11.520: A Workshop on Geographic Information Systems

11.188: Urban Planning and Social Science Laboratory

Homework 2

Extracting Census Data, Simple Queries, and Suitability Maps

Distributed:	Lecture 5
Due: (at start of class/lab)	**Lecture 9 -** Question #1 **Lab 8** - Question #2

BEFORE STARTING ARCMAP, PLEASE PAUSE AND SKIM THROUGH THE *WHOLE* ASSIGNMENT FIRST.
Getting the 'big picture' first will help you develop a better GIS strategy and will reduce time and energy wasted.

A local non-profit group is interested in locating a site for building a senior center in Cambridge. Given your expertise in GIS, you are hired as a GIS analyst by this company to help them locate the best site. After a long meeting with the organization and the community you agree to run some numbers in order to get a handle on the locations and characteristics of potentially suitable Cambridge sites. You settle on the following criteria to get rolling with your site selection process:

1. The minimum area of land needed for the project is 1 contiguous hectares (1 hectare = 10,000 square meters = 2.471 acres, and 1 acre = 43,560 square feet).

2. Ideally, the site should be located near, but not in, a residential neighborhood.

- The 1999 land use data is available from [MassGIS](#) or from class data folder (`M:\data\camb_area_lu_1999.shp`) to classify land

use. Recall from Lab #2 that you can refer to [MassGIS Datalayer Descriptions and Guide to User Services](#) and, specifically, the [documentation for the land use datalayer](#) on the web for further discussion of land use codes. The non-profit organization has been promised financial assistance from a number of wealthy philanthropists and has the support of the Cambridge City Council and Mayor. Thus they are capable of buying any land even if it is already developed. (In other words, don't worry about whether or not the land is vacant or about how much it would cost to acquire the land.) You decide to consider any land that:

- Was classified as open (excluding urban public space and cemeteries), commercial, or industrial in 1999.
  (Hint: Read the land use code definitions carefully -- all of them.)
- Abuts, but is not actually in, a residential area.
  (Hint: Using "Select by Location" with "Are Within Distance Of" is a way to accomplish this.)

3. Accessibility to the project is a major concern for the organization, especially given the often limited mobility of seniors. You determine that the project should be located within 200 meters of a major road.

- MassGIS major road layer, `majmhda1.shp.is` available to assist you in determining which places meet this accessibility requirement. The shapefile is located in the class data directory (**M:\data**\). (Consult the [MassGIS metadata for this layer](#) for the correct set of attribute information.)

4. The organization is also worried about health risks. They decide that they want the site to be far (more than 300 meters away) from [Toxic Release Inventory (TRI)](#) sites as identified by the [Environmental Protection Agency (EPA)](#) data from their toxic release inventory databases.

- Data from the [Envirofacts](#) database are available to identify relevant TRI site locations. A shapefile of TRI facilities for Massachusetts, `mass_tri_facilities.shp`, is stored in the **M:\data** directory. Note that this constraint applies to **all** TRI

facilities, not just those located in Cambridge or the abutting towns north of the Charles River (i.e., TRI facilities in Brookline or the Boston neighborhoods of Allston and Brighton count too).

- We use the EPA Toxic Release Inventory database for illustrative purposes only -- to show how well-documented external databases can be incorporated into our own local analysis. In reality, not all TRI sites are health risks to nearby residents and case by case analysis (based on the specific nature of a site's toxic releases) is warranted before rejecting locations proximate to a site.

- The TRI facilities shapefile includes all the TRI facilities for Massachusetts in the EPA's database (as of October 26, 2000 and for which a latitude-longitude location was provided) that fell within the borders of Massachusetts. Not all of the facilities in the EPA's database included latitude-longitude locations, and some of those with location data did not map to locations within Massachusetts (as determined by intersecting the points with a theme circumscribed by the Massachusetts border).

5. Accessibility by seniors with limited financial means for joining private clubs is deemed especially important. Therefore, you decide to narrow the criteria to census block groups where:

- The percentage of the below-poverty-level senior (aged 65 or over) in the senior population with known poverty status is high. For the purpose of this assignment, a block group is considered to have a high percentage of below-poverty-level senior if the percentage of below-poverty-level seniors is at least 10 % above the average for the five towns including Cambridge, Arlington, Belmont, Somervillle and Watertown. For example, if you found that the overall five-town percentage of impoverished seniors was 15%, then you would be looking for block groups with at least 15% * (1+ 10%) = 16.5% impoverished seniors.

- Use the **Massachusetts Block Groups 2000** layer from **MIT Geodata Repository** as we used in lab 5. Set the coordinate system of the coverages to Massachusetts State Plane, NAD 1983 meters and set the 'map units' to be meters.

- Use the **ma_towns00.shp** layer in **M:\data** directory as the town boundary.

- As a tip, you may want to select a sub set of features in both the Block groups layer and the ma_towns00.shp layer that are relevant to this problem, export them to your own working directory, and add them to ArcMAP. In this way, the operation in ArcMAP is

faster and at the same time, you get the writing permission on these files so that you can modify them when needed. More importantly, you will be able notice the "sliver" problem.
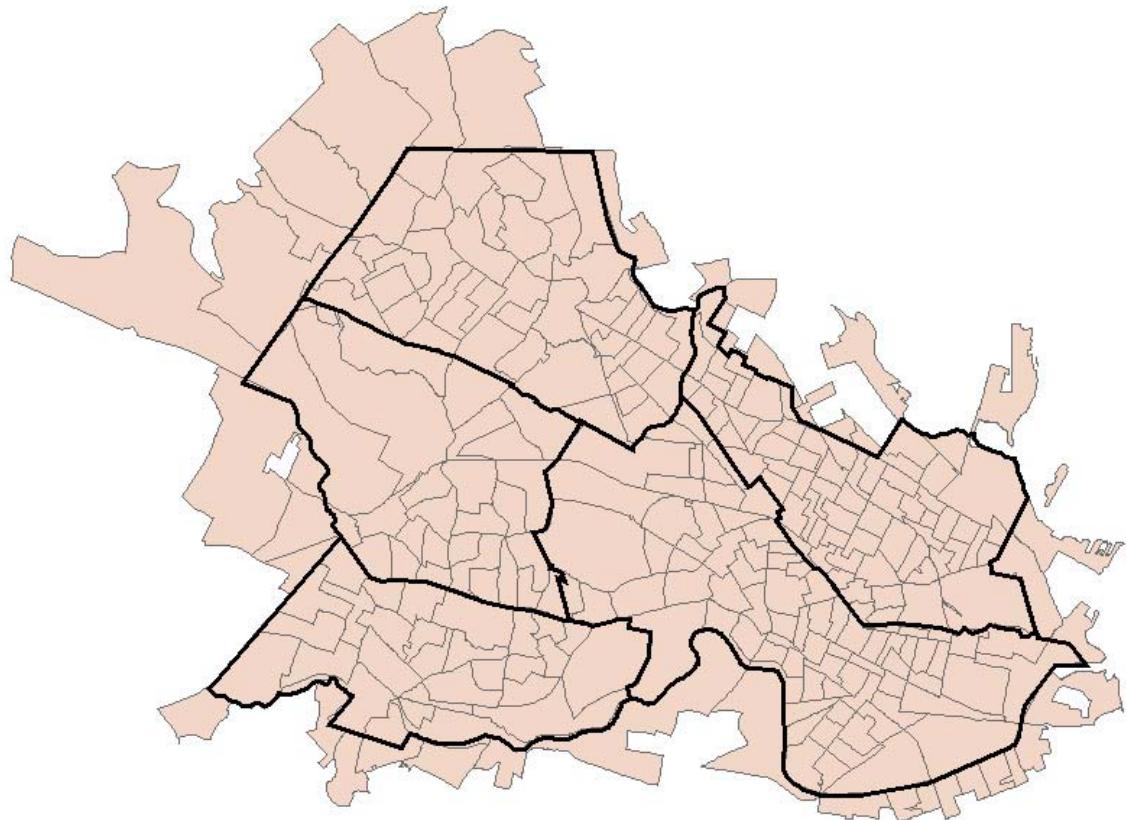
The graph above shows the selected five town and the block groups that are selected using "Select by Location" function with "intersect" specified as the topological relation between two layers. Many block group outside the five town boundary are also selected. If you zoom into a small area across town boundary as shown below, you will see why those block groups are selected. The two themes have different levels of details and you would have a 'sliver' problem if you tried to reconcile the common boundaries.
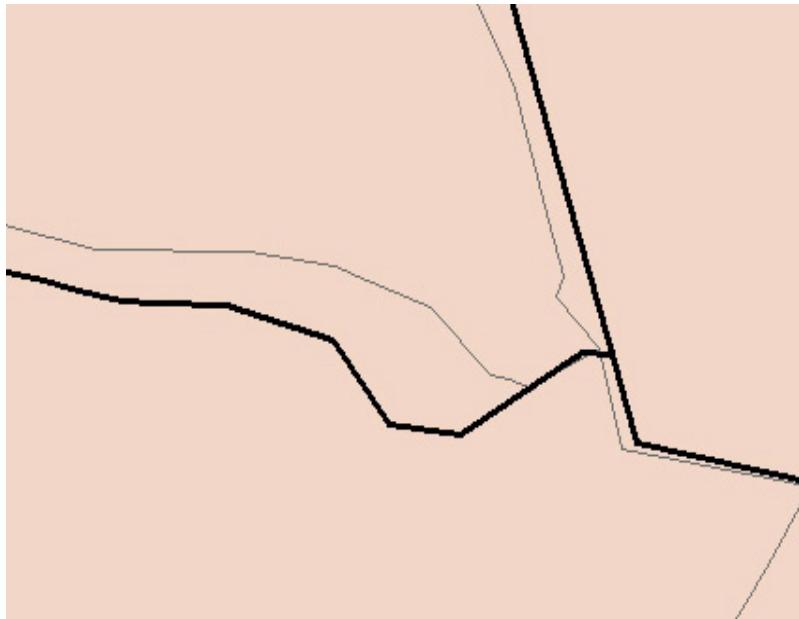
- ▪ Find the relevant demographical data from Census 2000. Open Summary File 3 : Technical Documentation (.pdf) and search the key word "POVERTY STATUS". On Page 470, you will find "**P87. POVERTY STATUS IN 1999 BY AGE [17]**". There are 17 columns in this section. Read the description of the variables in this section carefully and decide which variables you want to use to calculate the percentage of the below-poverty-level senior. Then go to page 12 of the file and find the corresponding ".uf3" file is "st00007.uf3". This file for Massachusetts is "ma00007.uf3", which has been downloaded to M:\data directory for your convenience.

6. In order to assure access to the senior center to a large number of senior citizens, we are interested in looking at the data for the four towns (north of the Charles River) that border Cambridge -- Arlington, Belmont, Somerville and Watertown, as well as Cambridge itself. **However**, because the political support for the facility is so strong in Cambridge (the Mayor and Council being behind it), the facility should be located within the Cambridge city limits.

Question 1 [50 points]

Fill inthe following table and answer the questions below. Instead of using ArcMap, you are free to use other database management software (such as MS-Access). Note that you will have to make careful use of the 'summarize' command in ArcMap (or equivalent 'total' or 'group by' commands in other database managers) in order to compute the correct values for many of the columns. Note, also, that all five towns are in Middlesex County.

Town	Number of BG's			Number of Seniors with Known Poverty Status			Average of *Percentages Calculated for each Block Group*
	Total	Populated Block Groups	Block Group's with seniors with known poverty status	Total	Below the Poverty Level	Percent	
Arlington	44				387		
Belmont	27	27					5.41
Cambridge	80			9,051		12.88	
Somerville	67				1,063		
Watertown	29	29	29	5,159	388	7.52	6.77
Overall	247			32,967		9.74	

Take a look at the last two columns and note that the percentage of impoverished seniors in a town is different than the town average of this percentage calculated for each block group.

1. Are they same or different?
2. What is the explanation for the discrepancy?
3. What about the direction (positive/negative) of these differences?
4. Which set of numbers are most appropriate for use in our analysis?

Most of these statistics should be calculated only for the populated block groups.

1. Which Block Group (or Block Groups) has no population?
2. Which Block Group (or Block Groups) has no seniors with known poverty status?

Hand in a discussion of the answers to these questions along with this table completely filled in.

Tips: Filling in the table looks simple enough -- just run some of those 'Field' commands for the requisite block groups. But one of the tricky parts of this assignment is figuring out which block group is within which town. The block groups can span a town (i.e.,

place) boundary. For this exercise, we define a block group to be in a town if the centroid of the block group is inside the town boundary. Using ArcMap's "Spatial Join" functionality sounds like a good idea, but with spatial-join, the polygon-over-polygon options don't allow 'centroid of one within bounds of the other. To facilitate your work, we have created a match-up table (**M:\data\blkgrp2t.dbf**) that specifies the containing town name for each of the relevant block groups. You can join the table using "BKG_KEY" field to the block groups layer to figure out which block group is within which town. If you are interested, the **blkgrp2t.dbf** table is created by doing a "Select by Location" operation for each of the five towns, with "have their center in" as the topological relation.. For example, the graph below shows the block groups that "have their center in" Cambridge. A new field "Town" is created and its value is set to be "Cambridge" for those block groups.

This method works when only a handful of towns are considered. If you want to do analysis on hundreds towns, you don't want to do it one by one by hand. There are more advanced tools to solve this problem such as creating a new shapefile that contains points representing the centroid for each block group in Middlesex County and then doing a spatial join between this new Centroid layer and the town layer. Unfortunately, creating the centroid layer from the blockgroup shapefile is overly complicated and involves using a VBA script to add the X,Y coordinates of the centroids to the attribute table, then saving these columns as a new data table, and then using the Tools/Add-XY-data option in ArcMap to make a point shapefile of the centroids. Doing all this is too complicated and distracting to be part of this homework.
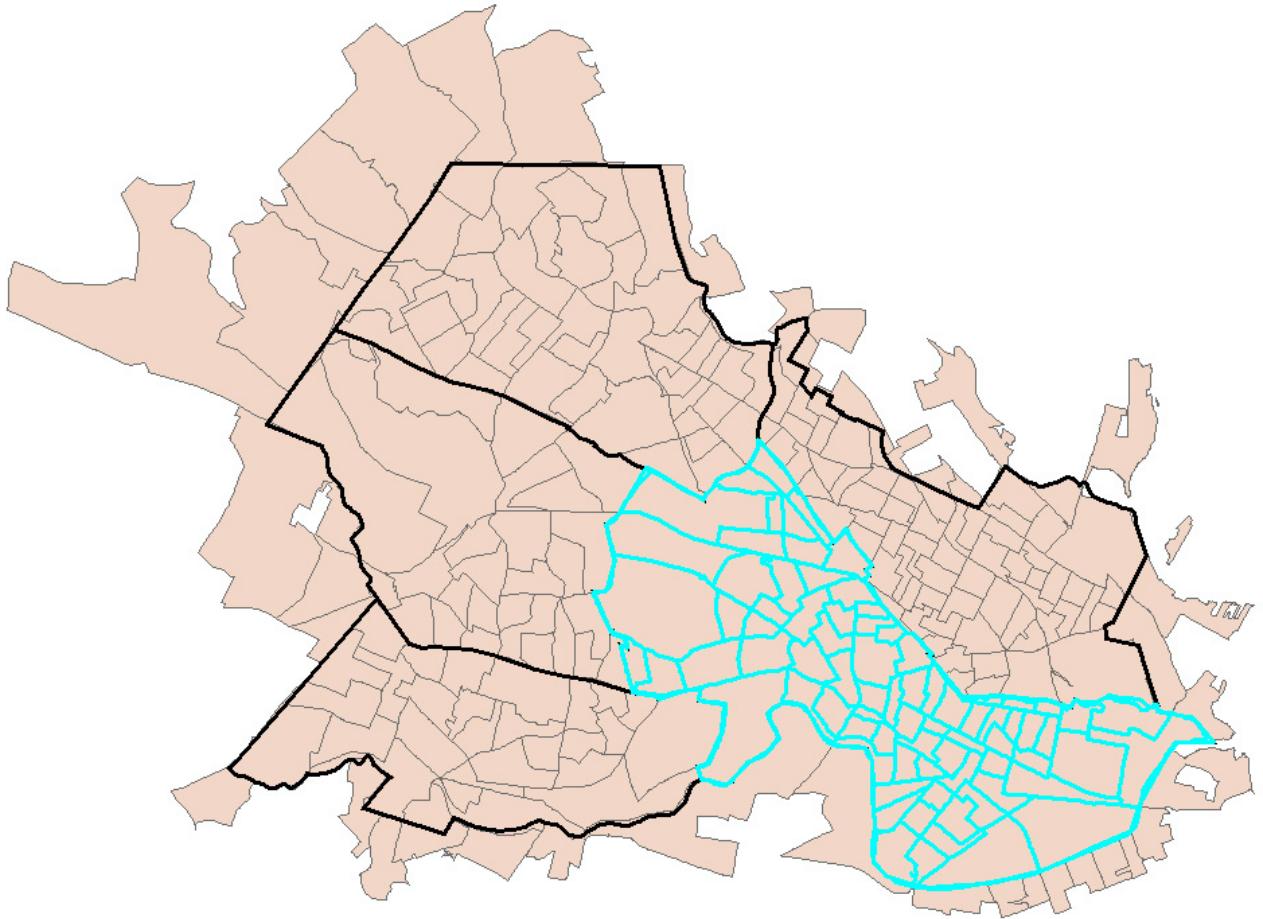
Question 2 [50 points]

Use ArcMap and the various layers that we have used in class exercises to undertake a basic site suitability analysis. Feel free to augment your maps with some other map layers stored in the class data directory.

First, ***prepare 3 maps*** showing the locations that are acceptable based on the criteria identified above. Map these criteria **separately**:

1. proximity to major roads and distance from the TRI Facilities.
2. appropriate land use characteristics
3. block groups having a high percentage of seniors below the poverty level.

Next ***prepare a fourth map*** that shows those Cambridge locations that meet all these criteria as well as the 1-continguous-hectare constraint.

Along with the maps, **provide a page or two** (*not* a treatise) **of discussion** concerning:

1. any choices or interpretations that you made in generating the suitable locations and which you feel bear some explanation, and
2. your conclusions from this initial analysis regarding suitable sites. Regarding your conclusions, don't just pick one site--there isn't a definitive 'best site' given the criteria we've suggested. Be sure to include some interpretation regarding the extent to which the analysis helps you focus on one part of town, on one or another criterion, on a set of proximity issues that might forecast special interest concerns/complaints, etc. Would you suggest tightening or relaxing some of the criteria? Can you suggest important considerations that are not well captured in this suitability analysis?

Hand in a discussion of the answers to these questions along with the four maps.

FAQ and other Suggestions

- You are required to put all the maps electronically (in PDF format) under your www directory *and* also turn them in on paper.
- You will be graded on the quality of *both* your analysis *and* your visual presentation.
- Doing a good job on this assignment will require that you look critically at the results you get. Using your common sense is a must!
- Think hard about the order in which you need to undertake various steps **before** you attempt them in ArcMap. You may save yourself considerable time depending upon the order in which you do the various overlay and site suitability steps. This homework contrasts with the labs in that we do not spell out the step-by-step procedures you need to follow to reach the result. Hence, this assignment will require more thinking on your part.
- This homework set is the major out-of-class assignment during the next few weeks and you should spread the work out during that time. We expect it to take between 15 and 20 hours total. Try a different approach, and reflect on what you've done and what it means. There isn't a single right answer to the site suitability question and there is more than one way to do many steps (as well as a few dead ends). If you do spread out the work, you'll find the exercise to be much more enjoyable.

Need Help?

Please don't be shy about asking for help if you are struggling with understanding the assignment or having trouble with a particular aspect of ArcMap. We've heard many stories of students struggling for many hours over minor issues that had easy solutions. We'd like to avoid these misadventures, so please contact us sooner rather than later if you get stuck. Also, we recognize the value of group work and encourage study group discussion but we want you to turn in individual work that reflects your own learning and hands-on discovery.