

January 2007

Volume 7, Number 1

Accessing and Visualizing Census Data using ArcGIS

by Z. Aslıgül Göçmen, Ph.D.

This publication intends to provide guidance for individuals who might be interested in accessing census files and using them to visualize their communities based on census variables. Census data that are available in the format for use with ArcGIS can be accessed through various sources. Environmental Systems Research Institute (ESRI), the company that provides ArcGIS, is one of the sources that host files necessary to perform visualization and basic spatial analyses with census data. A more comprehensive reference to (population and housing) census data can be obtained from the U.S. Census Bureau (www.census.gov).

Some Basics about Census Data

Census data are geographic data that encompass both boundaries of places and the characteristics of population and housing within them. In other words, census data include both spatial and attribute data that are essential to the functioning of any Geographic Information Systems (GIS). Spatial data refer to the graphics or the shapes that define geographic entities (i.e., delineated census units), whereas attribute data refer to the characteristics of these geographic entities (i.e., any information that describes the population and housing within census units). One of the things that make GIS powerful as an analytical tool is its ability to link spatial and attribute data together, and visualize and analyze attribute data for different communities or other geographic entities.

In the broadest use of the term, Census data consist of the count and description of population and housing in delineated census units that are gathered every ten years. These units are generally based on the number of people living within them in the previous census counts or on major physical and man-made features within an area. The data on individuals and housing units are reported by various geographic units including census blocks (smallest geographic unit census data are reported), census block groups, census tracts, counties, and statistical areas (e.g., metropolitan statistical area). Take note that certain information is not provided at all levels in order to protect individual privacy. A good example is that household income is reported in census block groups and census tracts, but not in the smallest unit: census blocks.

The steps below are intended to guide the reader to access spatial and attribute Census data, and also relate these two types of data for visualization.

Accessing Census Data on the Web

As mentioned, census data can be accessed on the web at no cost through different venues. We will concentrate only on one of them: ESRI. To access census tracts for Dane County (or any county of your choice):

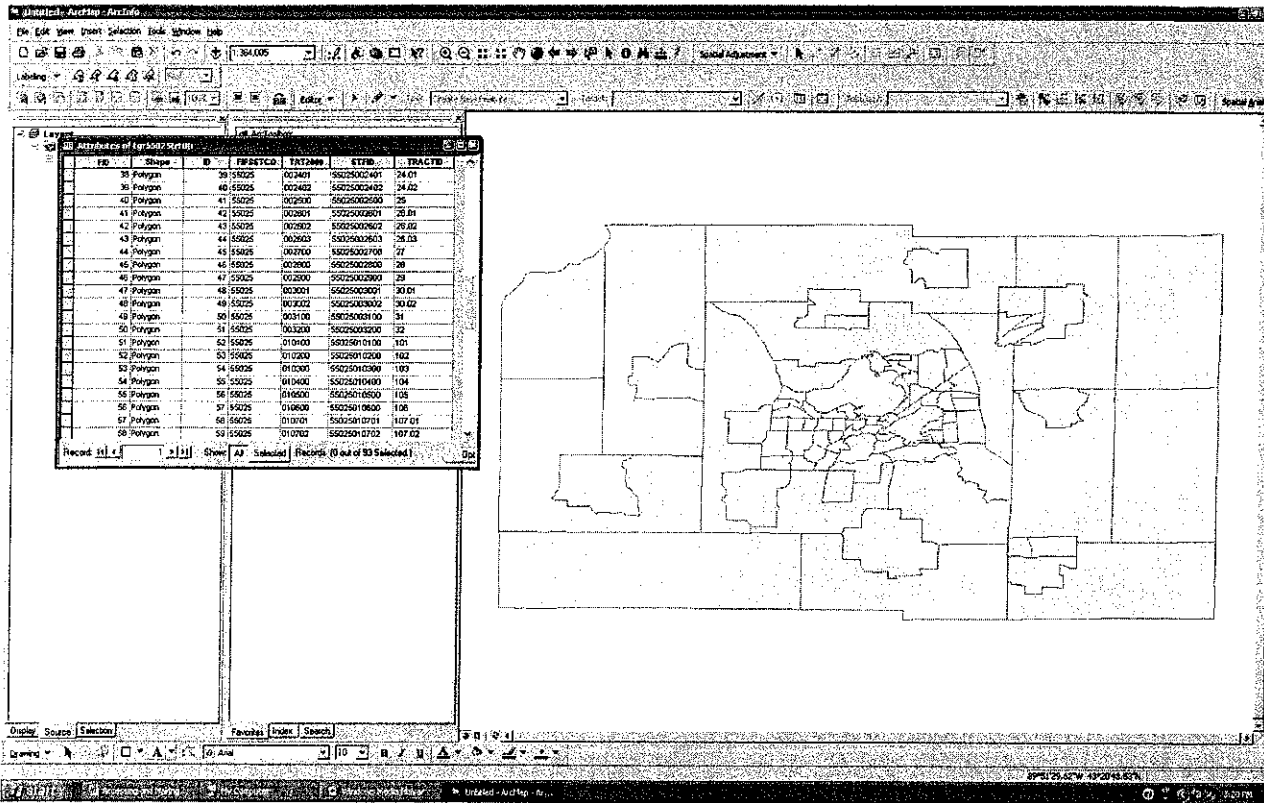
1. Go to http://www.esri.com/data/download/census2000_tigerline/index.html
2. Click on "Preview and Download" on the left.
3. Click on Wisconsin on the map.
4. Select Dane County in "Select by County" and click "Submit". You may choose another county, however, please note that the directions below refer to the Dane County files.
5. You will notice that there is a range of spatial files that you can download. We are interested in census tracts, so click the box next to "Census Tracts 2000" and also "Census Tract Demographics SF1" under "Available Statewide Layers", and then click "Proceed to Download". We need both files since one of them provides the demographic information we need to visualize and analyze the spatial data.
6. Click on "Download File". If you are given a chance to select where to download, make a note of it. If you are not given a choice, it is likely that the files were downloaded to your desktop.
7. You will notice that these files are provided in compressed formats. You will need to extract from these files to be able to use them. As you will have downloaded multiple files, you may need to extract more than once. So, first open the folder that the zip files are located.
8. Click to open one of the zip folders and you will notice the option to "extract all files" on the left under folder tasks. Please note that Microsoft Explorer may provide different options depending on the different versions it uses. Click on "Extract all files". Once you finish extracting these files to the desired folder, you will see the destination folder pop up with the new files added. You will also notice that these files are still in a compressed format, which suggests that you will have to repeat the "extract" operation again. You will need to repeat these steps for the second set of files that you have just downloaded.
9. You should be all set to use the census data. Thus, you can proceed by starting ArcGIS.

Visualizing Census Data in ArcGIS

The following steps will guide you through how to visualize these downloaded census files in ArcGIS.

1. Start ArcMap with a "New empty map". Once you get an empty map view, click on the + sign on the tool bar (add data) and then browse to the folder you have downloaded the census files into. (If you cannot find your files in the folder or access the folder, click on the "Connect to Folder" icon, which is the second from the left.) Add both the `tgr55025trt00.shp` and `tgr55000sf1trt.dbf` files (in either order). The `.shp` file stands for "shapefile" indicating that it is a spatial file, whereas `.dbf` file stands for "database" file, indicating that it is an attribute file, without graphic components. Please note that the shapefile also has its own attribute data file.
2. Once you add the files to your view, the shapefile will display. In order to see the attributes within this file, right click on the layer on the left side (where it says `tgr55025trt00`) and click "Open Attribute Table". (Note that the window on the left where the data files are listed is called the "Table of Contents" window.) You will notice that the table has 6 different attributes (columns) and 93 records (rows). You will also notice that the attributes do not contain any population and housing data, but rather basic identification information for different tracts (Figure 1).

Figure 1: Attribute table and shape file for Dane County census tracts layer

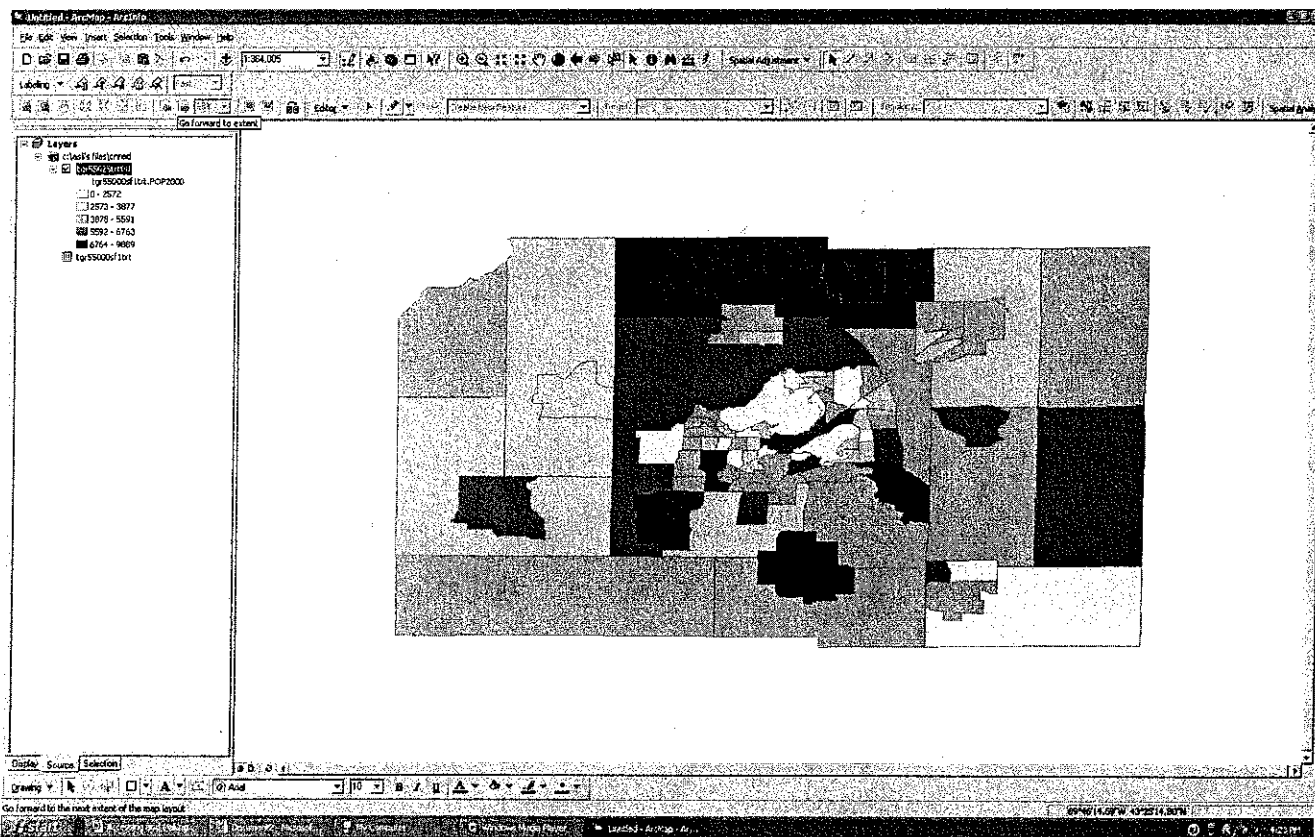


- Without the information on housing and population, the census shapefile is not very useful to display. The second downloaded file (tgr55000sf1trt.dbf) contains the information that can be visualized. In order to see what it contains, repeat the steps above (right click on the tgr55000sf1trt in the table of contents in the left side and click "Open Attribute Table"). You will notice that the file contains many more attributes (columns) and many more records (rows). In fact, there are 1333 records, in other words, 1333 census tracts. The discrepancy between the number of census tracts within this file and the previous one is because this new file contains census tract information for every census tract in every county in the State of Wisconsin and not Dane County alone.
- Visualization of the information within the state-wide data file in a map format requires an establishment of a relationship between these two attribute files. ArcGIS uses a relational database management system, which allows the establishment of such a relationship. The type of relationship we want to create is called a "join", and more specifically, a "join by attribute". A "join by attribute" requires that each table has a unique and common attribute field to be joined to each other. It also requires that each record be unique. These requirements allow the establishment of a one-to-one relationship among different data tables. The ability to form such relationships is another strength of GIS, which allows the visualization and analyses of external datasets as long as they can be matched with spatial information found in the GIS data layers. In order to perform the "join by attribute" between these two data tables, right click on the shapefile in the Table of Contents window and click on "Joins and Relates", and then "Join".
- A "join data" window will appear. Make sure to select "Join attributes from a table" and proceed as following: in 1). Select STFID as the field that join will be based on; in 2). Select tgr55000sf1trt with which the attribute information will be joined from (you will not necessarily need to do anything with the checkbox below); and in 3). Select STFID again. Then, proceed by clicking OK. STFID stands

for the identification number for any census tract, which is determined by the U.S. Census Bureau. Remember that we need to have a common attribute in each table to perform a join based on attribute information of the external database file to the attribute table of the shapefile. Here, the common field is STFID.

6. Click “yes” to creating an index.
7. Re-open the attribute table of the shapefile to check that the join has worked properly. If all the information about population and housing is displayed in the shapefile’s attribute table, then the join was successful and the information can be visualized in a map.
8. Close the attribute table. Right click again on the shapefile in table of contents and click “Properties”.
9. Click on the “Symbology” tab in the layer properties window.
10. Select “Quantities” on the left side, and “graduated colors” within quantities. In the “Fields” area, select POP2000 (or tgr55000sf1.trt.POP2000) for “Value”, and accept the default settings. Click OK.
11. Enjoy the new census map (Figure 2) you have created!

Figure 2: Dane County’s Population by Census Tract in 2000



Please note that the steps described above utilize ArcGIS 9.1. Also, note that there are many considerations that go into analyzing and visualizing census data. Those will be discussed in the upcoming publications.

Aslı Göçmen is an Assistant Professor and UWEX GIS Specialist at the UW-Madison Department of Urban and Regional Planning. She can be reached at: gocmen@wisc.edu or 608-265-0789.