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

GIS AND PUBLIC HEALTH EXERCISE 7 – DEFINING NEIGHBORHOODS (ArcGIS 10)

PREPARATION

Download the **exer7** 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 **exer7**. 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 ArcGIS10 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 10** to start ArcCatalog.

Find the button labeled **Connect to Folder** and click the button. Navigate to **c:\exer7** 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.

PREPARING TO DEFINE NEIGHBORHOODS

Go to **Start** ⇒ **Programs** ⇒ **ArcGIS** ⇒ **ArcMap 10** to start ArcMap.

In the “ArcMap – Getting Started” window, close the window you would use to open an existing map document or make a new map using a template.

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 Neighborhoods. Click OK. The name of the Data Frame in the Table of Contents window should now appear as Neighborhoods.

Add a Database of County Subdivisions

To begin, add a database of county subdivisions in New Hampshire. The database was downloaded from the U.S. Census Bureau TIGER® web site. Initially, a shapefile of New Hampshire county subdivisions from the 2010 TIGER database was downloaded. Two county subdivisions along the coast covering water areas were deleted. Then, the data were projected using NAD_1983_StatePlane_New_Hampshire_FIPS_2800 coordinates. Map units are meters.

County subdivisions are the primary divisions of counties and equivalent features for the reporting of Census Bureau data. This feature includes the primary governmental or administrative divisions of a county and other independent places. In New Hampshire, towns are the primary governmental divisions of a county. Townships, grants, locations, and purchases in New Hampshire are county subdivisions which are not part of any town and have limited or no self-government. Many of these entities are uninhabited. There are 259 county subdivisions in New Hampshire.

A table of data was prepared from data downloaded from the Census Bureau's American FactFinder web site for the purposes of this exercise. The data are from the 2010 Redistricting Data Summary File (P.L. 94-171), among the first data released from the 2010 census. The data in this table provide information on the total population within different racial groups by county subdivision. These data were joined to the shapefile of county subdivisions.

Find the button labeled **Add Data** and click the button. You should find the **c:\exer7** folder in your catalog. If not, please connect to the folder using the **Connect to Folder** button. Navigate to **c:\exer7\data\agency\shapes** and add the **tl_2010_33_cousub10_Project_Pop.shp** shapefile.

Right-click the tl_2010_33_cousub10_Project_Pop data layer and select **Open Attribute Table** from the menu. Scroll to the right and explore the data fields. The state FIPS code for New Hampshire is 33. Each county also has a FIPS code. FIPS and ANSI county subdivision codes are also provided. The GEOID10 field concatenates the state, county, and county subdivision FIPS codes. Note that this is a text field.

The table also contains fields (ALAND10 and AWATER10) with data on the surface land and surface water areas in each county subdivision in square meters.

For the 2010 Census, individuals could identify themselves with one or more of six racial groups: White, Black or African American, American Indian and Alaska Native, Asian, Native Hawaiian and Other Pacific Islander, or Other. For this exercise, we will be using data on the total population (D001), the total population identifying themselves as being of one race alone (D002 through D008), and the total population identifying themselves as two or more races (D009). The values in D002 and D009 sum to the value in D001.

Field Name	Field Type	Description
D001	Numeric	Total population
D002	Numeric	Total: Population of one race
D003	Numeric	Total: Population of one race: White alone
D004	Numeric	Total: Population of one race: Black or African American alone
D005	Numeric	Total: Population of one race: American Indian and Alaska Native alone
D006	Numeric	Total: Population of one race: Asian alone
D007	Numeric	Total: Population of one race: Native Hawaiian and Other Pacific Islander alone
D008	Numeric	Total: Population of one race: Some Other Race alone
D009	Numeric	Total: Two or More races

Close the table.

Use the **Save** button or go to **File ⇒ Save** to save your map document. Navigate to **c:\exer7\mapdocs** and save the file as **exer7.mxd**.

Add a Shapefile of Air Quality Monitoring Station Locations

Next, add a shapefile of air quality monitoring stations for ozone in New Hampshire. This database was developed from data on criteria pollutant networks downloadable from the Map

Monitoring Sites page of the U.S. EPA's AirExplorer site. The active sites in New Hampshire in the ozone monitoring network were selected for this exercise.

Use the **Add Data** button and navigate to **c:\exer7\data\agency\shapes** to add the **nhaqsites_Project.shp** shapefile.

Right click the shapefile and select **Open Attribute Table** from the menu. Explore the fields. In addition to AQSID, the air quality station identifier, there is a field for the name of the station, the county and the town where the station is located, the datum, the longitude, the latitude, the elevation in meters, and the year established.

Close the table.

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

Defining neighborhoods and using them in spatial analyses are important GIS functions. Next, you will define neighborhoods around the air quality monitoring stations using two different methods. First, you will define neighborhoods using buffers. Second, you will define neighborhoods using Thiessen polygons.

Once the neighborhoods have been defined, you will use the neighborhoods defined by buffering to interpolate the population within the neighborhood of each monitoring station using area interpolation methods.

DEFINING NEIGHBORHOODS

Defining Neighborhoods Using Buffers

Click the **Geoprocessing ⇒ Buffer** menu to open the "Buffer" window.

Use the Browse button and select nhaqsites_Project as the "Input Features".

Use the Browse button and specify the "Output Feature Class" as:

c:\exer7\data\agency\shapes\Buffer_of_nhaqsites_Project.shp

Under "Distance [value or field]" click the radio button for "Linear unit" then enter 25 as the number of units and select Kilometers from the pull-down menu of units.

Set the "Dissolve Type (optional)" to ALL so that neighborhoods around each air quality site will be merged.

Click OK to buffer the sites.

Close the "Buffer" window.

The new shapefile of buffer areas will be added to the Data Frame automatically.

There were 12 air quality monitoring stations and you should see 1 buffer area because you chose to dissolve ALL of the overlapping buffer zones. Right click the Buffer_of_nhaqsites_Project.shp layer and select **Open Attribute Table** to verify this.

The table has 1 attribute field, Id.

Edit the attribute table to assign unique identifiers to each buffer zone. Use the **Customize ⇒ Toolbars** menu to open the list of Toolbars and check the **Editor** toolbar if it has not already been checked. This will add the **Editor** toolbar to the interface.

Select **Editor ⇒ Start Editing** to open an editing session.

Make sure the **Buffer_of_nhaqsites_Project** is highlighted in the "Create Features" window so that you are editing the correct layer.

Right click the **Buffer_of_nhaqsites_Project** layer and select **Open Attribute Table** from the menu to open the attribute table.

Click in the Id field for the first record in table (the record Id of 0), and edit the Id to be 1. Then click any place in the blank record below to register the edit.

Use the **Editor ⇒ Save Edits** menu to save your edits. When you see that the edits are correct, use the **Editor ⇒ Stop Editing** menu to stop editing.

Close the table.

Save the map document.

Defining Neighborhoods Using Thiessen Polygons

Next, define neighborhoods using Thiessen polygons. If you do not have the function for creating Thiessen polygons in your version of ArcGIS 10, read through this section until you find instructions for adding the **nhaqsites_Project_CreateThie** shapefile which has already been created for you.

Otherwise, click on the **ArcToolbox window** button to open ArcToolbox.

Click on **Analysis Tools ⇒ Proximity** to show the various proximity tools available in ArcToolbox.

Double click **Create Thiessen Polygons** to open the "Create Thiessen Polygons" window.

Select the nhaqsites_Project layer from the pull-down menu under "Input Features" so that the Thiessen polygons will be formed around the monitoring station sites.

Under "Output Feature Class", set the output shapefile to be:

c:\exer7\data\agency\shapes\nhaqsites_Project_CreateThie.shp

Under "Output Fields (optional)" select "ALL" from the pull-down menu.

Click OK. Close the "Create Thiessen Polygons" window when processing is complete.

The shapefile of Thiessen polygons will be added to the Data Frame.

If you do not have the function for creating Thiessen polygons in your version of ArcGIS 10, navigate to **c:\exer7\data\agency\thiessen** and add the **nhaqsites_Project_CreateThie** shapefile which has already been created for you.

Right click on the Thiessen polygon layer and select **Open Attribute Table** from the menu to open the attribute table. You should have 12 polygons, one for each monitoring station site. The attributes of the associated site have been used to describe each polygon.

The Theissen polygon shows the area closer to a particular monitoring station site than to any other of the 12 sites.

Click the symbol in the legend for the Thiessen polygon layer in the Table of Contents to open the "Symbol Selector" window. Set the fill to Hollow and click OK. This will allow you to see how the county subdivisions are arranged in relation to the Thiessen neighborhoods. Then turn off the visibility of the Thiessen polygon layer.

Save the map document.

ESTIMATING THE POPULATION IN A NEIGHBORHOOD USING AREAL INTERPOLATION

Performing an Intersect Operation

Click on the **ArcToolbox Window** button to open ArcToolbox.

Click on **Analysis Tools ⇒ Overlay** to show the various overlay tools available in ArcToolbox. We will be using the Intersect operation in this polygon overlay analysis.

Click on **Intersect** to open the "Intersect" window.

Click on the **Buffer_of_nhaqsites_Project** layer and then click on the **tl_2010_33_cousub10_Project_Pop** layer in the pull-down menu to add them as the input layers for intersection.

Edit the file path and name to save the result of the Intersect operation as:

c:\exer7\data\agency\shapes\Buffer_of_nhaqsites_Inters.shp.

Click OK.

After the result of the Intersect operation is added to the Data Frame, close the "Intersect" window and open the attribute table of the Buffer_of_nhaqsites_Inters.shp layer and check to see that you have 190 polygons in the intersection of the two layers. Note that the ALAND10 and AWATER10 of the county subdivisions shown for each record are not accurate areas for the records in the table. They describe the areas of the county subdivisions, not the areas of the part of the county subdivision in a buffer zone.

Close the table.

Then, turn off the visibility of the **Buffer_of_nhaqsites_Project** and **tl_2010_33_cousub10_Project_Pop** layers so that only the **Buffer_of_nhaqsites_Inters.shp** shapefile resulting from the Intersect operation is visible.

Save the map document.

Areal Interpolation by Area Weighting

In order to perform the areal interpolation of population from the county subdivision (source) zones to the air quality monitoring station buffers (target zones), we will need to add three fields to the **Buffer_of_nhaqsites_Inters.shp** data layer.

- TOTAREA will sum the land and water areas from the county subdivision layer.

- INTER_SQM will calculate the area of the polygon in square meters.

- WEIGHT will calculate the area of the polygon in square kilometers as a percent of the area of the county subdivision in square meters.

- ESTPOP will calculate the population of the polygon by multiplying the weight and population of the county subdivision.

These estimated populations can then be summed by air quality monitoring station buffer zone to ascertain how many New Hampshire residents live in each zone.

Open the **Buffer_of_nhaqsites_Inters.shp** attribute table and use **Table Options ⇒ Add Field** and add a field called **TOTAREA** to the table. Make it a Double field with a precision of 12 and a scale of 0. The precision indicates the number of significant digits and the scale indicates the number of decimal places.

Right click on the TOTAREA field header and select **Field Calculator** from the menu. A window will open telling you that you are calculating the value of a field outside of an edit session and that you will not be able to undo the calculations. Click Yes to continue.

In the “Field Calculator” window, make sure the radio button for VB Script is clicked under “Parser”. Scroll down the list of variables under “Fields:” and double click the [ALAND10] field to add it to the window below under TOTAREA=. Then click the “+” button. Then double click the [AWATER10] field to add it to the string as shown below:

[ALAND10]+[AWATER10]

Then click OK. This will add the value of the land area of the county subdivision and the value of the water area of the county subdivision and populate the TOTAREA field with that value. When you are sure the values are correct, add the INTER_SQM field.

Open the **Buffer_of_nhaqsites_Inters.shp** attribute table and use **Table Options ⇒ Add Field** and add a field called **INTER_SQM** to the table. Make it a Double field with a precision of 12 and a scale of 0.

Right click on the INTER_SQM field header and select **Calculate Geometry**. A window will open telling you that you are calculating the value of a field outside of an edit session and that you will not be able to undo the calculations. Click Yes to continue.

In the “Calculate Geometry” window, select “Area” as the property from the pull-down menu and make sure the “Units:” are “Square Meters [sq m]”. Then click OK.

The field should be populated with values that describe the area in square meters of each polygon resulting from the Intersect operation.

Next, use **Table Options ⇒ Add Field** and add a field called **WEIGHT** to the table. Make it a Double field with a precision of 4 and a scale of 1.

Right click on the field header and select **Field Calculator**. Make sure the radio button for VB Script is clicked under “Parser”. Type the following statement in the box under “WEIGHT =” exactly as shown to calculate the percent of the county subdivision area that is in a particular air quality monitoring site buffer zone:

[INTER_SQM]/[TOTAREA]*100

Click OK.

The field should be populated with a value that shows the percent of the county subdivision area that is in a particular air quality monitoring site buffer zone. Some county subdivisions are completely within a buffer zone (100.0%) while others are only partially within a buffer zone.

Finally, use **Options ⇒ Add Field** and add a field called **ESTPOP** to the table. Make it a Double field with a precision of 7 and a scale of 0.

Right click on the field header and select **Field Calculator**. Make sure the radio button for VB Script is clicked under "Parser". Type the following statement in the box under "ESTPOP=" exactly as shown to calculate the estimated population from a county subdivision that is in a particular air quality monitoring site neighborhood:

[D001]*[WEIGHT]/100

Click OK.

The field should be populated with a value that shows the estimated population of the county subdivision area that is in a particular air quality monitoring site neighborhood.

Right click on the ESTPOP field header and select **Statistics** from the pull-down menu.
You should find that there 1,062,018 people living in New Hampshire who are within 25 km of an air quality monitoring site for ozone. The total population of all county subdivisions is 1,316,470 so about 80.7% of the state's population is within the 25 km neighborhood of an active site.

Your could use the same techniques to estimate the number of people of different races within different buffer zones or to estimate the number of people within neighborhoods defined as Thiessen polygons.

Use the **Save** button to save the map document and then use the **File ⇒ Exit** menu to close ArcMap.