

11.520: A Workshop on Geographic Information Systems

11.188: Urban Planning and Social Science Laboratory

Lab Exercise 5: Working with 2000 Census Data & MIT Geodata Repository

Due in Lab 6.

Overview

In this exercise you will use the Census STF3 files to create a thematic map indicating the percentage of workers in block groups in Middlesex County who drive alone to their jobs. You may find the notes from the "Making Sense of the Census" lecture a useful reference.

Your goal is to create a table that contains the percentage of workers who drive alone to work (by car, truck, or van) for each block group in Boston Metropolitan Area. Note that unlike the simple example shown in class, to complete this task you must *normalize* the data. You may want to look at the Lab #5 notes for more discussion of normalizing and selecting the appropriate "universe" variable.

The geographical layers used in this lab are available from the MIT Geodata Repository, instead of using the data from M: drive as we usually do.

Necessary Resources :

View the "Making Sense of the Census" lecture notes for background information on this topic.

Summary File 3: Technical Documentation ([PDF](#))

- Chapter 3: Subject Locator

This index is designed to quickly identify the STF3 tables you may need. Each table variable is listed alphabetically.

- Identification Section : Table (Matrix) Section

Use this section to find the data dictionary number for the STF3 table that you are interested in.

- Chapter 2 : How to Use This File

One of the most important parts of this chapter is the explanation of how to extract the correct geographic summary level.

2000 Census Variable Locator

- Once you find the census variables you are interested you can use this file to locate the 2000 Census data file that holds your variable.

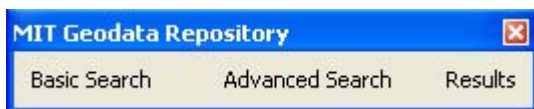
Hints

The find tool may come in helpful when trying to find a variable in the *2000 Census Technical Documentation*. Search for keywords like "Means of Transportation to Work" and "Persons 16 years and over".

Part I: Getting data from MIT Geodata Repository

We are going to use two geographical layers from MIT Geodata Repository-- Massachusetts Block Group 2000 and Massachusetts County Boundaries 1991.

First we need to turn on the MIT Geodata Repository Toolbar in ArcMAP. Right click in the toolbars areas in the ArcMap window and choose "MIT Geodata Repository Toolbar" in the context menu.

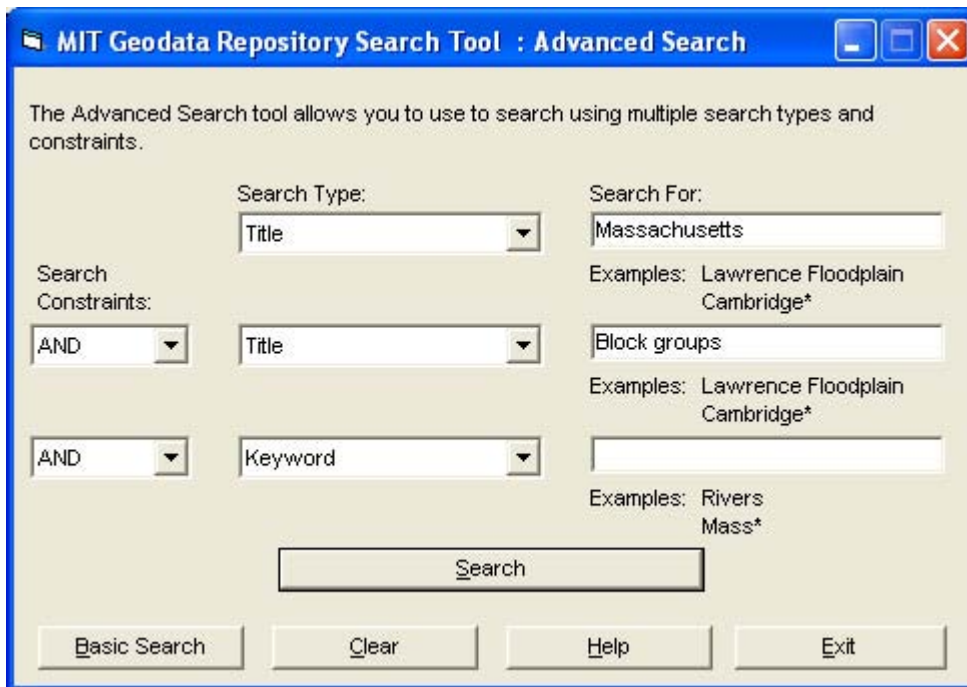


Click on "Advanced Search", and a Login window should show up as follows,

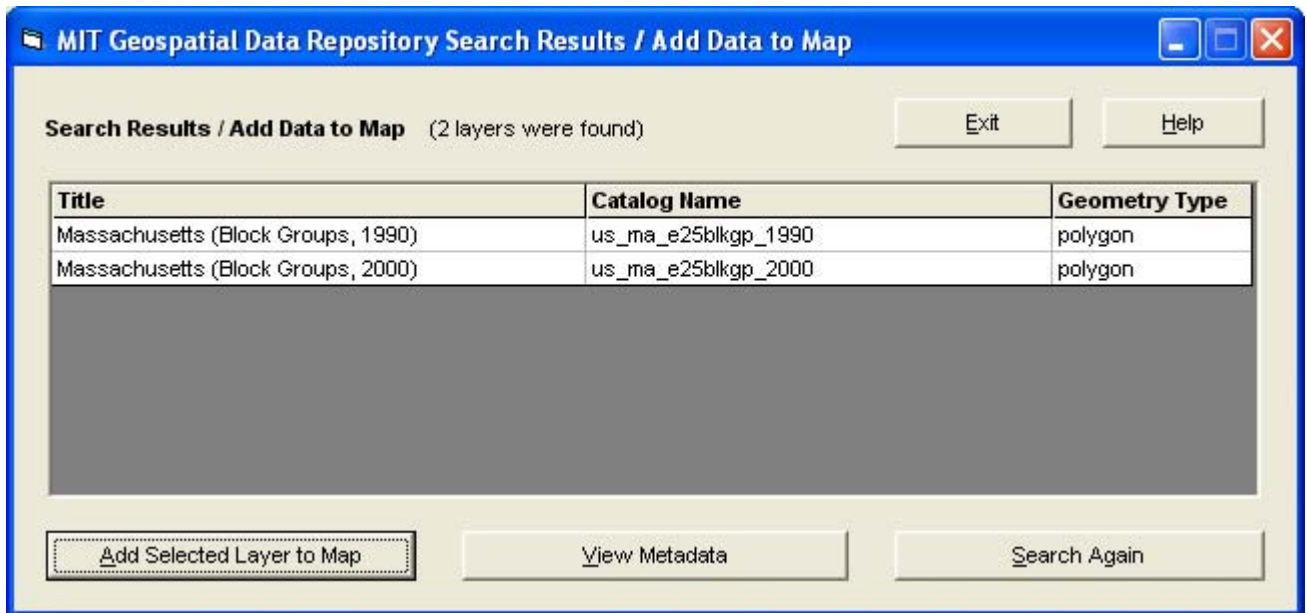


You need to register for an account in MIT Geodata Repository before you can use it. Click on "No Account? Register Now" and follow the instructions in the pop up Internet Explorer window to create a user account.

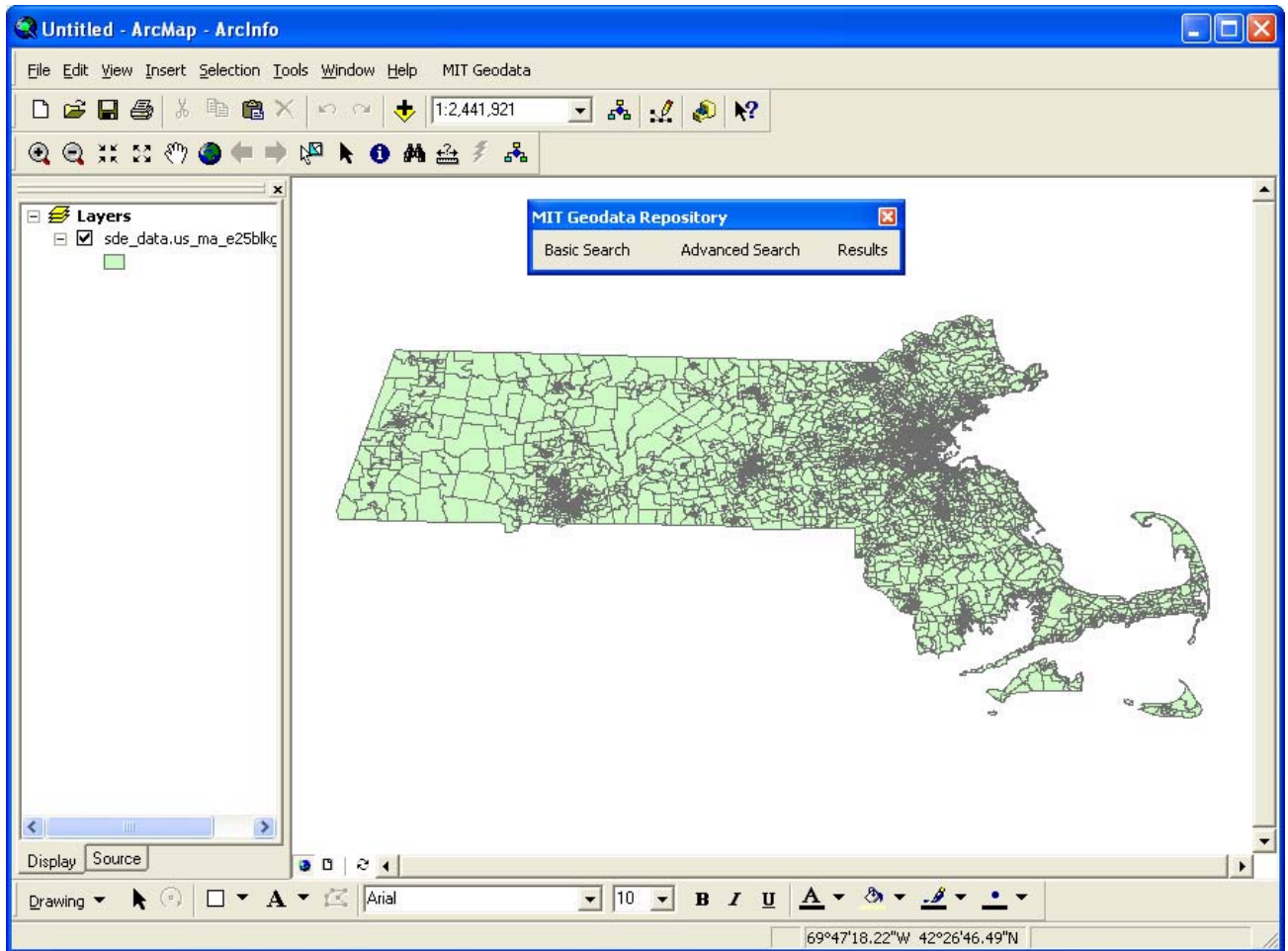
After the registration, log in using your Kerberos ID and the Geodata Repository Password you choose in the registration and click on "Login to server".



In the Advanced Search window as above, choose "Title" and Search for "Massachusetts" and "Block groups". Click on "Search" to get the search result as follows,

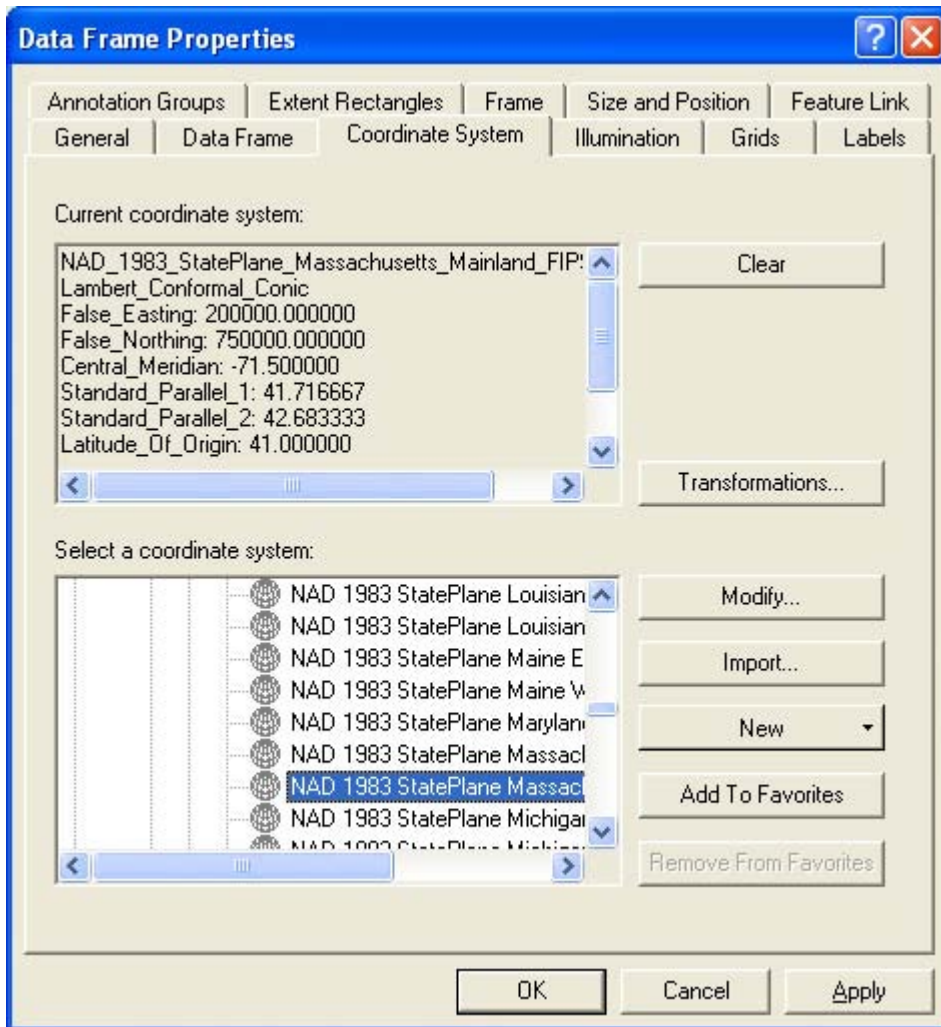


Click on the second row (Block Groups 2000) and click "Add Selected Layer to Map". It will take a while for the layer to be added from the MIT Geodata Repository server. The ArcMap window now should look like,

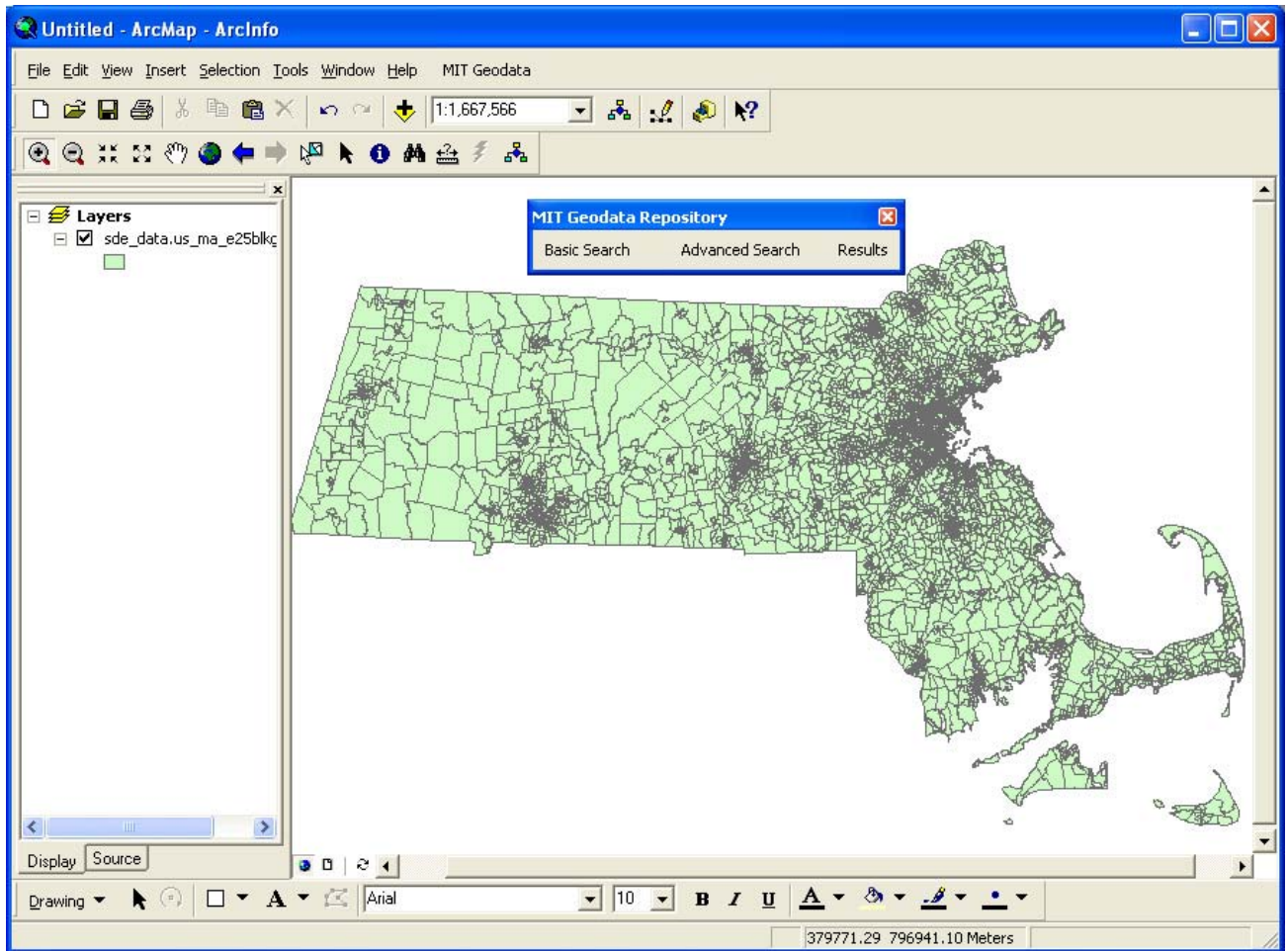


Part II: Set the Coordinate System

The Block Groups layer is not displayed in the correct projection. Change the projection to "State Plane / NAD 1983 / NAD_1983_StatePlane_Massachusetts_Mainland_FIPS_2001" from the "Coordinate System" tab in the Data Frame Property window. In "Select a coordinate system" window, expand "Predefined", then "Projected Coordinate Systems", followed by "State Plane", "NAD 1983", and "NAD 1983 StatePlane Massachusetts_Mainland_FIPS_2001".



After the correct coordinate system is specified, ArcMap will redisplay the Block Groups layer as follows,




Add in the "Massachusetts County Boundaries 1991" layer from MIT Geodata Repository in the same way as we add the "Block Groups" layer.

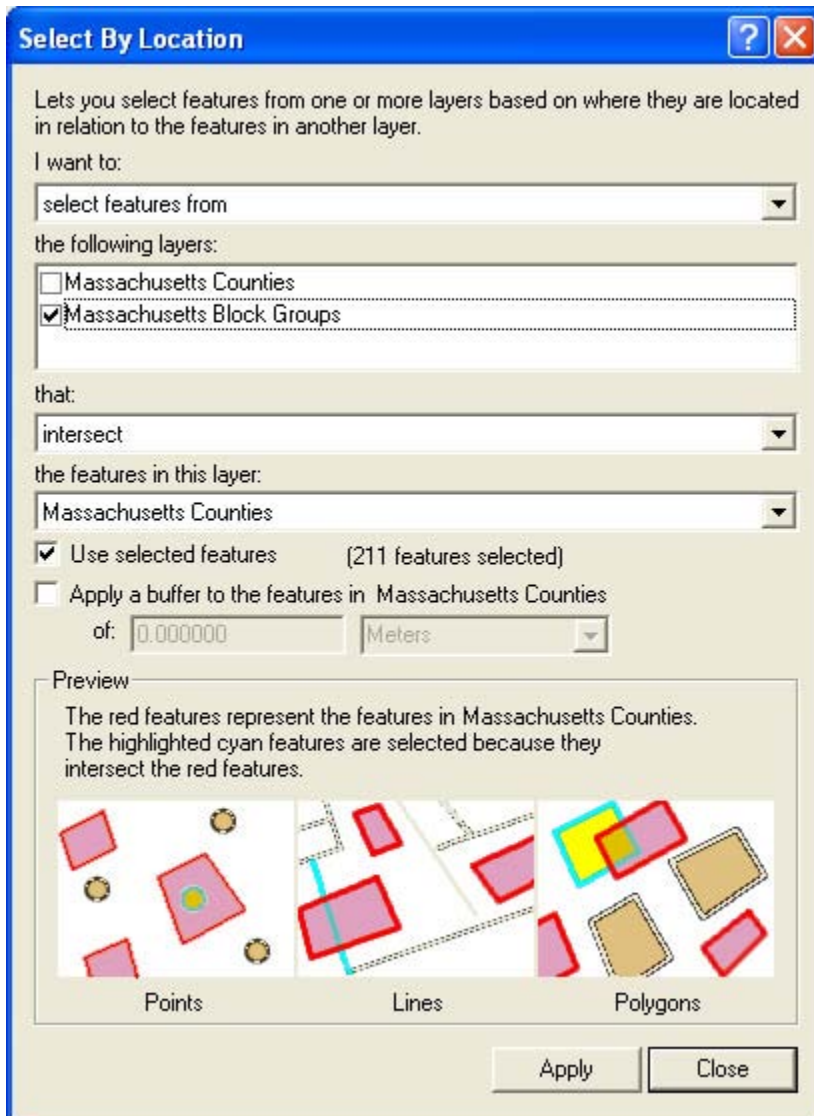
Part III: Select the Block Groups in the Boston Metropolitan Area and Export Shapefile

In order to work with a smaller file, we want to create a new Shapefile that represents the Block Groups that fall within major Boston area counties (Essex, Middlesex, Worcester, Suffolk, Plymouth, Norfolk, Bristol, and Barnstable). We can select these block groups using the **Select By Location** tool in ArcMAP.

STEP 1. In the Massachusetts County Boundaries layer, select Essex, Middlesex, Worcester, Suffolk, Plymouth, Norfolk, Bristol, and Barnstable counties. If you are familiar with counties in Massachusetts, you can simply use the **graphical selection tools**

 to select the eight counties. If you aren't, use the **Selection > Select By Attribute tool** to select these eight counties from the county layer.

STEP 2. With the counties selected you are ready to use the **Select By Location** tool. This tool works on the currently active theme, so you will need to make the Massachusetts Block Groups 2000 layer is active. Go to the **Selection** menu and select **Select By Location**. When the dialog window appears, choose to select features from your active layer (Massachusetts Block Groups 2000) that **Intersect** with the selected features of the U.S. Counties layer. Refer to Lab 3 for how to use the **Select By Location** tool.



STEP 3: All the Block Groups that intersect with Essex, Middlesex, Worcester, Suffolk, Plymouth, Norfolk, Bristol, and Barnstable should now be highlighted. You are ready to

create a new file based on these selected features using ArcMap's **Data > Export Data** tool (Refer the images below). Call this new layer **b_blkgrp.shp**, save it to your working directory and add this layer into ArcMAP.

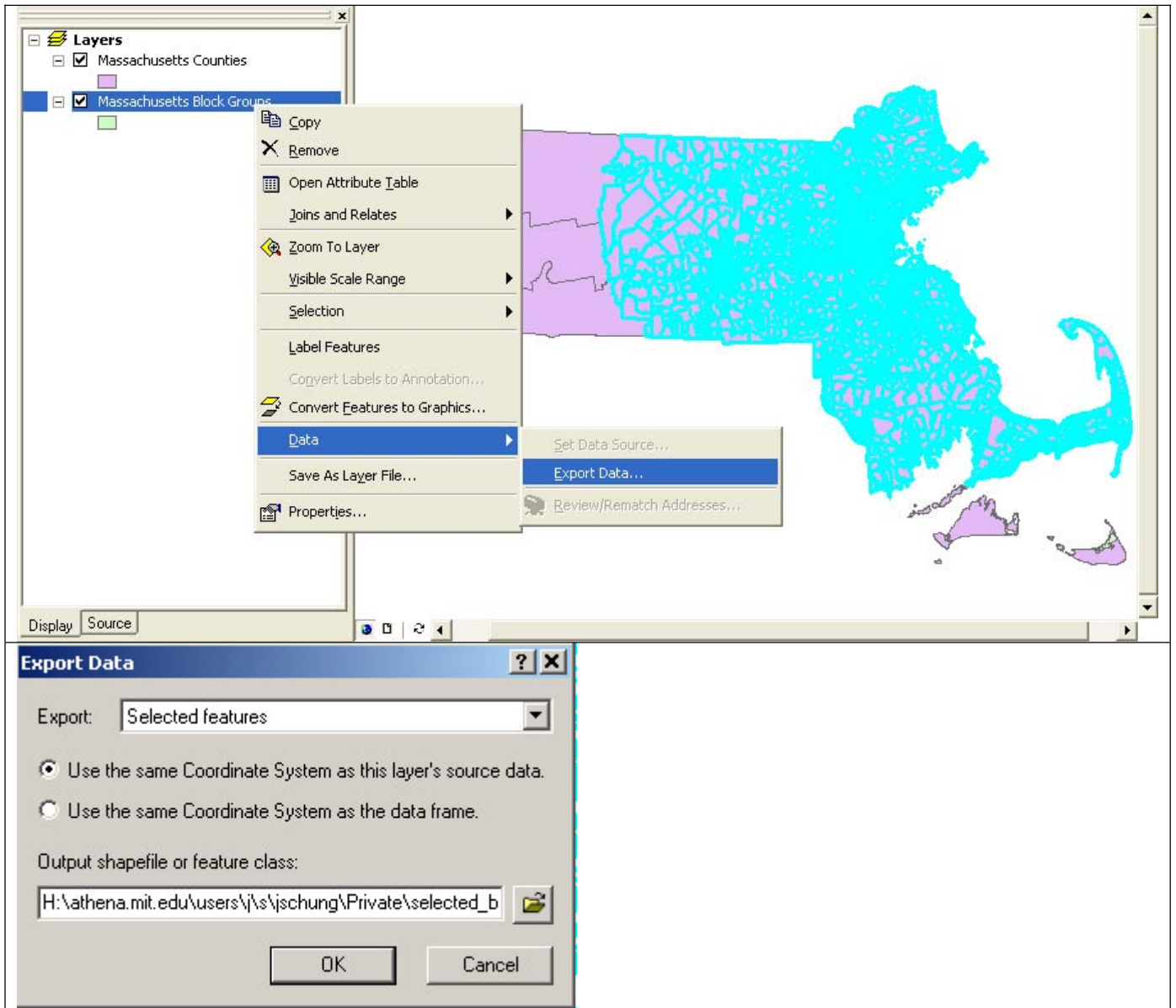


Fig. 1. Export to a new shape file

STEP 4: In your data frame, remove the two layers from MIT Geodata Repository. We no longer need these file and their size could affect processing speed. From this point you will only need to use the **b_blkgrp.shp** layer.

Part IV: Working with Census Data : Finding The Variables

(Note: Census Bureau provides similar content, which is listed in: <http://www.census.gov/support/SF1ASCII.html>. Please choose your preferred instruction for use.)

STEP 1: To find the corresponding variables in the Census and locate the correct ".uf3" file.

Open Summary File 3 : Technical Documentation ([PDF](#)) and search the key word "means of transportation to work". On Page 439, you will find "**P30. MEANS OF TRANSPORTATION TO WORK FOR WORKERS 16 YEARS AND OVER [16]**". There are 16 columns in this section, in which P030001 is the total workers 16 years and over--the universe, and P030003 "Drove Alone" is the workers who drove to work alone. So the fields P030001 and P030003 are the two of our interest in this lab. Then we have to identify the ".uf3" file which contains these two fields.

On Page 12, 13 of the Summary File 3 : Technical Documentation ([PDF](#)), the fields matrix numbers and their corresponding ".uf3" files are listed. We can see that the file "st00003.uf3" contains the fields from P25 to P37. The P030001 and P030003 of our interest are in this file. The variable locator of year 2000 is linked here: [2000 Variable locator](#):

STEP 2: Copy the access template file, the geographical header file, and the data file "st00003.uf3" to your own working directory H:\11.520\lab5

The st00003.uf3 file for Massachusetts is **ma00003.uf3**, available at http://www2.census.gov/census_2000/datasets/Summary_File_3/Massachusetts/. For your convenience, it has been downloaded to our class data locker: **M:\data\census2k**.

The geographical header file for Massachusetts is **mageo.uf3**, available at http://www2.census.gov/census_2000/datasets/Summary_File_3/Massachusetts/. It has also been downloaded to our class data locker: **M:\data\census2k**.

The access template file "**sf3.mdb**", is available at <http://www.census.gov/support/2000/SF3/>. It has also been downloaded to **M:\data\census2k**.

In order to operate on the three files, you have to copy them from **M:\data\census2k** to your own working directory such as **H:\11.520\lab5**.

STEP 3: Import ma00003.uf3 and mageo.uf3 to sf3.mdb

Before you import **ma00003.uf3** and **mageo.uf3**, change their extension name to ".txt" so that MS Access can recognize them. Open the access file **sf3.mdb** in your own working directory, you will see dozens of tables empty but with table structures setup.

Then click Menu File-->Get External Data-->Import.

In the Import window, change the file type to ".txt" files, browse to your working directory and find the "mageo.txt" file.

In the Import Wizard window, click "Advanced". After the Import Specification Window pops up, you click "Specs." And in the Import/Export Specification window, scroll to the end to find "SF3GEO Specification" and click Open. Then click OK to return to the Import Text Wizard and click "Finish". The sequences are illustrated as below. (Note that the check boxes on the first few graphic below don't matter since loading in the "SF3GEO Specification" will change the file format type from however it starts out to become 'Delimited' and the code page choice will become '...ASCII'. That is what you want to match the format of the text files.)

Once the "mageo.txt" file is imported, a new table "Mage" will appear in the Access window. You need to import "ma00003.txt" in the same way except using "SF30003 Import Specification".

STEP 4: Build a query to construct a table that contains the percentage of workers who drive alone to work for each block group in Boston Metropolitan Area

Click on "Queries" in the left pane of MS Access and choose "Create query in design view".

Now the query window looks like the following.

Follow the instructions below to complete the query.

1. Double Click "SUMLEV" in "Mageo" and enter "150" in the Criteria box since the summary level is block group.
2. Double Click on P030001 and P030003 in "Ma00003" so that both fields show up in the bottom tables.
3. In one empty column of the bottom table, right click and choose "build". In the expression builder, type in "Ratio: [Ma00003]![P030003] / [Ma00003]![P030001] " and click OK.

4. In another empty column of the bottom table, right click and choose "build". In the expression builder, type in "BLK_ID: [Mageo]![STATE]+[Mageo]![COUNTY]+[Mageo]![TRACT]+[Mageo]![BLKGRP]" and click OK.

5. Link the two tables by the field LOGRECNO.

The final query design should look like the following,

6. Under Menu Query, choose "Make-table query", enter the table name "**BlockGroupDriveAloneRatio**" and click OK.

7. Under Menu Query, click "Run!". Two message windows will pop up as follows, click **Yes** to confirm both. (Note that, when you use summary level = '150' rather than '090' in the Access query, you will paste 5053 rows rather than 6728 and the second warning message will not appear. Summary level '150' has a single row for each block group. Summary level '090' may have more than one row per block group - it has one row for each unique block group within a 'place'. Since place boundaries sometimes divide block groups into two or more parts, those splits block groups will show up more than once in the summary level = '090' rows.)

STEP 5: Clean up the MS Access database.

A new table **BlockGroupDriveAloneRatio** has been created from above steps. This table has all the information we need to continue the following lab. In order to save disk space on your own directory, you can export the table to another database and delete the "sf3.mdb" following the steps below.

In Windows Explorer, browse to your working directory, create an empty MS ACCESS file and name it as "DriveAloneRatio.mdb". Then in the left pane of MS Access, click on "Tables". The **BlockGroupDriveAloneRatio** table should show at the top of the list. Right click on the table, choose Export.

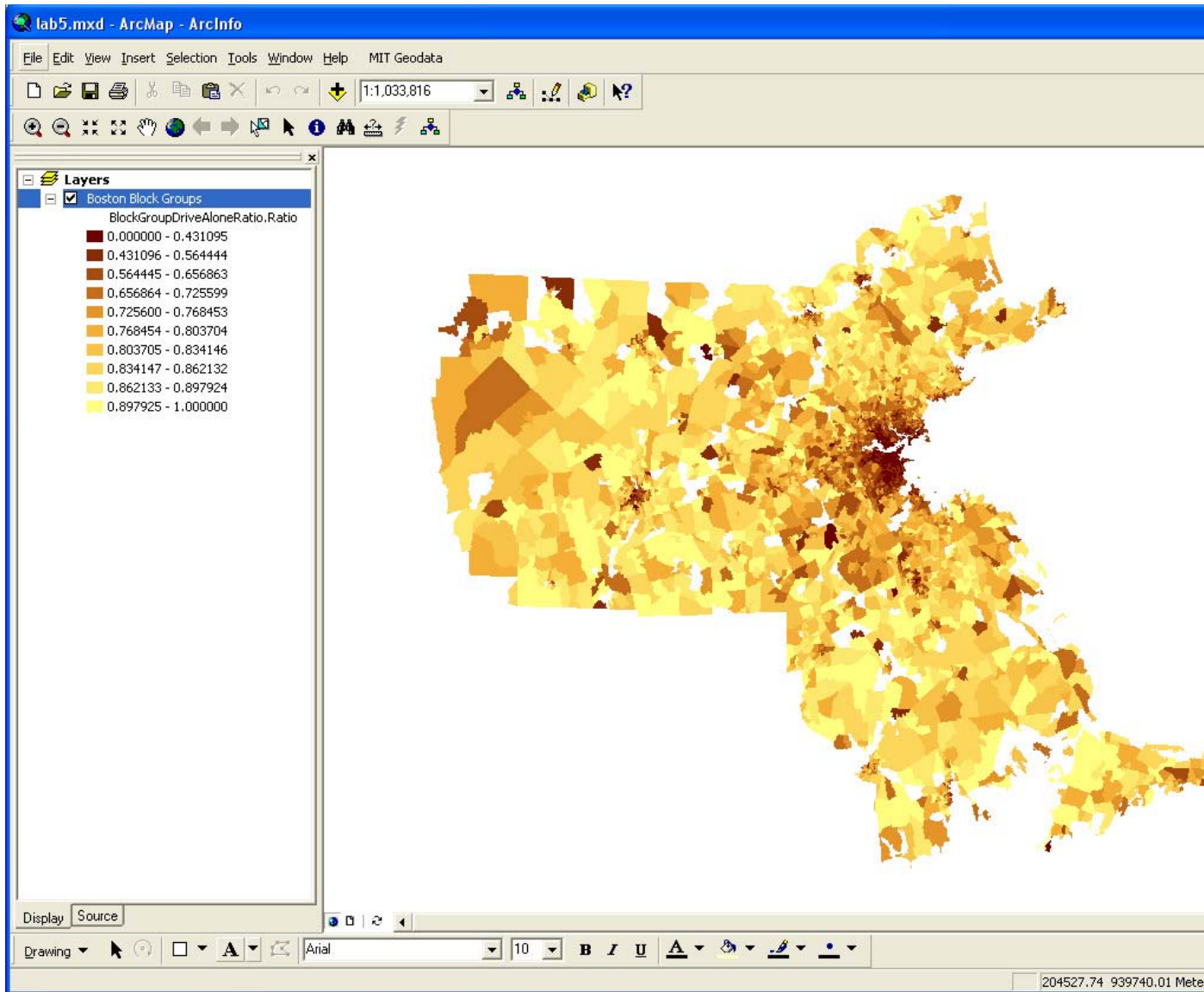
Browse to your working directory, select the database "DriveAloneRatio.mdb" and click Export

Click OK to confirm the export and then close the MS Access "sf3.mdb" file. In the Windows Explorer, you can check the file size--the sf3.mdb is about 55mb while the DriveAloneRatio.mdb is only 504kb. Open the file DriveAloneRatio.mdb to check if it contains the correct data. Delete the file "sf3.mdb" to save disk space if you want. Later in the lab, we only need to use the file "DriveAloneRatio.mdb".

Part V: Join the BlockGroupDriveAloneRatio table from "DriveAloneRatio.mdb" to the Block Group Layer in ArcMAP and Create a Layout.

In ArcMAP, add in the table **BlockGroupDriveAloneRatio** from the MS Access database **DriveAloneRatio.mdb**. Join this table to the **Boston Block Group** Layer using the field **BLK_ID** in the **BlockGroupDriveAloneRatio** table and the field **"BLK_KEY"** from the **Boston Block Group** Attribute table.

Make a thematic map showing the ratio of drive-alone workers by block group using **"Ratio"** field in the table (Choose a reasonable symbology method). The map should be similar to the following one (but with fewer no-data polygons. The map that is shown was generated using summary level = '090' and then ignored those blockgroups that appeared more than once in the table as a result of being split by 'place' boundaries).



As always, include the appropriate cartographic elements in your map and create a layout.

Export the final layout into a ".pdf" file and save it..
