Massachusetts Institute of Technology
Department of Urban Studies and Planning

11.188: Urban Planning and Social Science Laboratory

11.205: Intro to Spatial Analysis (1st half-semester)

11.520: Workshop on GIS (2nd half-semester)

Homework 1: Mapping of Community Characteristics

Due (online) on Friday, March 12, 2021, before the start of extra lab - 12:30 PM.

For your convenience, you may want to download a PDF version of Homework 1.

NOTE: Homework assignments will take longer and be much more important to your grade than any one lab exercise, so devote your energy accordingly.

INTRODUCTION

In this exercise, you will explore the spatial patterns of the housing and socioeconomic characteristics of communities in and around Boston. To assist in this task, we provide: (1) demographic data at the census tract level from the 1990 US Census, and (2) boundary files for cities and towns, major roads, shopping center locations, and census tracts.

We ask you to use these data to prepare a short report with textual discussion plus three maps for Part I and two maps (with a table) for Part II.

Before starting the hands-on work, read through the entire assignment to get a sense of the datasets, analytic approach, and processing steps. Then, make sure you can access the datasets on Stellar. [*Look for 'Data for Homework #1' under the 'All Data' header in the 'Materials' tab.]*

DATA

- **Census Tract Boundaries**

  The census tract boundaries are saved in a 'shapefile' that contains only the boundary geometry and a few geographic identifiers (like county, track number, etc.). This shapefile must be joined with a dbf-formatted table in order to relate the census tract attribute data to specific census tract geometry. The census tract shapefile is called **`msa5_tr90.shp.`** It is located in the homework data folder and contains all the 1990 census tracts in the five Eastern Mass counties in and around Boston.

- **Socioeconomic Data for the Census Tracts**

  The socioeconomic data for these tracts have been pulled from the 1990 Census SF3A datasets and are stored in the same homework data folder in a dbf-formatted table called **`msa5_tr90_data.dbf`**. This file must be 'linked' or 'joined' to the attribute table for the tract boundaries by a common field called "**STCNTYTR**" before you will be able to generate thematic

maps using the census data. *(STCNTYTR is the abbreviation for STate-CouNTY-TRact.)* Use the QGIS help files to see how to 'join' the data table to the attribute table if you want to get started with the homework before we show you how to do this in class.

The **msa5_tr90_data.dbf** table includes 60+ variables from the much longer list of all variables in the decennial Census. Take a look at the dictionary for the specific census data fields in msa5_tr90_data.dbf. (**Note:** this list is a subset of the full Census Bureau's listing and technical documentation for the hundreds of population and housing variables from the 1990 census. This technical document is archived in the homework data folder as msa5_tr90_data_dictionary.txt. More details about the US Census data products are available here. For this exercise, you will need ONLY the shorter data dictionary of 60+ variables mentioned above in order to do the homework. Over the next few weeks we will have additional exercises using 2010 US census data. We use 1990 data for this homework as working with historical data is critical to being able to understand demographic changes over time. It is important to note that the socioeconomic data we use in this homework is, in part, no longer contained in the Decennial Census. The Census Bureau replaced one component of the Decennial Census (Summary File 3 commonly called the "Long Form") with another survey called the American Community Survey. When working with historical Census data you need to keep in mind that small-area comparison (tract, block group, or block) of census data across decades can be tricky due to changes in census tract and block group boundaries. MIT's Rotch Library has CDs from a third-party firm, GeoLytics, which has reconciled past census data to year 2000 and 2010 geographic boundaries. Should you wish to analyze trends in census data for your individual project later in the semester, you will likely want to use the GeoLytics CD and/or other online sources that we will help you acquire comparable data across the decades.)

- **Shopping Centers and Major Roads**

Besides the census data, which will be used primarily in Problem #1, you will need a map of major roads and shopping centers for Problem #2. The shopping center shapefile (for the Boston metro area) is called ***shopcntrs*** and is also stored in the homework data folder. The major roads layer is called ***majmhda1*** and can also be found in the homework data folder. All of these coverages use the following coordinate system: Massachusetts State Plane, Mainland Zone, NAD 1983, meters. **Be sure to set the map units and distance units in the ***Data Frame Properties*** window so you can measure distances in your Data View window and be sure that the distances you compute in Problem #2 are reasonable.**

Data Sources: The roads coverage comes from the Mass Highway Department via MassGIS and the shopping center coverage is proprietary data provided by SSR Research (circa 1995) for internal MIT educational use.

- **Building a project folder**

The datasets needed for this homework are manageable in size, so we suggest that, before you start work, you copy all the necessary datasets into a single working directory which you then save in your personal network locker and/or on a flash drive or USB drive. Then, you can copy the entire folder into C:\temp for local-drive work on the WinAthena and CRON machines and, when you finish a working session, you can copy the folder back to your locker or other device. In fact, it is good practice to keep two copies of the folder - e.g., 11.188_hw1a and 11.188_hw1b - so that your always have a recent backup of everything in case things go far wrong during a working session. Here is a summary of the datasets with their pathname (beginning at the top of the homework data directory):

- **.\msa5_tr90\msa5_tr90.shp** - eastern Mass census tracts (1990) in Mass State Plane (mainland) NAD83 coordinates
- **.\msa5_tr90\msa5_tr90_data.dbf** - selected census data for eastern Mass census tracts (1990)
- **.\msa5_tr90\msa5_tr90_data_dictionary.txt** - data dictionary for selected 1990 census tract variables
- **.\shopcntrs*shopcntrs*.shp (et al.)** - shapefile of eastern Mass shopping center locations
- **.\majmhda1*majmhda1*.shp (et al.)** - MassGIS+MassDOT shapefile of major roads in Mass
- **.\matown00\matown00.shp(et al.)** - MassGIS shapefile outlining Mass cities and towns
- **.\msa_water\msa_water.shp (et al.)** - MassGIS shapefile of Mass water bodies (convenient overlay for visualization)

SUMMARY

A map should always have a purpose. A good map should deliver the information that you want readers to understand. Therefore the map should be very intuitive without requiring reading of the discussion of the map in your paper or report. Try to give the map to your friends who have no training in GIS to see if they can recognize the message you were trying to deliver and ask them whether they find the evidence to be compelling.

- Your goal is to gain some understanding of housing and socioeconomic patterns in metro Boston (circa 1990!).
- Try to have fun and explore QGIS while doing this homework. Use the on-line help to experiment with QGIS's capabilities. Here's another mapmaking hint: you can specify different colors for your foreground and background features, you may want to turn off the outlines of polygons in a layer, and you may want to choose different widths and sizes for lines and symbols. Exploit this functionality when overlaying different data layers.
- To make your maps more interpretable, you can overlay the political boundaries of Massachusetts towns. Also consider using the water bodies in Eastern Massachusetts. The Mass boundaries are located in the data folder as **matown00.shp**. A shapefile of Eastern Massachusetts water bodies is available in the data folder as **msa_water.shp**.

Problem 1: Exploratory Mapping
Metropolitan Area Census Data (60 points)

1. [20 points] Create a thematic (or chloropleth) map showing the population density of the MSA.

- You should calculate density as persons (**P0010001**) per acre (**landacre**). Normalizing by the 'landacre' variable in the census data is more reliable than using the 'area' variable since the tracts extend into Boston harbor, the Charles River, etc. whereas the 'landacre' variable is the census estimate of **land** acreage within each Tract.
- Be sure to include only those polygons (census tracts) where **landacre** > 0 and the relevant census data are not missing. (For example, census data are missing from a cluster of tracts north of Boston.) Use the **Select > Select by Attribute** menu option to select the tracts you want to include.
- Classify the data into approximately five categories. The classification method you use is up to you, and can include customized category breaks. Experiment with the available methods and pick the one you feel best clarifies your exploratory interest (for example, do you want to differentiate within high density areas, or characterize the full range of densities?) ***In a sentence or two, explain your focus,***

your choice of classification, and why your cartographic technique makes sense.

2. [20 points] Map the homeownership ratio--the ratio of owner occupied housing units to the total occupied housing units.

- Remember that the '**tenure**' variables count the number of owner-occupied and renter-occupied housing units. Be careful about what the numerator is and what the denominator is.
- Just as for the previous map, you will also need to exclude tracts that lack adequate data. Again, **include a few sentences explaining your choice of classification scheme and any pattern that you want to show.**

3. [20 points] Map another Census attribute of your choosing with interesting spatial patterns using the same process as described in number two.

- For all three maps, be sure to include a title, source, logo, north arrow, scalebar, and legend (with indication of classification choice). The quality of your presentation **does** count!
- In separate text explain the "story" that your map tells. What spatial patterns, if any, are suggested? (Don't try to overly explain the map!) ***Explain the reasoning behind your classification choices, how you handled missing values, and any other relevant judgments and assumptions.*** Please limit this discussion to a maximum of ***one page.*** We do **not** want lengthy explanations. There is more than one reasonable choice for your classification and normalization choices. We simply want to be sure that you've thought about the issues and made reasonable choices.

Problem 2: Introductory Spatial Analysis
Relationships between Roads, Shopping Centers, and Residences

For this problem, you are asked to investigate the relationships among the locations of shopping centers, major roads, and residential clusters. After doing some exploratory mapping as you did in the first problem, you are asked to dig a bit further into the data, develop a few specific measures that carefully exclude incomplete or inapplicable data, and then develop maps that successfully visualize the results and the reasoning behind your analysis.

The shopping center data are stored in the data folder as **`shopcntrs.shp`**. (These data are proprietary and not to be used or redistributed for non-MIT purposes.) Included in these data are characteristics such as square footage of retail space (**totalsf** as a text field and **squarefeet** as a numeric integer) and type of center (**propertysu**).

Explore these two variables to try to determine if a relationship exists between them. To do this you may want to calculate the average size of each type of shopping center. Note that not all observations include a value in the **totalsf** or **squarefeet** fields. Note that these shopping 'centers' do *not* include places like Central Square (Cambridge) where commercial/retail activity is present among individually owned parcels and buildings along a city street. This dataset focuses on shopping center developments where a large tract of land or strip mall under common ownership is divided up into clusters of businesses.

1. [20 Points] Create one map showing the relationship(s) between shopping center location and the location of major roads and population centers.

- Be sure to use different symbols and/or sizes on your map for the different types of shopping centers. Likewise differentiate major and not-so-major roads -- use the **class** field in the ***`majmhda1`*** attribute table.
- Use high population density as an indicator of where population centers are located and shade the tracts based on population density (as you computed it in problem 1)

- Be sure to turn the tract outlines off so they don't clutter the map. In fact, it will take some effort to develop maps with good symbolization and cartographic choices so that they are both readable and informative.

2. [20 Points] Buffer the major roads and create a second map that examines whether certain types of shopping centers tend to be inside the buffer.

- Select the Interstate Highways and Routes 2 and 3 (not 3A) from the major roads layer.
- Use the buffer tools to create an 800 meter buffer around these selected roads and then calculate the share of shopping centers by type (among those within the 5-county region) that lie within the buffer.
- Map your results and include in your "layout" a map of the entire area, as well as a more detailed map zoomed into an area of the region with interesting spatial patterns.
- In addition to your two map views, create and display a **table** showing, for each type of shopping center (**propertysu**), the number of such shopping centers within the 5-county **msa5_tr90** region, and the number and percentage of each type of shopping center that falls within the buffer areas.

  Explain in a couple of paragraphs, separate from the maps, (a) the steps you took to select those roads and shopping centers that you included when computing your statistics, and (b) your interpretation of any general pattern that you observe regarding the location of shopping centers, major roads, and population centers. In particular, be sure that your discussion covers:

  - What conceptual relationships are these maps intended to portray?
  - Which classification schemes did you choose to use, and why?
  - Which shopping centers and tracts have you excluded from the analysis, and why?
  - Do there appear to be any interesting spatial relationships shown on the map?

Homework Requirement

Don't just turn in the maps! You should turn in a short report that integrates the maps and tables together with the explicit answers to both questions. Use the maps and tables in the paper to illustrate and amplify your verbal reasoning rather than simply to produce maps without a stated context and purpose.

Please submit your homework (in PDF or WORD format, preferably PDF) using the Stellar homework turn-in capability at https://stellar.mit.edu/S/course/11/sp21/11.188/index.html.

Created and Modified:1993-2015 by Raj Singh, Thomas H. Grayson, Annie Kinsella Thompson, Joseph Ferreira, Myounggu Kang, Jschung\ Chung , Jinhua Zhao, Mike Flaxman, Yi Zhu, Lulu Xue, Shan Jiang, and Eric Schultheis

Last modified: *February 26, 2021 by rounaq*

Back to the CRON Home Page.
Back to the 11.188 Home Page.