11.520: A Workshop on Geographic Information Systems

11.188: Urban Planning and Social Science Laboratory

Lab Exercise 6: Vector Spatial Analysis

Due: in Lab 7

In this exercise, you will perform several different types of spatial analysis using the TIGER files (**cambtigr**), census block group data (**cambbgrp.shp**), land use data (**landuse85.shp**), and various retail store locations for Cambridge.

Preliminaries

1) Launch ArcMap and add the following layers in the directory **M:\data**.

* **M:\data\landuse85.shp**   Cambridge land use in 1999 per MassGIS

* **M:\data\cambbgrp.shp** (not a coverage)   U.S. Census 1990 block groups for Cambridge

* **M:\data\cambtigr** (coverage)   U.S. Census 1990 TIGER file for Cambridge

* **M:\data\cambridge_ice_cream.shp**   Ice cream stores in Cambridge

* **M:\data\cambridge_bookstores.shp**   Bookstores in Cambridge

* **M:\data\cambridge_record_stores.shp**   Record/CD/Tape stores in Cambridge

The source for the **cambridge_ice_cream.shp**, **cambridge_bookstores.shp**, and **cambridge_record_stores.shp** layers is the http://www.bigyellow.com/ web site on November 1, 1999. The locations were downloaded from the web page, then geocoded in ArcMap. We'll be exploring address geocoding in a subsequent lab.

Recall the usual preliminaries:

- Since the datasets we'll be using are in Mass State Plane coordinates, select **View > Data Frame Properties** in the menu bar and set the map units to **meters** and the display units to **miles**. If you want, you can also change the name to something more informative than "Layers".

2) Go to the `M:\data\chapter11` folder.

   You can find ArcGIS map document files, **ex11a.mxd** and **ex11c.mxd** and directories, "**Data**" and "**MyData**".

   Please copy all files and directories under the **chapter11** folder and save them in your personal working space.

Part 1: Point in Polygon Spatial Join and Cluster Analyses

In this part of the lab, you will examine the demographic characteristics of the block groups in which various types of bookstores are located. The bookstore data are contained in the shape file called `cambridge_bookstores.shp`.

ArcMap can do 'polygons to point' overlay operation using a spatial join. This time, we'll explicitly set up a spatial join between `cambridge_bookstores.shp` and `cambbgrp.shp` so that block group attributes can be added to bookstore table.

Spatial join

- Click the **cambridge_bookstores** layer from the data frame window.
- Click the right mouse button to select **joins and relates > join**. *Join Data* window will come up.
- Select **"Join data from another layer based on spatial location"** for "*What do you want to join to this layer*" option.
- Choose `cambbgrp.shp` for a layer to join to this layer.
- Check the "it falls inside" option
- Specify the name (**bookstore_bgrp.shp**) of the new file and the location (**your working directory**) where it is stored.
- Click "OK"

The new shape file you just created (**bookstore_bgrp.shp**) will be added in your data frame. If you examine the attribute table for `bookstore_bgrp.shp` you should now see the attributes of `cambridge_bookstores` and `cambbgrp.shp`. The "Shape" columns aren't true attributes; they serve as placeholders that represent the geometry of layers that allow us to perform the spatial joins this way.

Now we can make a thematic map of `cambridge_bookstores.shp` using the attributes of `cambbgrp.shp`. Open the *layer property window* of **bookstore_bgrp.shp** to make a "Graduated symbol" map using the "Med_hh_inc" field. You should end up with the bookstore locations sized by the median household income of the block group containing

them. To provide additional context for interpreting these locations, create a thematic map of land use using the `landuse85.shp` layer. This time you can name the file as "**bookstore_lu99**". Follow the same instructions as in the Explanatory Mapping portion of Lab 2, except this time ignore the directions regarding the layer `sales89.` Make the Cambridge TIGER file visible as well.

- For the definition of the land use code, please check the Massgis (Massachusetts Geographic Information Systems) land use definition at http://www.state.ma.us/mgis/lus.htm

Take a close look at the pattern of bookstores. Does anything look interesting? Write your observations as requested in the Question 1 of the assignment. Make a layout of your map and submit this as the answer for Question 2. For extra credit, or just for fun, redo the spatial join, this time tying the block groups to the ice cream stores and the record stores (this requires a spatial join for each of the store layers) and display them on the map too. Do the additional store locations help you to see a pattern?


Part 2: Buffer Analysis

In this part of the lab, you will analyze the demographics of the neighborhood around MIT's biological research facility on Ames Street. You will build on what you learned in the "Simple Buffering" part of Lab 3 to do a more elaborate analysis. Here, you will use ArcMap to calculate the number of children 5 years of age and under who live within half a mile of the facility. To begin, you will draw a half mile buffer around Ames Street. First you will select Ames Street from the `cambtigr` coverage and then draw a buffer only around that street.

Select the `cambtigr` layer, then use the **Selection > Select by Attributes...** menu item to select the arcs where the **"Fname"** field is **"Ames" and the "Ftype"** field is **"St"**. You may have trouble spotting the arcs you just selected; use the **View > Zoom Data > Zoom To Selected Features** menu item to help you find them. You should end up with **three** selected arcs.

Now let's draw a 1/2 mile buffer around Ames Street. Use the **Analysis Tools > Proximity>buffer** (please see following figure) in toolbox to start the buffer tool. In the first step, make sure to buffer only the selected features of `cambtigr`. In the second step, specify a distance of 0.5 miles. Note that you *must* type the leading zero! In the third step, tell it to *dissolve* the barriers between the buffers by selecting "ALL" and to save the results in a new layer called `amesbuf` in your working directory. A new data layer called **amesbuf** will appear in your data Frame. Now move down the **amesbuf** layer so that `cambtigr` and the point layers display above the buffer. You should be able to clearly see the selected arcs in `cambtigr` at the center of the buffer.
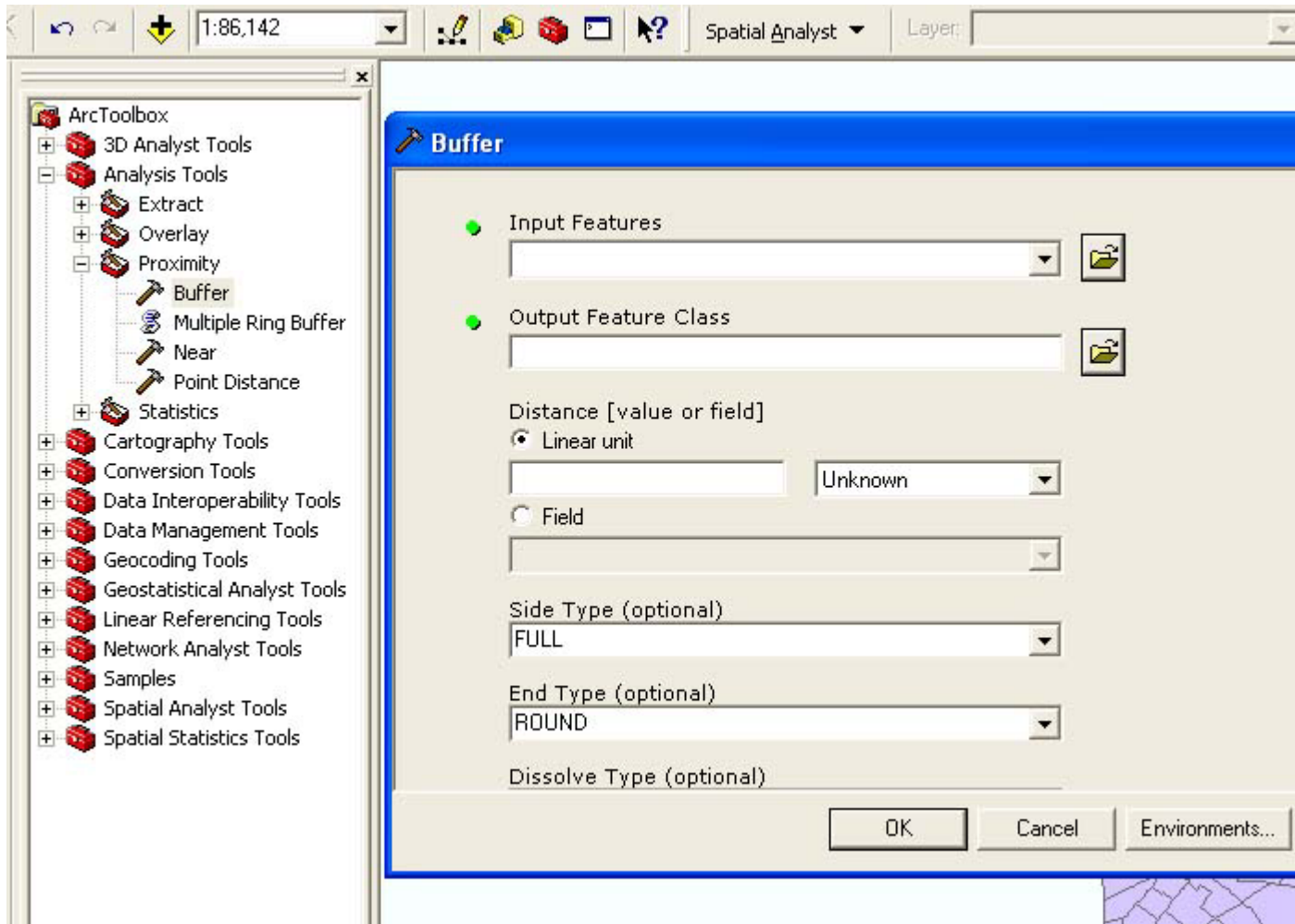
Number of Children within a Half Mile of Ames Street

Since you are interested in finding the number of children that live in the buffered area, your database must include the relevant age variables. Take a look at the attribute table of `cambbgrp.shp` in ArcMap and note that there are several age-related variables that contain numeric counts:

Age Fields in the Attribute Table of `cambbgrp.shp`	
Field	**Description**
Age_lt_1	Number of children less than 1 year old
Age_1_2	Number of children 1 or 2 years old
Age_3_4	Number of children 3 or 4 years old
Age_5	Number of children 5 years old
Age_6	Number of children 6 years old

Let's take a look at the buffer relative to the block groups. Adjust the display properties of the `cambbgrp.shp` layer so that it appears only with a thick black border: set the foreground color to transparent and the outline width to 2. Display the layer on top of both `cambtigr` and **amesbuf**. You can see that a very small fraction of many block groups falls within the buffer area. What if you really want to know what proportion of each block group falls within the buffer? The **union** and **intersect** operations are good tools for this analysis.

Before using any of these commands let's look at the **amesbuf** coverage attributes created by the buffer command. When you open the buffer's attribute table, you should see that the buffer command has created a table with one row (since we only produced one buffer polygon) and one field, **BufferDis**, which contains the buffer distance we specified when we created the buffer.

The **union** and **intersect** operations can be used to "overlay" the block group layer with the buffer layer so that the combined layer tags each feature (polygon) with attributes that indicate the original block group and whether the polygon is inside of or outside of the buffer region.

Now let's explore how the **union** and **intersect** operations differ. We shall use both operations, **union** and **intersect,** to combine all the information attached to the `cambbgrp.shp` layer with the ones attached to the **amesbuf** coverage. The **union** operation computes the geometric intersection of two polygon coverages. All polygons from both coverages will be split at their intersections and preserved in the output coverage. The **intersect** operation, on the other hand, preserves only those features in the area common to both coverages in the output file. Visually, the difference between these two commands is:
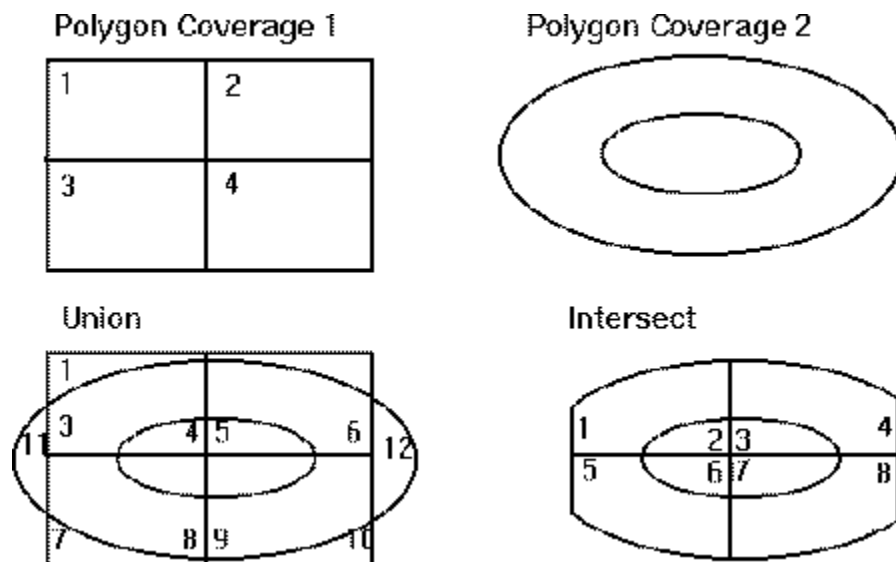
Note that in the **Tools > GeoProcessing Wizard**, "Polygon Coverage 1" in the diagram above is called the "input layer" and "Polygon Coverage 2" is called the "overlay layer."

Polygon Overlay: Intersect

In the **Arctoolbox,** find intersect tool (Please see following figure). we will use the
intersect option to create a new layer, `amesbgbuf-i`, that combines both layers. Here is
the procedure:
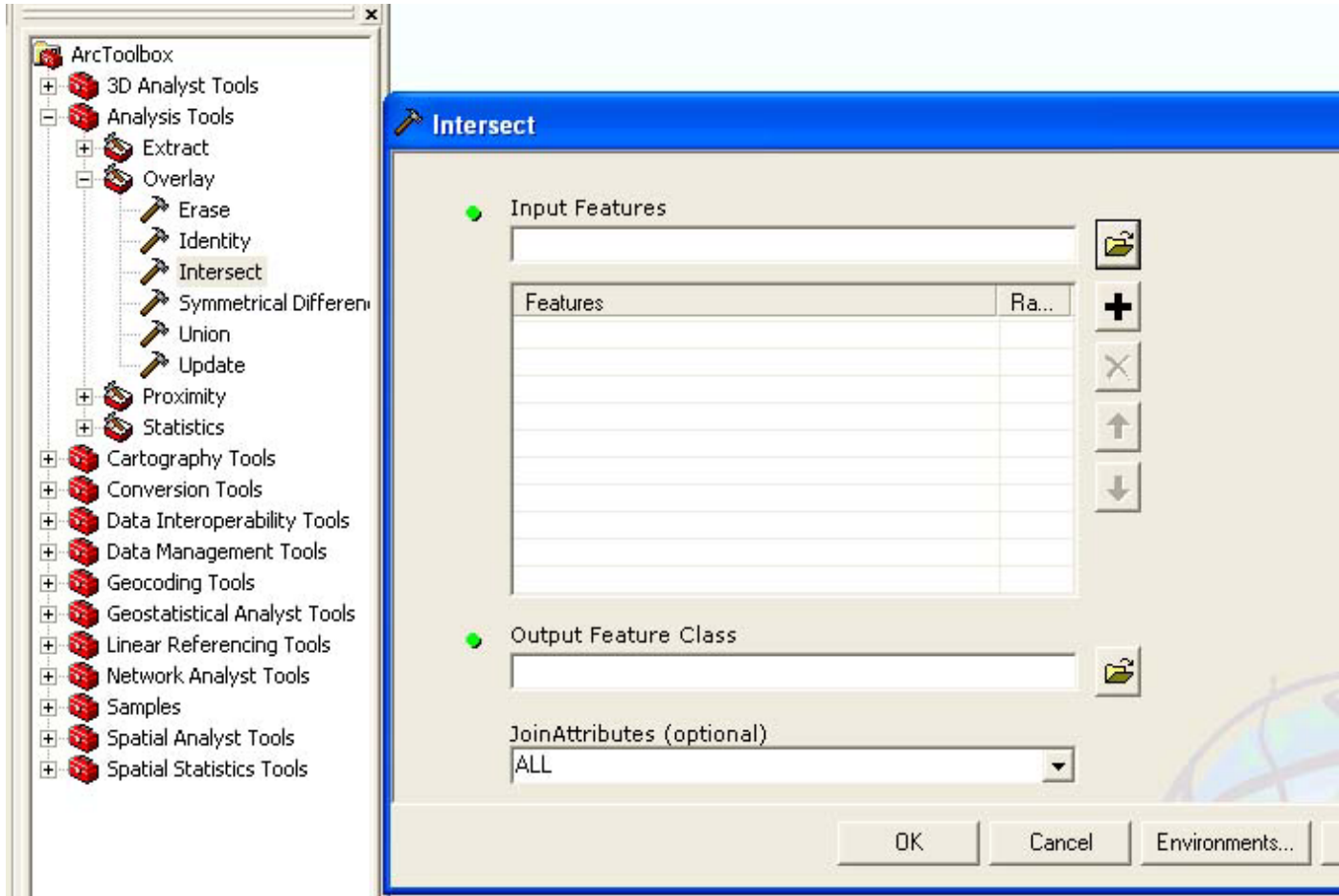
Procedures:

- Start up the intersect window with the **Toolboxl>Analysis Tool>Intersect...**
  menu item
- Select "Input features"
- Locate `cambbgrp.shp` for the "Input features: "
- Locate **amesbuf** for the "Input features:" (after this step, you will see the above
  two layers in "features" window)
- Specify the output file **[your working directory]/amesbgbuf-i.shp**
- Click Finish

The new layer, **amesbgbuf-i,** will appear in your view. Take a look at its attribute table. Notice that there are fields there named "Area" and "Perimeter". These contain the area and perimeter from the **full** block group polygon from which these were derived. ArcMap did **not** recalculate the area and perimeter for these polygons. Hence, you need to calculate the new area and store it in the table, plus calculate the ratio of the old to new areas.

Open the attribute table of **amesbgbuf-i**, and select the **Option > Add Field** menu from the bottom of the **amesbgbuf-i** attribute table and add the following fields to the table:

- **Newarea** (Type: Float)
- **Arearatio** (Type: Float)
- **Popupto5** (Type : Float)

Click on the "NewArea" field
and click the right mouse
button.

- Select **Calculate
  Values...**
- Check Advanced
- Type the following
  VBA statement in the
  first text box

  Dim dblArea
  as double
  Dim pArea as
  IArea
  Set pArea =
  [shape]
  dblArea =
  pArea.area

- Type the variable
  dblArea in the Text
  box directly under the
  area field name

  NewArea = **dblArea**

  Simply what the code
  does is to obtain a
  new area value for
  each polygon and
  update the old area
  value, which is the
  area before
  intersecting polygons,
  with the new area
  value. You can find
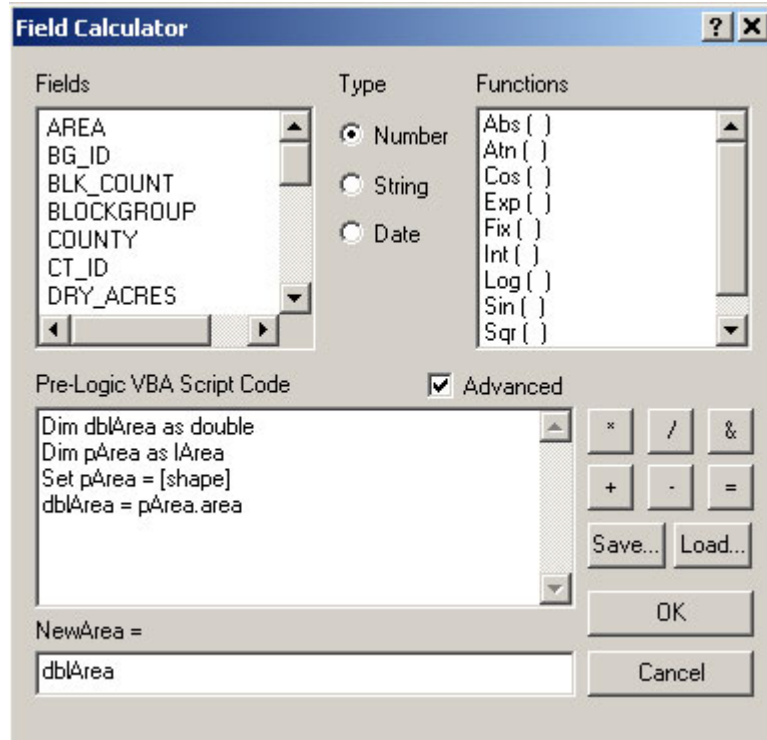 the VB code from
  ArcGIS help pages.

Fig.1. Updating Area for a Shape file

You will basically calculate a weight for each record in the overlay coverage. The weight corresponds to the percentage of the original coverage's area that is carried over to the overlay coverage. Using the assumption that attribute values are distributed evenly throughout each block group polygon in the coverage (e.g., that population is distributed evenly throughout each block group in Cambridge), the weights (arearatio) can be used to estimate new values for the overlay coverage for any attribute.

Now we'll calculate the ratio of the old to new area. Click on the heading for "Arearatio" and use the **Calculate Values...** menu item again to set

Arearatio = **[Newarea] / [Area]**
Now we are ready to calculate our estimate of the number of children within the buffer aged up to 5 years, adjusted for the relative portion of the block groups inside the buffer area. We are assuming that people are evenly distributed across each block group and hence the number of people falling within a buffer is proportional to area of the polygon with the buffer area. Click on the heading for "Popupto5" and use the **Calculate Values** menu item again to set
Popupto5 = (**[Age_lt_1] + [Age_1_2] + [Age_3_4] + [Age_5]**) * **[Arearatio]**
At this point, stop editing the table and save your results. Now you can use the **Selection > Statistics** menu item to calculate the sum of the estimates across all the block groups. This sum is your estimate of the number of children under 5 living within 1/2 mile of the biological research facility at MIT. This is the answer to Question 3 of the lab assignment. Question 4 asks you to make a thematic map that documents your efforts.

Polygon Overlay: Union

Now let's see how **union** is different from **intersect**. Let's do the same operation of combining both coverages and apportioning people along the buffer boundary, but this time using the **union** option in the **Analysis Tools>Overlay**. We shall call the output coverage `amesbgbuf-u`.

Repeat all the steps in the "Intersect" section above, with these changes:

- In the Analysis Tools, choose "Union" rather than "Intersect."
- Make the name of the output layer `amesbgbuf-u.shp` instead of `amesbgbuf-i.shp.`
- Before you calculate the statistics on `amesbgbuf-u.shp`, select only those block groups that lie **within** the buffer. You can identify these because the "BufferDis" field (the buffer distance, which you specified when you created the buffer) that ArcMap automatically adds to buffer layers is **nonzero**. When you union the layers, all the attributes from both layers are carried over into the new layer. Logically, if a polygon is outsize the buffer, then it is not possible to assign any fields from the buffer layer to that polygon. To compensate, ArcMap will populate these fields in these records with zeroes (for numeric fields). You will need to use the **Select by Attribute** menu item to restrict the rows to those within

the buffer before performing the statistics calculation. If you don't, your answer will be very different than that in the "intersect" part of the exercise.

Refer to the information in the "About Union" box to help you decide what the input and overlay layers are. You should obtain the same numerical results for the child count with either the union or the intersect operation. The geography, however, will look considerably different. For your calculations to work, however, you'll need to keep track of which polygons in the `amesbgbuf-u.shp` layer were originally inside the buffer; these will be the polygons that have a nonzero value for the "BufferDis" field. Why is this not an issue with the layer you made with the intersect operation? Think about how union is different from intersect even though you can use either as a step toward the same end. Write your answer in the spot for Question 5 in the assignment.

Part 3: Other Spatial Analysis preparation tools (Optional)

- Dissolving features and clipping layers

Please copy and paste all files and directories under `M:\data\chapter11` into your personal working space if you didn't do it yet. In the chapter11 directory, you can find two ArcGIS map document files, ex11a.mxd and ex11c.mxd. Open the each document file using ArcMap and do the exercise follow the instruction in the text book "Getting to know ArcGIS" page 270 (ex11a) and page 282 (ex11c).

Assignment

Please use the assignment page to complete your assignment.