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


# GIS AND PUBLIC HEALTH EXERCISE 8 – MODELING ACCESSIBILITY (ArcGIS 9.3.1)

## PREPARATION

Download the **exer8** 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 **exer8**. 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:\exer8** 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.

## ASSEMBLING THE DATABASES

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

In this exercise, you will be modeling accessibility of women 40 to 74 years of age residing in New London County, Connecticut, to mammography facilities in New London County certified by the U.S. Food and Drug Administration as shown on the FDA’s web site. In order to perform the analysis, you will model the women’s origins based on the centroids of the 21 towns in New London County, the destinations based on the locations of the 9 mammography facilities in the county, and the distance to mammography facilities based on an origin-destination cost matrix created from a street network database for the county.

**Add Databases of Towns, Town Centroids Origins, and Mammography Facilities**

To begin, add a database of towns in New London County, Connecticut. The database was developed from data downloaded from the U.S. Census Bureau TIGER® web site from the 2010 TIGER database.

The data were projected using NAD_1983_StatePlane_Connecticut_FIPS_0600_Feet coordinates. Map units are feet.

Find the button labeled **Add Data** and click the button. You should find the **c:\exer8** folder in your catalog. If not, please connect to the folder using the **Connect to Folder** button. Navigate to **c:\exer8\data\agency\shapes** and add the **towns.shp** shapefile.

Right-click the towns data layer and select **Open Attribute Table** from the menu. There are 21 towns in New London County. Each town has a numeric ID from 1 to 21. The NAME field provides the name of the town. The TOWN_NO field is the town number used by the State of Connecticut to identify the 169 towns in the state.

**Close** the table.

Use the **Save** button or go to **File ⇒ Save** to save your map document. Navigate to **c:\exer8\mapdocs** and save the file as **exer8.mxd**.

Next, use the **Add Data** button to add a database of town centroids used to model the origins of women needing mammography services. Navigate to **c:\exer8\data\agency\shapes** and add the **origin.shp** shapefile.

Right-click the origin data layer and select **Open Attribute Table** from the menu. There are 21 town centroids in the database, one for each town in New London County. Each town centroid has a numeric ID from 1 to 21. The NAME field provides the name of the town. The FEM4074 field shows the 2010 population of women 40 to 74 years of age in a town based on the 2010 Redistricting Data Summary File (P.L. 94-171) downloaded from the Census Bureau's American FactFinder web site. For the purposes of this exercise, we will assume these women need an annual screening mammogram.

Next, add a shapefile of mammography facilities in New London County certified by the U.S. Food and Drug Administration, geocoded from address data from the FDA's web site. Navigate to **c:\exer8\data\agency\shapes** and add the **facility.shp** shapefile.

Right-click the facility data layer and select **Open Attribute Table** from the menu. There are 9 facilities in the database. Each facility has a numeric ID from 1 to 9. The NAME field provides the name of the facility. The ADDRESS, TOWN, STATE, and ZIP fields provide the facility's address. The ESTCAPAC field gives the estimated capacity of the facility to provided screenings on an annual basis. ESTCAPAC is not information distributed by the FDA; this field was added by the authors for the purposes of the exercise.

**Close** the table and **Save** the map document.

**Symbolize the Origin and Facility Data Layers**

Right click the origin layer and select **Properties** from the pull-down menu. Then click the **Symbology** tab. Under "Show:" select **Quantities ⇒ Graduated symbols**. Under "Fields:" in the "Value:" section, select FEM4074 from the pull-down menu as the field containing the values to be mapped. Under "Classification" use the Natural Breaks (Jenks) classification and set the number of classes to 3.

Click the **Template** button to open the "Symbol Selector" window and choose the Circle 1 symbol. Then click OK. Set the Symbol Size from 6 to 18 and click Apply and OK.

Next, right click the facility layer and select **Properties** from the pull-down menu. Then click the **Symbology** tab. Under "Show:" select **Quantities ⇒ Graduated symbols**. Under "Fields:" in the "Value:" section, select ESTCAPAC from the pull-down menu as the field containing the values to be mapped. Under "Classification" use the Natural Breaks (Jenks) classification and set the number of classes to 3.

Click the **Template** button to open the "Symbol Selector" window and choose the Square 1 symbol. Then click OK. Set the Symbol Size from 6 to 18 and click Apply and OK.

**Save** the map document.

Look at the Data View to compare the distribution of facilities of different capacity with the distribution of women who need service across the towns.

**Add a Network Database**

Finally, add a network database of the street network in New London County. This database was created in Network Analyst from data on all lines in New London County in the 2010 TIGER databases downloaded from the Census Bureau's web site. The data were projected based on the Connecticut State Plane Coordinate system and the road segments were selected from all lines. Map units are feet. A field, LENGTHFT, was added to each record giving the length of the street segment in feet.

Use the **Add Data** button and navigate to **c:\exer8\data\agency\shapes** and add the **tl_2010_09011_streets_Projected_ND.nd** network dataset. When you are asked whether you want to add all feature classes that participate in tl_2010_09011_streets_Projected_ND.nd to the map, click No.

**Save** the map document.

**CREATING THE ORIGIN-DESTINATION COST MATRIX**

**Enable the Network Analyst Extension and Choose Origin-Destination Cost Matrix**

If you do not have the Network Analyst extension, read through this section, then complete the analysis of accessibility by following the steps outlined in the **Explore the OD Cost Matrix** section below.

If you have the Network Analyst extension, use the **Tools ⇒ Extensions** menu to make sure that the "Network Analyst" extension is checked. Then use the **View ⇒ Toolbars** menu and check Network Analyst to make sure that the Network Analyst extension toolbar is visible.

Click the **Show/Hide Network Analyst Window** button in the toolbar to open the Network Analyst window. Make sure that **tl_2010_09011_streets_Projected_ND** is shown as the "Network Dataset:" in the pull-down.

From the Network Analyst pull-down menu, select **New OD Cost Matrix**. You should see that the Network Analyst window is updated to show "OD Cost Matrix" and to list Origins (0), Destinations (0), Lines (0), and Barriers (0).

In addition, the Table of Contents has been updated with OD Cost Matrix added as a layer in the Data Frame.

**Load the Origins and the Destinations**

Right click “Origins (0)” in the Network Analyst window and select **Load Locations** from the menu. This will open the “Load Locations” window.

Next to “Load from:”, select the **origin** database from the pull-down list. Then click OK.

The Origins entry in the Network Analyst window should be updated to show that there are 21 origins or “Origins (21)”. All origins should appear as located in the Data View, based on the symbol for Located Origins in the Table of Contents.

Next, right click “Destinations (0)” in the Network Analyst window and select **Load Locations** from the menu. Next to “Load from:”, select the **facility** database from the pull-down list. Then click OK.

The Destinations entry in the Network Analyst window should be updated to show that there are 9 destinations or “Destinations (9)”. All destinations should appear as located in the Data View, based on the symbol for Located Destinations in the Table of Contents.

Right click the Origins layer under OD Cost Matrix in the Table of Contents and select **Open Attribute Table** from the menu. Look at the attribute table. Then close the table. Do the same with the Destinations layer and then close the table.

**Save** the map document.

**Create the OD Cost Matrix**

To create the OD Cost Matrix, click the **Solve** button in the Network Analyst toolbar. You will see that the “Lines (0)” entry in the Network Analyst window is updated to “Lines (189)” and there are 189 line segments in the Data View such that every origin is connected to every destination.

**Explore the OD Cost Matrix**

Right click on the Lines layer under OD Cost Matrix and select **Open Attribute Table** from the menu. The Name Field gives the name of the town followed by the name of a facility. The OriginID field gives the ID of the town centroid of origin. The Destination ID field gives the ID of the facility destination. The DestinationRank field gives the rank from closest to farthest for each destination from a given origin. The Total_LENGTHFT field gives the travel distance (cost) measured from the origin to the destination along the street network. This is represented as a straight line in the Data View but the distance is the network distance.

Right click on the Lines layer under OD Cost Matrix and select **Data ⇒ Export Data** from the menu. In the “Export Data” window, export All features. Under “Use the same coordinate system as:”, check the radio button for “this layer’s source data”. Set the path and name for the “Output shapefile or feature class” to:

**c:\exer8\data\agency\shapes\Lines.shp**

and click OK.

If you do not have the Network Analyst extension, use the **Add Data** button and navigate to **c:\exer8\data\agency\lines** and add the **Lines.shp** shapefile. This database was exported from the origin-destination cost matrix created using Network Analyst.

Right click on the Lines.shp shapefile and select **Open Attribute Table** from the menu. The Name Field gives the name of the town followed by the name of a facility. The OriginID field gives

the ID of the town centroid of origin. The Destination ID field gives the ID of the facility destination. **Note that this field name is truncated to Destinatio in the Lines.shp shapefile.** The DestinationRank field gives the rank from closest to farthest for each destination from a given origin. **Note that this field name is truncated to Destinat_1 in the Lines.shp shapefile.** The Total_LENGTHFT field gives the travel distance (cost) measured from the origin to the destination along the street network. This is represented as a straight line in the Data View but the distance is the network distance. **Note that this field name is truncated to Total_LENG in the Lines.shp shapefile.**

**Save** the map document.

**MODELING ACCESSIBILITY**

To model accessibility from each origin, we will use the formula:

$$PA_i = \sum_j A_j / d_{ij}^{\beta}$$

where

$PA_i$ is the potential accessibility of women in town *i*

$A_j$ is the attractiveness of facility *j*

$d_{ij}^{\beta}$ is the distance from town *i* to facility *j* with the exponent $\beta$ serving to model distance decay.

The ratio of attractiveness to distance from each origin to each destination is summed over all destinations for a given origin.

The Lines.shp layer provides us with the distance from every origin to every destination in feet. For this analysis, we will use the estimated capacity of the facility as a measure of attractiveness, convert distance from feet to miles, and assume that $\beta$ equals 2.

Right click the Lines layer and select **Join and Relates ⇒ Join** from the menu.

In the “Join Data” window, select “Join attributes from a table” from the pull-down menu under “What do you want to join to this layer?”

Choose the DestinationID field named Destinatio from the pull-down as the field that the join will be based on under “1. Choose the field in this layer that the join will be based on:”.

Choose the facility layer from the pull-down as the table to join from this layer under “2. Choose the table to join to this layer, or load the table from disk”. Make sure that the box next to “Show the attribute tables of layers in this list” is checked so that you can use the attribute table from the facility layer.

Choose the ID field from the pull-down as the field from the table on which to base the join under “3. Choose the fields in the table to base the joins on.”

Make sure that the “Keep all records” radio button is selected under “Join Options”. Then click OK. When you are asked whether or not you would like to create an index for the field join, click Yes.

Right click the Lines layer and select **Open Attribute Table** from the pull-down menu. Scroll to the right. You should see that you have joined data about the facilities from the facility layer to

each of the lines linking an origin to a facility destination based on the facility ID. Make sure that you have 189 records in the table and that the data were joined correctly.

You now have access to the ESTCAPAC field to include it in the calculation of accessibility.

Next, click the **Options** button and select **Add Field** from the menu. Enter ACCESS as the field name in the "Name:" section and select **Double** as the field type from the pull-down next to "Type:". Then, double click in the Precision line and set the precision to 12 and double click in the Scale line and set the scale to 0. Then click OK.

The Lines.ACCESS field will be added after the Total_LENG field. Scroll through the table to find the field. The value of the field should be 0 for every record.

Right click the header of the Lines.ACCESS field and select **Field Calculator** from the menu.

In the window under Lines.ACCESS =, enter the following formula to calculate the value of the field:

[facility.ESTCAPAC]/(([Lines.Total_LENG]/5280)^2)

This expression takes the estimated capacity for the destination i and divides it by the distance from the origin to the destination squared. Note that the length in feet is converted to miles by dividing the Lines.Total_LENG by 5,280. The conversion to feet is within parentheses, then that quotient is squared within parentheses to form the denominator of the ratio.

Once you have the correct formula entered, click OK.

The Lines.ACCESS field should be populated with values measuring accessibility from each origin to each destination based on estimated capacity of the facility as a measure of attractiveness and distance in miles squared as a measure of the cost of using the facility from a particular destination.

The final step in calculating accessibility for each origin is to sum the values in the Lines.ACCESS field by origin.

Right click the Lines.ACCESS field and select **Summarize** from the menu.

Under "1. Select a field to summarize:" select OriginID from the pull-down menu.

Under "2. Choose one or more summary statistics to be included in the output table:" click on the "+" sign in front of Lines.ACCESS and check the box in front of "Sum".

Set the path for the output dBase table to **c:\exer8\data\agency\dbase\Sum_Output.dbf** and click OK.

This will create a table with one row for every origin and sum the Lines.ACCESS field values for that origin yielding the potential accessibility measure.

When you are asked whether or not you want to add the result table in the map, click Yes.

**Close** the attribute table for the Lines layer.

Then click the **Source** tab at the bottom of the Table of Contents and scroll down to the Sum_Output table. Right click the table and select **Open** from the menu. The table should have 1 record for each origin from 1 to 21. A field counting the number of destinations summed for each

origin has been created. This should equal 9 for every origin. Then, the value of the potential accessibility measure is given in the Sum_ACCESS field.

You should see that women in the town of Voluntown have the lowest potential accessibility (174) because the town is located far away from facilities of any size and women in the town of Groton have the highest potential accessibility (218,277) because they have a fairly large center located in town and the town is centrally located with respect to other large facilities.

**Close** the table and **Save** the map document.

**MAPPING POTENTIAL ACCESSIBILITY**

To map the potential accessibility values, perform a table join joining the Sum_Ouput table to the towns layer.

Click on the **Display** tab at the bottom of the Table of Contents and turn off the visibility of the OD Cost Matrix, if you created it in Network Analyst, by making sure the box in front of the layer name is unchecked. Also, make sure the Lines.shp layer is not visible. Finally, turn off the visibility of the tl_2010_09011_streets_Projected_ND layer of streets.

**Save** the map document.

Right click the **towns** layer and select **Join and Relates ⇒ Join** from the menu.

In the “Join Data” window, select “Join attributes from a table” from the pull-down menu under “What do you want to join to this layer?”

Choose the ID field from the pull-down as the field that the join will be based on under “1. Choose the field in this layer that the join will be based on:”.

Choose the Sum_Output table from the pull-down as the table to join from this layer under “2. Choose the table to join to this layer, or load the table from disk”.

Choose the OriginID field from the pull-down as the field from the table on which to base the join under “3. Choose the fields in the table to base the joins on.”

Make sure that the “Keep all records” radio button is clicked under “Join Options”. Then click OK. If you are asked whether you would like to create an index, click Yes.

Right click the towns layer and select **Open Attribute Table** from the menu and scroll to see that you have correctly joined the Sum_Output table to the towns layer. Then **Close** the table.

Right click the towns layer and select **Properties** from the menu and click the **Symbology** tab.

Select **Quantities ⇒ Graduated colors** from the “Show:” window.

Select the Sum_ACCESS field from the pull-down menu next to “Field:” as the field to be mapped.

Use 5 class intervals and Natural Breaks (Jenks) classification and select a color ramp. Then click OK.

Examine the map to explore the pattern of accessibility based on your analysis.

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