School of Geography & Earth Sciences
McMaster University

EARTH SC/ENVIR SC/GEOG 2GI3: GEOGRAPHIC INFORMATION SYSTEMS

EXERCISE 3: VECTOR AND RASTER DATA

Introduction

ArcGIS works with both vector and raster data. Vector data can be handled using the basic ArcMap interface. However, in order to work with raster data, the Spatial Analyst extension must be added to the program. In this exercise, you will use both types of data. At the same time, you will learn other useful skills: (1) selecting data, (2) creating and populating fields, (3) editing shapefiles, (4) creating shapefiles, and (5) creating grids (ESRI's raster data format).

Due:	Beginning of your lab period on Week 9.
Instructions:	Answers must be typed using MS Word, OpenOffice, or some other word-processing package.
Grading:	Style and format is worth 20% of your mark. 1 mark is deducted for each mistake up to a total of 5 marks since your exercise is worth 25 marks. Please consult lab notes for Exercise 3 (Exercise 3: Overview, which is found under Exercises) for style and format.

Exercise (25 marks)

The data for both parts of this exercise are located on Avenue in a zipped file called EX3. The files for Part A are located in a subfolder PARTA and the files for Part B are located in a subfolder called PARTB.

Part A: Vector Data

You are an employee of the City of Hamilton. Administratively, the City is divided into 15 Wards. These wards are shown in the following map: (http://map.hamilton.ca/Static/PDFs/WardMaps/AllWards_Statistics.pdf). Your boss has asked you to create a shapefile of Wards 6 to 9, which are the focus of a new study pertaining to land-use intensification. Eventually, your shapefile will be intersected with other shapefiles to inventory land uses and businesses in the study area. To help you create the shapefile, the Planning Department has given you a shapefile containing 2006 census tracts (HamiltonCTs.shp), which you will use to create wards. Unfortunately, the Planning Department does not have a shapefile for wards.

A. You have already completed some work. Your work is saved under a map document called Topology. Open Topology. The data frame contains three layers: Nodes, Arcs, and Census Tracts. If necessary, reestablish linkages between these layers and their data sets.

B. Set the current and scratch workspaces to appropriate locations for data input and output. At this point, be sure to save your map document.

Questions (3 marks)

1 What is the coordinate system of the data frame? (1)
2. If the layer Census Tracts had been added to the data frame first, what would have been the coordinate system of the data frame? (1) Why? (1)

C. The Arc layer outlines the ward boundaries (this layer was created using the ward map provided by the City of Hamilton). Creating a ward is quite simple and involves selecting all census tracts within a specific ward's boundary and merging them to form a new "ward" polygon. This is done through an edit session.

D. To begin your edit session, right click on Census Tracts and select Edit Features | Start Editing. Ignore the spatial reference warning and click Continue. The Editor toolbar appears if it is not already on the screen. If it does not appear, you can choose it from a drop down menu by right clicking near the top of the screen.

E. Several options are available for selecting census tracts. They are found under the Select Features tool on the Tools toolbar. Choose an option under Select Features (experiment with them first – see Question 3) and select all census tracts within Ward 8's boundary. If you make a mistake, you can click on the Clear Selected Features tool to deselect all selected census tracts. Alternatively, you can hold down the Shift key and click on individual census tracts to add or remove them.

Questions (2 marks)

3. For Ward 8, what Select Features options can be used to select only census tracts within its boundary without requiring further addition or removal using the Shift key? (1)
4. How many census tracts are within Ward 8's boundary? (1)

F. Next, merge the selected census tracts to form a polygon for Ward 8. This is done through the Editor toolbar. Click on Editor | Merge | OK.

G. Open the attribute table for Census Tracts. At the bottom of the table, click on the Show Selected Records button. The record for Ward 8 is shown. Under the field called Name, change the name of the polygon to "Ward 8". Close the attribute table. Use the Identify tool to verify that the polygon's name has been changed.

H. Click on the Clear Selected Features button to deselect Ward 8.

I. Save your edits by clicking on Editor | Save Edits.

J. Repeat the steps above to create polygons for Wards 6, 7, and 9. When all wards have been created, click Stop Editing and save your edits. This ends your edit session.

K. Select Wards 6 to 9 and create a new shapefile from the selection. The shapefile must be called Wards.shp, and the coordinate system must be the same as that of the data frame. Add the shapefile to the data frame and deselect the polygons (if they are selected). Remove Census Tracts from the data frame.

L. Your boss has asked for a screen capture of the newly created wards. Her instructions were very specific. The wards are to be colored as shown on the map provided by the City, the wards are to be labeled by name using a 16 pt. bold Arial font, and the Node and Arc layers are to be shown. Color and label the polygons as per instructions (you can do this through Layer Properties). After doing so, save the Topology map document.

Question (4 marks)

5. Provide a screen capture of your work upon completing step L. Your screen capture must not be cropped in any way – that is, all parts of the screen must be shown – otherwise, you will receive a 0 on this question. (4)

M. Your boss knows that the shapefile is a standard nontopological vector data format. However, she knows that other vector data formats, such as the coverage, are topological. She is interested in seeing what the data structure would look like if the wards polygons had been saved as a coverage rather than a shapefile. To this end, she has asked you to prepare a series of topological tables for the Wards layer corresponding to a node file, an arc file, and a polygon file. The Node and Arc layers that you have prepared are instrumental in completing this task.

N. Rather than write down X and Y coordinates for the node file, you have decided to add these to the Node layer's attribute table. Open the attribute table for Nodes. Under Table Options, select Add Field. In the popup window that appears, name the field "XCoord" and set its type as Double with a precision of 12 and a scale of 2. Click OK. This means that XCoord will be a decimal number with 12 digits, 2 of which will be to the right of the decimal point. Right click on XCoord and select Calculate Geometry. If a warning appears, ignore it. In the popup window that appears, next to Property select X Coordinate of Point. Make sure that Units is set to meters. Click OK. You have now computed the X coordinates of the nodes.

O. Using the procedure in step N, calculate the Y coordinates of the nodes. Make sure that you name the new field "YCoord" and that next to Property you select Y Coordinate of Point.

Question (2 marks)

6. Provide a screen capture of the Nodes layer's attribute table upon completing step O. **To receive full marks, the attribute table must not be maximized.** Your screen capture must not be cropped in any way – that is, all parts of the screen must be shown – otherwise, you will receive a 0 on this question. (2)

P. Close the attribute table for Nodes.

Question (5 marks)

7. With reference to the map created from the Nodes, Arcs, and Wards layers, complete the Arc and Polygon topology tables found at the end of this exercise. You may print out the page, write your answers by hand, and submit the page as part of your work. Note that Ward 8 = polygon 1, Ward 7 = polygon 2, Ward 6 = polygon 3, Ward 9 = polygon 4, and the external polygon = 0. To receive full marks, nodes, arcs, and polygons must be referred to by their numbers, and arcs must be listed in ascending order separated by commas in the polygon file. (5)

Part B: Converting Vector Data to Raster Data (Rasterization)

In this part of the exercise, you will make use of land use data to examine the mixed pixel problem and the impact of rasterization. A simple way to measure the effect of such conversion is to calculate the area for the different land-use types.

A. Open the map document called PartB. It contains a map of land uses in the City of Hamilton. The current coordinate system for the data frame is GCS_North_American_1983. Remember to set your workspaces.

B. Change the projection of the data frame to NAD 1983 UTM Zone 17N.

C. Open Land Use's attribute table. Under Table Options, select Add Field. In the popup window that appears, name the field "VecArea" and set its type as Double with a precision of 15 and a scale of 3. Click OK.

D. Use Calculate Geometry to compute the area of each land-use polygon in square meters.

Question (1 mark)

8. When calculating area *for this example*, the system defaults to using the coordinate system of the data frame. Why is this default chosen? (1)

E. The field in the attribute table is updated and contains the area in square meters for each land-use polygon. However, you are interested in the total area for each land-use category. This means that you must create a summary table. Right click on the field name Category and select Summarize. In the popup window that appears, Box 1 should be set to Category. In Box 2, expand VecArea and check Sum. In Box 3, call the output table VecLU.dbf (make sure that the output table is saved as a dBase table – that is, it has a .dbf extension). Click OK. When prompted, add the table to the map document and examine its contents.

F. Add Grid1000m.shp to the data frame. Ensure it is above all the other layers in the Table of Contents. This is a vector layer of square polygons (1000m × 1000m). You will use the layer to simulate a raster. Change the layer's symbology so that its polygons are not filled with a color. This can be done by double clicking on the layer's symbol in the Table of Contents and selecting Hollow from the Symbol Selector. Zoom into the map to a scale of 1:24,000 and use the Pan tool to move around the map. You will notice that cell may contain multiple land use types, but when a raster is created, only one value is assigned to a cell. This can create problems since the raster may omit or generalize the original values. It is one of the risks of rasterization.

G. Now add SampleArea.shp to the data frame. Right click on the layer and select Zoom to Layer. Change the layer's symbology so that its polygons are not filled with a color. Turn off Grid1000.

Questions (3 marks)

9. Encode the sample area raster using the "cell center" approach. Begin your encoding in the upper left-hand corner of SampleArea. Your answer will consist of three rows with three numbers in each row. The numbers correspond to the various land uses as follows: 1 = commercial, 2 = government and institutional, 3 = open area, 4 = parks and recreational, 5 = residential, 6 = resource and industrial, and 7 = waterbody. You may want to use the Measure tool to identify the center of cells. (2)
10. Comment on the resulting raster in terms of land-use omissions and generalizations that occur. (1)

Next you will convert the Hamilton land use data set to raster format (in this case a grid – ESRI's raster format). The functionality for raster data sets is stored in the Spatial Analyst extension. Extensions are subsets of tools that deal with specific types of analyses.

H. On the menu, select Customize | Extensions and click the check box next to Spatial Analyst. Click Close.

I. The tools for this extension are found in various toolboxes in ArcToolbox. To convert the land use layer to a raster, you will utilize the Polygon to Raster tool. This tool is located in the Conversion Tools toolbox, and can be accessed by expanding the To Raster toolset. Double click on the tool. A popup window appears.

J. At the bottom of the popup, click Environments. Here you can set your workspaces if you have not done so already. This is particularly important when using the Spatial Analyst extension since temporary grids are often created and stored in the scratch workspace. After this, fill in the inputs to match Figure 1 (the one exception might be the location of the output raster dataset; however, make sure that it is named LU750). The cell assignment type is set to Maximum Combined Area, which is the "winner takes all" approach. When completed, click OK.

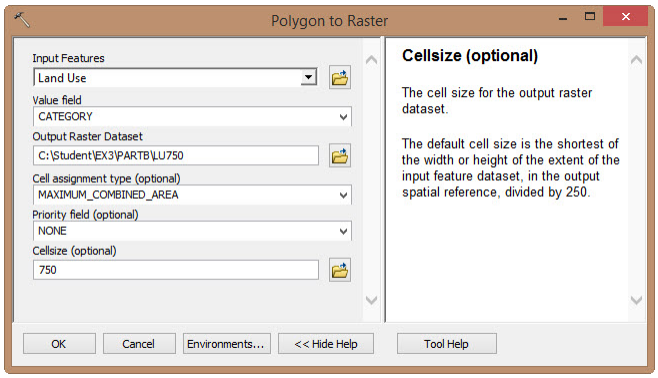

Figure 1. Polygon to Raster tool.

K. After the raster is created change the colors of the land uses to match those of the vector data (Land Use). The reason for this is that when a raster is loaded into a data frame, a random color palette is assigned to it.

Question (1 mark)

11. Provide a screen capture of LU750 upon completing step K. Your screen capture must not be cropped in any way – that is, all parts of the screen must be shown – otherwise, you will receive a 0 on this question. (1)

L. Redo the vector to raster conversion using a cell size of 250m. Name your raster LU250.

To quantify the impact of rasterization, for each newly created raster, you will compute the total area of each land-use category. These totals will then be compared to those for the vector layer.

M. Open 750LU's attribute table. Under Table Options, select Add Field. In the popup window that appears, name the field "Area750" and set its type as Double with a precision of 15 and a scale of 3. Click OK.

N. To calculate area, right click on Area750 and select Field Calculator. Click Yes to acknowledge the warning. The Calculate Geometry option is not available for raster layers. Instead, the formula to use is: area = # of cells × area of each cell. So for this raster, the formula to use is:

AREA750 = [COUNT]*750*750

O. Repeat the same procedure for 250LU. Name the area field "Area250" and change the cell area in the above calculation so that it is 250 * 250. All land use layers should now have fields with areas computed in square meters that can be placed next to each other and compared. However, you will take it one step further and calculate the percent change in area for each land-use type in the two raster layers and put everything in one table.

P. Open VecLU and open 250LU's attribute table. Both have the same Category field. You will use this field to join the tables together in order to calculate the percentage change in area for the 250m raster. In the VecLU table, create two fields named PC250 and PC750, both being of type Double, precision 15, and scale 3. These fields will store the percentage change in area for each category.

Q. You will now join the 250m land use raster to VecLU. Right click on VecLU in the data frame and select Joins and Relates | Join. In the popup window that appears, set Box 1 to Category, Box 2 to 250LU, and Box 3 to Category. Check Keep All Records and click OK. Click No to the Create Index box.

R. To calculate percentage change, you will use (raster area − vector area)/vector area × 100 since you are using the vector layer as the baseline to detect change. This can be executed in the Field Calculator (right click on PC250 in the VecLU table and select Field Calculator). Now you can construct your equation by double clicking on the appropriate field names in the Calculator. After the calculation is complete, remove the join by right clicking VecLU and selecting Joins and Relates | Remove Joins | Remove All Joins.

S. Using the same procedure, compute values for PC750. Upon completion, remove all joins. The VecLU table now contains the percentage change for both rasters.

Questions (4 marks)

12. Provide a screen capture of the VecLU table upon completing step S. **To receive full marks, the table must not be maximized.** Your screen capture must not be cropped in any way – that is, all parts of the screen must be shown – otherwise, you will receive a 0 on this question. (2)
13. After reviewing the VecLU table, comment briefly on the differences that you observe. (2)

ARC FILE
(Nodes for FNODE and TNODE must be listed as 1 to 12; Polygons for LPOLY and RPOLY must be listed as 0 to 4)

ARC	*FNODE*	*TNODE*	*LPOLY*	*RPOLY*
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

POLYGON FILE
(Arcs must be listed as 1 to 15 in ascending order separated by commas)

POLY	*ARCS*
1 (Ward 8)	
2 (Ward 7)	
3 (Ward 6)	
4 (Ward 9)	