Online Supplement to *GIS and Public Health, 2nd Edition*
Ellen K. Cromley and Sara L. McLafferty

GIS AND PUBLIC HEALTH EXERCISE 3 - TYPES OF SPATIAL DATA (ArcGIS 9.3.1)

PREPARATION

Download the **exer3** folder you will need for this exercise from the online supplement.

All of the databases and files used in the exercise will be stored in various subfolders within the folder called **exer3**. The following instructions are written for this folder to be located on the **c:** drive. If the folder is located on another drive, the path names shown below should be modified accordingly. Some of the folders are empty. They have been included because you may need to save the results of an operation to one of these folders.

The map documents created using ArcGIS 9.3.1 reference the spatial databases and tables in the application based on the directories and paths where the data are stored. Changing the locations of databases in the system can prevent a GIS application from working properly.

Connecting to the Exercise Folder

Go to **Start** ⇒ **Programs** ⇒ **ArcGIS** ⇒ **ArcCatalog** to start ArcCatalog.

Find the button labeled **Connect to Folder** and click the button. Navigate to **c:\exer3** then click OK and look at the Catalog tree in the left window to see that the folder has been added.

Within the data folder, data can be organized in folders identifying the agency that produced the data and then by the format of the data. For these exercises, you will consider yourself to be working for the organization called “agency” that is creating the GIS.

As you work through the exercises, you will be retrieving data from and saving data to specific folders. Please make sure you understand the System Design for the exercises.

Use the **File** ⇒ **Exit** menu to close ArcCatalog.

TYPES OF SPATIAL DATA

Go to **Start** ⇒ **Programs** ⇒ **ArcGIS** ⇒ **ArcMap** to start ArcMap.

In the “ArcMap Start using ArcMap with” window, click the radio button labeled “A new empty map” and then click OK.

Rename the Layers data frame by right clicking the word Layers and selecting the **Properties** item in the menu. Then select the **General** tab and enter the name Spatial Data. Click OK. The name of the Data Frame in the Table of Contents window should now appear as Spatial Data.

Add Vector Spatial Data

To begin, add a shapefile of acute care hospitals in Massachusetts distributed by MassGIS. MassGIS is the Commonwealth of Massachusetts’ statewide resource for geospatial technology and data. The hospitals database and other geographic data can be downloaded from the MassGIS web site. Find the button labeled **Add Data** and click the button. You should find the **c:\exer3** folder in your catalog. If not, please connect to the folder using the **Connect to Folder** button.

Navigate to **c:\exer3\data\massgis\shapes** and add the **HOSPITALS_PT.shp** shapefile. Right click on the Spatial Data Frame (not the file) and select the **Properties** item in the menu. Click the **General** tab and note that the map units are meters then click the **Coordinate System** tab and note that the Coordinate System is now:

NAD_1983_StatePlane_Massachusetts_Mainland_FIPS_2001

Click OK to close the "Data Frame Properties" window.

Use the **Save** button or go to **File ⇒ Save** to save your map document. Navigate to **c:\exer3\mapdocs** and save the file as **exer3.mxd**.

The hospitals shapefile is an example of vector data, specifically, point data. Each hospital is represented by a point. Vector data can also use lines or polygons to represent features.

Use the **Add Data** button and navigate to **c:\exer3\data\massgis\shapes** and add:

biketrails_arc.shp a shapefile of lines representing trails for bicycling and other physical activity

and

regemsp1.shp a shapefile of polygons representing emergency medical service (EMS) regions in the state. If the "Unknown Spatial Reference" window opens to warn you that the data source is missing spatial reference information, click OK. For certain types of processing, you might need to define the spatial reference of the data before proceeding.

Now use the **Add Data** button and add:

regemsa1.shp a shapefile using lines to represent the boundaries of EMS regions in the state.

Compare the regemsp1 and regemspa1 shapefiles by turning their visibility on and off and by exploring their attribute tables. Areas can be represented as polygons or as lines forming the boundaries of the areas. Both types of representation may be useful in a GIS application.

Notice in the Table of Contents that the default draw order of point, line, and area shapefiles is set so that lines will be drawn on top of areas, and points will be drawn on top of lines and areas. It is possible to change the draw order. Click the **Display** tab of the Table of Contents, then click and hold the name of the file you wish to draw in a different order and drag the file name to the desired position in the Table of Contents.

Use the **Save** button or go to **File ⇒ Save** to save your map document.

Vector data from other agencies can be combined with the data from MassGIS.

Use the **Add Data** button and navigate to **c:\exer3\data\census\shapes** and add:

tl_2009_25_county.shp

This is a shapefile from the U.S. Census Bureau of Massachusetts counties represented as polygons. Note that the county boundaries extend offshore. The shapefile of EMS regions from MassGIS shows the coastline. Use the panning and zooming tools to explore other differences in the representations of the state boundary in the two databases. Then click the Full Extent button to display the full extent of the data in the Data View.

Right click the tl_2009_25_county.shp shapefile and select **Properties**. Click the **Source** tab and see that the spatial reference for the Census data is GCS_North_American_1983. The GIS

software has performed an on-the-fly projection to display the Census data. It has not changed the underlying spatial reference of the shapefile. For some types of processing, it might be necessary to project the Census data to the Massachusetts State Plane Coordinate System. Use the **Save** button or go to **File ⇒ Save** to save your map document.

Add Raster Spatial Data

In addition to vector data, health applications of GIS often incorporate raster data.

Use the **Add Data** button and navigate to **c:\exer3\data\massgis\raster** and add:

imp_se5.img

If you receive a prompt asking you whether or not you want to create pyramids for more rapid display, click Yes. Wait while the pyramids are built and then look at the database when it is added to the display.

This database shows areas of impervious surface – surfaces water cannot penetrate such as paved areas, roofs of building, and human-compacted soil – which are important in many environmental studies. Surface water areas, vegetated areas, and naturally occurring barren areas are not considered impervious surfaces.

This database was developed by using semi-automated image classification techniques to classify areas in 50 cm Vexel UltraCam near infrared orthoimagery acquired in April, 2005. The resolution of the imp_se5.img raster is 1 meter.

The data are in grid format. Use **Help** to find and display information about working with grids in ArcGIS.

Note that the default draw order places raster data below vector point, line, and area data.

Because raster databases are so large, they are often distributed in tiles and users must consult an index to discover which tile or tiles they need for their analyses. This tile covers the area near Fall River, Massachusetts.

Right click the imp_se5.img file and select **Zoom to Layer** to zoom the display to the extent of the raster image. Turn off the visibility of the tl_2009_25_county.shp and the regemsp1.shp shapefiles so that you can see the entire impervious surface raster.

It appears that there are some areas of impervious surface lying outside the state of Massachusetts. This tile covers part of the neighboring state of Rhode Island including land and surface water.

In the default symbolization, the color ramp shows areas of impervious surface in a lighter shade and areas of vegetation in a darker shade. Double click the color ramp in the legend for the imp_se5.img raster and check **Invert** to make the areas of impervious surface dark and the areas of vegetation light. The click OK.

Use the **Map Scale** window in the interface to zoom to a scale of 1:250 so that you can see the pixels. Pan and find an area where there is a single pixel of impervious surface.

Click the **Measure** tool to measure the distance along one edge of the pixel. Click the **Measure Line** button in the “Measure” window interface to depress it. Then use the pull-down **Choose Units** menu in the window to select **Distance** and to check **Meters** as the measurement unit. Then measure the length or width of the pixel by clicking at one corner of the pixel, dragging the

cursor, and clicking at the other corner along the length or width. This should be approximately 1 meter. Double click to stop measuring. Then close the “Measure” window.

Use the **Go Back to Previous Extent** button to return to the extent of the raster image.

Add Orthophoto Data

The impervious surface data were classified so the raster does not look like a photograph. Use the **Add Data** button and navigate to **c:\exer3\data\massgis\orthophoto** and add:

22828270.sid

This orthophotograph is in MrSID format. Make sure that you add the 22828270.sid file. You can also add individual Red, Green, and Blue bands. In this case, add the single SID file to add all bands for the entire orthophotograph. You should see a single data layer named 22828270.sid with three bands in the Table of Contents when you add the orthophotograph.

The orthophoto should draw on top of the raster grid. If it does not, change the draw order in the Table of Contents so that you can see the orthophoto. The orthophoto images cover smaller areas than the impervious surface grids. Right click the orthophoto and select **Zoom to Layer** from the menu to zoom to the orthophoto extent.

Use the **Map Scale** tool to zoom to a scale of 1:1000. Pan and zoom and compare the raster and orthophoto image. Turn the visibility of the orthophoto on and off to compare the two.

Save the map document.

Add Cadastral Data

Use the **Add Data** button and navigate to **c:\exer3\data\massgis\cadaster** and add:

par95.shp

This is a shapefile of property parcels for the town of Fall River. It is a polygon shapefile. It should draw over the orthophoto and raster data.

Double click the rectangle in the legend for the par95 shapefile to open the “Symbol Selector“ window. Click the Hollow symbol and then increase the outline width to 2 using the arrows in the Options box to the right. Then change the Outline Color to Solar Yellow. Click OK to change the symbolization of the parcel layer so that you can see through the polygons.

Compare the parcel boundaries to the buildings you see in the photo.

Right click the par95.shp and select **Open Attribute Table** from the menu. Look at the attribute information for parcels. Then close the table.

CREATE A NETWORK DATABASE

The final type of data to explore is network data. You will first create a shapefile of streets.

Create a Shapefile of Streets

Use the **Add Data** button and navigate to **c:\exer3\data\census\shapes** and add:

tl_2009_25005_edges.shp

This is a shapefile of line segments for Bristol County, Massachusetts, from the 2009 TIGER® files distributed by the U.S. Census Bureau. The state code for Massachusetts is 25 and the county code for Bristol County is 005. The shapefile contains lines representing street segments and other linear features. To create a street network database, first select the lines representing streets and then save the data to a new shapefile.

Right click tl_2009_25005_edges.shp and select **Open Attribute Table** to open the attribute table.

Press the Options button at the bottom of the table and select **Select by Attributes** from the menu.

With the Method set to “Create a new selection” scroll the window below until you find the variable “ROADFLG” in the list.

Double click “ROADFLG” to add it to the SELECT box below then click the button for the “=” sign to add the equal sign after “ROADFLG” then click the “Get Unique Values” button to display a list of all the values in the ROADFLG field. Double click ‘Y’ to select all of the lines which are flagged as road features.

When the query string in the box looks like the string below,

"ROADFLG" = 'Y'

click Apply and Close to execute the query. You should have 44,331 records selected out of 71,374 records in the shapefile. The selected street segments should be highlighted in the table and in the map display.

Close the attribute table.

Right click the tl_2009_25005_edges.shp data layer and select **Data ⇒ Export Data** from the menu. Make sure that you are exporting Selected features in the “Export:” window and click the radio button for exporting the selected records to a new database using the same coordinate system as “the data frame” so that the street data will be projected using Massachusetts State Plane Coordinates.

Output the new shapefile to:

c:\exer3\data\agency\shapes\streets.shp

and click OK.

In the message box asking you if you want to add the exported data to the map as a data layer, click Yes. Right click the streets.shp shapefile and select **Properties** from the menu. Click the **Source** tab and check to make sure the data are in Massachusetts State Plane Coordinates.

Turn off the visibility of all the other data layers in the display except tl_2009_25005_edges.shp and streets.shp. Then use the **Selection ⇒ Clear Selected Features** menu to clear the selection.

Pan and zoom to compare the two layers. Open the attribute tables for the two shapefiles and compare them to see that the attributes are the same for both databases. Then close the attribute tables. Remember, you should update the metadata for streets.shp because you have processed the data, including changing the map projection.

Turn on the visibility of the orthophoto layer and click on the line segment in the legend for the streets.shp layer in the Table of Contents to open the "Symbol Selector" window. Change the color of the line segments in the streets.shp layer to bright yellow or orange. Click OK. Look at the street segments in relation to the orthophoto view of the landscape to see how well the street database matches up with the orthophoto showing the roads.

Turn off the visibility of all layers except streets.shp.

Save the map document.

Add a Field for the Length of the Street Segment

Right click the streets.shp shapefile and select **Open Attribute Table** from the menu. Click the Options button and select **Add Field** from the menu.

In the "Add Field" window, enter the name LENGTH in the "Name:" window. Use the pull-down to select Double as the field type. This means the LENGTH field will be a double precision numeric field. In the "Field Properties" box, change the precision (the number of significant digits) to 10 and the scale (the number of decimal places) to 1. Then click OK.

Scroll to the right of the attribute table to see that the LENGTH field has been added. The value of the field should be "0" for every record.

Right click the field header LENGTH and select **Calculate Geometry** from the menu. You will receive a warning that the calculation you make cannot be undone. Click Yes to continue.

Select Length as the Property you wish to calculate and check the radio button to Use the coordinate system of "the data source." Set the Units to Meters [m]. Then click OK.

Based on data stored in the SHAPE field, the GIS software will calculate the length of the street segment in meters to a precision of 0.1 m and add the value to the record in the LENGTH field. Close the attribute table.

Save the map document.

Enable the Network Analyst Extension and Create a Network Database

You can use the streets.shp shapefile and the Network Analyst extension of ArcGIS to create a network database. If you do not have the Network Analyst extension, read through this section to learn how you can use it to create a network database. At the end of the section, you will be directed to add a network database already built and included with the data for this exercise.

Use the **Tools ⇒ Extensions** menu and check the box to enable the Network Analyst extension. Then click Close.

Enabling the extension does not automatically add the Network Analyst toolbar to the user interface. Use the **View ⇒ Toolbars** menu and check Network Analyst to add the Network Analyst toolbar to the user interface.

Then **Save** the map document.

Open **ArcCatalog**. If you have not already done so, connect to the Network Analyst extension in ArcCatalog, too, using the **Tools ⇒ Extensions** menu.

If you are creating a shapefile-based network dataset, navigate to the shapefile location in the Catalog tree. In this case, navigate to **c:\exer3\data\agency\shapes** because we are creating a shapefile-based network.

Right-click the feature dataset **streets.shp** and select **New Network Dataset** from the menu.

Enter a name for the network dataset as streets_ND. Click Next.

Accept the default connectivity settings or click Connectivity and make changes to the connectivity settings. In this case, accept the defaults. Click Next.

We do not have elevation field data or elevation field settings to modify the connectivity so click the radio button “No” and then click Next.

Click the radio button for “No” if you are not using turns. For this exercise, click the radio button for “Yes” and check the Turn Sources check box to use Global Turns. Click Next.

Specify the attributes for the new network dataset. Click the Add button and enter LENGTH as the Name of the field. Select Cost as the Usage Type because the LENGTH field will be used to determine the distance as a measure of the cost of traversing the street segment. Select Meters as the Units from the pull-down list and Double as the Data Type. Then click OK. After the LENGTH field has been added to the window, click Next.

Click the radio button for “No” because you do not want to establish driving directions. Then click Next.

Review the New Network Dataset Summary. Click Finish.

Click Yes to build the network dataset. Optionally, click No if you want to build the network dataset at another time. Click Yes and wait while the network builds.

You should see that two files have been created in the folder with the streets.shp shapefile: streets_ND.nd (a Shapefile Network Dataset) and streets_ND_Junctions.shp (a shapefile of points which are the nodes in the network).

Close ArcCatalog and restore ArcMap.

View the Street Network Database

Use the **Add Data** button and navigate to **c:\exer3\data\agency\shapes** and add the **streets_ND.nd** file to the view. When you are asked if you want to add all of the files, click Yes. This will add another copy of the streets.shp shapefile used to create the network, the network database, and the shapefile of nodes. If you want to add just the network edges and not all of the network nodes and another copy of the streets.shp shapefile to the Data Frame, you can click No.

You should see that the Network Analyst tools are enabled when a network database is added to the Data Frame.

If you do not have the Network Analyst extension, use the **Add Data** button and navigate to:

c:\exer3\data\agency\network

and add the **streets_ND.nd** network. When you are asked if you want to add all of the files, click Yes. This will add another copy of the streets.shp shapefile used to create the network, the network database, and the shapefile of nodes.

Compare the streets.shp shapefile and the streets_ND.nd network. The line segments are identical but the databases are different. Right click the streets_ND.nd network and select **Properties**. Notice that there is no Attribute table to open. Then close the "Properties" window.

Use the **Save** button to save the map document.

DISSOLVE, SELECT BY LOCATION, CLIP, AND MASK

In the exercises so far, you have created new data from existing GIS databases by changing the map projection, selecting features by attribute and exporting selected features, and creating a network database. In the last section of this exercise, you will use several other techniques for creating spatial databases.

First, turn on the visibility of the HOSPITALS_PT.shp, biketrails_arc.shp, and regemsp1.shp shapefiles.

Dissolve

You can use dissolve to create a shapefile of the state. First, you must create a "dissolve" field.

Right click the regemsp1.shp shapefile and select **Open Attribute Table** from the menu.

Click the Options button and select **Add Field** from the menu. In the "Add Field" window, enter STATE as the Name, select Text as the Type from the pull-down list, and set the Length of the field to 5 in the Field Properties window. Then click OK.

Right click the STATE field header in the table and select **Field Calculator** from the menu. In the window, set the value of the field equal to "MA" (use the double quotes when you enter the text string). Click OK. Then click Yes to acknowledge that you will not be able to undo this edit. Close the attribute table.

Click the **Show/Hide ArcToolbox Window** button to open Arctoolbox. Double click **Data Management Tools ⇒ Generalization ⇒ Dissolve** to open the "Dissolve" window. Use the button to select regemsp1.shp as the Input Features to be dissolved.

Set the path and file name for the Output Feature Class to:

c:\exer3\data\agency\shapes\regemsp1_Dissolve.shp

Check the box in front of STATE in the list of fields to use that field as the dissolve field. Because the state includes islands, check the box to include multipart features. In some cases, you might wish to process some of the attribute fields like the area of the polygons, too, so that the areas of the individual polygons will be summed to the area of the state and added as an attribute to the new shapefile. For now, just click OK. Close the "Dissolve" window after the dissolve operation is completed successfully.

The new shapefile is automatically added to the Data Frame. Once you have looked at the new layer, turn its visibility off.

Select by Location

You can select by location using the **Select Features** tool to select features manually or you can use the **Selection ⇒ Select by Location** menu.

Click the **Select Features** tool in the interface to depress it and then click inside the polygon representing part of Region 5, the South Eastern Massachusetts Regional EMS Council region, in

the seemsp1 layer. This is the region where Fall River is located. The boundary of the region you selected is highlighted. Right click the regemsp1.shp shapefile and select **Open Attribute Table** from the menu. Then press the Select button at the bottom to show just the record for the selected polygon. You should have the record with REG_EMS_ value of 110 selected. Close the table.

Right click the regemsp1.shp shapefile and select **Data ⇒ Export Data** to create a new database including just the selected region. Make sure you are exporting Selected features from the pull-down list and that you are Using the same coordinate system as "this layer's source data" and save the Output shapefile in:

c:\exer3\data\agency\shapes\seemsp1.shp

then click Yes to add the exported data to the Data Frame.

Use the **Selection ⇒ Clear Selected Features** menu to clear all selections. Turn off the visibility of the regemsp1.shp shapefile. Right click the seemsp1.shp shapefile and select **Zoom to Layer** from the menu.

Save the map document.

Next, we will select all of the hospitals located within the Region 5 area. Use the **Selection ⇒ Select by Location** menu. You want to select features from the HOSPITALS_PT layer so check the box in front of that layer in the window. Use the pull-down list to select the hospital features that "are completely within" the features in the "seemsp1" layer. Click Apply and Close. You should see that you have used the Region 5 area to select 9 hospitals located within it.

Right click the HOSPITALS_PT.shp shapefile and select **Data ⇒ Export Data** to create a new database including just the selected hospitals. Make sure you are exporting Selected features Using the same coordinate system as "this layer's source data" and save the Output shapefile in:

c:\exer3\data\agency\shapes\reg5hosp.shp

then click Yes to add the exported data to the Data Frame.

Use the **Selection ⇒ Clear Selected Features** menu to clear all selections. Turn off the visibility of the HOSPITALS_PT layer.

Save the map document.

Clip to Create a New Database

Now look at the seemsp1 layer and the biketrails_arc layer. Some of the trails cross the boundary of Region 5 and run into adjacent regions. To create a new database containing only the portions of the trails located in Region 5, use the **Clip** function in ArcToolbox.

Click the **Show/Hide ArcToolbox Window** button to open Arctoolbox. Double click **Analysis Tool ⇒ Extract ⇒ Clip** to open the "Clip" window. Click the button next to Input Features, navigate to the correct folder, and use the biketrails_arc layer as the input feature to be clipped using the Add button. Then click the button next to Clip Features, navigate to the correct folder, and use the seemsp1 layer as the layer whose features will be used to clip the trails. The Output Feature Class Name should be:

c:\exer3\data\agency\shapes\biketrails_arc_Clip.shp

Then click OK.

After the operation has been performed successfully, close the “Clip” window. The clipped shapefile has been automatically added to the Data Frame. Turn off the visibility of the biketrails_arc layer.

Now right click the biketrails_arc_Clip layer and select **Open Attribute Table**. Scroll to the right and notice that there is a field storing the length of each segment. When you clipped the trails, this field was not updated. You can update the length by right clicking on the SHAPE_LEN field and selecting **Calculate Geometry** from the menu. Calculate the Length of each segment in Meters [m] to update the field. Then close the table.

Mask a Raster Database

You can also use software functions to extract part of a raster database. Use the seemsp1.shp to clip the impervious surface raster.

Use the **Tools ⇒ Extensions** menu and check the box to enable the Spatial Analyst extension. Then click Close. If you do not have this extension, read through the exercise until you find instructions for what to do if you do not have the Spatial Analyst extension.

Enabling the extension does not automatically add the Spatial Analyst toolbar to the user interface. Use the **View ⇒ Toolbars** menu and check Spatial Analyst to add the Spatial Analyst toolbar to the user interface.

Then **Save** the map document.

In ArcToolbox, double click **Spatial Analyst Tools ⇒ Extraction ⇒ Extract by Mask**.

In the Extract by Mask window, click the button and select imp_se5.img as the Input raster.

Then click the button and select seemsp1.shp as the Input raster or feature mask data.

Click the button and set the path for the output file to:

c:\exer3\data\agency\raster\Extract_imp_1

Then click OK. This process will take about 6 minutes. The masked raster will automatically be added to the Data Frame. Turn off the visibility of imp_se5.img. Click the polygon in the legend for the seemsp1 layer in the Table of Contents and change the symbol to Hollow in the Symbol Selector. Then zoom in to see that the raster of impervious surface has been clipped to the EMS region boundary.

If you do not have the Spatial Analyst extension, use the **Add Data** button and navigate to:

c:\exer3\data\agency\masked

and add the **extract_imp_1** image file for which pyramids have already been built.

These are only a few of the many ways you can create new data from existing data in a GIS application. Make sure that you understand how the data are changing as you process them and whether or not attributes need to be updated. Also, remember to update metadata for the new databases.

Click the **Save** button and then use the **File ⇒ Close** menu to close the map document.