Spatial Analysis for Institutional Research

NEAIR Pre-Conference Workshop
New Brunswick, New Jersey
November 4, 2007

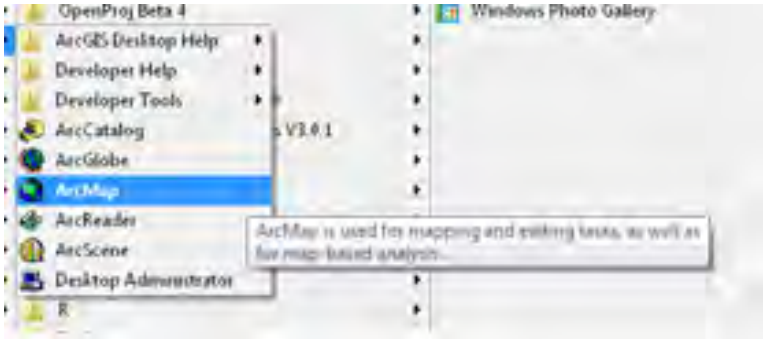
Allan Joseph Medwick
E-mail: medwick@alumni.upenn.edu
Telephone: 267.872.0336

Table of Contents

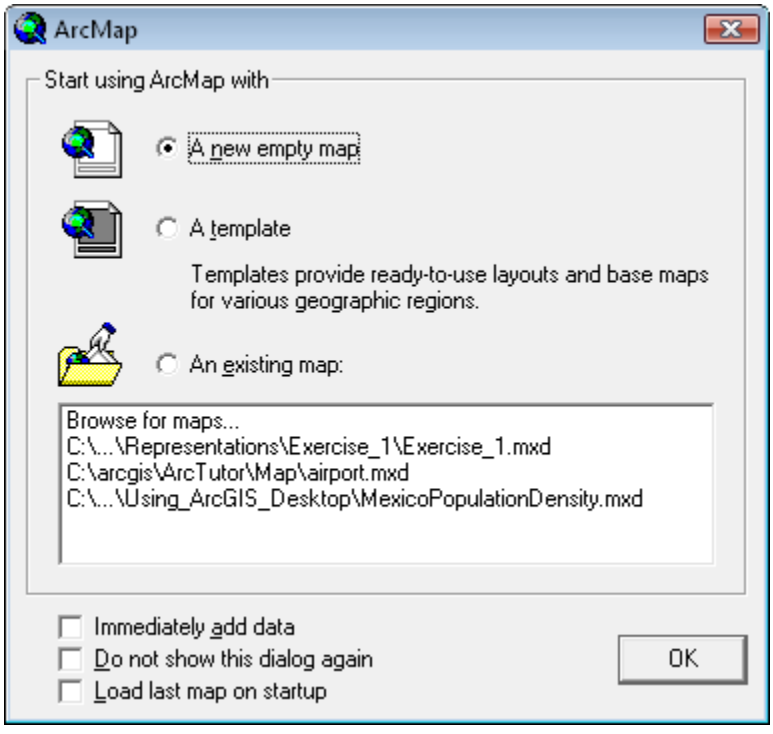
Exercise One: México is Not Just for College Students

Goal: The purpose of this exercise is to introduce you to some of the basic features of ArcMap.

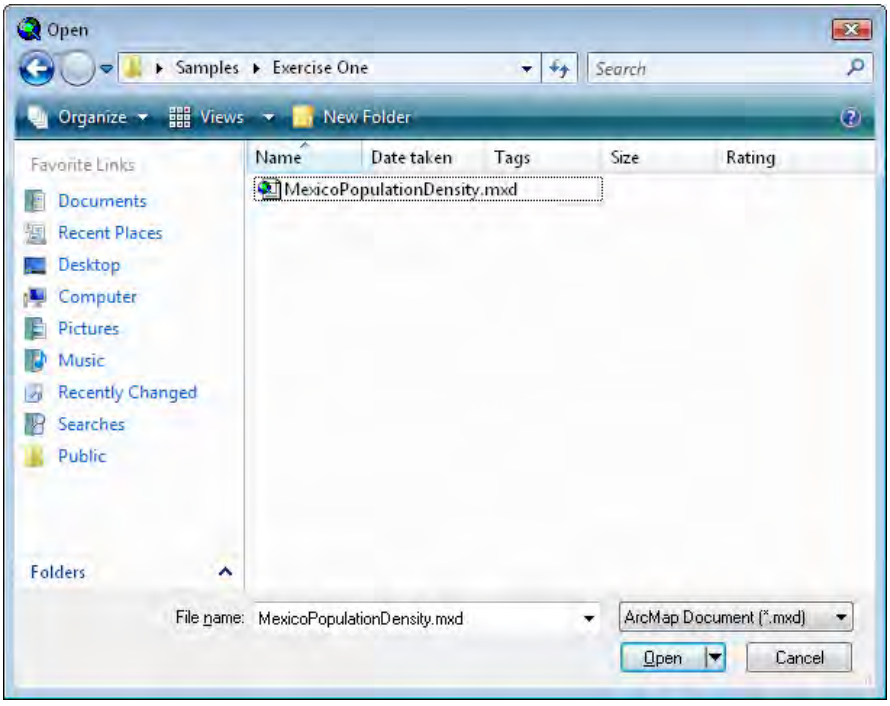
1. Go to Start, Programs, ArcGIS and click on ArcMap.



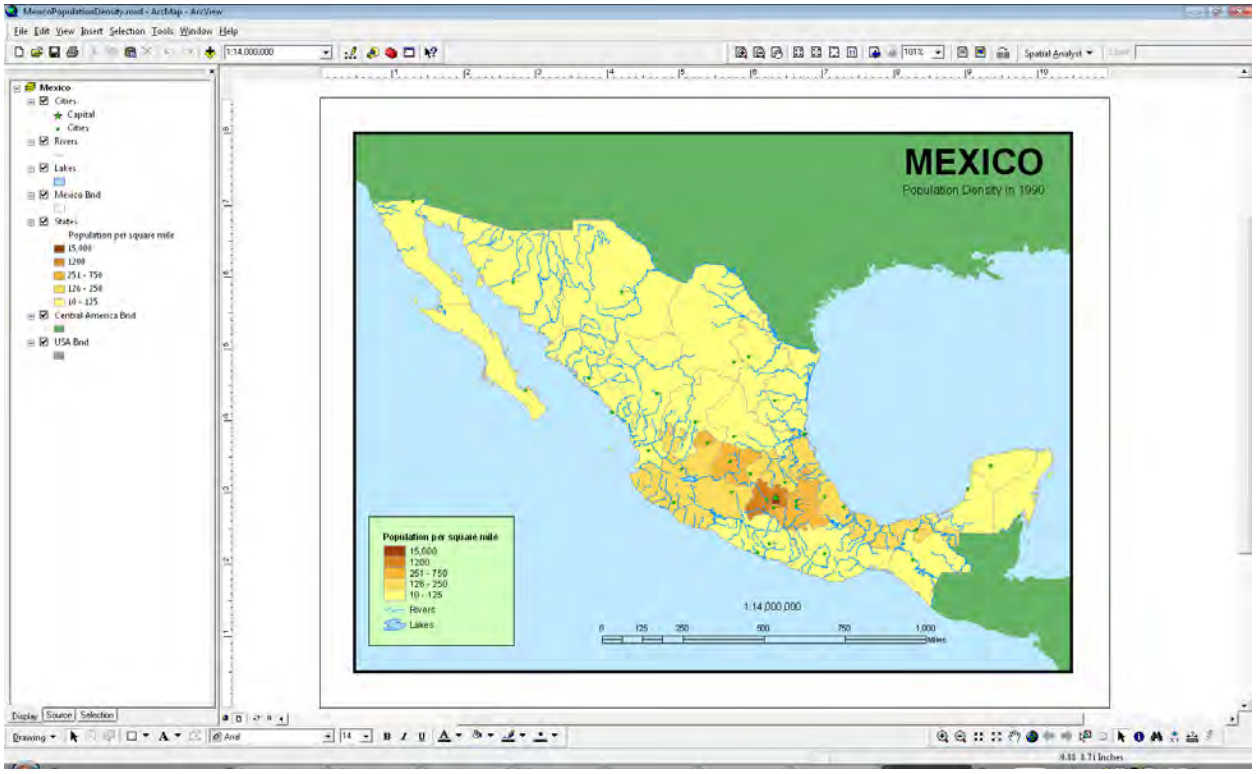
2. Select "An existing map", "Browse for maps…", and click "OK".



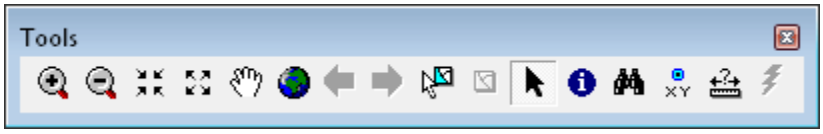
3. A File Open Dialogue box will open. Navigate to where you stored the sample exercises. Within the samples folder, there is a folder named "Exercise One", open it, select MexicoPopulationDensity.mxd, and press Open. MXD is the extension in ArcGIS for map files.



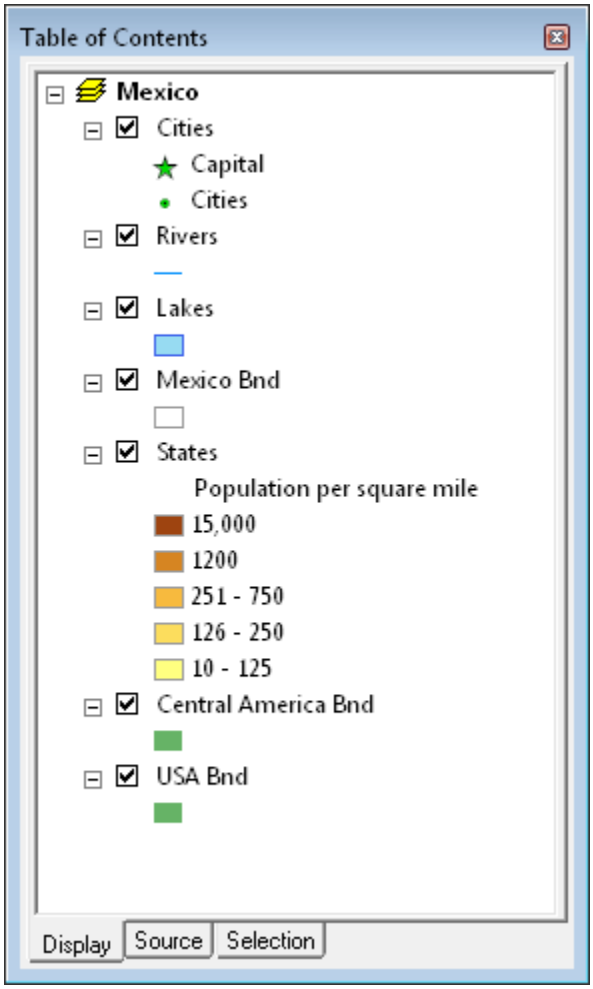
4. A map of México showing population density will now be open. This is called the Layout View.

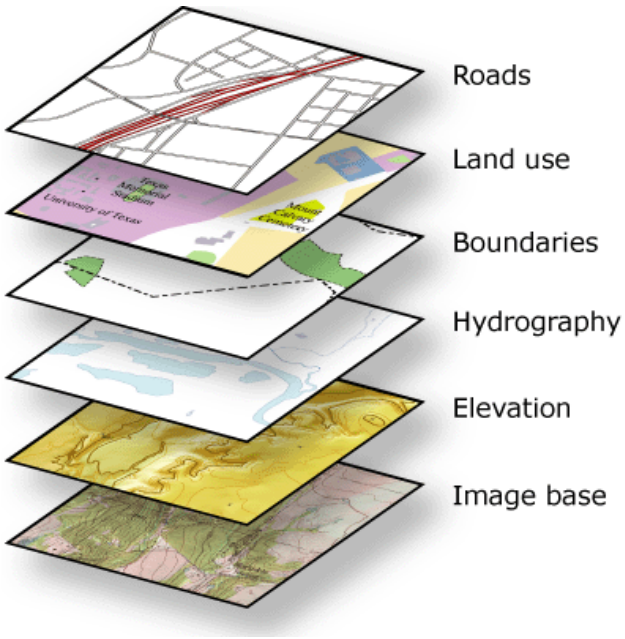


5. Notice the following toolbars. We will talk about what each of the buttons does in the workshop.



6. Take a look at the Table of Contents on the left hand side. Notice how each part of the map is listed separately. These are called layers. The yellow stack of notes with the word "Mexico" next to it represents a data frame. There may be more than one data frame in a map document.





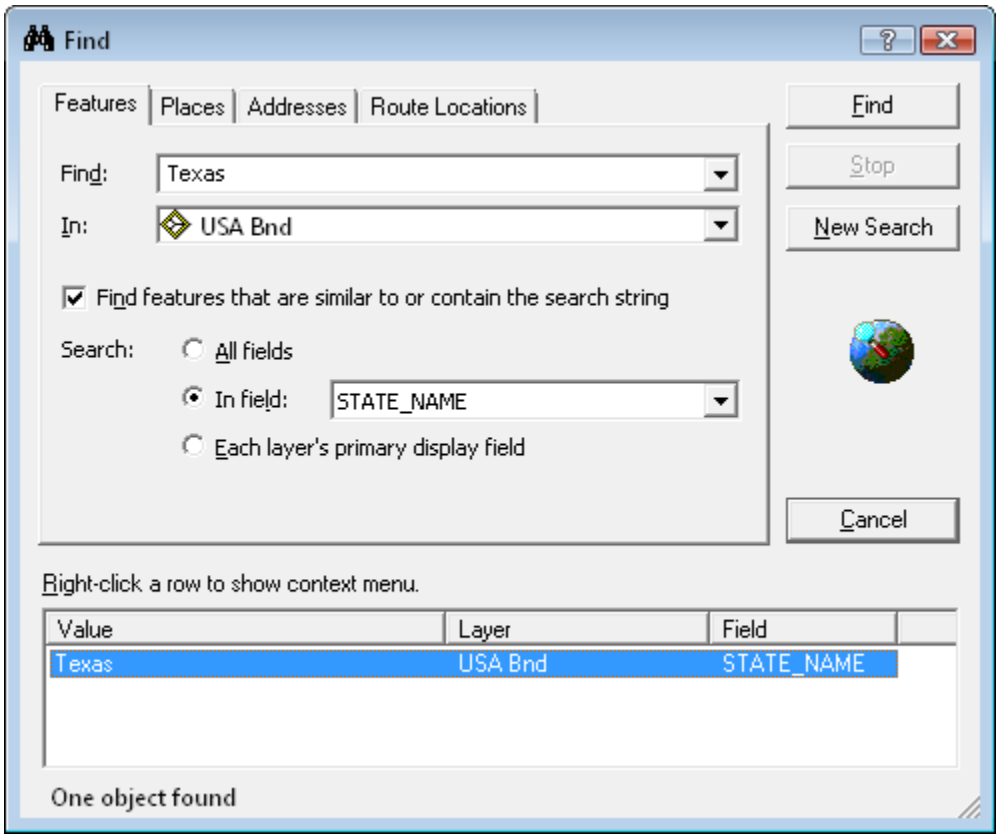
7. Using the check boxes next to the names of the layers, examine what happens to the map when you "turn off" each layer. When you are finished, turn all of the layers back on.

8. Now click on the word "Cities" and drag it to the bottom of the list. What happens to the symbols on the map? The order to the layers determines in which order they are drawn. Placing the Cities layer on the bottom resulted in it being covered by the States layer effectively hiding it from view.

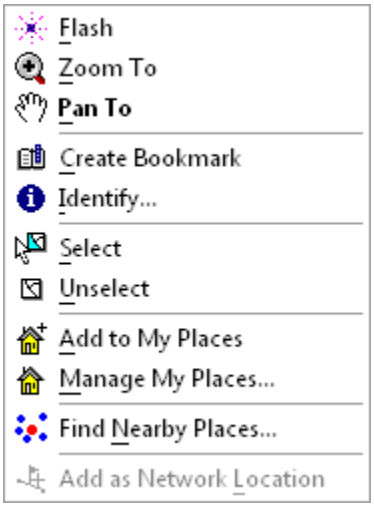
9. Using the Magnifying Lens with the Plus inside of it, zoom in on Mexico City. Notice how the scale bar has now extended to off the page. We will ignore this for now.

10. Click on the Earth icon to view the full extent of the map. Notice how much more of Central America and the United States was present. Remove the Central America Bnd and USA Bnd layers using the "X" button. Now reclick the Earth button. Notice that the full extent view is only Mexico. Click the undo button twice to return these two layers to the map.

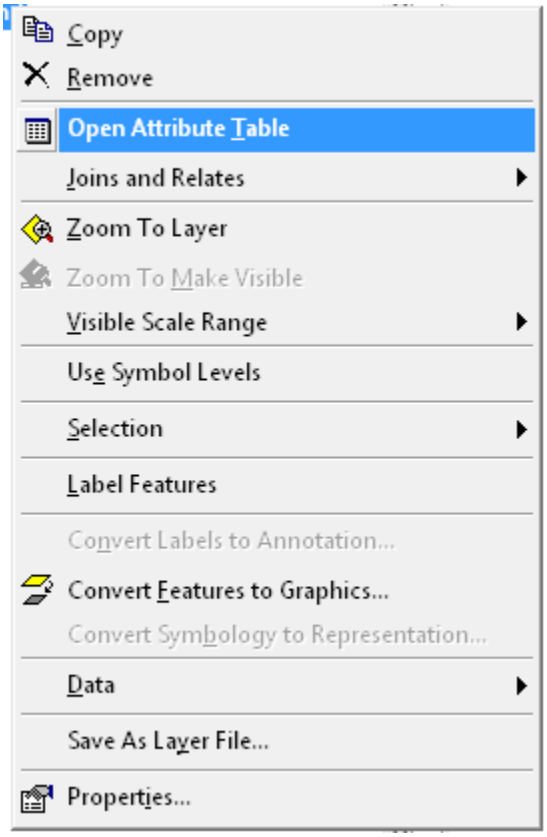
11. Use the binoculars icon to search for Texas in the USA Bnd layer. Click on the Find button.



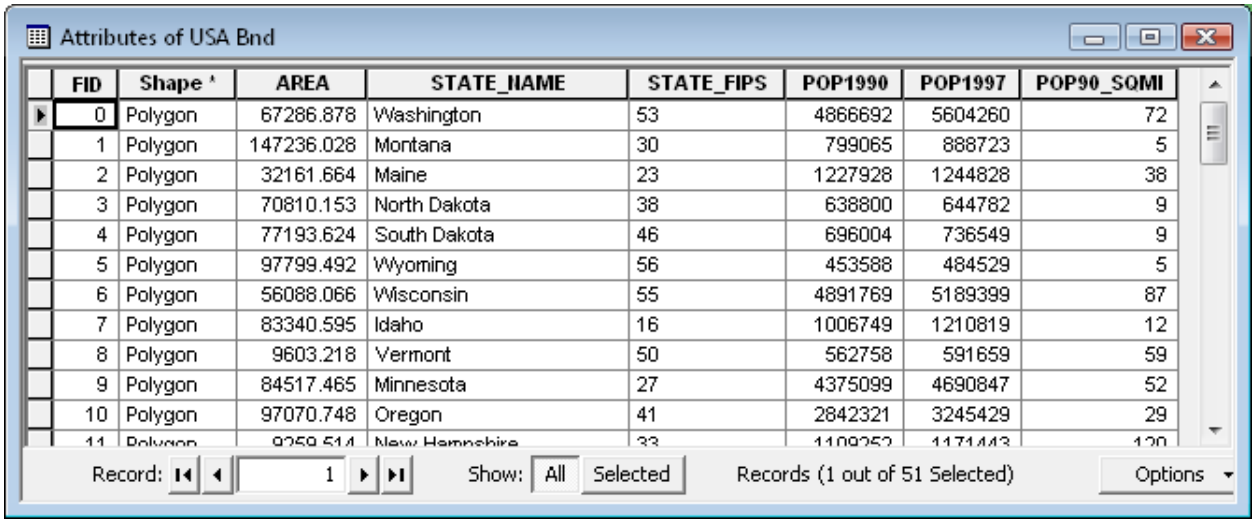
12. Right click on the result row and press Select. Notice all the options available in the context menu.



13. Texas should now be outlined in a light blue color on your map. Close the find window. If Texas is present, it means that all of the states are in that particular layer, but they are all "colored in" or symbolized the same way. Right click on USA Bnd in the Table of Contents and select Open Attribute Table. Notice all of the options. We will explore more of these later in the workshop.

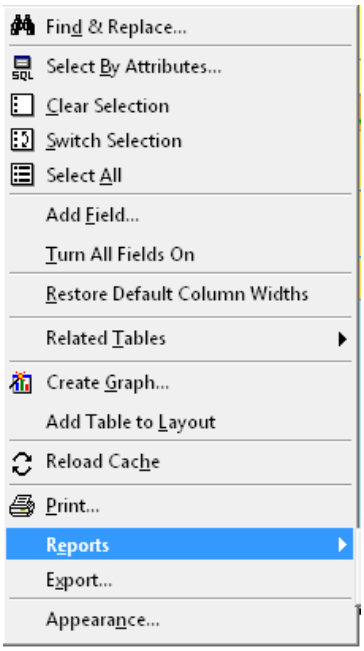


14. The Attributes Table contains all of the information that is available on objects in that particular layer. Notice that there are 51 records (one for each state and the District of Columbia). If you scroll down, you will notice that the row for Texas is highlighted, since it is selected in the map.

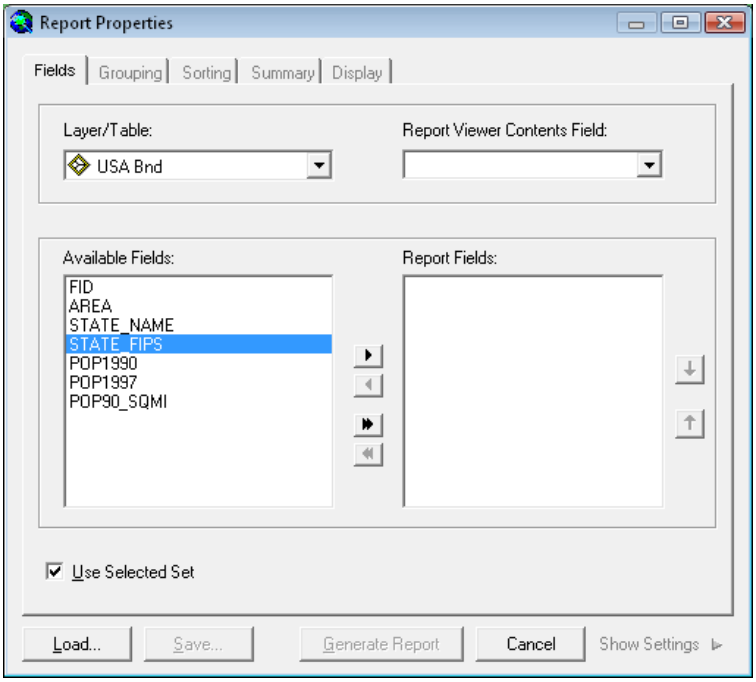


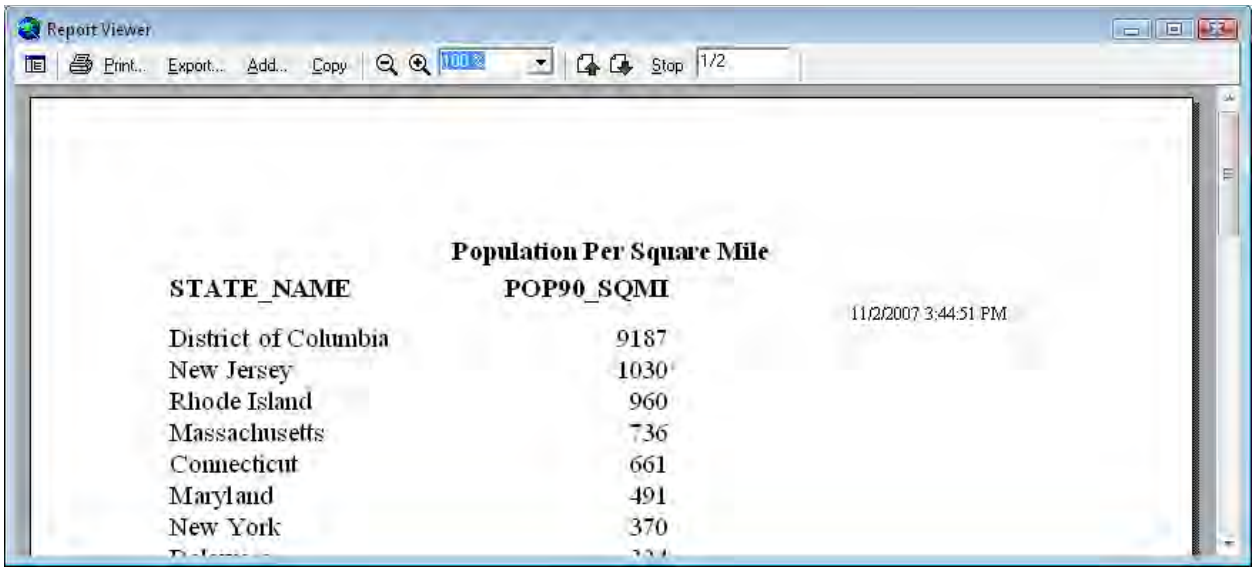
FID	Shape *	AREA	STATE_NAME	STATE_FIPS	POP1990	POP1997	POP90_SQMI
0	Polygon	67286.878	Washington	53	4866692	5604260	72
1	Polygon	147236.028	Montana	30	799065	888723	5
2	Polygon	32161.664	Maine	23	1227928	1244828	38
3	Polygon	70810.153	North Dakota	38	638800	644782	9
4	Polygon	77193.624	South Dakota	46	696004	736549	9
5	Polygon	97799.492	Wyoming	56	453588	484529	5
6	Polygon	56088.066	Wisconsin	55	4891769	5189399	87
7	Polygon	83340.595	Idaho	16	1006749	1210819	12
8	Polygon	9603.218	Vermont	50	562758	591659	59
9	Polygon	84517.465	Minnesota	27	4375099	4690847	52
10	Polygon	97070.748	Oregon	41	2842321	3245429	29
11	Polygon	9259.514	New Hampshire	33	1109252	1171443	120

15. Click on the Options button. Notice all of the options. Click on the Reports item and select create new report.

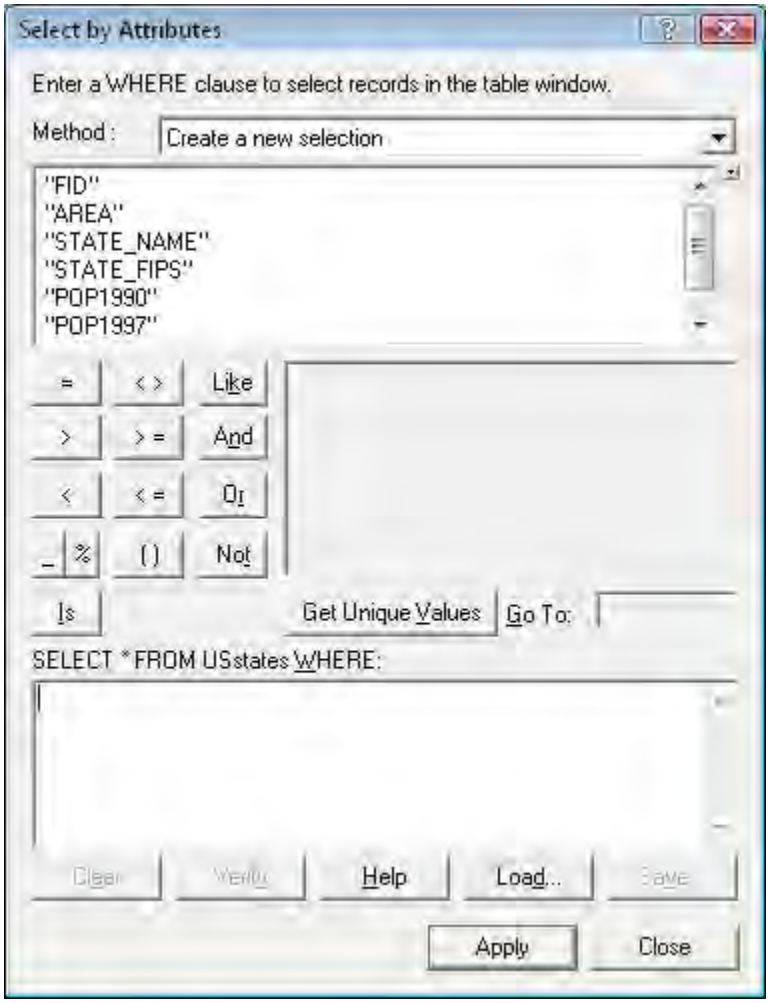


16. A Report Properties Window opens. Select State_Name and Pop90_SQMI and move them to the Report Fields box using the single arrow button. Unclick "Use Selected Set". Add some sorting and summary options. Then click generate report.



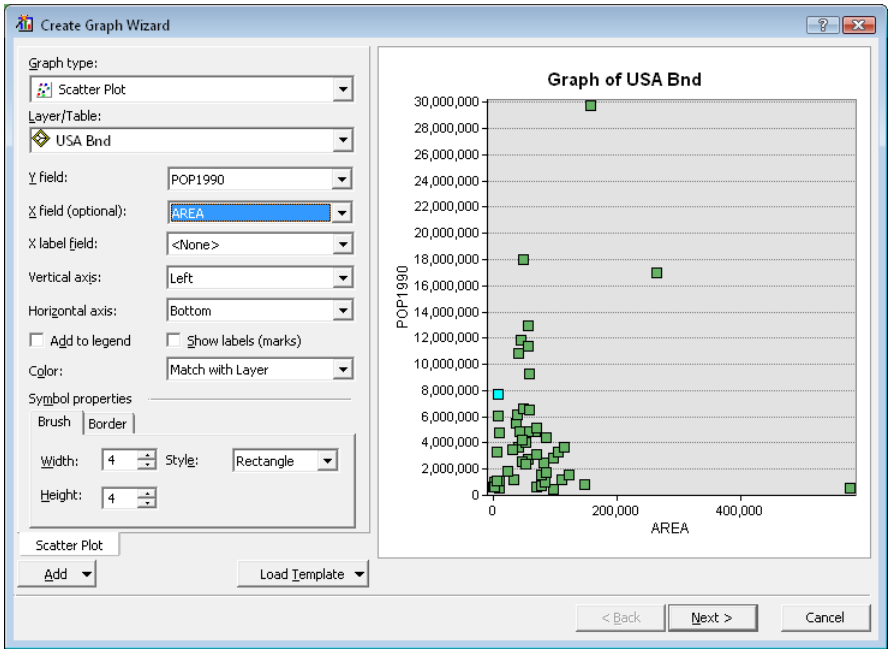


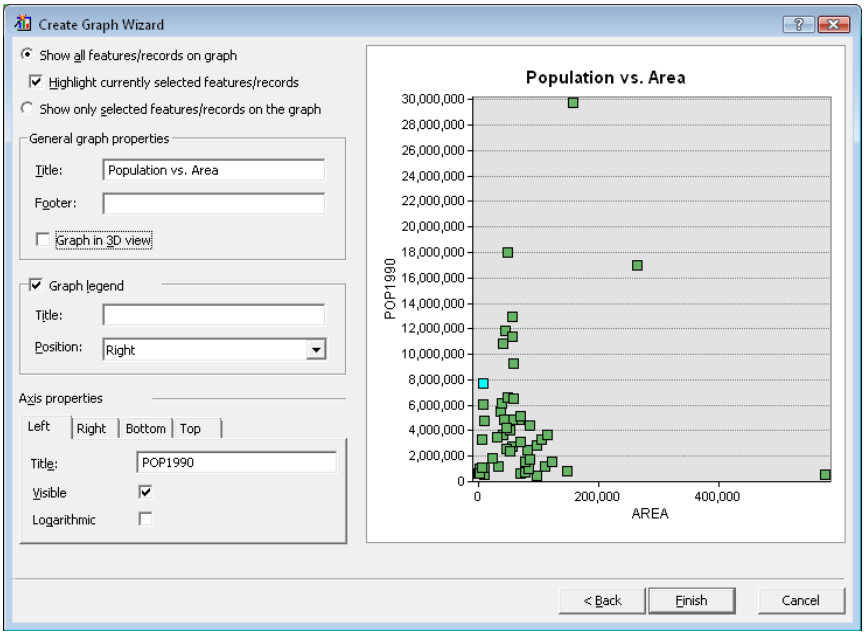
17. Close the report without saving. Now let's select a new state meeting certain requirements. It should have an area greater than 7000 but less than 8000. What is the name of the state?



Notice how ArcGIS allows you to query the database using SQL statements. You may also select a state based on name. Strings must be surrounded by single quote marks as double-quotation marks denote field names. You can use the Boolean Operators and the other operators to build your statement.

18. Go back to the Options button an click on Create Graph. Use the following options.

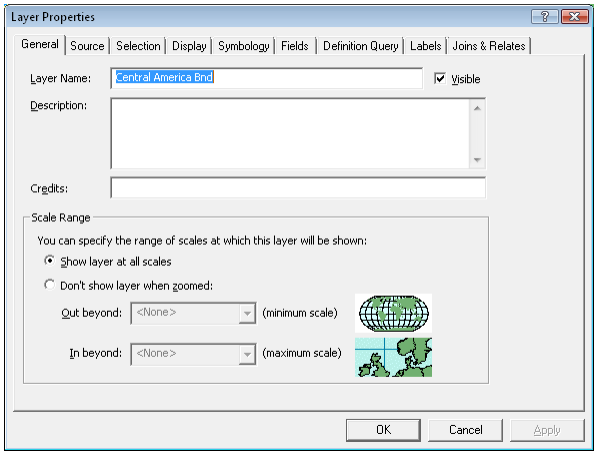




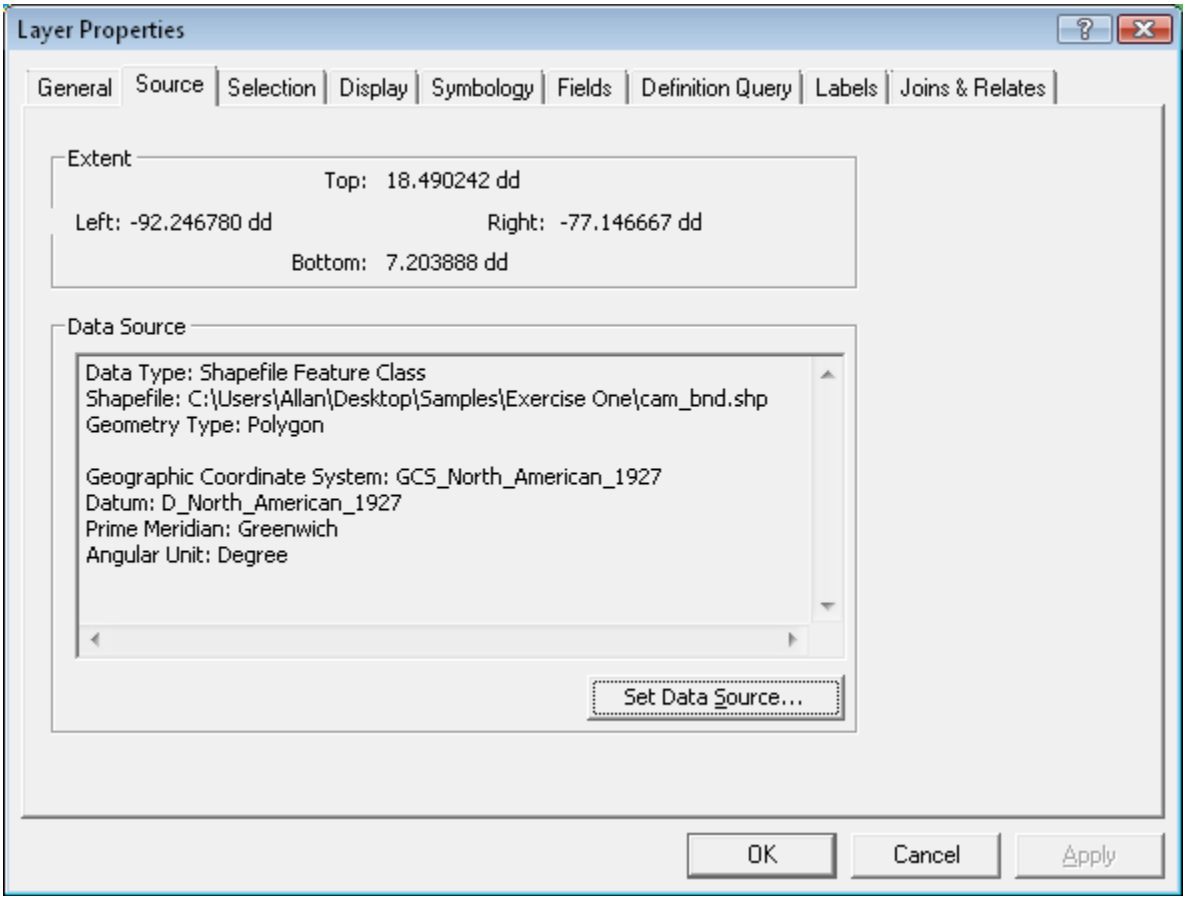
The graph now appears in its own window. You can export the graph, print it, or add it to the layout.

19. Close the Attribute Table window.

20. Double-click on Central America Bnd. This opens the Layer Properties window. Notice all of the different tabs. Click on the General tab. Rename this layer as "Countries South of Mexico" and click Apply. Notice that the name of the layer changes in the Table of Contents.

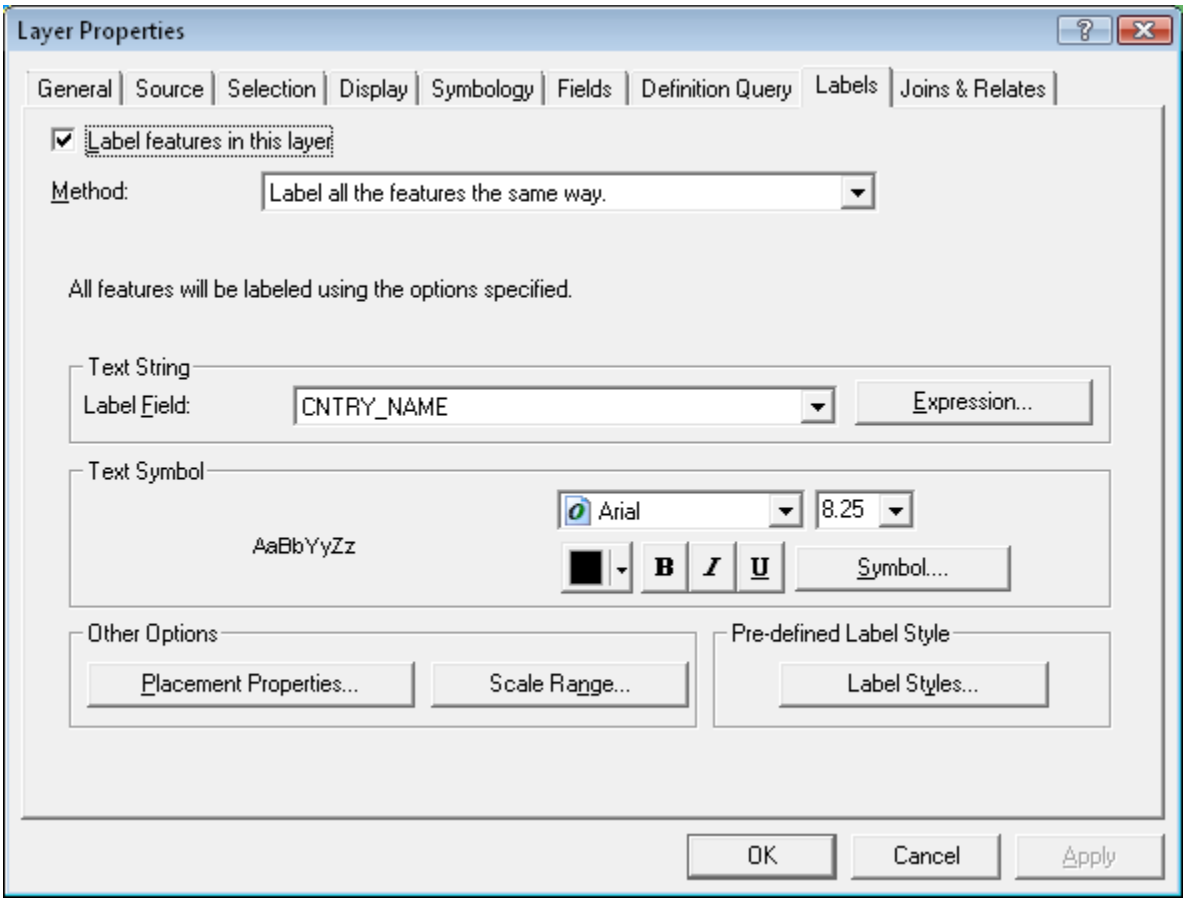


21. Click on the Source tab. Notice the information contained in the tab.

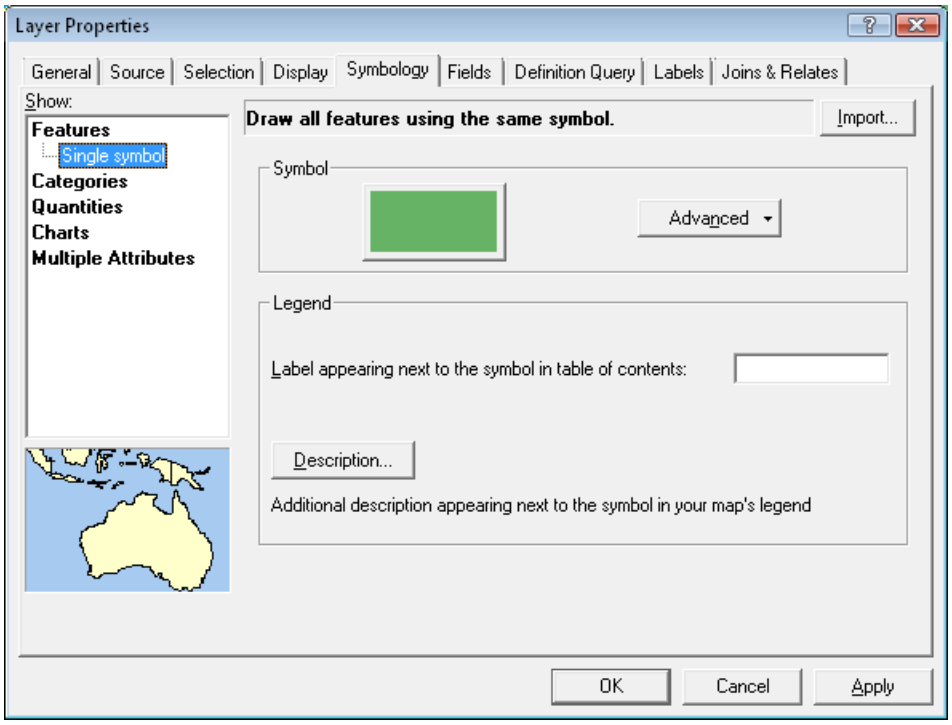


22. Examine the Selection and Display tabs. Then select the Definition Query tab. Here you can write a SQL statement that will determine which features get shown. Use the Query Builder to select only the county Belize. What happened? Remove the query text from the window and click apply to return to normal.

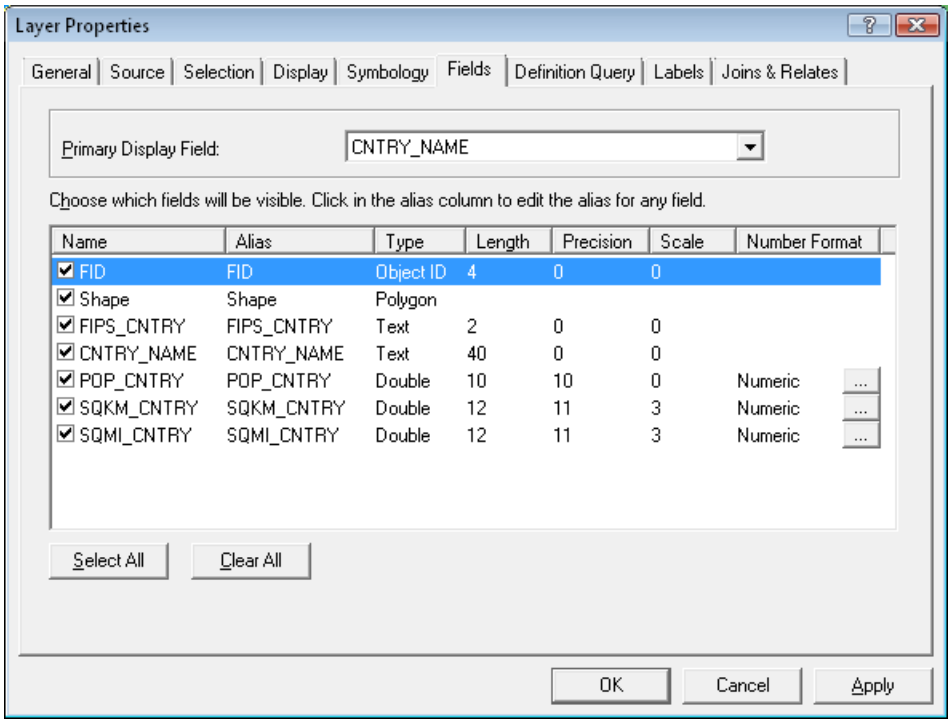
23. Click on the Labels tab. Click the box by "Label Features in this Layer." Click Apply. What happens?



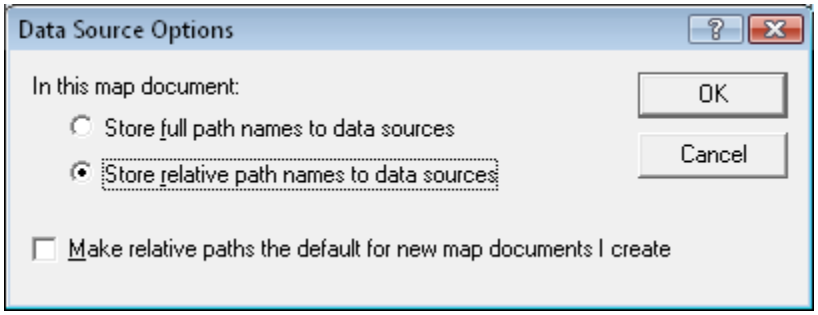
24. Click on the Symbology tab. Notice all of the options available. We will discuss this more later.



25. Click on the Fields tab. Notice how you can change the length, type, and number formats.



26. Click on file, Document Properties, Data Source Options. Notice how ArcGIS gives you the option to save full path names or relative path names. This is something to be aware of in case you open a map and all your data has gone missing.



27. Exit ArcMap.

Exercise Two: New Ideas, New Energy, New Brunswick

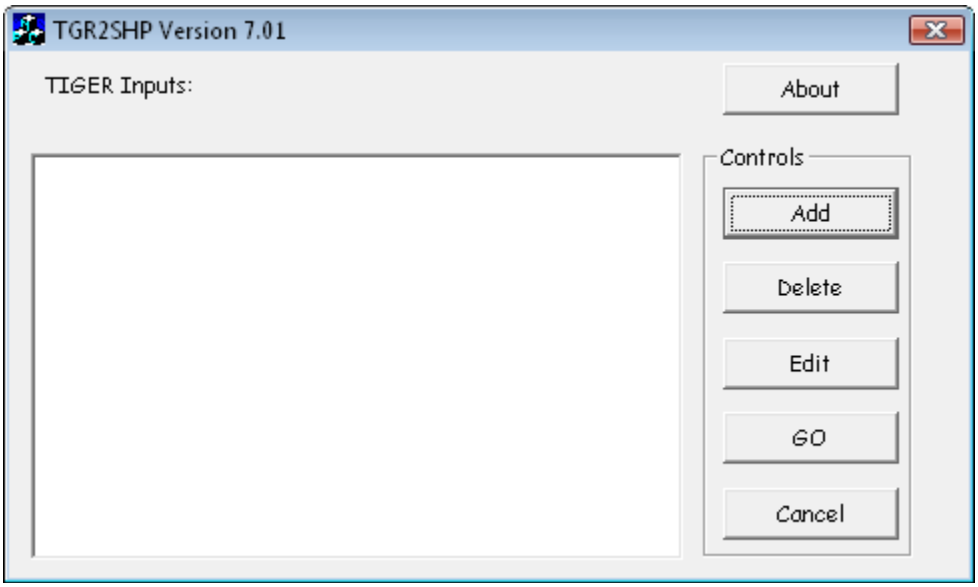
Goal: Using data from the Bureau of the Census's Topologically Integrated Geographic Encoding and Referencing (TIGER) system create a map with layers for streets, railroads, and rivers. Use this map to find the location of our hotel and mark it. Calculate the distance from the hotel to my home.

Preliminary Steps (You don't have to do these in the workshop)

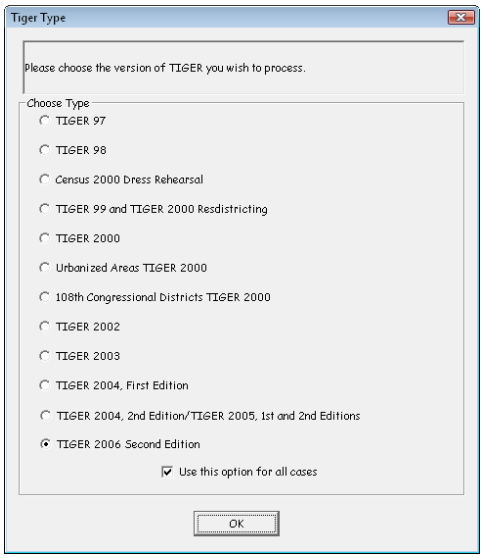
a. Visit http://tnatlas.geog.utk.edu/tea/ and download a program called TGR2SHP.
b. Install TGR2SHP.
c. Visit http://www.census.gov and click on TIGER.
d. Click on **2006 Second Edition TIGER/Line® Files**
e. Click on New Jersey
f. Download TGR34023.ZIP (Middlesex County).

Steps

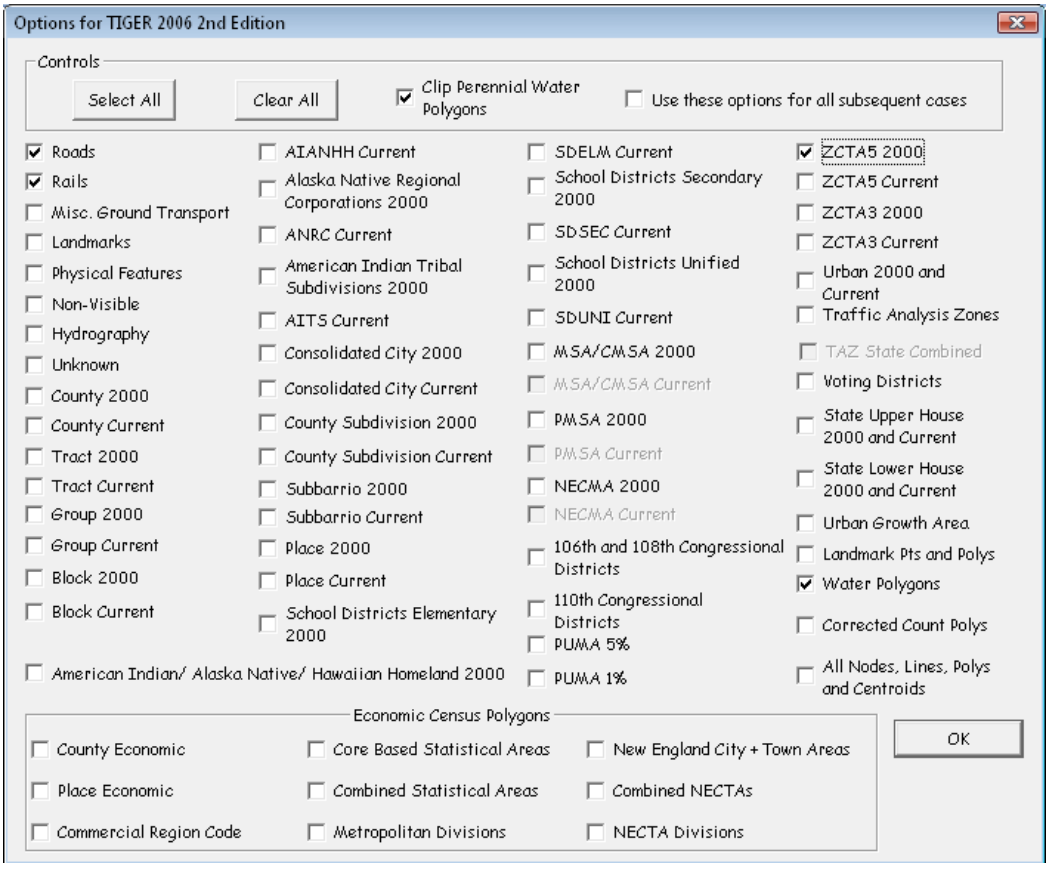
1. Run C:\...\Samples\Exercise Two\TGR2SHP\ tgr2shp.exe



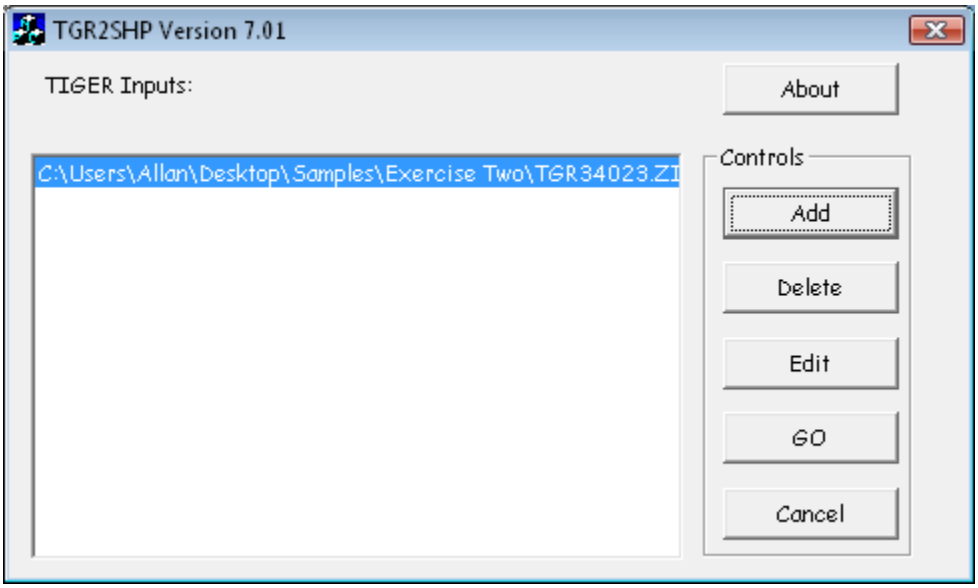
2. Click the Add button and select C:\...\Samples\Exercise Two\TGR2SHP\ TGR34023.ZIP. This is a TIGER 2006 Second Edition file. Click OK.



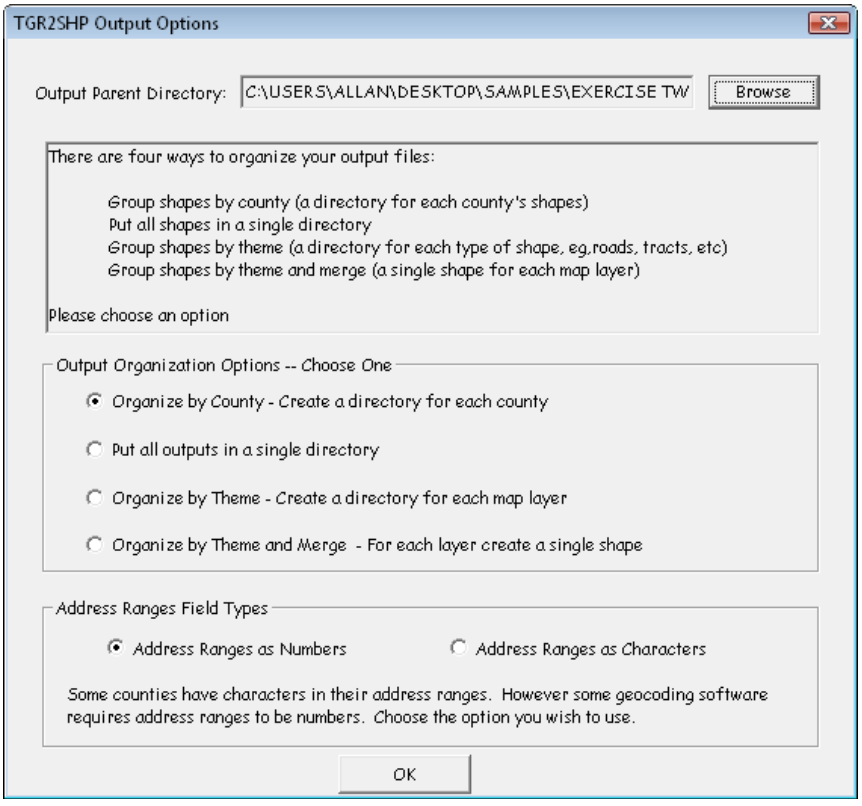
3. Select the following options and click OK.



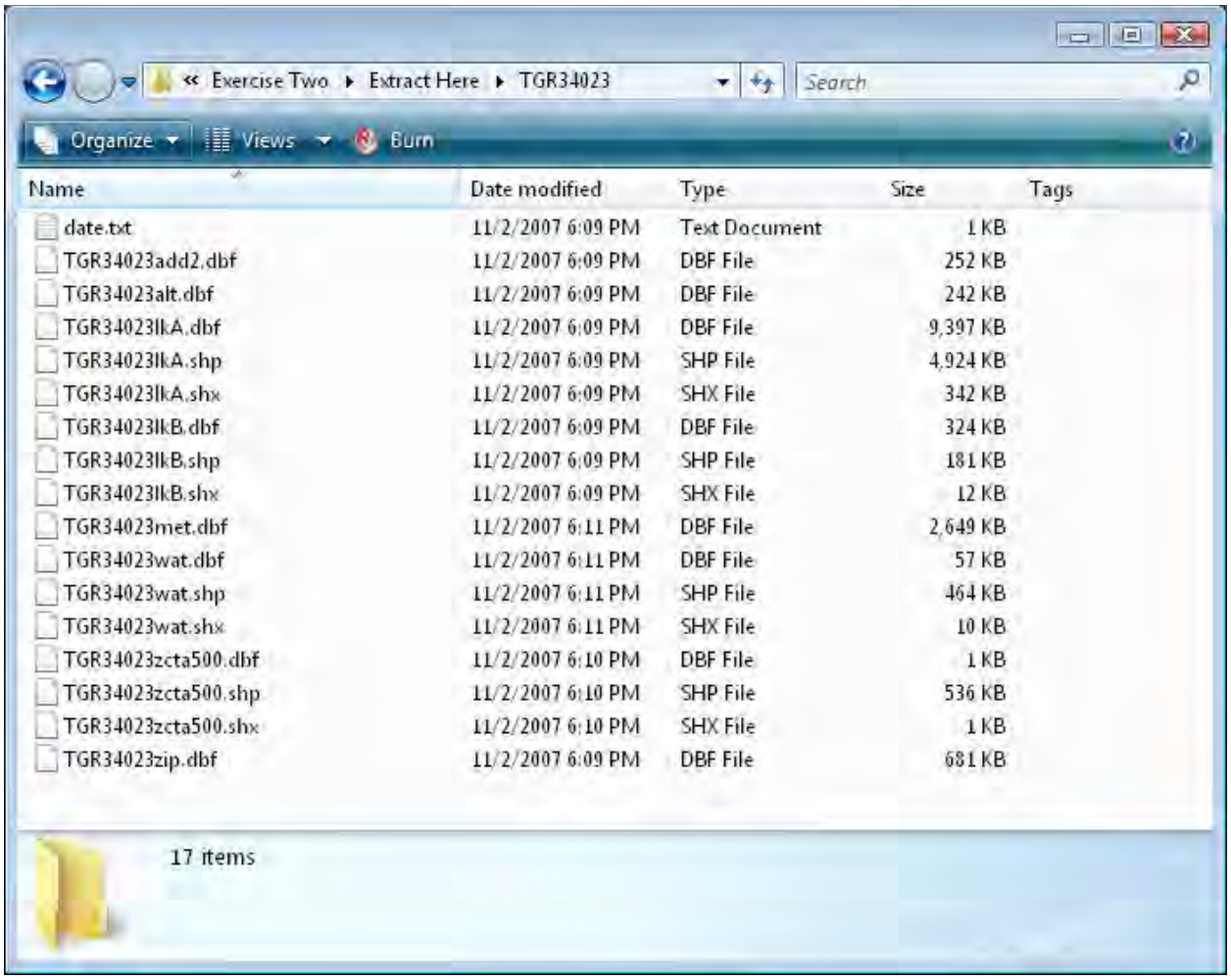
3. Click GO.



4. Place the files in the Extract Here folder. All of the other default options are fine. Click OK.



5. Wait for the process to finish. Then naviagate to the TGR34023 folder in Extract Here.

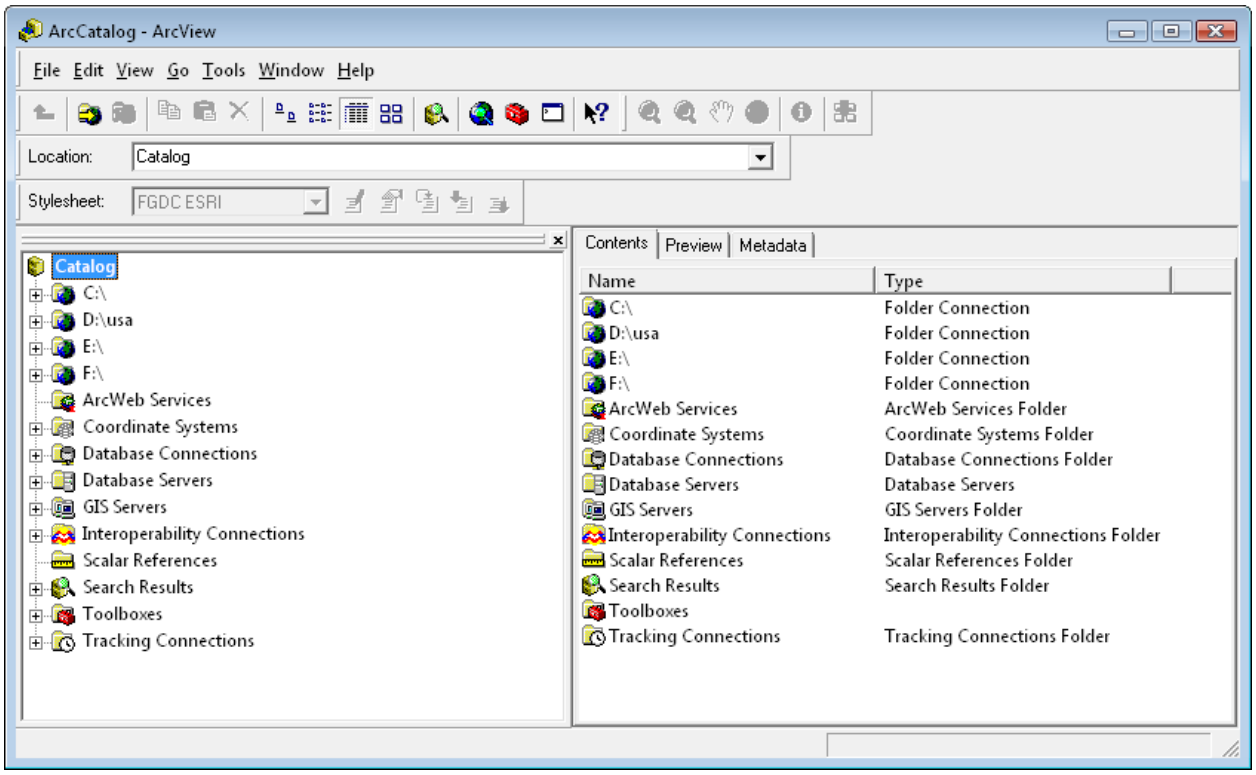


ArcGIS File Types

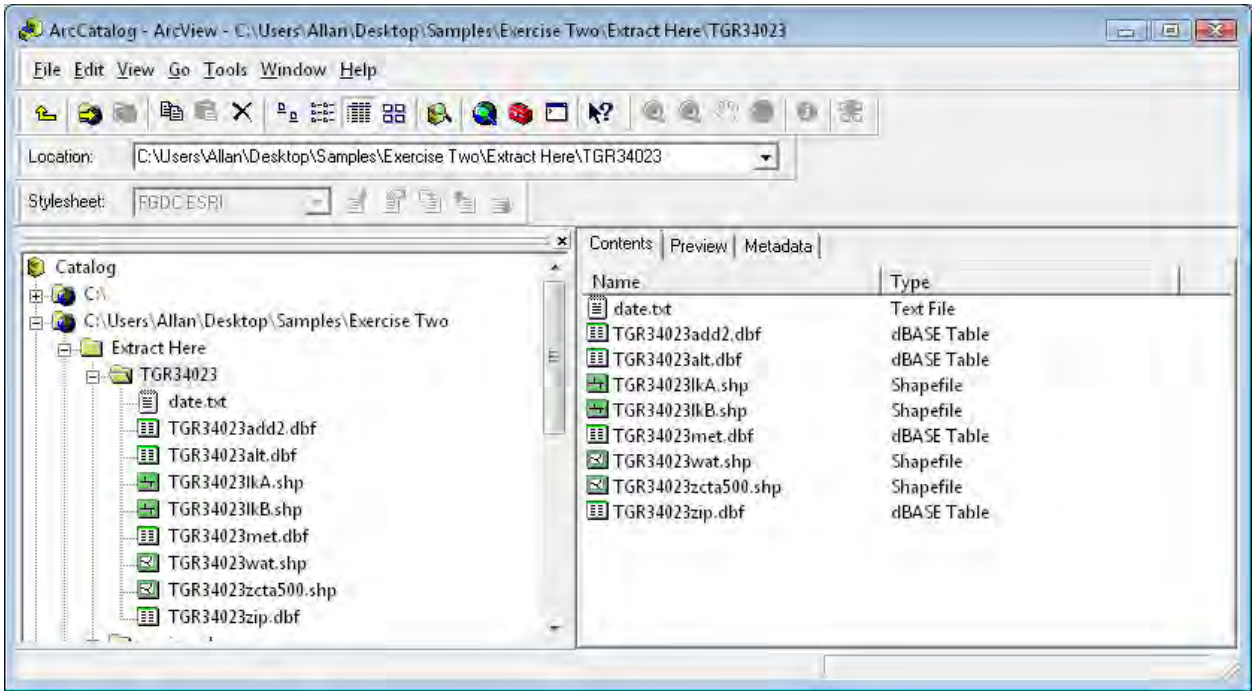
.shp	The main file that stores the feature geometry. Required.
.shx	The index file that stores the index of the feature geometry. Required.
.sbn and **.sbx**	The files that store the spatial index of the features.
.prj	The file that stores the coordinate system information. Used by ArcGIS.
.xml	Metadata for ArcGIS—stores information about the shapefile.
.dbf	The dBASE table that stores the attribute information of features. Required.

There is a one-to-one relationship between geometry and attributes, which is based on record number. Attribute records in the dBASE file must be in the same order as records in the main file.

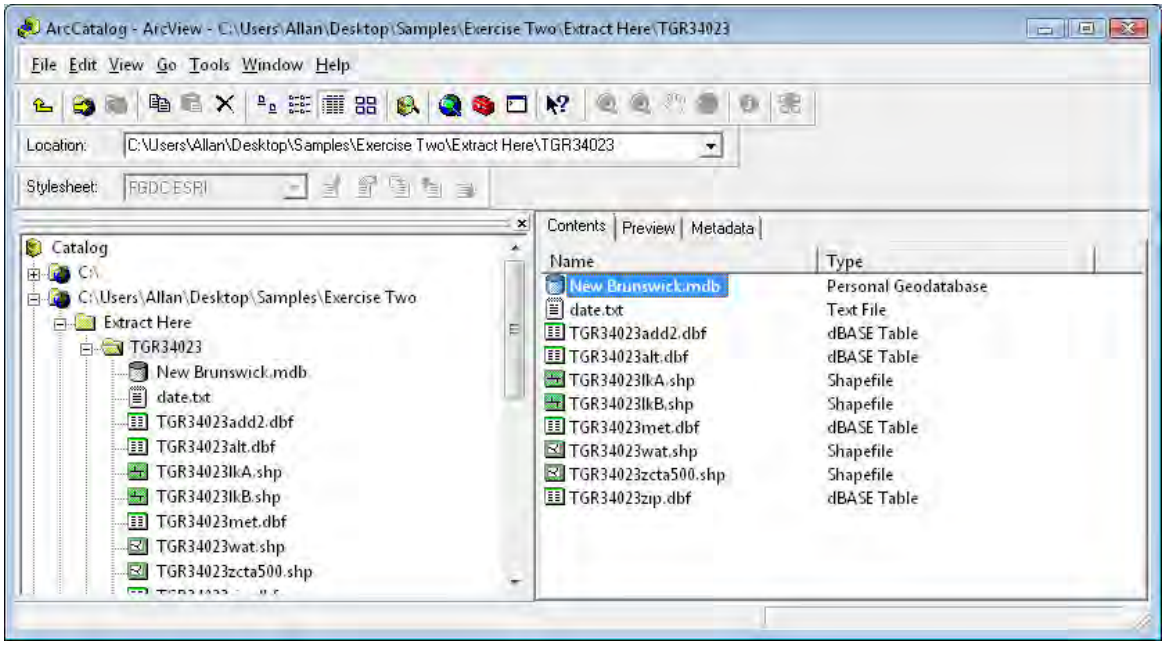
6. Open ArcCatalog. Click on the icon with the yellow arrow and folder (Connect to Folder). Browse to your sample data folder.



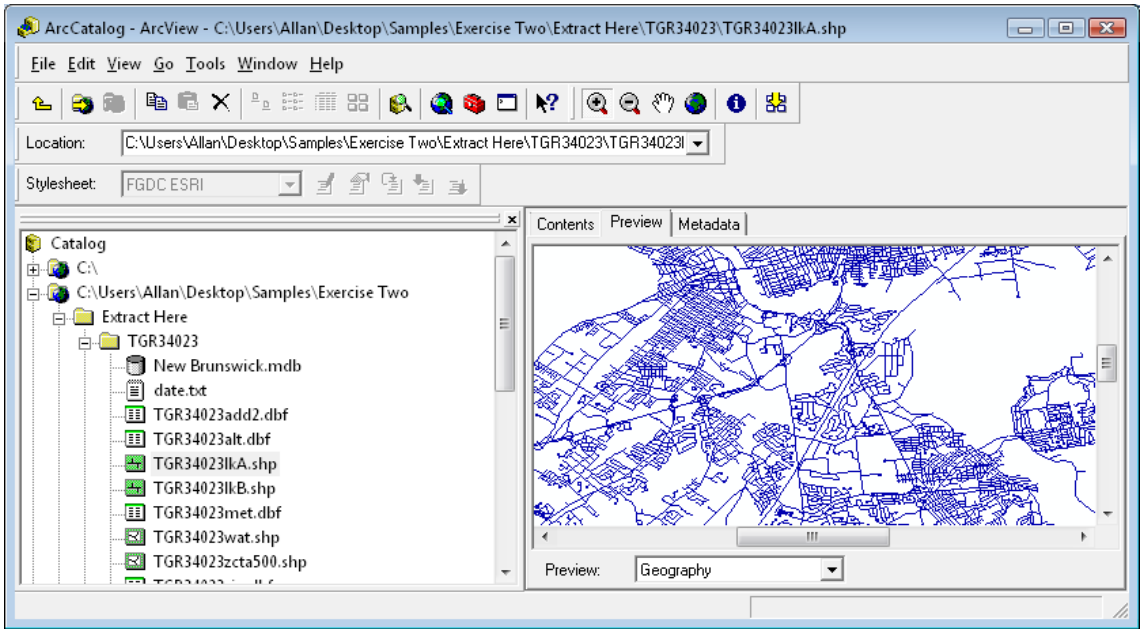
7. Now you can see the files we created. The different icons denote different data types. ZIP Code and Water are polygons. Roads and Railroads are lines. The remaining items are tables.



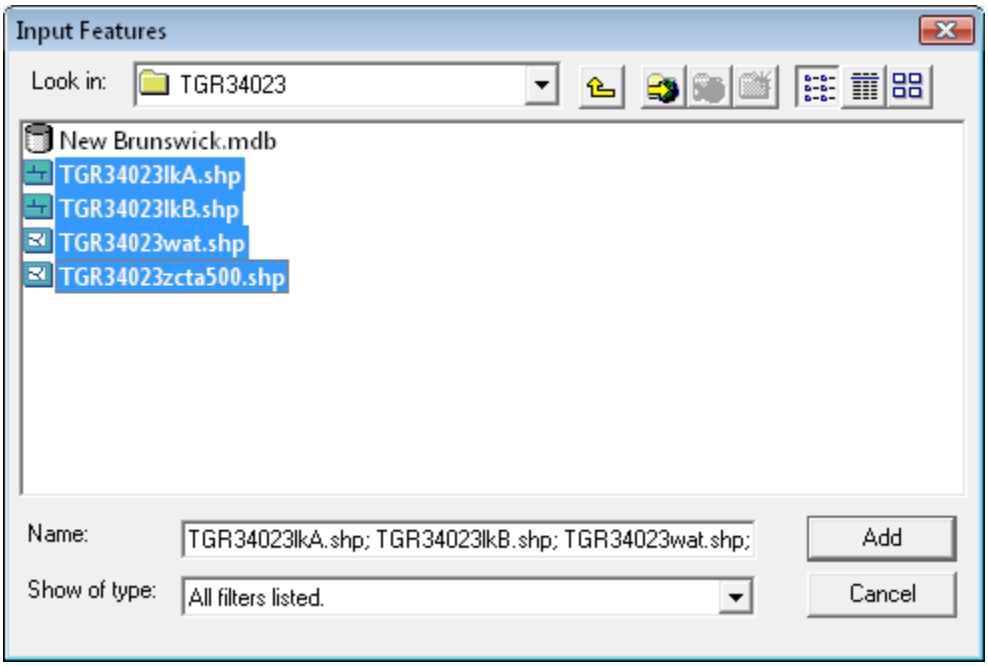
8. Create a Personal Geodatabase. This is not a necessary step, but I included it here to introduce the concept. A personal geodatabase is a Microsoft Access file that stores all of the layers and associated files. Right click in ArcCatalog, select New, then click Personal Geodatabase. Name the file "New Brunswick".

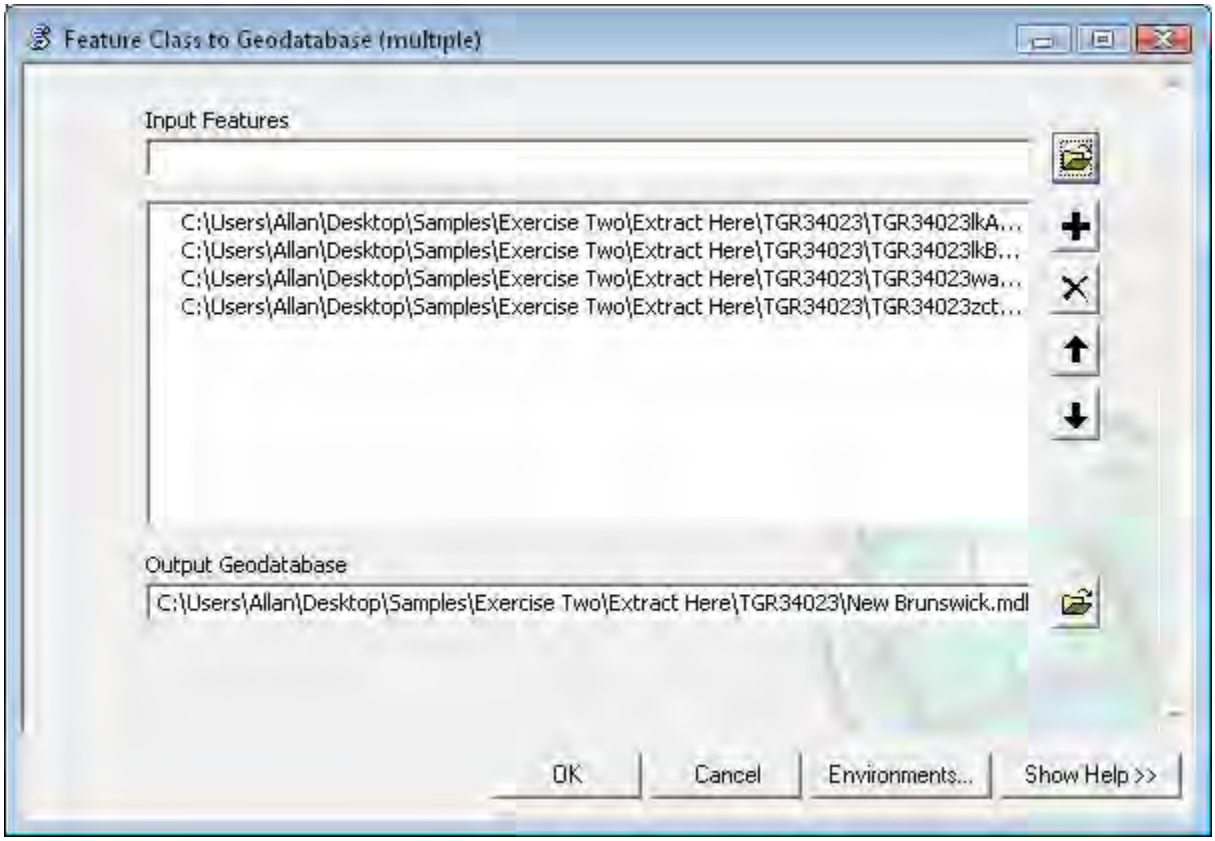


9. Explore the preview and Metadata tabs. Select "TGR34023lkA.shp" and click on the preview tab. This layer contains all of the streets in Middlesex County, New Jersey. Do the same for "TGR34023lkB.shp" and "TGR34023wat.shp". Now examine the Metadata tab. What kind of information is contained there?

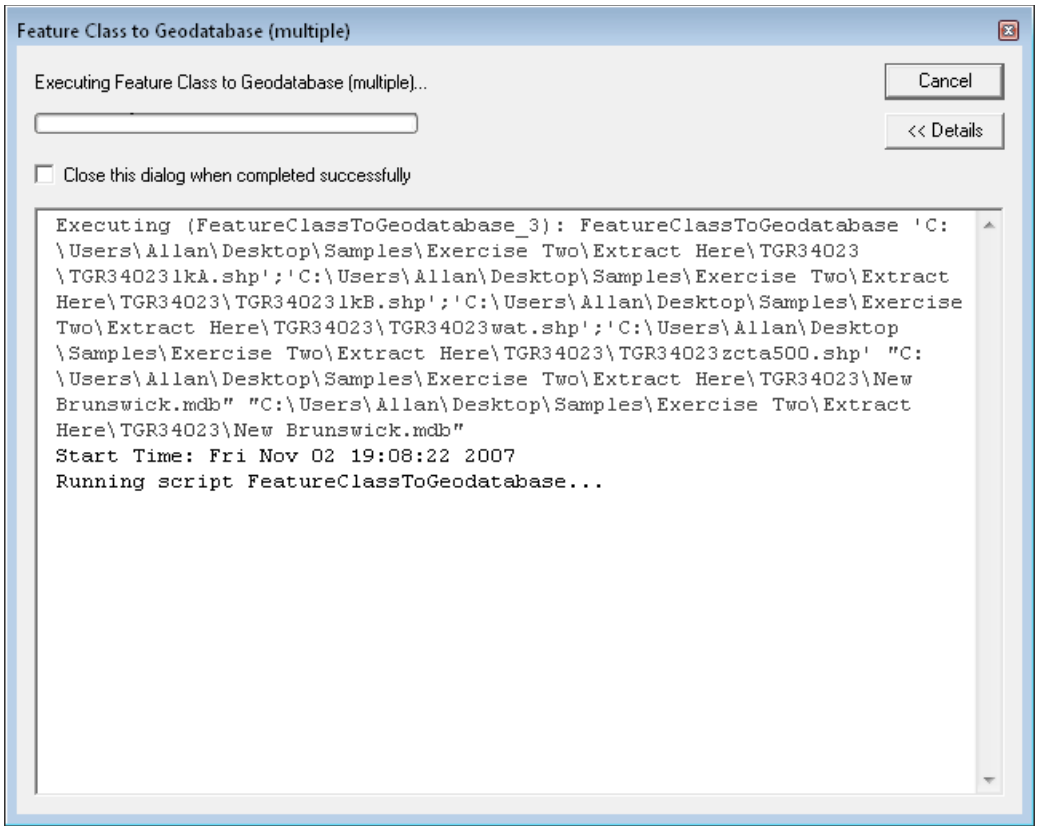


10. Let's import these shapefiles into our personal geodatabase. Right click on New Brunswick.mdb and click on Import → Feature Class (multiple). Select the four shape files in the sample folder. Click OK.

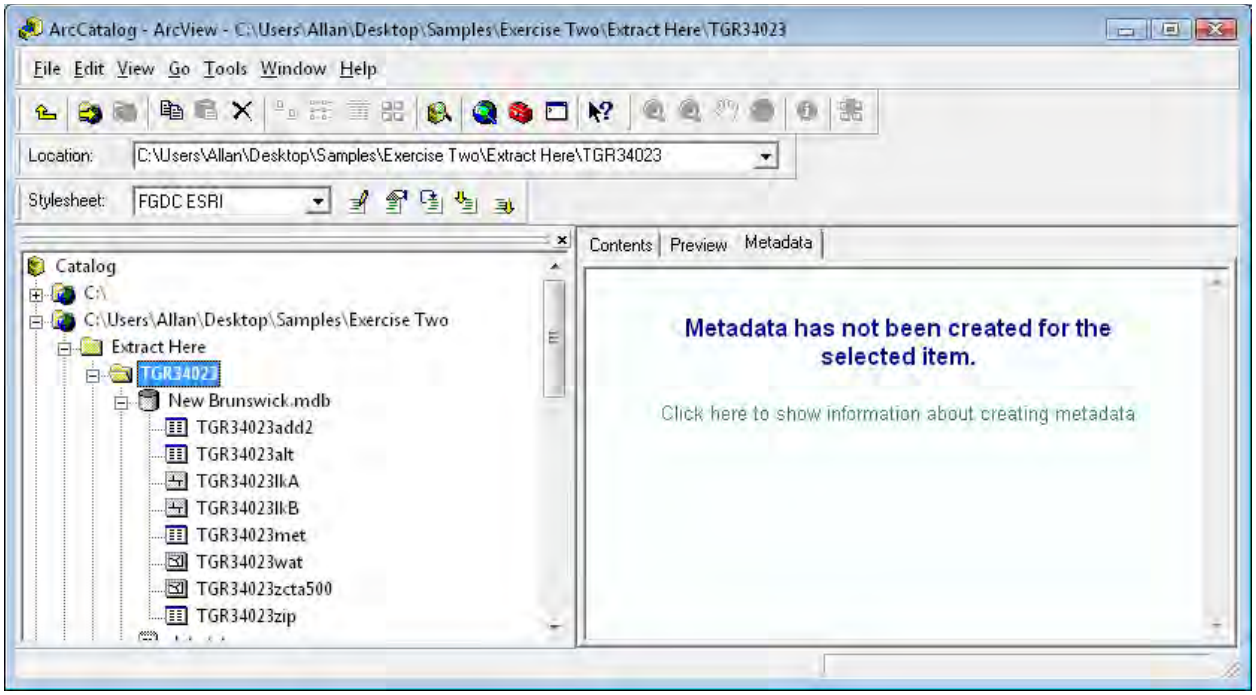




ArcGIS will then add the shapefiles to the geodatabase

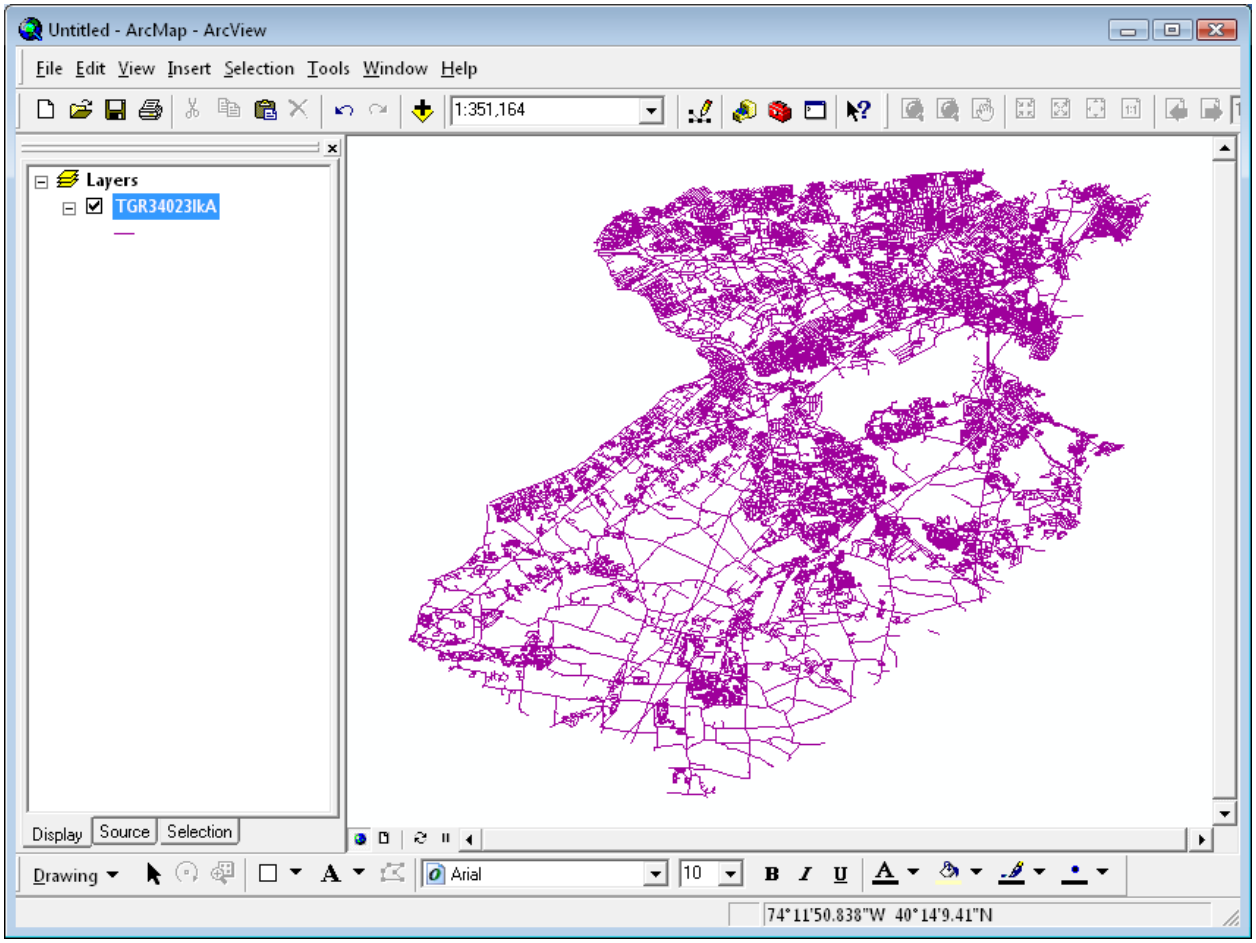


11. Do the same but for Table (Multiple). When it finishes, all of the shapefiles and tables will appear within the geodatabase.



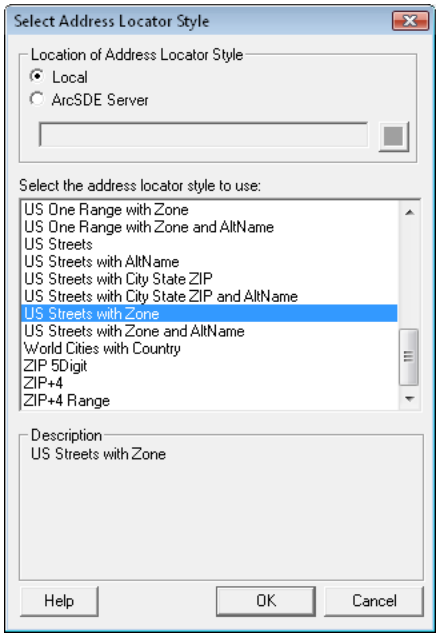
12. Explore the Stylesheet toolbar. Metadata in ArcGIS is data about the data. How to organize metadata is beyond the scope of this workshop, but you should consider maintaining the metadata for your own projects. Click on the small globe and magnifying lens icon to open ArcMap and then close ArcCatalog.

13. In ArcMap, click the Add Layer button. Select "TGR34023lkA.shp" from within the geodatabase we just created. Take a few moments to explore the layer. Open up the Attribute Table. Use the identification tool (the "i" within a blue circle) to select one of the roads.

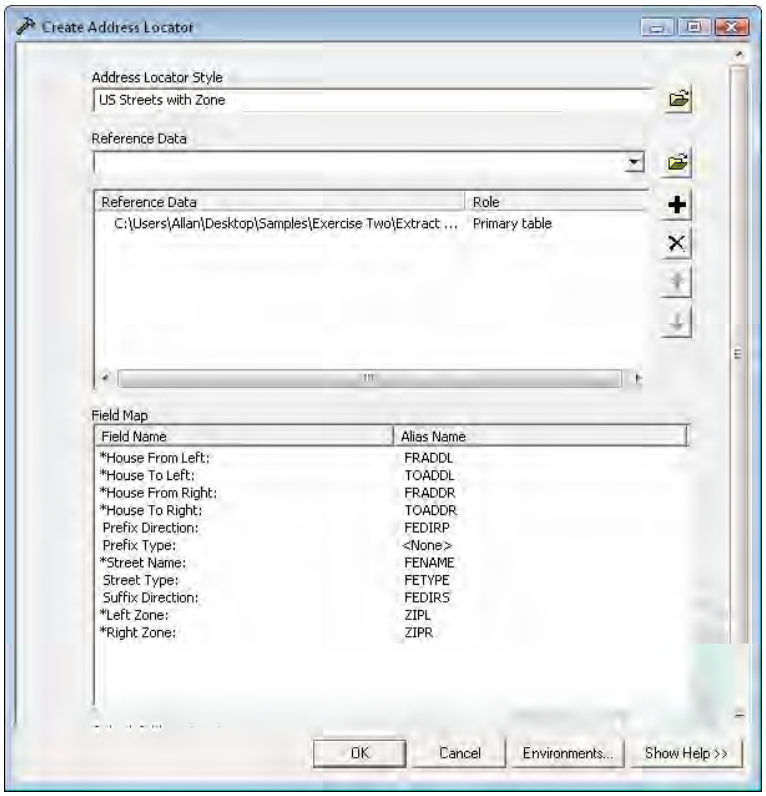


14. Add the railroad ("TGR34023lkB.shp") and water layers (""TGR34023wat.shp"). Using the General tab under Properties, rename each layer ("Streets", "Railroads", "Rivers"). For rivers, use the Symbology tab to change the color of the water to Blue. Save the work you have done so far.

15. Creating an Address Locator. Click on the Red Toolbox. Expand the Geocoding Tools box. Click on Create Address Locator. Select "US Streets with Zone" and click OK.

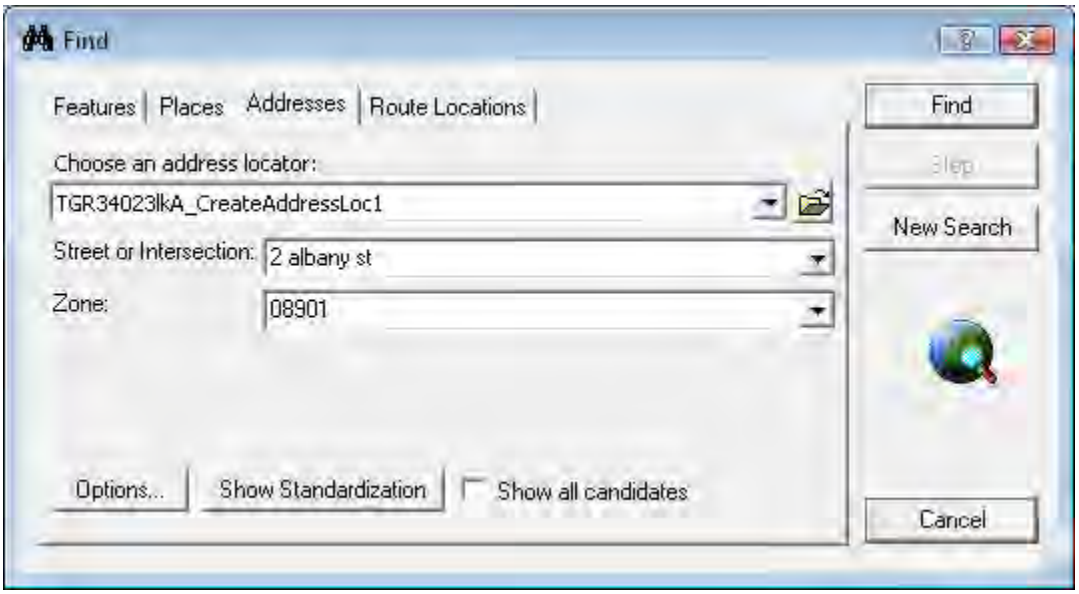


For the reference data, select the streets layer. In the role column, select "Primary Table". All of the remaining items should populate automatically. Click OK.



16. Now let's find out where we are. Click on the binocular's icon, and select the address tab. Using our new address locator, type in "2 Albany Street" in the address line and "08901" in the Zone line.

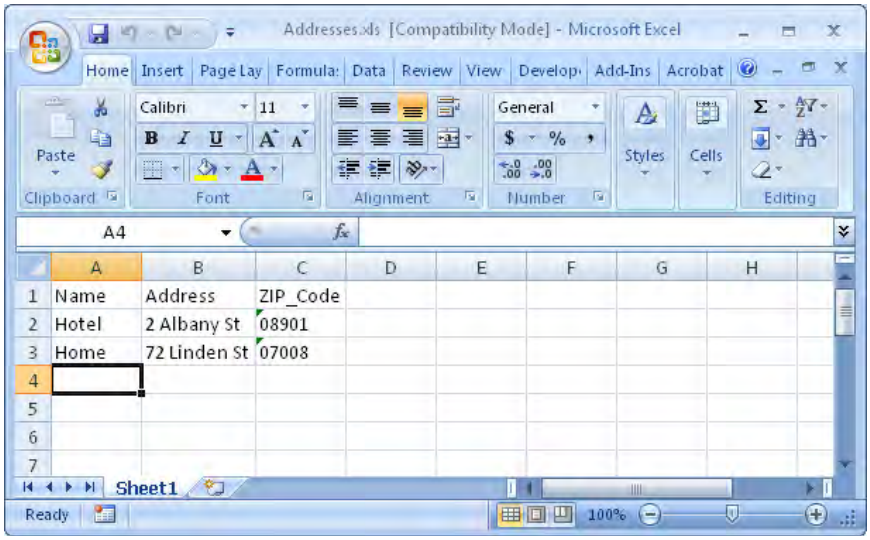
Hyatt Regency New Brunswick
2 Albany Street, New Brunswick, New Jersey, USA 08901



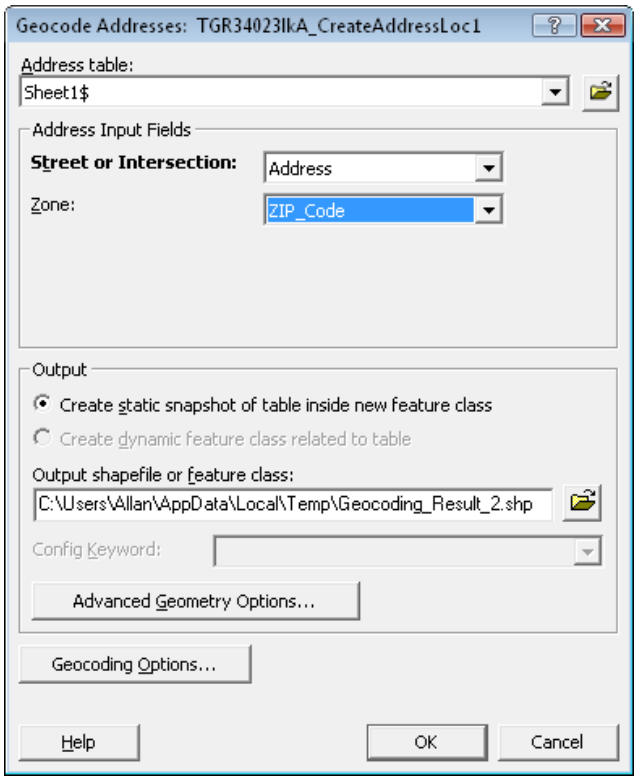
Note: You can rename the address locator by renaming it in ArcCatalog. It is located in the geodatabase we created.

Right click on the first result. Zoom to this point. Does it look correct?

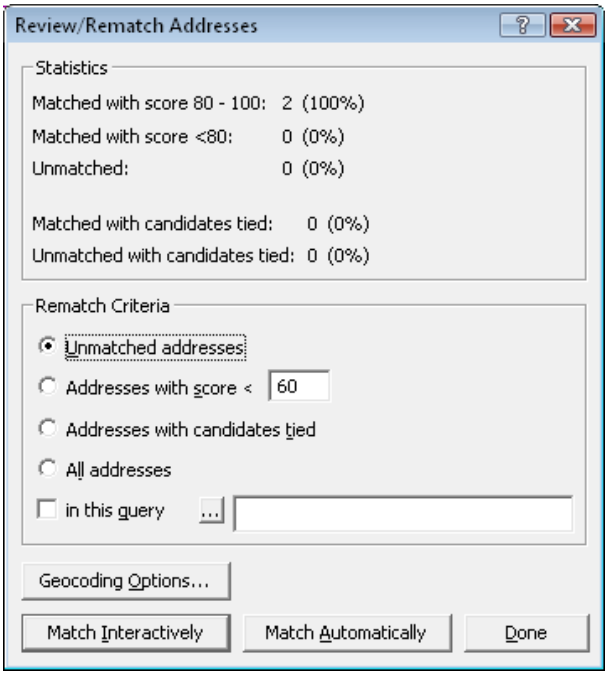
17. Geocoding. Navigate to the Exercise Two folder. Inside of that folder is an Excel spreadsheet ("Addresses.xls") with my home address and the address of the hotel.



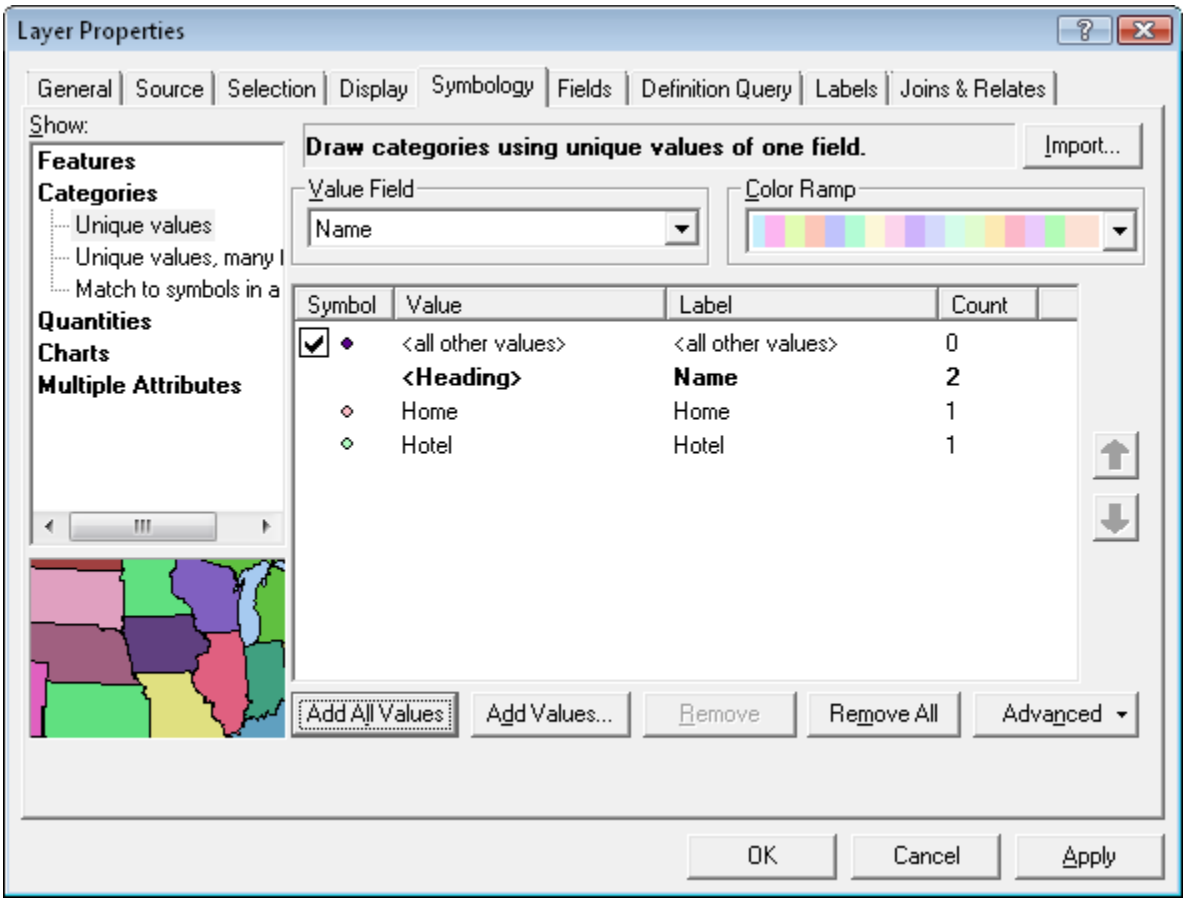
18. Under Tools, select Geocoding, then select Geocode Addresses. Add the address locator we created if it is not there. Browse to the Excel file and then select Sheet1$. Set Zone to ZIP_Code.



Click OK when finished.



19. A new layer appears on the map. Rename this layer as "Addresses". Then click on the Symbology tab. Select "Categories" and "Unique Values". Change the Value Field to "Name". Click on "Add All Values".



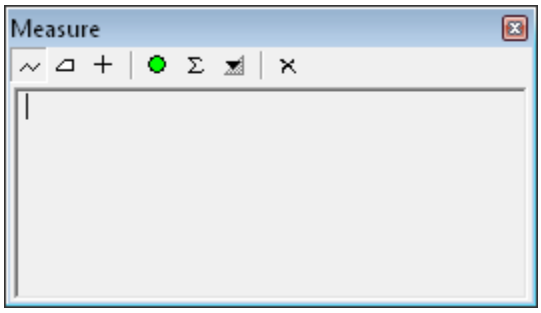
Double click on Home. Click on "More Symbols" add "3-D Residential". Now select "Cape Cod 1".

Double click on Hotel. Click on "More Symbols" add "3-D Buildings". Now select "Hotel 1".

Resize the hotel icon, if necessary.

Unclick. "All Other Values".

20. Turn off all other layers except for addresses. Zoom in on the hotel and my home. Click on the distance icon. A new window will open.



Using the down arrow icon, select "miles" as the measurement. Draw a line between my home and the hotel. Using Google Maps, the distance is approximately 17.5 miles. Since this is a straight line, it is a little shorter.

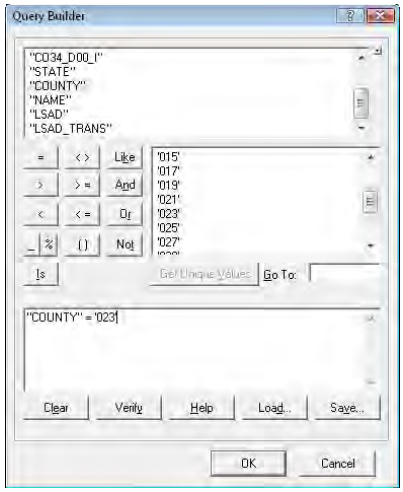
72 Linden St
Carteret NJ 07008
(40.578624, -74.227389)

2 Albany St
New Brunswick NJ 08901
(40.496680, -74.441452)

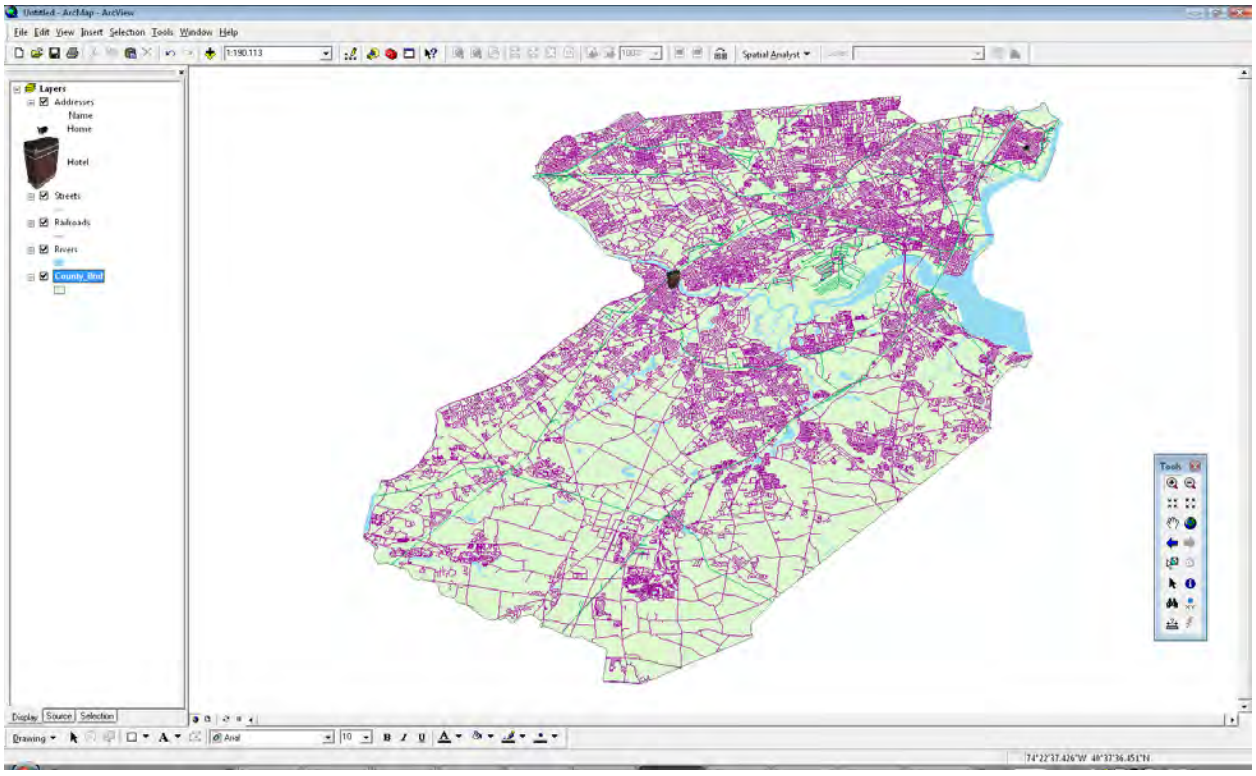
Using the geocoded addresses from (http://geocoder.us/) and the distance calculator at (http://www.movable-type.co.uk/scripts/latlong.html), the distance was 20.25 km, which is 12.58 miles.

This is the "great circle" distance, which is the shortest distance over the Earth's surface. It is an "as the crow flies" method that assumes the Earth is a perfect sphere and ignores elevation.

21. This map is missing something. The county does not have a border. The Bureau of the Census also posts cartographic boundary files. Download the "Census 2000 County and County Equivalent Areas in ArcView Shapefile (.shp) format" file for New Jersey. For this workshop, it is already in the Exercise Two folder ("co34_d00_shp.zip"). "34" is the Federal Information Processing Standards (FIPS) code for New Jersey. Add the co34_d00.shp file to the New Brunswick geodatabase in the same way we added the shapefiles before. This time use a SQL query to select only Middlesex County (County = 23). Call this "County_Bnd".



22. In ArcMap, add the new County_Bnd layer to the map and move it to the bottom of the Table of Contents.

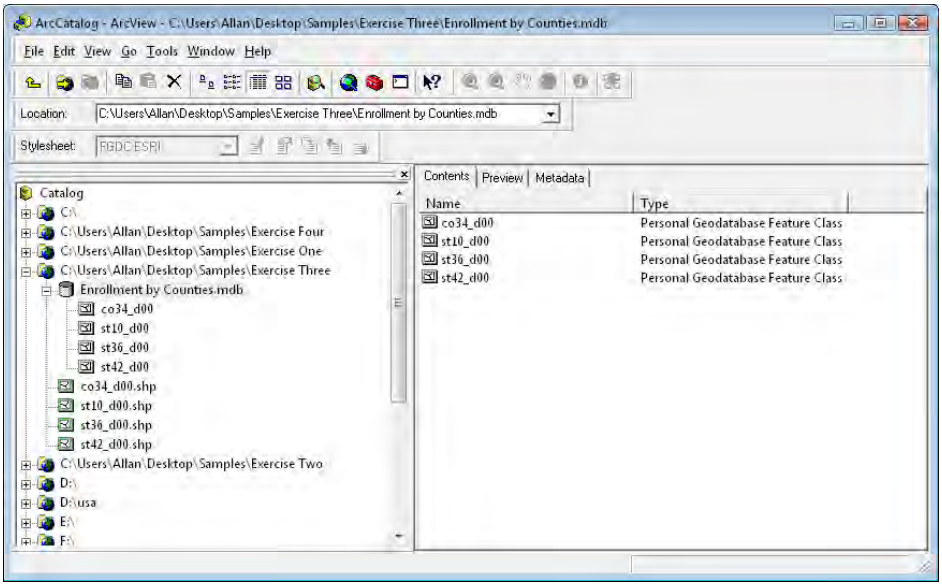


The water doesn't line up perfectly. We won't cover how to fix this in this workshop. Alternatively, you can add in the "TGR34023zcta500" layer, which is the 5-digit ZIP Code Tabulation layer from the geodatabase. Since it was extracted from the same TIGER file, it has a much better fit.

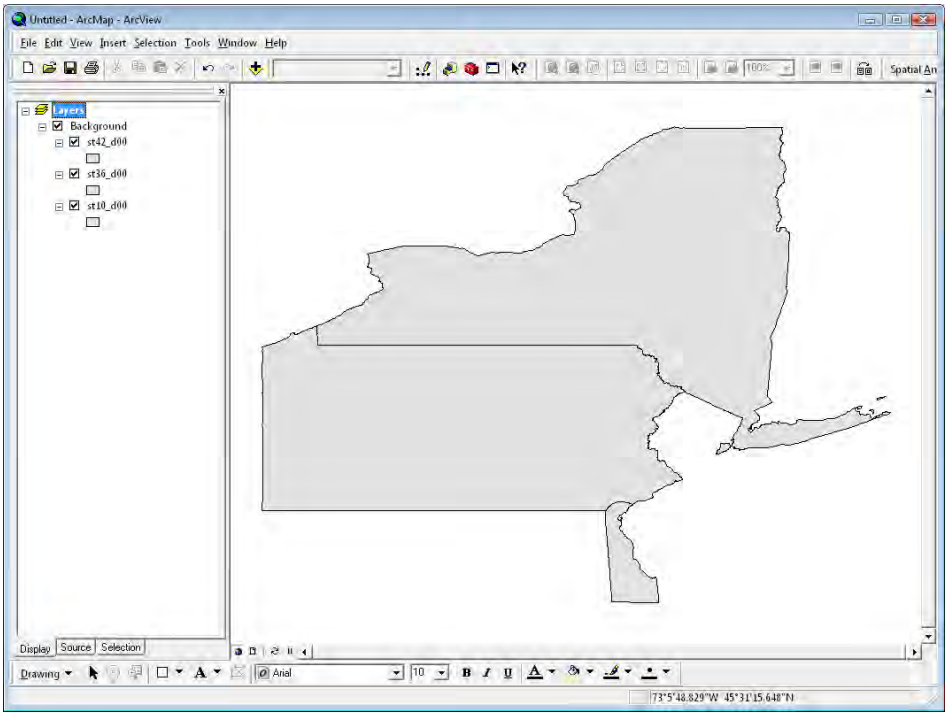
Extra Credit: Properly label the street names using the name and the street type. Hint, you need to create an expression to join the two fields.

Exercise Three: Enrollment by Counties

1. Open ArcCatalog and connect to the Exercise Three folder. Create a Personal Geodatabase called "Enrollment by Counties" and add all of the shapefiles to it.



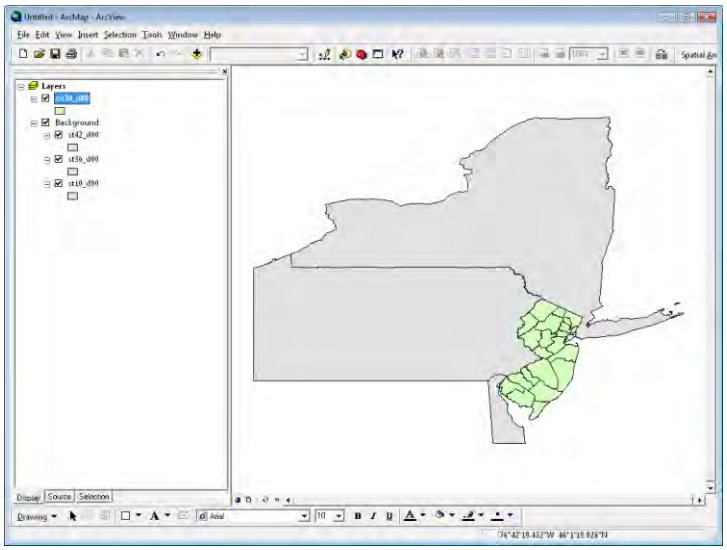
2. Open ArcMap and create a new map. Add all of the layers that start with "st" to the map. Color these layers light grey. Create a layer group. Move all of the states into this group. Call this group "Background".



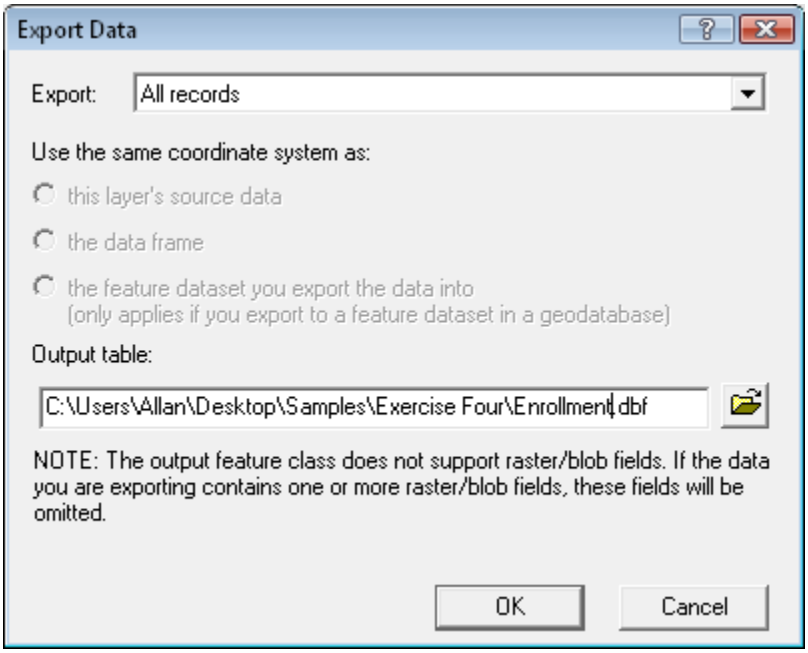
3. Next add the county layer to the map.

4. Add the data from the Excel spreadsheet to the map.

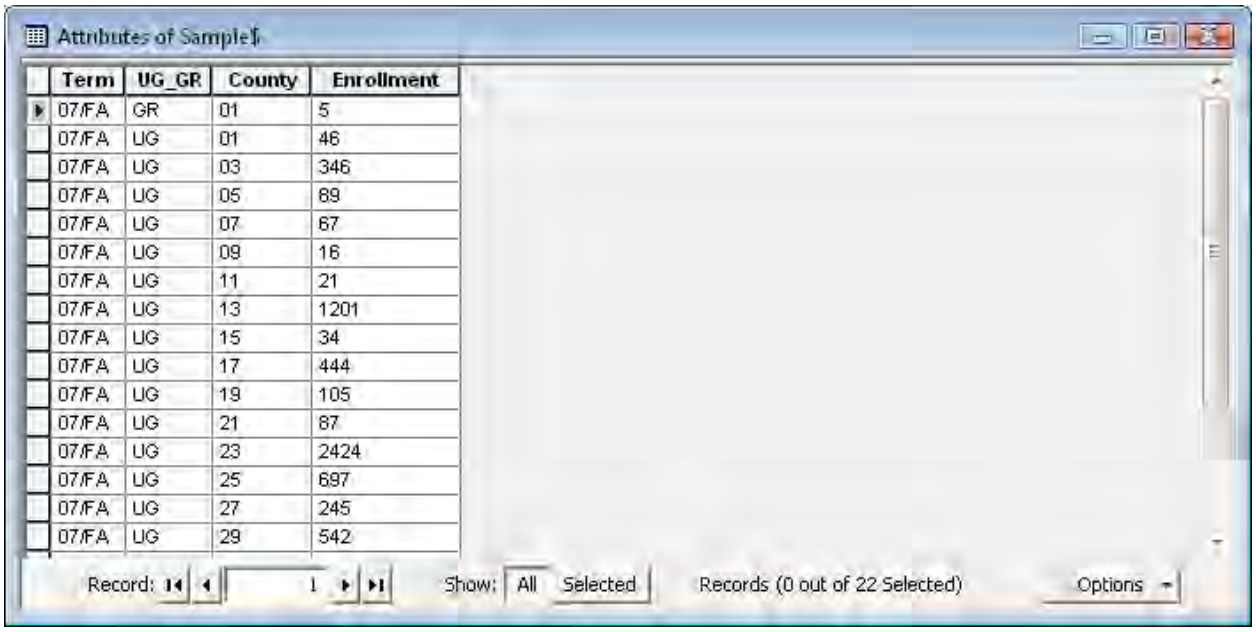
5. ArcMap should now look like this:



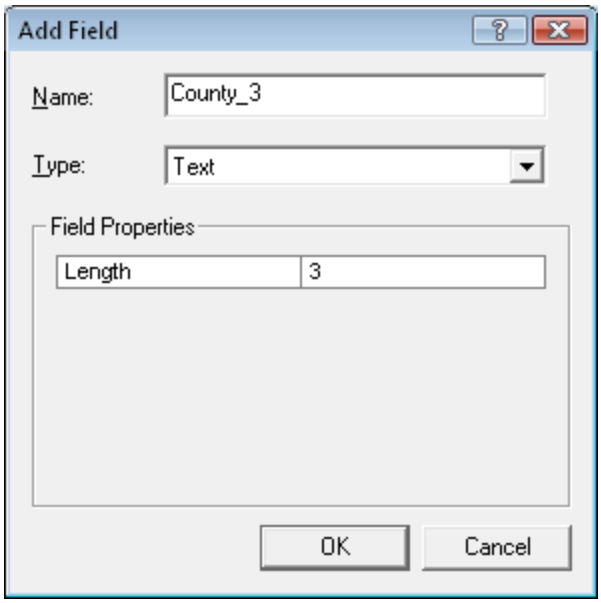
6. Export all of the data from the Excel file into a new table and add it to the map. Delete the original Excel file.



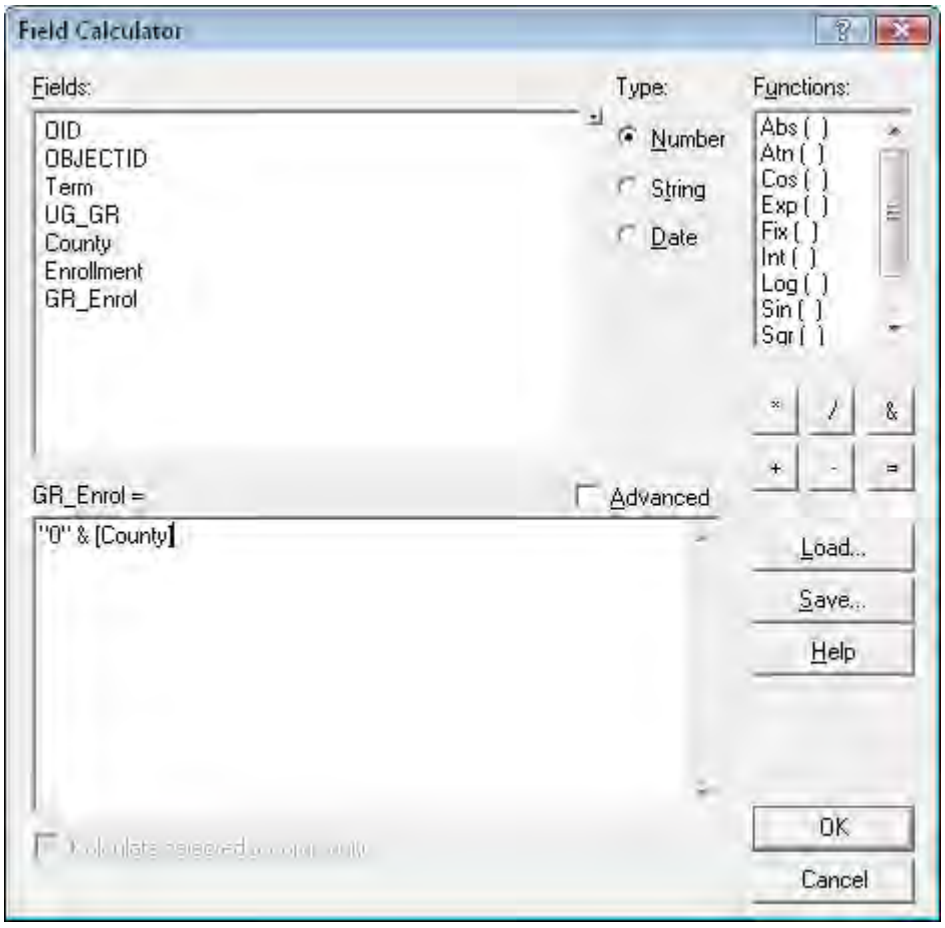
7. County from the Excel file only has two characters, but the County shapefile is three characters. We need to add a new column to the database and add a zero before the county number.

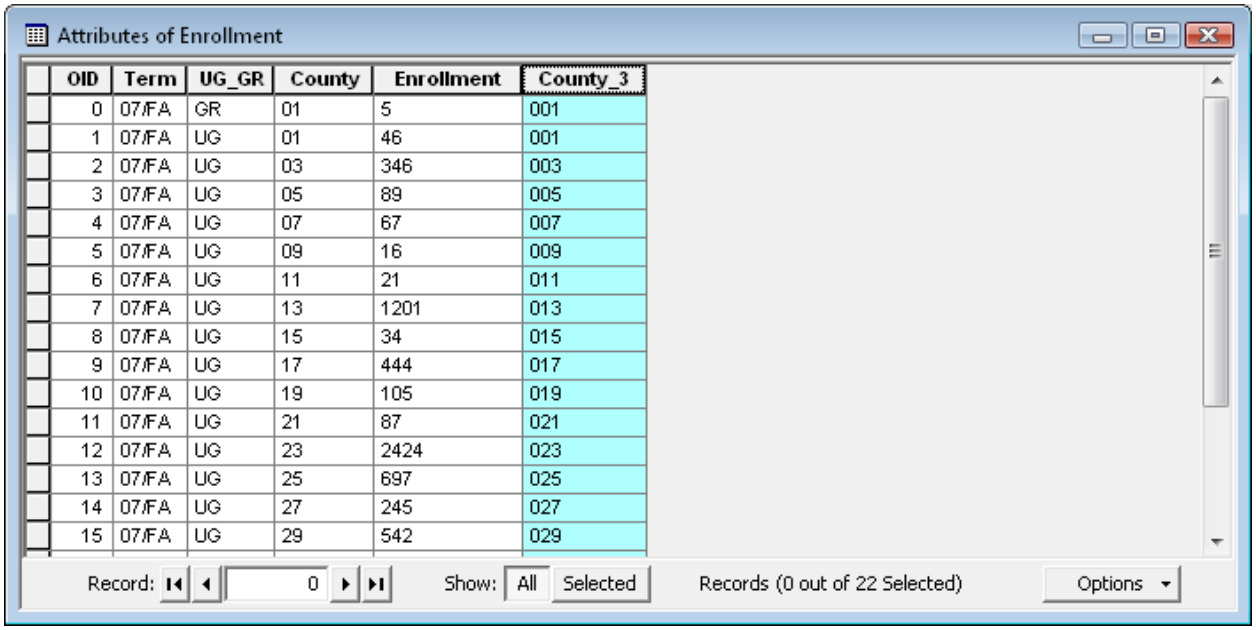


Click on Options, Add new Field. Then create a new text field with a length of 3.

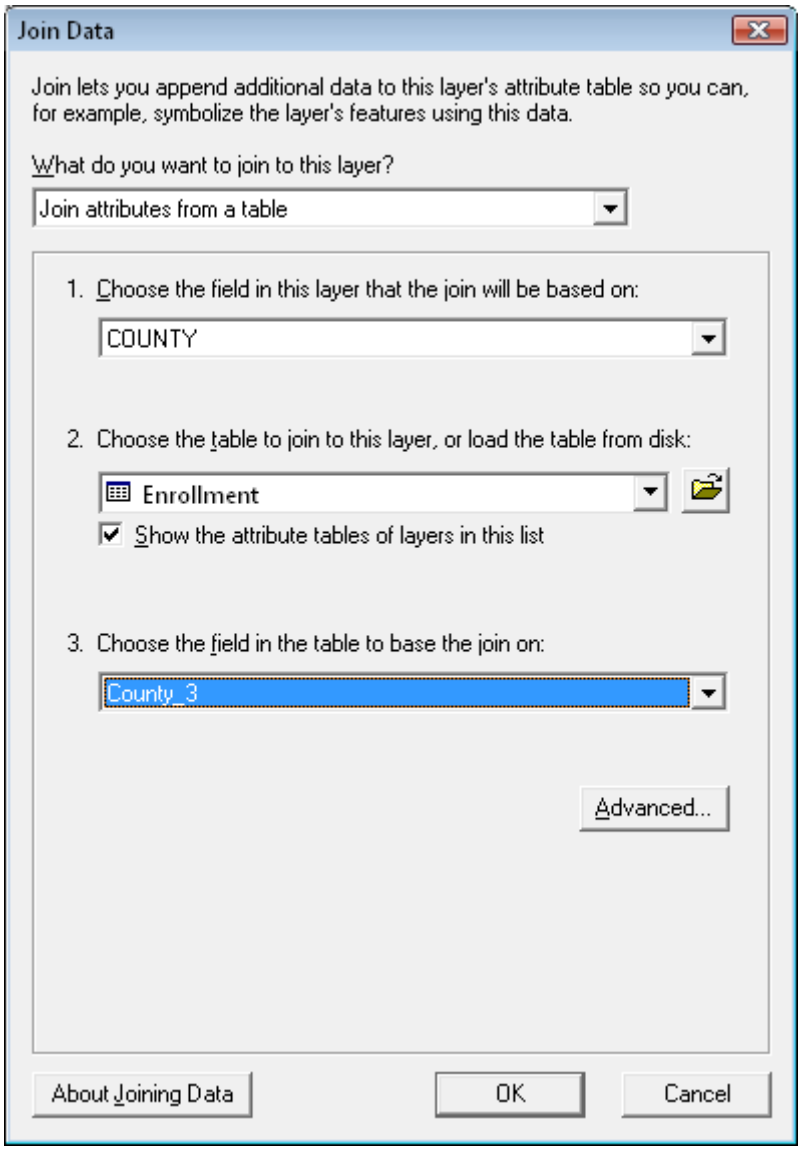


8. Right click on the new field and select Field Calculator.





9. We are now ready to join the enrollment data to the county layer. Right click on county layer → joins and relates → join.



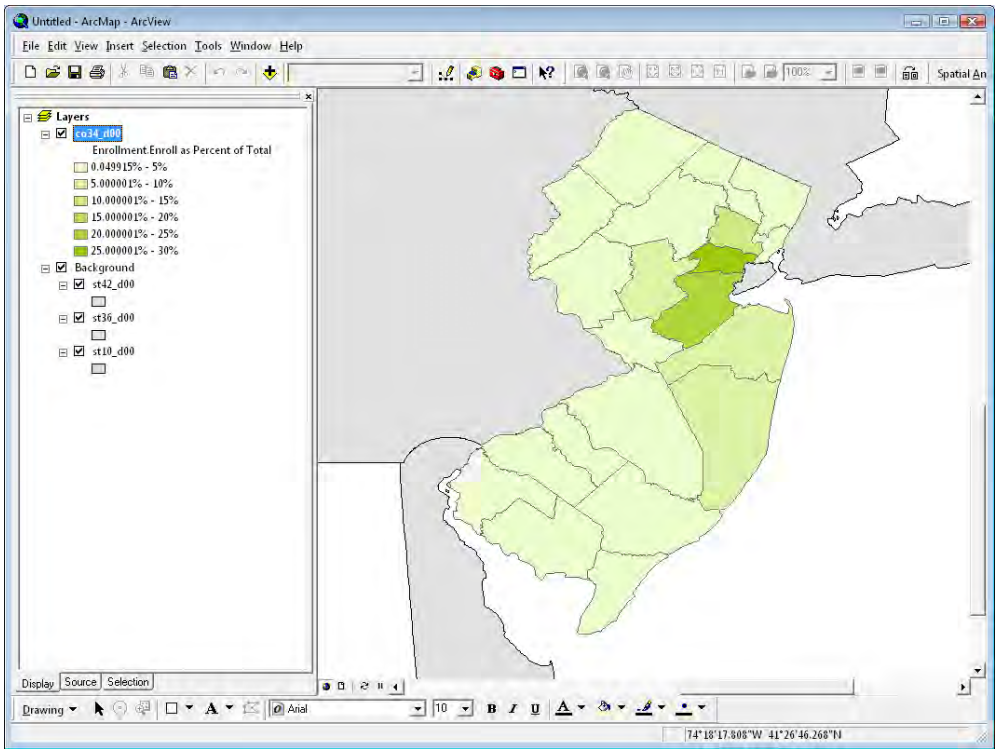
The enrollment data is now in the county layer and can be used to create a choropleth map.

10. Try to map enrollment through the Symbology tab. Notice how it is not an option. The enrollment file is stored as text. We will need to convert this to a number in order to map it. Create a new field, and then store the enrollment in it. Now try to create a choropleth map again. Try making manual breaks every 5% from 5% to 30%.

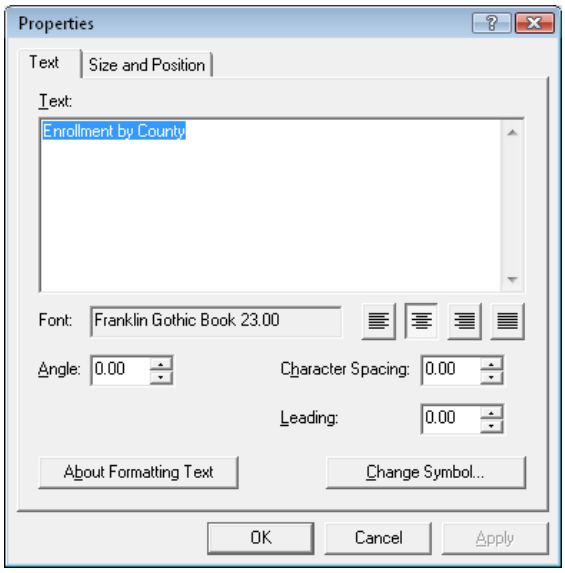
Attributes of Enrollment

OID	Term	UG_GR	County	Enrollment	County_3 *	Enroll
0	07/FA	GR	01	5	001	5
1	07/FA	UG	01	46	001	46
2	07/FA	UG	03	346	003	346
3	07/FA	UG	05	89	005	89
4	07/FA	UG	07	67	007	67
5	07/FA	UG	09	16	009	16
6	07/FA	UG	11	21	011	21
7	07/FA	UG	13	1201	013	1201
8	07/FA	UG	15	34	015	34
9	07/FA	UG	17	444	017	444
10	07/FA	UG	19	105	019	105
11	07/FA	UG	21	87	021	87
12	07/FA	UG	23	2424	023	2424
13	07/FA	UG	25	697	025	697
14	07/FA	UG	27	245	027	245
15	07/FA	UG	29	542	029	542

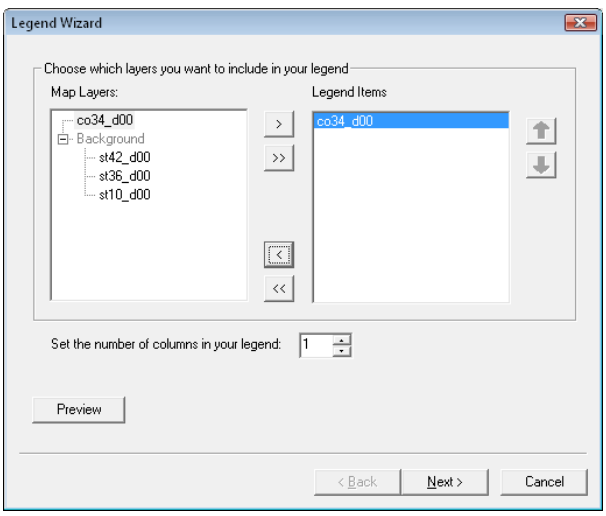
Record: 0 Show: All Selected Records (0 out of 22 Selected) Options

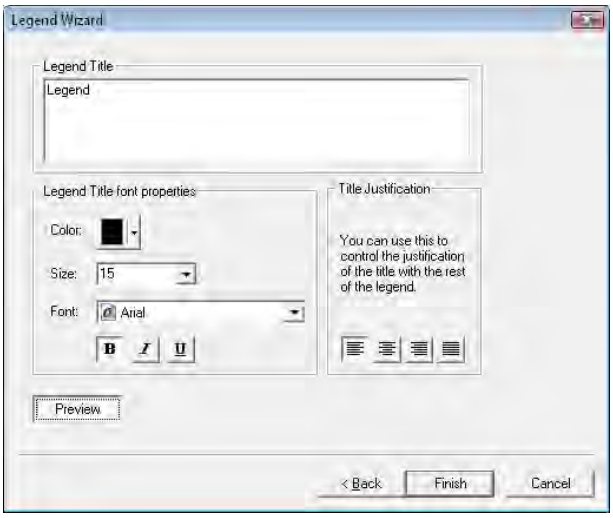


11. Under View, click on Layout View. Click on Insert, Add Title. Enter Text. Click "Change Symbol" to adjust the font and size.

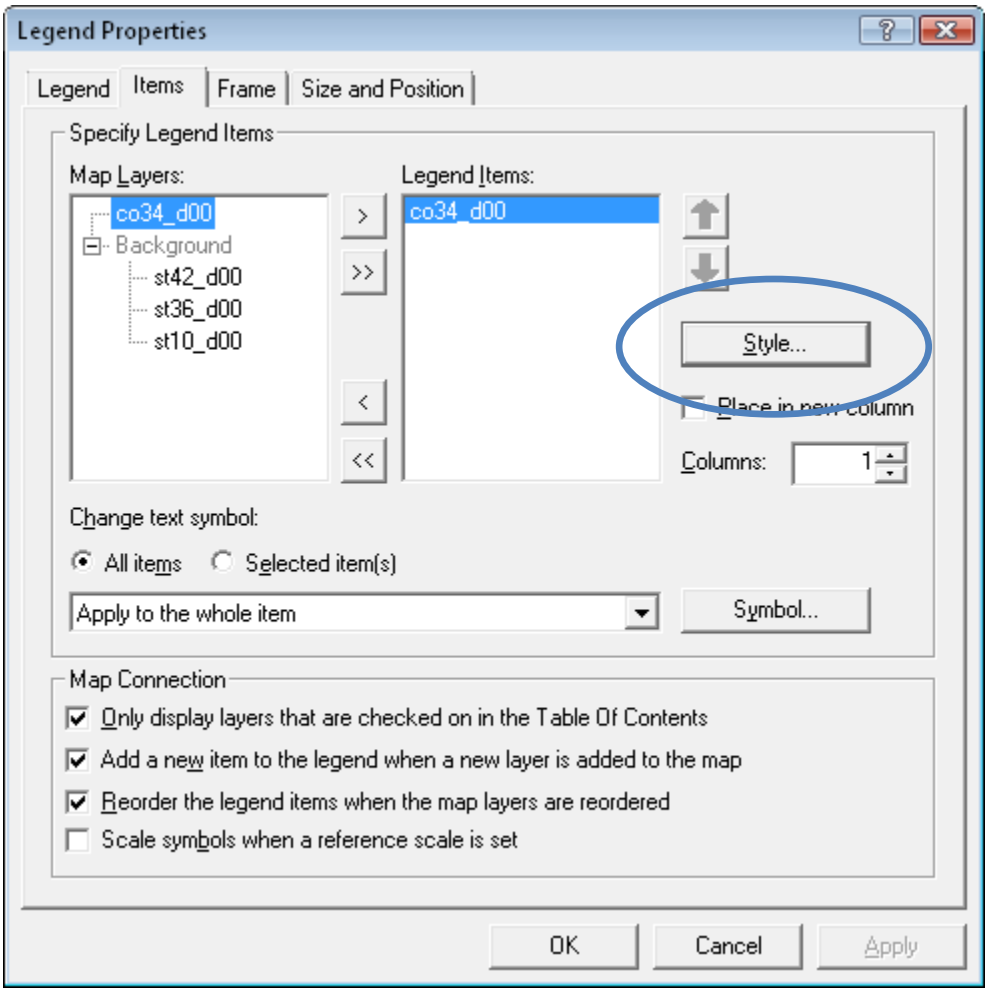


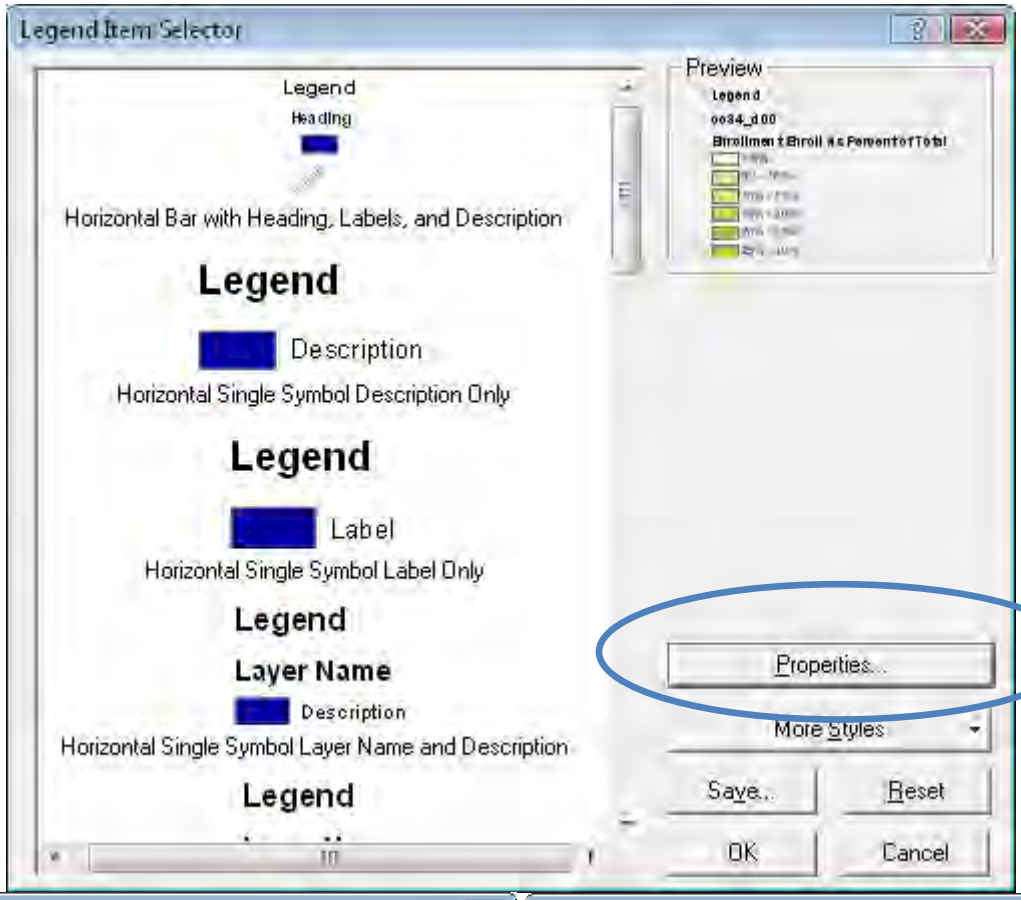
12. Insert a legend. Adjust the labels in the Symbology tab to fix the labels in the legend.

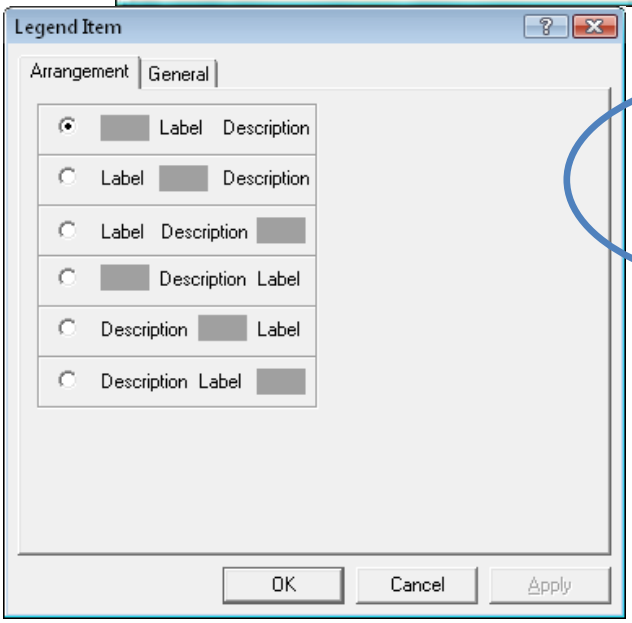


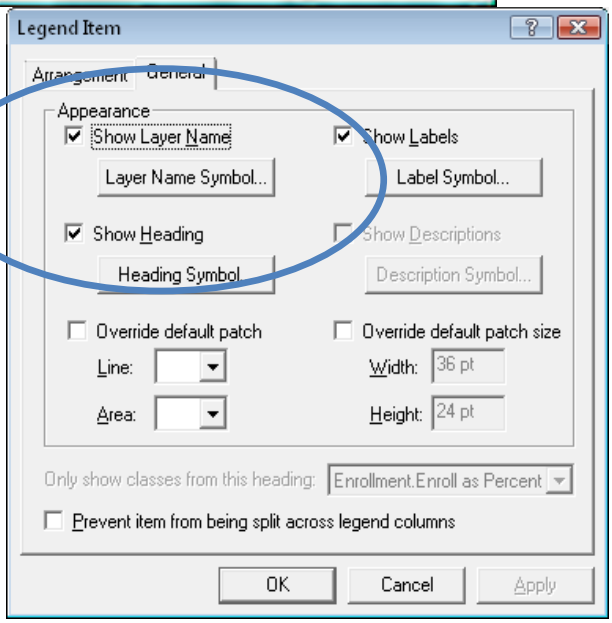


13. Edit the legend properties.




Legend Item Selector
Legend
Heading
Horizontal Bar with Heading, Labels, and Description
Legend
Description
Horizontal Single Symbol Description Only
Legend
Label
Horizontal Single Symbol Label Only
Legend
Layer Name
Description
Horizontal Single Symbol Layer Name and Description
Legend
Preview
Properties...
More Styles
Save...
Reset
OK
Cancel

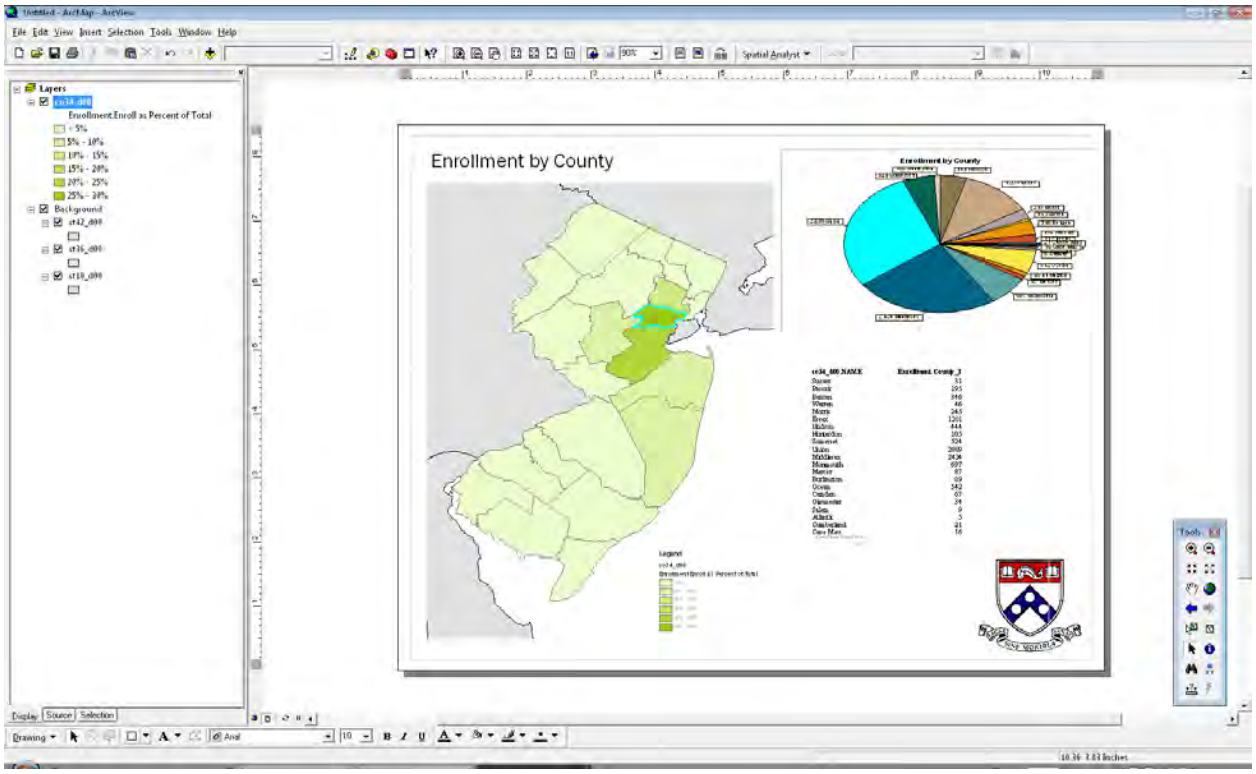

Legend Item
Arrangement
General
Label Description
Label Description
Label Description
Description Label
Description Label
Description Label
OK
Cancel
Apply


Legend Item
General
Appearance
Show Layer Name
Layer Name Symbol...
Show Heading
Heading Symbol
Override default patch
Line:
Area:
Show Labels
Label Symbol...
Show Descriptions
Description Symbol...
Override default patch size
Width:
36 pt
Height:
24 pt
Only show classes from this heading:
Enrollment.Enroll as Percent
Prevent item from being split across legend columns
OK
Cancel
Apply

14. Add the logo picture to the layout.

15. Right click on the county attribute table. Under options, create a graph (pie chart). Add this to the layout.

16. Change the page layout to landscape instead of portrait. Try adding a report to the layout.



Exercise Four: Pittsburg Schools

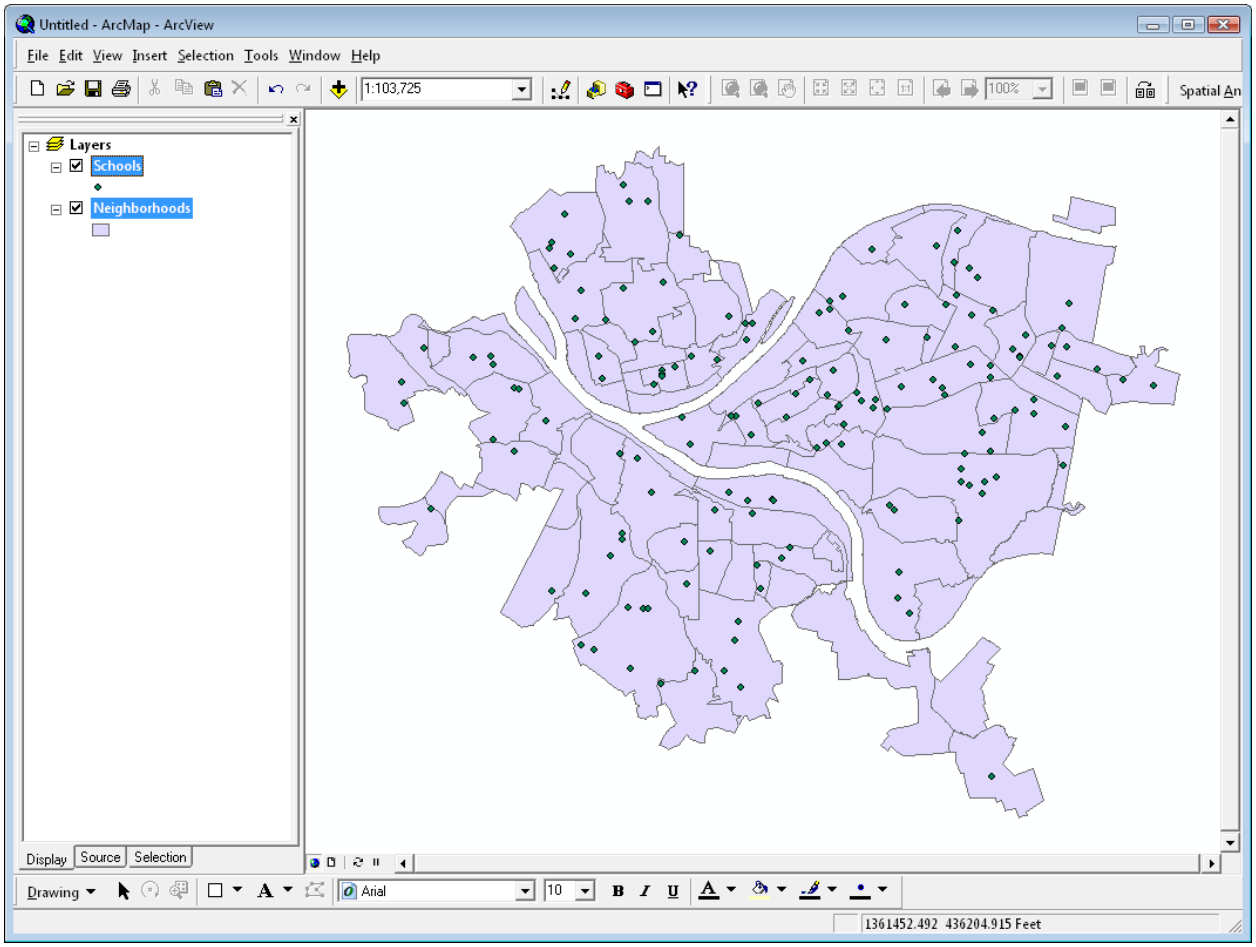
Goal:

1. **Create a map showing schools in the city of Pittsburg by enrollment, school type. Only include schools that are open.**
2. **Create a map showing K-12 population for the Commonwealth of Pennsylvania. Include a map zoomed into the city of Pittsburg that also includes public school enrollment.**
3. **Create a map showing school enrollment compared with school age population.**

Part I: Pittsburg Schools by Enrollment and Type

1. Connect to the Exercise Four folder in ArcCatalog. Examine the contents of the folder. Create a new map in ArcMap.

2. Add the Neighborhoods and Schools layers to the map. The white lines going through the city is a river.



3. Examine the Attribute Table for the Schools layer. Notice the following fields:

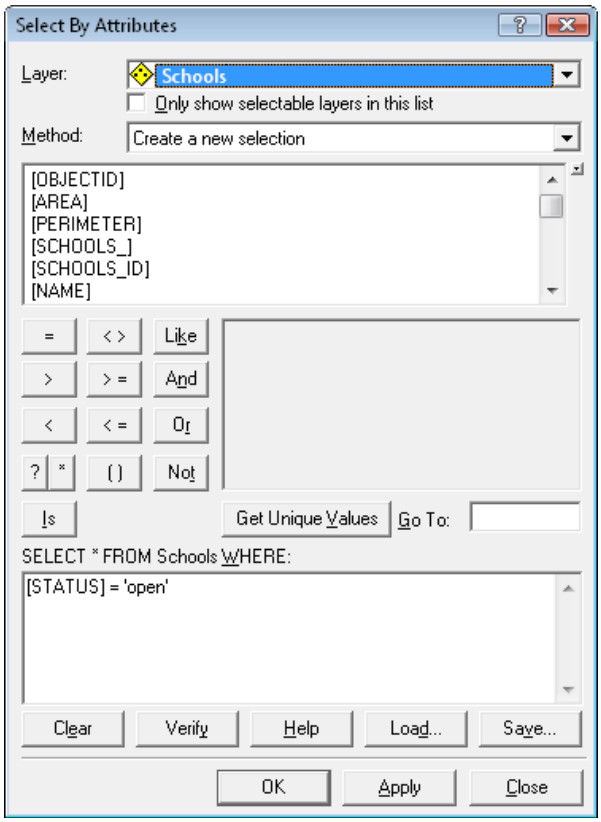
DISTRICT = school type

City of Pittsburg	Public School
Pittsburg Dioceses	Parochial School (private)
Private School	Private School (private)

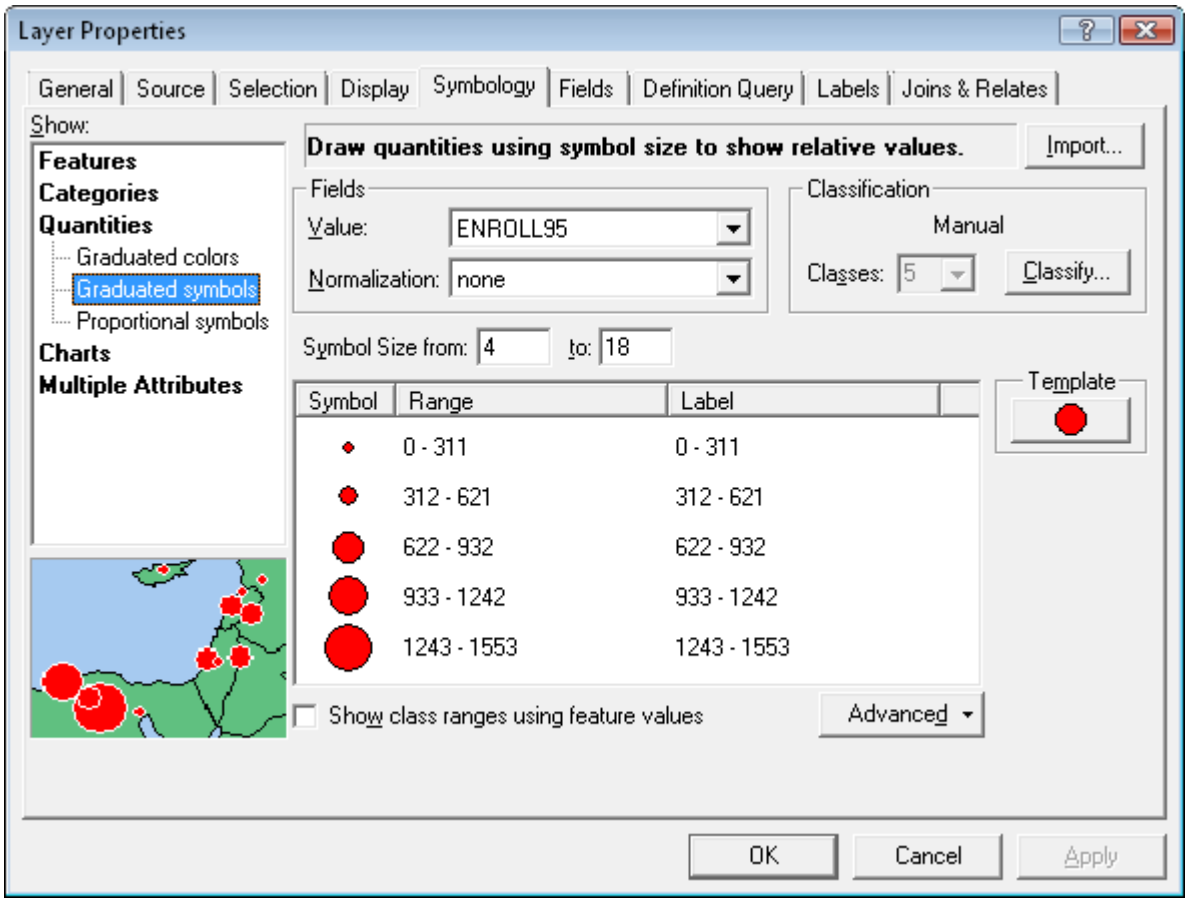
STATUS = open or closed

ENROLL95 = number of students enrolled in 1995

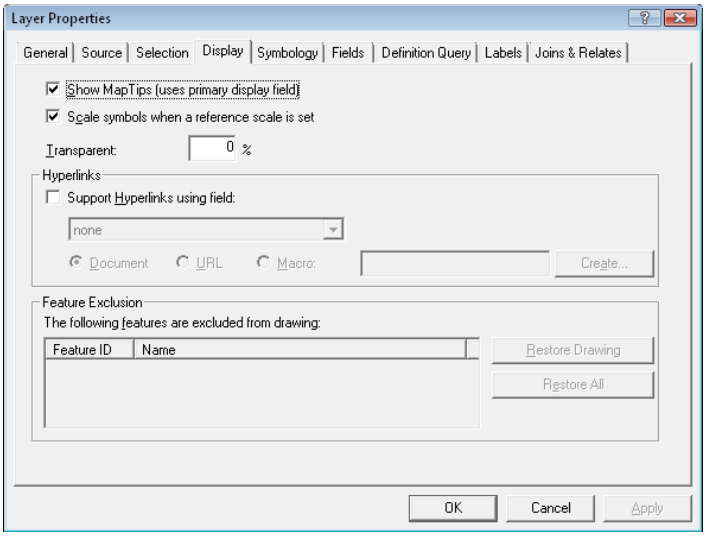
4. Click on Tools, Select by Attribute. Select Schools that are open and that are public. Right click on the Schools layer in the Table of Contents. Under Selection click on Create Layer from Selected Features. Rename this layer as Public Schools. Do the same for non-public schools. Hint: use the "<>" operator. Rename this layer as Private Schools. Remove the Schools layer from the map.



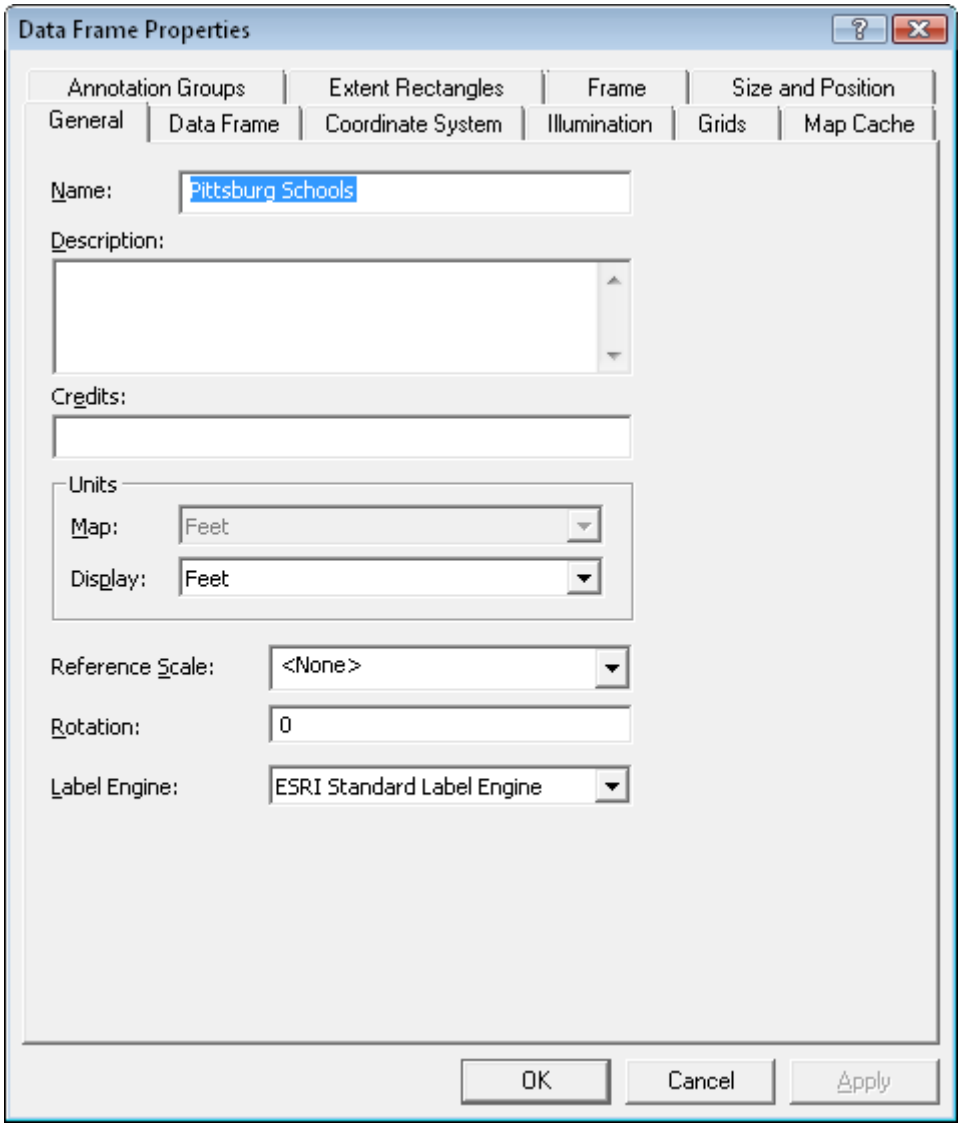
5. Use the Symbology tab to create graduated symbols based on ENROLL95. Use the same range for public and private schools. Make sure the symbols are the same, but there are two different colors.



6. Add MapTips. Open the Layer Properties window for Private Schools. In the Fields tab, make sure that NAME is the Primary Display Field. Now go to the Display tab and click Show MapTips. Do the same thing for the Public School layer. The names of the schools will now appear when you hover over them.



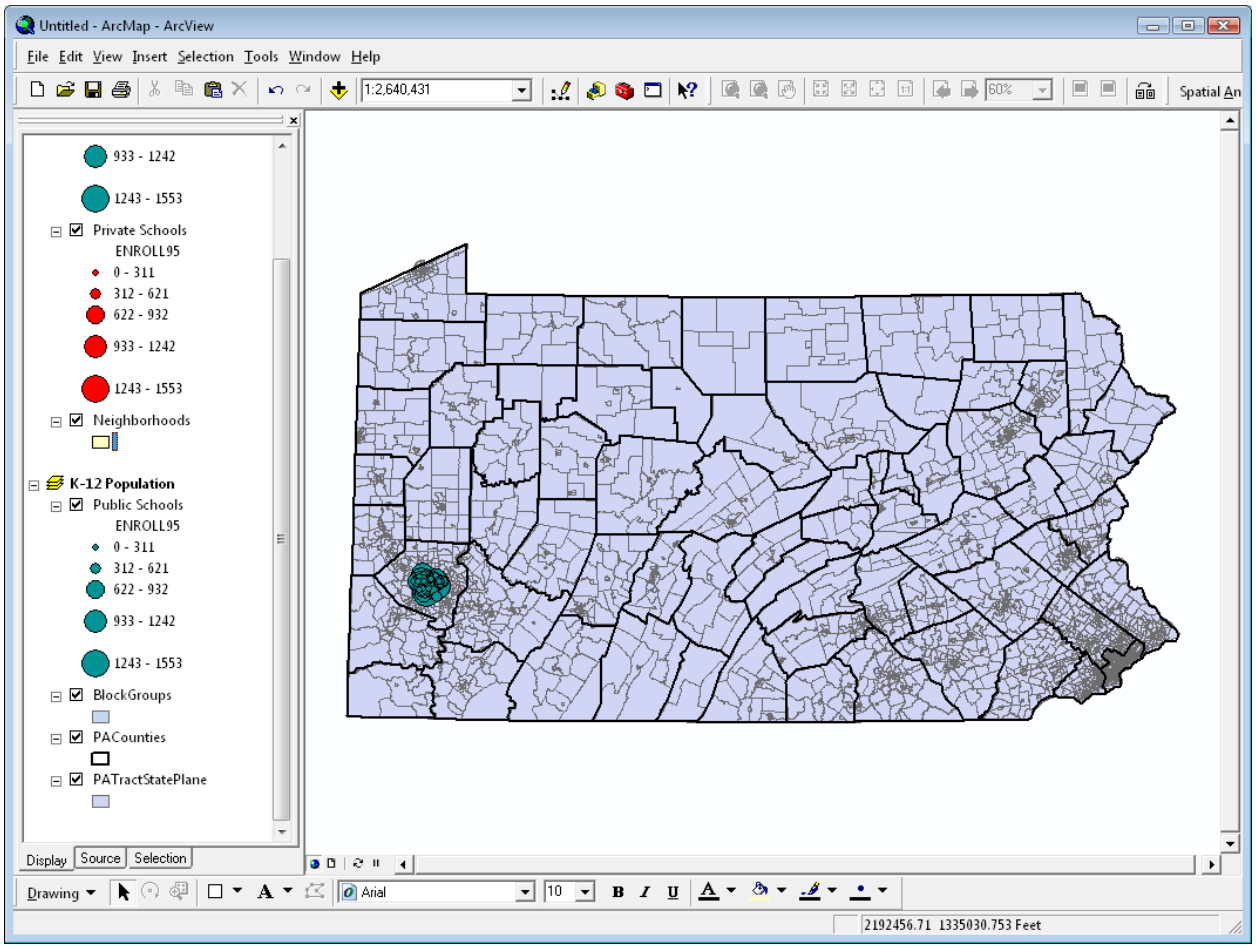
7. From the Insert Menu, click Data Frame. A New Data Frame will appear in the table of contents. Rename the current data frame as Pittsburg Schools, by right clicking on it and clicking on properties. Notice some of the other tabs available. We will not cover those in this workshop.



Part II: K-12 Population in the Commonwealth of Pennsylvania

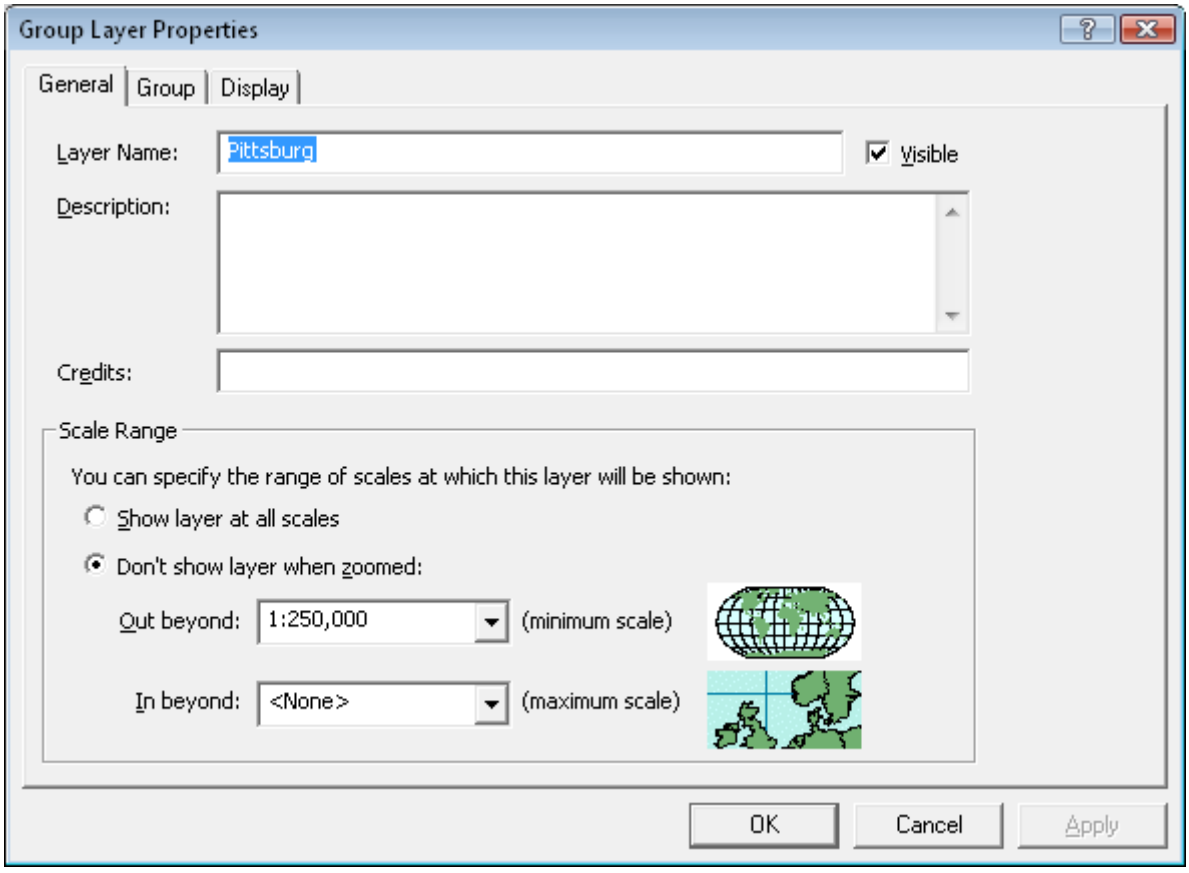
1. Rename the "New Data Frame" as "K-12 Population"

2. Add the Pennsylvania counties layer and the PATractStatePlane layer. Notice how the counties layer covers the census tracks layer. On the Symbology tab for the Counties layer, set the color as Hollow. Set the outline width as 1.5. Add in the BlockGroups layer and copy and paste the Public School layer from the above frame to K-12 Population frame.



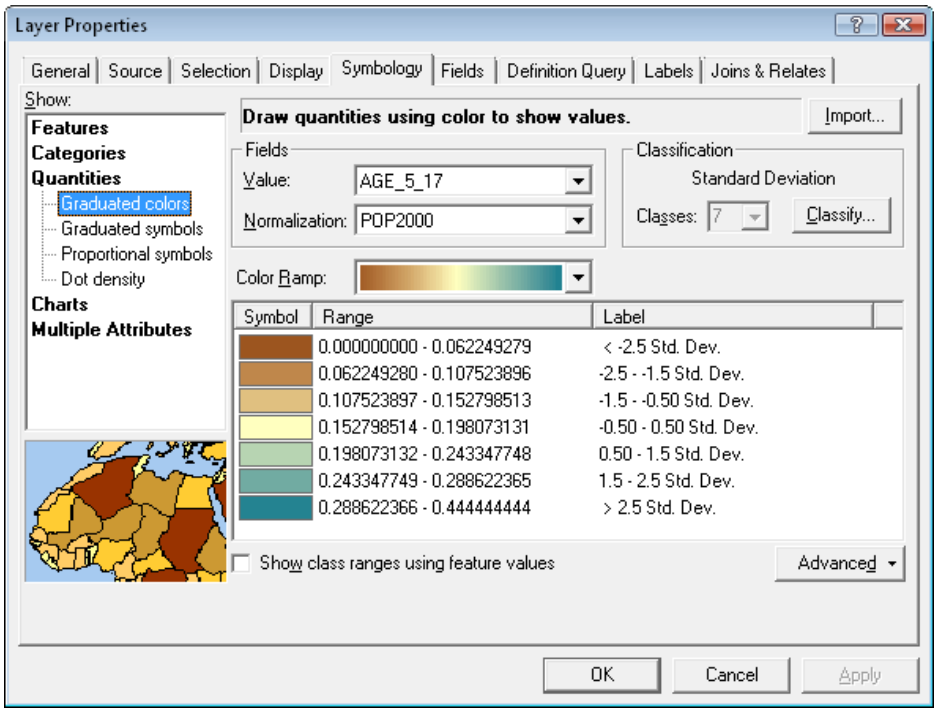
3. Adjust the Public Schools layer to only show schools with low enrollment (over 0 and less than 200). Use a definitional query. Change the symbology of this layer to a single feature label (use School 1).

4. Right Click on the K-12 Population frame and create two layer groups. One for Pennsylvania and one Pittsburg. Move the layers into the right groups. Rename these group. Notice the opton to not show the group when zoomed. Use 1:250,000 as the scale. Do the same for the Pennsylvania group, but use "In Beyond."

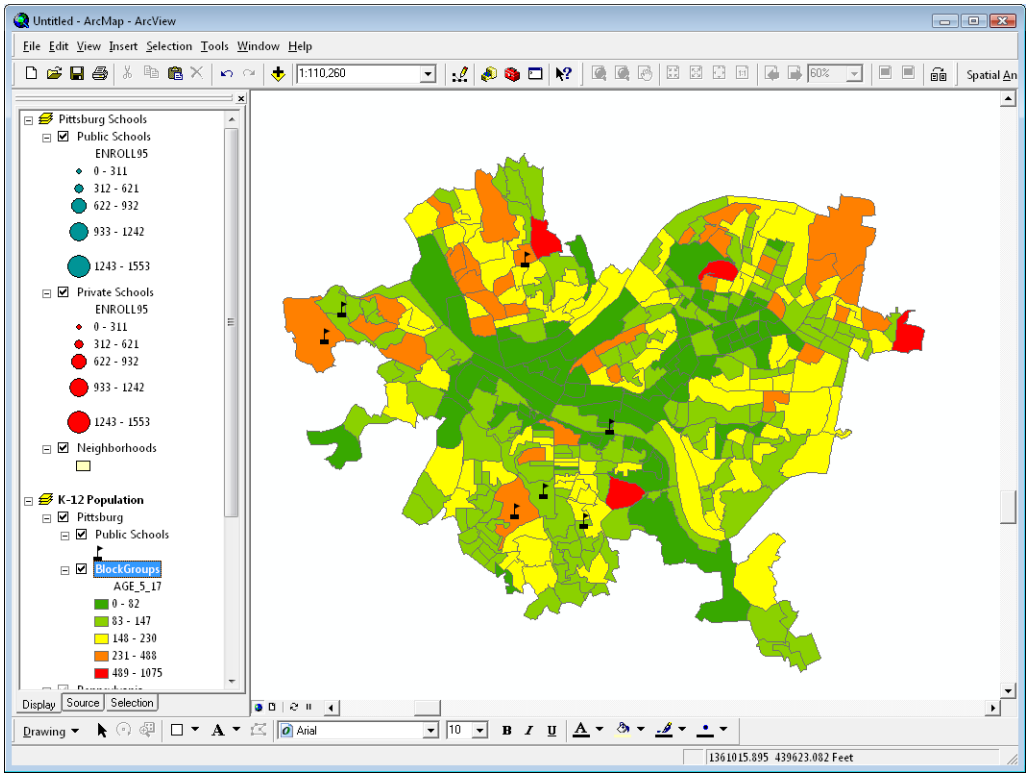


5. Zoom in on Pittsburg. Using View → Bookmark → Create Bookmark, create a bookmark named Pittsburg. Use the Globe icon to zoom out to the full extent. Now use the bookmark you just created to return to Pittsburg. You can create bookmarks for any areas you regularly work on.

6. Using the Symbology tab, create a choropleth map of K-12 school-age population normalized by 2000 population. Explore different options. How does this change how you interpret the map?



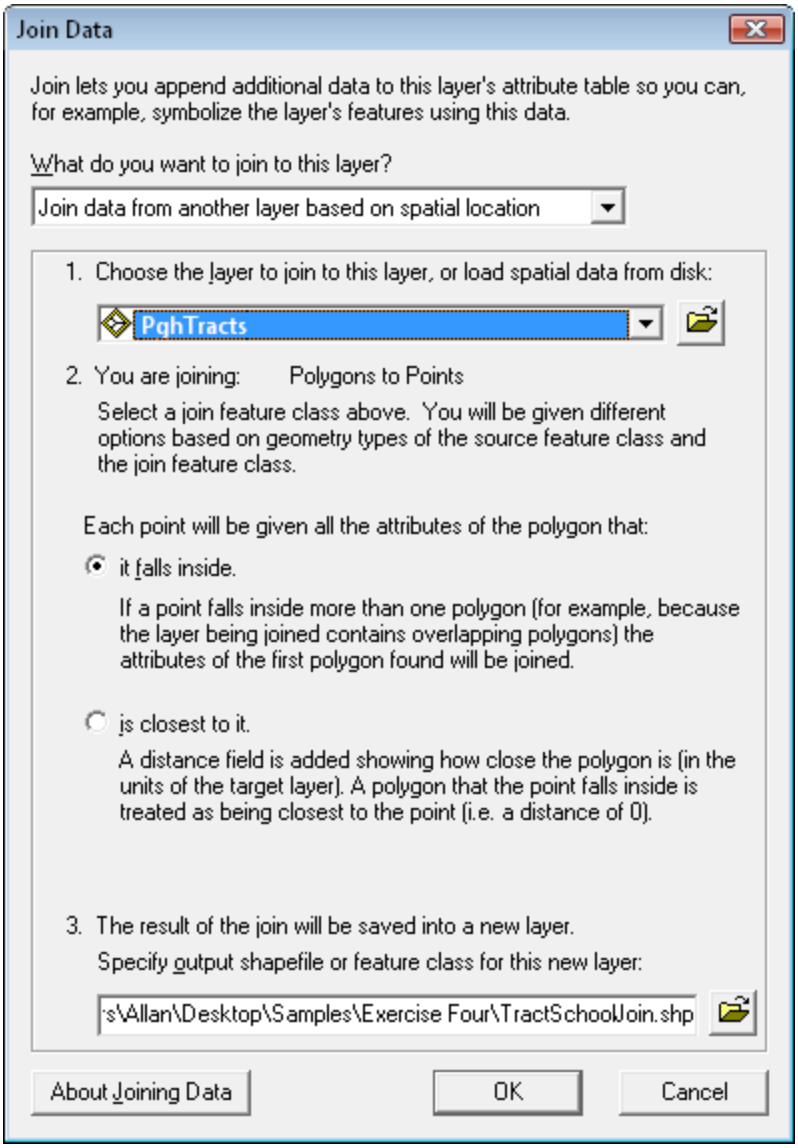
7. Create a choropleth map for the BlockGroups layer. Why might these schools have less than 200 students?



Part III: School Enrollment compared with School Age Population

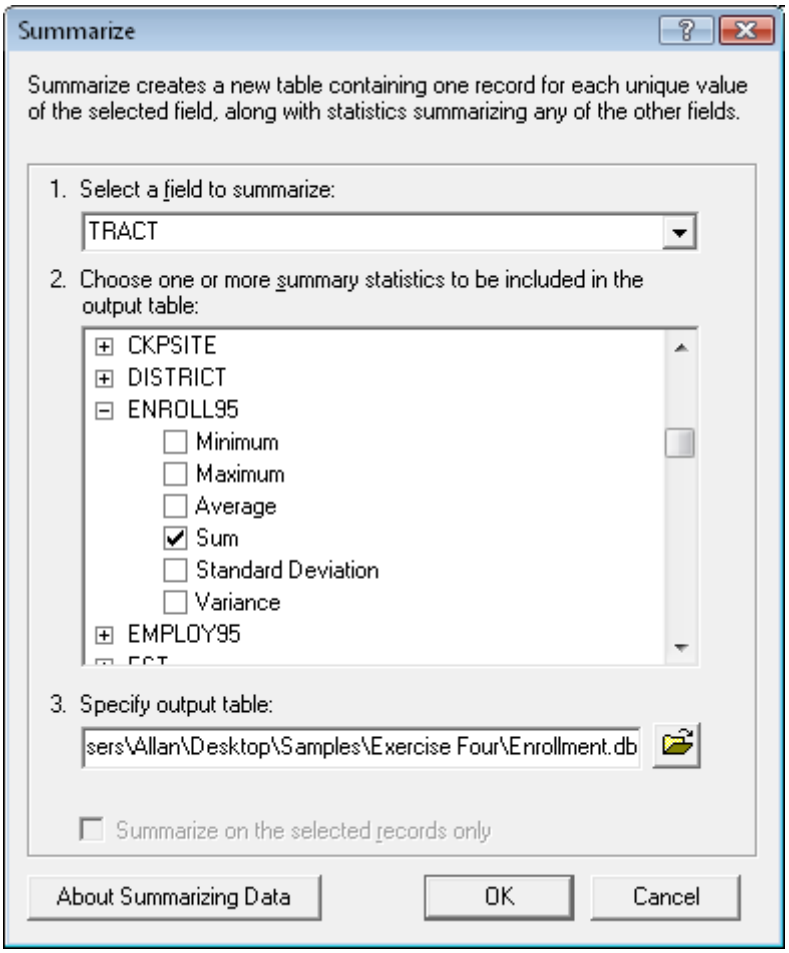
1. Create a new map in ArcMap and add a copy of the PghTracts and Schools layers.

2. Right-click on the Schools layer in the Table of Contents, click Joins and Relates, and click Join. In the Join Data dialog box, from the "What do you want to join to this layer?" drop-down list, chose Join data from another layer based on spatial location. Save it as a new layer called TractSchoolJoin.shp.

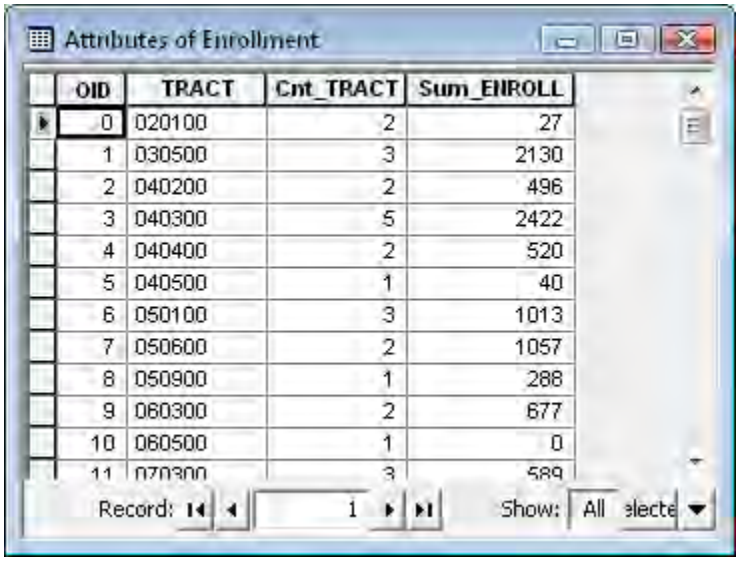


3. Examine the Attributes Table of the new layer. Notice how each school has the information from the PghTracts layer it belongs in.

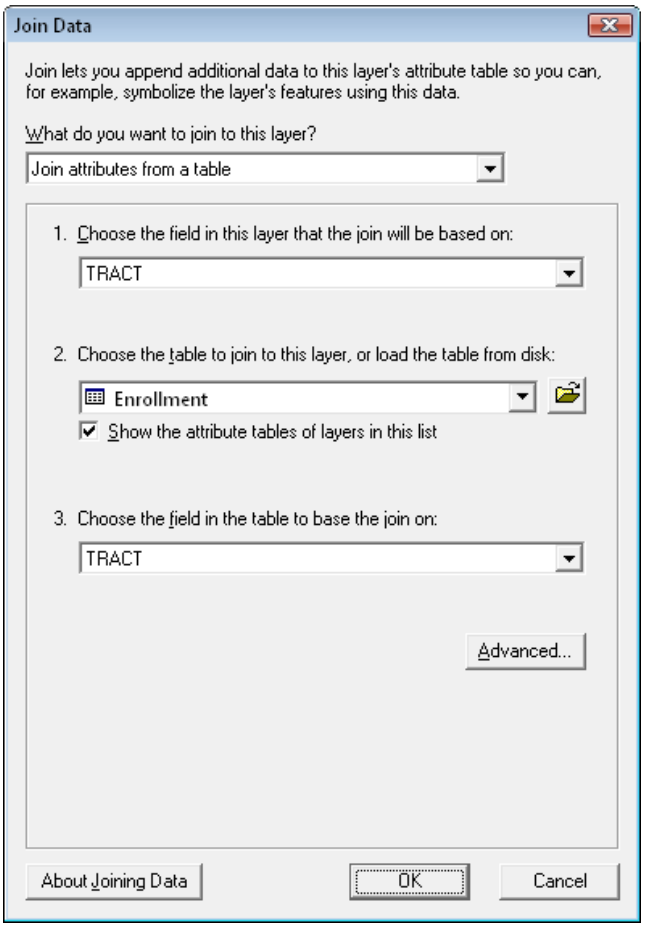
4. Scroll to Tract, right click on the column heading, and click Summarize. Change the Output Table name to Enrollment.dbf. Add the table to the map. Notice the change to the "Source" tab in the Table of Contents.

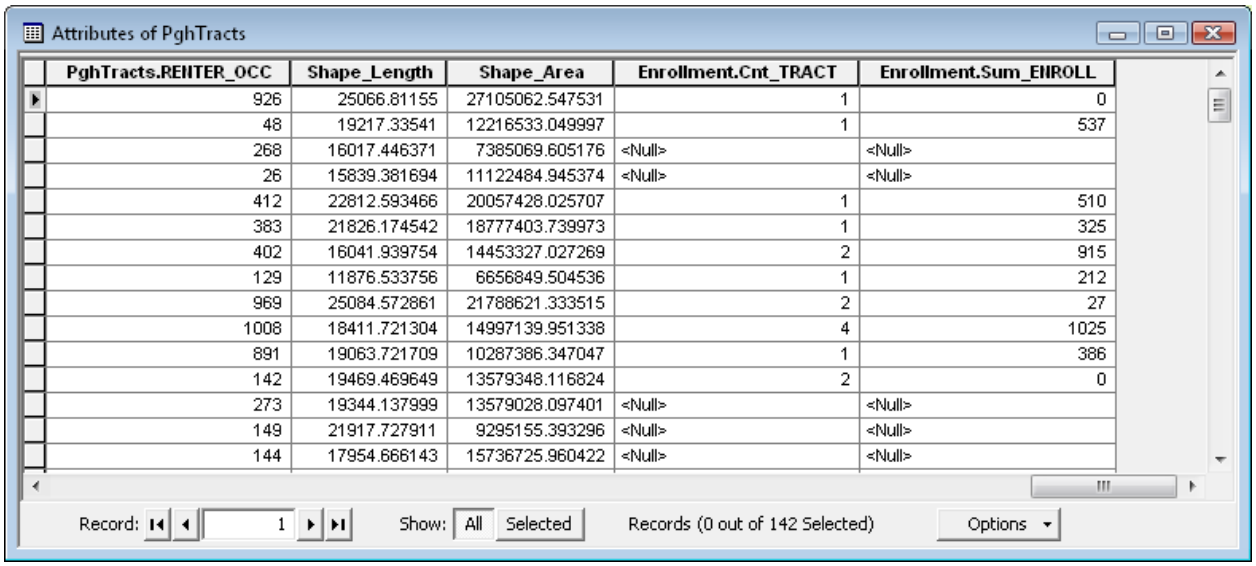


5. Open up the Enrollment table. Notice the Cnt_TRACT field. This is the number of schools in the tract.



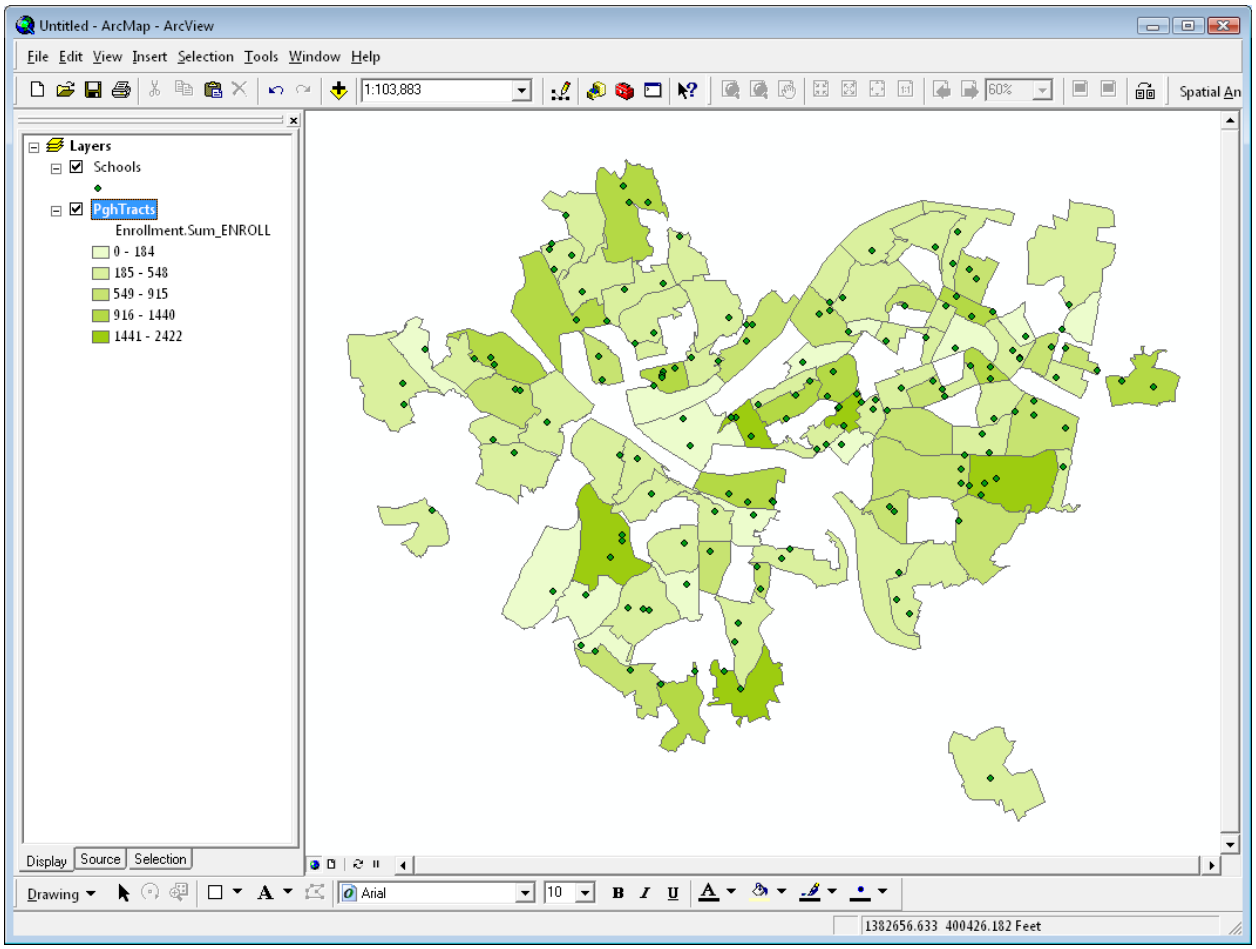
6. Join the count data to the tract data. Open up the Attribute Table for PghTracts. Notice how the fields from the Enrollment table are now at the end of the file. Some are null because they do not contain any schools.





PghTracts.RENTER_OCC	Shape_Length	Shape_Area	Enrollment.Cnt_TRACT	Enrollment.Sum_ENROLL
926	25066.81155	27105062.547531	1	0
48	19217.33541	12216533.049997	1	537
268	16017.446371	7385069.605176	<Null>	<Null>
26	15839.381694	11122484.945374	<Null>	<Null>
412	22812.593466	20057428.025707	1	510
383	21826.174542	18777403.739973	1	325
402	16041.939754	14453327.027269	2	915
129	11876.533756	6656849.504536	1	212
969	25084.572861	21788621.333515	2	27
1008	18411.721304	14997139.951338	4	1025
891	19063.721709	10287386.347047	1	386
142	19469.469649	13579348.116824	2	0
273	19344.137999	13579028.097401	<Null>	<Null>
149	21917.727911	9295155.393296	<Null>	<Null>
144	17954.666143	15736725.960422	<Null>	<Null>

7. Use the Symbology tab to create a choropleth map with the newly created data.



8. If time permits, use the layout view to create a printable map with a legend and title.

Additional Resources

Institutional Research and GIS

New Directions for Institutional Research, no. 120, Winter 2003

ArcGIS

Getting to Know ArcGIS Desktop, 2nd Edition, ESRI Press
GIS Tutorial: Workbook for ArcView 9, 2nd Edition, ESRI Press
The ESRI Guide to GIS Analysis, Volumes 1 and 2, ESRI Press
ArcGIS 9: Using ArcGIS Desktop, ESRI Press (included with software)

General Mapping & Design

Brewer, Cynthia. Designing Better Maps, ESRI Press
Monmonier, Mark. *Mapping it Out*
Monmonier, Mark. *How to Lie with Maps*
All of Edward Tufte's books, especially *Visual Explanations*.

Geographic Analysis

O'Sullivan, David, and David Unwin, *Geographic Information Analysis*

Websites

Color Brewer - http://www.personal.psu.edu/cab38/ColorBrewer/ColorBrewer_intro.html
Google Earth – http://earth.google.com (check out Keyhole Markup Langage, KML)