only display 2 numbers past the decimal.

To make these labels even more informative, let’s add the council district number and the council person’s name to the label. **Double click** the Boston\_city\_council\_districts layer and click the **Labels** tab. Click the **Expression** button. Delete the text in the Expression window. Then build the label expression using these steps:

Type in: **"Council District " &**  (make sure to include a space after District)

Double click **[Boston\_city\_council\_districts.DISTRICT]**

Type in **& Vbnewline &**

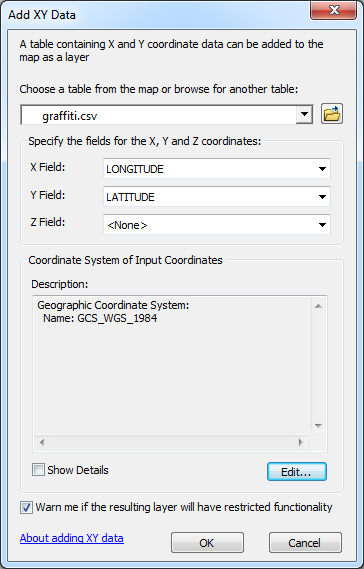
Double click **[Boston\_city\_council\_districts.Councillor]**

Type in **& Vbnewline &**

Double click **[Boston\_dist\_results.csv.Winning\_Perc]**

**Your final expression should look like this:**

"Council District " & [Boston\_city\_council\_districts.DISTRICT] & Vbnewline & [Boston\_city\_council\_districts.Councillor] & Vbnewline & [Boston\_dist\_results.csv.Winning\_Perc]

Click **OK** and **OK** to see the change on the map. As there is more information displayed, you might want to decrease the font size of the labels. Another trick to make the labels more readable is to add a “halo” around the text. Do this in the **Labels** tab of **Layer Properties** by clicking the **Symbol** button > **Edit Symbol** > **Mask** tab and clicking the button next to **Halo**. Click **OK** 3 times to invoke the change on your map.

**10) Map graffiti incidents in Boston**

The **\Intro\_to\_GIS\Tables\Boston\_graffiti.csv** file contains graffiti incidents between January of 2010 through August of 2013. Note there is Latitude and Longitude fields in this data. Next, you’ll use these fields to map graffiti incidents in Boston, and then summarize them by city council district.

Click **File > Add XY Data**. This is the command used to add any sort of data that contains map coordinates. In the Add XY Data dialog window, click the **Browse** button, and choose the **\Intro\_to\_GIS\Tables\graffiti.csv** file to add. Make sure the X field is specified as Longitude, and the Y field as Latitude. Click the **Edit** button and choose **Geographic Coordinate System > World > WGS 1984**. This World Geographic Coordinate System created in 1984 is the standard coordinate system nearly all latitude/longitude data is in. Your screen should look like the one on the right.

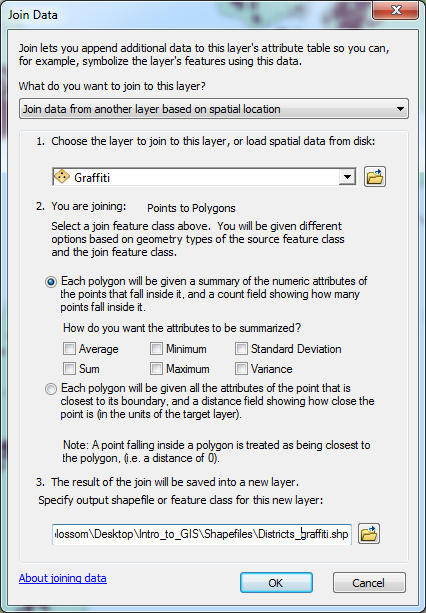
Click **OK** to add the data to the map . Click **OK** if you get a “Table does not have Object-ID field” message. Now the graffiti locations will appear on the map as dots. To save these to a shapefile, right click on the graffiti.csv Events layer and choose **Data > Export Data**. Specify the **\Intro\_to\_GIS\Shapefiles** folder to save it to, naming it **Graffiti.**  In the “Save as type:” dropdown box, click the arrow and specify **Shapefiles**. Click **Save** and **OK**, and add it to the map when prompted. You can right click the Events layer and choose Remove, as it is not needed anymore.

**11)** **Add a basemap layer**

ArcMap has several customized global base maps available to add as a layer (streets, terrain, satellite imagery, and more). These are often helpful to add to help visualize background, or general reference information. To add a basemap, click **File** > **Add Data** > **Add Basemap.** Click on any of the Basemaps available, and click **Add**. Drag this basemap layer to the bottom of the map legend, if it’s not at the bottom already. Using these basemaps can save a lot of cartographic compilation time. You must be online to use these basemaps.

**12) Summarize total graffiti incidents by City Council District.**

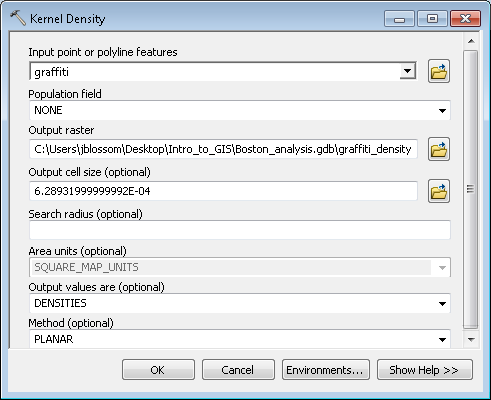
Now that the graffiti incidents are mapped on the city council districts one might ask: How many incidents happened in each district? This is a common type of question solved using GIS overlay analysis. To answer this question we’ll use a Spatial Join.

**Right click** on the **Boston\_city\_council\_districts** layer and choose **Joins and Relates > Join**. At the top, specify to “**Join data from another layer based on spatial location**”. For #1, choose Graffiti as the layer to join to. Leave #2 as is, and for #3 specify to save the output in the **\Intro\_to\_GIS\Shapefiles folder.** Name the new shapefile **Districts\_graffiti**. Your window should look like the one on below right:

Click **OK** to run the Join, and **Yes** to add the result to the map. Open the Districts\_graffiti attribute table. Notice now there is a Count field. This is the total count of graffiti incidents in that council district. Now the map can be symbolized on this statistic.

**13) Create a heat map of graffiti**

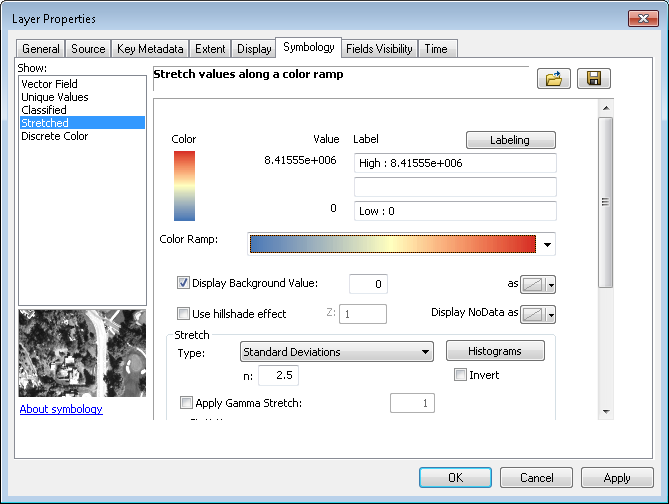
Looking in the attribute table of the graffiti layer shows that there are 7,954 locations. These locations show patterns on the map, but it is difficult to distinguish different density levels of the points, unless you zoom in to show less area and more detail. To better visualize the density we’ll make a density map, which are often called ‘heat maps’.

Make sure the spatial analyst extension is activated by clicking **Customize** > **Extensions** and checking the **Spatial Analyst** button. Then click **Close**. Open ArcToolbox by clicking **Geoprocessing > ArcToolbox**. Double click **Spatial Analyst Tools > Density > Kernel Density**. Choose **Graffitti** as the input point or polyline feature. Leave the population field at **NONE**. For the **Output Raster**, click the browse button, go to your **Intro\_to\_GIS** folder, and click the **New File Geodatabase** button: Name the geodatabase **Boston\_analysis**. Double click the Boston\_analysis geodatabase, and enter the Name of the output layer as: **graffiti\_density**. Geodatabase is a GIS data type that is able to store both raster and vector data layers (called ‘feature classes’). It allows for more advanced modelling (such as topology) between feature classes, than what is possible with shapefiles. It also allows up to 255 characters to name fields, whereas a shapefile’s limit is 10. Leave the rest of the options as the default for the Kernel Density command. Your screen should look like the one on the right:

Click **OK** to run the tool.

Examine the resulting density map.

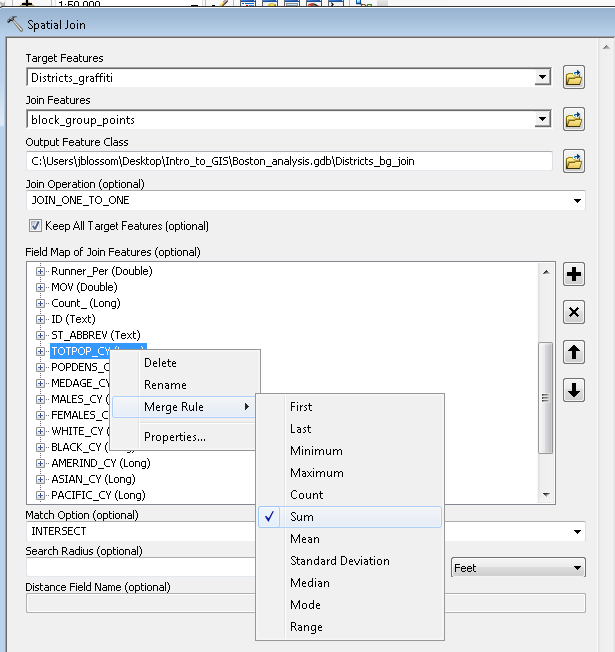
**14) Symbolize the density map to reveal more detail.**

The default Symbology showing 10 shades of blue can be improved on to show more detail in the data. Double click the **graffiti\_density** layer and choose the **Symbology** tab. Choose **Stretched** on the left, and choose a **blue to red** color ramp. Click the box next to **Display Background Value** (leave the value at 0) and click **OK**. How did this change the map? Experiment with different classification techniques.

**15) Calculate and analyze the demographics of the council districts.**

Race, age, and gender are often highly influential characteristics determining how people vote. Let’s now calculate and analyze these demographics for each council district. The council district shapefile does not contain this data, but the U.S. Census publishes demographic data at several levels of geography, the most refined being the block group level. The mapping company Esri assembles census data at the block group level on a yearly basis (based on the Census American Community Survey), combines it with consumer marketing and business data, and sells it as a product called Business Analyst. The **MA\_block\_groups** shapefile in your **\Intro\_to\_GIS\Shapefiles** folder is an export from the 2011 Esri Business Analyst data. Add this shapefile to your map with the **Add Data** button. Inspect the geographic extent of this layer by right clicking and choosing **Zoom to Layer**. This zooms the map to the extent of this data. Open the attribute table and note how block groups are in MA. Notice the attributes include population, race, gender, and age information (the CY in these fields stands for Current Year, which in this case is 2011). The ID field is the unique block group identifier (called the FIPS code, which stands for Federal Information Processing Standard). You will use the block groups to calculate district level demographics.

Symbolize the block groups with “No Fill” and a thick colored line. Zoom in, and take note that the block group boundaries mostly match the district boundaries, but not all the time. To transfer the demographic data from the block groups into the district boundaries, you’ll convert the block groups to points, spatial join them to the districts, and then summarize the demographics. **Note:** This is a quick and easy way to summarize demographic data, but it’s not perfect. After this exercise your instructor will discuss another way to do this called apportioning. Apportioning can be done by area or population. In **ArcToolbox**, double click the **Data Management Tools > Features > Feature To Point** tool. Choose **MA\_block\_groups** as the Input Feature, and specify to store the Output in the **Boston\_analysis.gdb**, naming the result **MA\_block\_group\_points**. Click the box next to “Inside” to ensure the points will fall inside each block group. Click **OK** to run the tool.

The new layer contains a point at the geographic center of each polygon. Now you’ll use a spatial join with statistics to calculate district level demographics. In **ArcToolbox**, double click **Analysis Tools > Overlay > Spatial Join.**

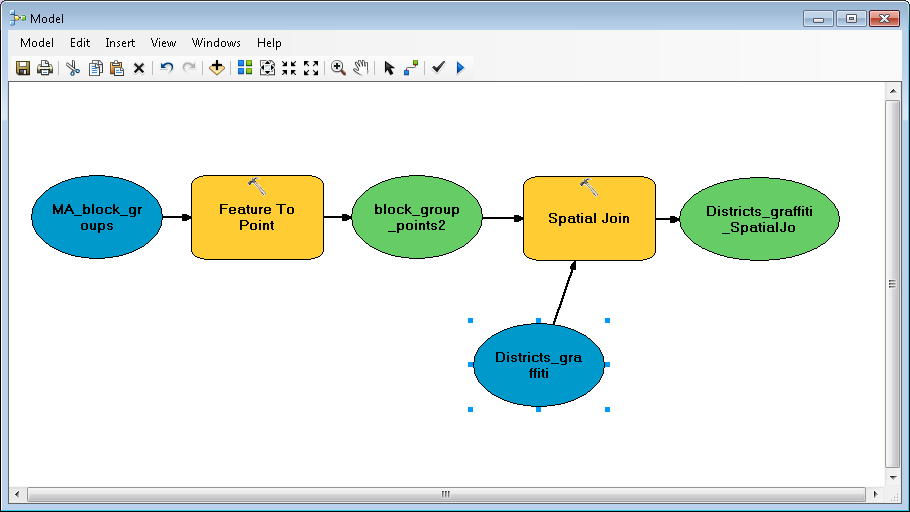
Specify **Districts\_graffiti** as the **Target Features**, and M**A\_block\_group\_points** as the **Join Features**. Specify the output as **Districts\_bg\_join**. Now you must set up a “merge rule” for the demographic attributes to specify what type of math will be applied to the multiple block group points that fall within each district. You’ll want to sum all of the total count variables (total population, race, gender), and calculate averages of the variables that are already averages (age). In the Spatial Join window under the **Field Map of Join Features**, right click the **TOTPOP\_CY(Long)** field and choose **Merge Rule > Sum** (see screenshot above). For the **MEDAGE\_CY** field choose **Merge Rule > Mean**. For the rest of the race and gender variables, choose a merge rule of Sum. Once your merge rules are specified, click **OK** to run the spatial join. Open the attribute table of the resulting Districts\_bg\_join layer and inspect the results. Now the districts contain summaries of this 2011 demographic data.

**16) Run the analysis in a model and export to Python.**

Commands can be scripted in ArcGIS using “Model Builder” or Python. First, we’ll make a model of the Features to Points and Spatial Join commands, and then export this model to the Python scripting language. Open a new model by clicking the ModelBuilder button: 

In ArcToolbox, go to the **Data Management Tools > Features** toolbox**.** Drag the **Features to Point** tool into the Model window**.** In **ArcToolbox**, go to **Analysis Tools > Overlay** and drag the **Spatial Join** tool into the model. **Double click** the **Feature To Point** tool within the model.

Specify **MA\_block\_groups** as the input, and **Boston\_analysis.gdb\block\_group\_points2** as the Output. Click the Inside option, and **OK**. Choose the **Connect** tool, and drag a line from the green “block\_group\_points2” oval into the Spatial Join tool. Choose **Join** Features when prompted. Click the **Select** tool, and double click the **Spatial Join** command. Specify the **Target Features** as **Districts\_Graffiti,** the **Output Feature Class** as **Districts\_bg\_join2**, and set up all of the **Merge Rules** like you did in step 15. Click **OK.** Your model should look like the one below.



Click **Model** and **Run Entire Model**. The commands will run in sequence. You can build as many command into a model as you want, and do programmatic tasks such as iterations and loops.

The commands and parameters in the model can be saved out into Python, which can also be used to batch process ArcGIS commands. Do this by clicking **Model** **> Export > To Python Script**. Name the file and save it to your Desktop. Find the python file on your Desktop in File Explorer, **right click** it, and choose “**Edit with IDLE**”. This is the Python editor that is installed with ArcGIS. However, any text editor can be used to edit Python code. More can be found on using Python for ArcGIS here: <http://resources.arcgis.com/en/communities/python/>

END OF EXERCISE

**Some general Info:**

Geographic Coordinates Systems are good for showing maps of a continent, or the world, but will not work for calculating distances or areas. Projected Coordinate Systems must be used when calculating distances or areas. In ArcGIS, the **Project** tool transforms a data layer from one projection to another, changing its coordinate space. The geographic location of the features usually does not change, but some projected coordinate systems can’t effectively model all features for an area. Defining a coordinate system in ArcCatalog, or using the ArcGIS **Define** tool just associates or sets that coordinate systems to the layer, it does not alter any coordinate values or geographic locations.

Common GIS (vector) file types to look for when searching for data: shapefile (.shp), coverage (.e00), mapinfo (.mif), geodatabase (.mdb and .gdb), kml (.kml or .kmz), AutoCAD (.dxf and .dwg).

Common image (raster) file types: .tif, .jpg, .jp2, .img, .sid, .bil.

Mapping axioms to follow and tips to minimize wrangling

* **When publishing a map, include: Title, Legend, Scale Bar, North Arrow, Citations.**
* **Don’t reinvent the map, if one exists, is applicable, and available for use, use it, and cite it.**
* **Organize, document, and backup your data.**

The entire contents of this course can be downloaded at:

<http://gis.harvard.edu/training/non-credit-training/past-workshops>

For GIS questions and help: <http://gis.harvard.edu/contactus>

**Appendix A: Adding Alaska and Hawaii inset maps**

1. Click **Insert > New Data Frame**. Drag the new data frame to the lower left of your page layout. In the Table of Contents you can rename your New Data Frame by clicking on it once, waiting a few seconds, and clicking on it again. The text will be editable, and you can type in “Hawaii”.
2. Drag the States layer into your Hawaii data frame on the Table of Contents.
3. Change the coordinate system to a Hawaiian system by following the directions listed in step 3 above, but enter Hawaii as the search term.
4. Use the Zoom In tool to zoom in to Hawaii.
5. Remove the frame outline by double clicking the Hawaii data frame in the Table of Contents, choosing the Frame tab, and choosing <None> for the Border.
6. Repeat these steps to add an inset map for Alaska.