

## ***Tufts University GIS Tutorial***

# **Working with 2000 Census Data in ArcGIS: Using the US Census Bureau Web Site for SF1 and SF3 Data**

Barbara Parmenter, PhD  
Revised October 2, 2007

### **Overview**

This tutorial shows you how to download census tract boundary data sets and associated SF1 and SF3 data sets for any county in the US, and to process them so as to be able to map the data in ArcMap.

Census GIS data and demographic/housing data are available from a number of sources, the US Census Bureau being the primary source. State, regional, and local agencies may also provide access to this data in various formats. Some of these formats are impossible to bring into a GIS (e.g., a .pdf file), others may need importing or some other manipulation and some may already be provided in a usable GIS format. In this tutorial, we use free Census GIS data from the Geography Network web site run by ESRI. This data is already in ArcGIS format. We then show how to access Census database tables from the US Census Bureau web site and join these to the ESRI GIS data layer with a small amount of data manipulation.

Before you begin this tutorial, please take a look at the US Census Bureau web site ([www.census.gov/](http://www.census.gov/)). This should be your first source for questions about the census, data, and data documentation.

### **Objective**

- Learn how to map SF1 data census tracts, block groups, blocks, or other level of tabulation data for any part of the US
- Learn how to download SF3 data at the tract level from the Census and to join this to tract boundaries for any part of the US
- Modify tables in Excel to bring them into ArcGIS
- Join tabular data to a geographic layer for mapping and analysis

### **Prerequisites**

You should already know how to add data layers, symbolize data, and create a map in ArcGIS. If you don't know this, go through the tutorial *Creating a Map with ArcMap*. It would be quite helpful if you have already gone through the tutorial *Working with Census 2000 Data from MassGIS* which contains an overview of the census and additional information on mapping. This tutorial uses ArcView 9.2 software.

## Getting Census Boundary GIS Data Sets

To make a map of census data, you need a GIS data set showing census boundaries plus a table or tables of the census data itself. In this tutorial, we are going to use tract boundaries. Tracts are collections of census blockgroups, which are collections of blocks (blocks are the smallest unit of geography used by the census).

One of the easiest places to obtain census tract (or other census geography) boundary data sets for any part of the US in ArcGIS format is at the ESRI's *Geography Network - Free* section (<http://www.geographynetwork.com/freeresources.html>).

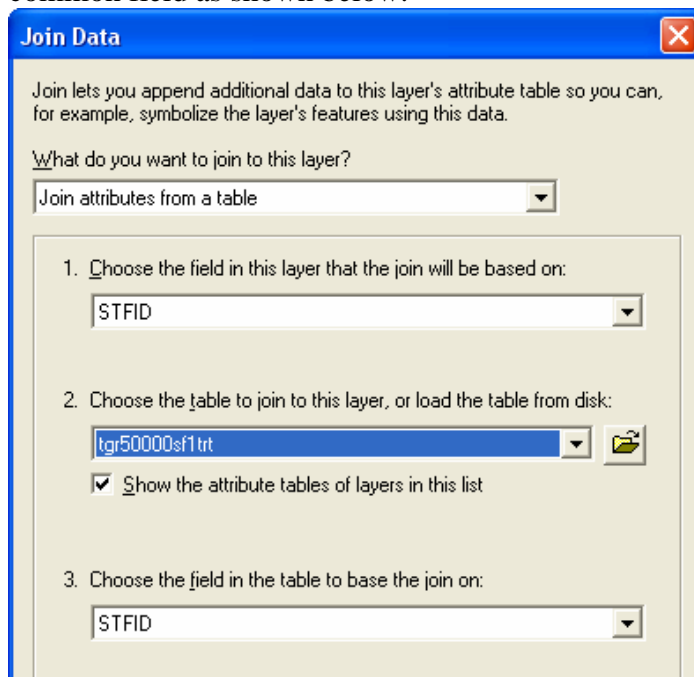
1. From this site choose the **Census TIGER/Line 2000 Data**
2. Read about the data then choose **Preview and Download** from the left-hand column
3. Select a **state**
4. Under *Select by Layer*, choose **Tracts 2000** (make sure you get this data layer and not one of the others)
5. Next check the *county* you want.
6. Scroll to the bottom and press **Proceed to Download**
7. Read the information and press **Download File**
8. Save the file to your personal folder and give it a name you will understand (but maintain the *.zip* extension) (Note: in the GIS lab, you can right-click on the download file to extract it on your H: drive folder)
9. Unzip the data set after download
10. The unzipped file contains another zipped file - unzip this one as well
11. You can now bring this data layer into *ArcMap*. Note that the data sets are in decimal degrees (GCS = Geographic Coordinate System).
12. Open the table for the *Tract* boundaries and look at what it contains. It only has fields identifying the geography (e.g., tract). It has no actual census demographic data. The *STFID* field will be the join field we use next. It is a *concatenation* (adding together) of the FIPS (Federal Information Processing Standard) codes for the state (e.g., Texas=48), county (e.g., Travis County=453), and tract (e.g., 000101).

## Getting Basic Census Data from the Geography Network

Now that you have a census GIS geography layer (tract boundaries), you want to find and join actual census demographic and/or housing data to it. In this section of the tutorial, we are going to download basic (SF1) demographic data from ESRI's web site. It is already formatted to be easy to join and easy to read.

1. From ESRI's *Geography Network - Free* section (<http://www.geographynetwork.com/freeresources.html>), go back to the **Census TIGER/Line 2000 Data** link, and go to the Preview and Download link again.
2. Go to the same state/county for which you have the tract boundaries

3. For select by layer, choose **Census Tract Demographics (SF1)**. This provides the basic data from the Summary File1 portion of the census in an easily readable format. **Be very careful to choose this table, not one of the many other choices!**
4. Download this data set and unzip it until you have the end product, a *.dbf* file.
5. Add this table to ArcMap and open the table to see what it contains. Nothing will appear in your map because it is a table of information, not a geographic layer. Note that it has data as well as geographic codes, including the *STFID* field.
6. With the Tract boundaries and the Tract demographics table in ArcMap, you can now perform a join using the *STFID* field. A join relates the attribute data in the table to the geographic data layer by using a common attribute field. **Right-click** on the **Tract geography** data layer and choose **Joins and Relates - Join...**
7. Create a join to the Tract demographics table (*.dbf*) using the *STFID* field as the common field as shown below:



8. If asked if you want to create an index, say Yes or OK.
9. With this join in place, you should be able to create a map of any of the census demographic fields - try it and see (**right-click** on the Tract geographic data layer, choose **Properties-Symbology** and create a *quantities* map using one of the fields. Note that most of the data is raw numbers. If you want to see, for example, Hispanics as a proportion of the total population, choose *Hispanic* as the value field, but for *Normalization*, choose *POP2000* (the total population of each census block group). *Normalization* divides the mapped value by the normalization value. **Note that the Normalization option called "percent of total" means percent of the chosen field value's total sum, so if you chose this for the Hispanic population, the result would show the proportion that each tract held of the total Hispanic population for a county, not the percent Hispanic of the tract's total population.**

## Creating a permanent shape file with the census data

A join between a geographic layer and a table is temporary - the join does not create a permanent new file. If you save a map file with the join in place, the next time you open that map file, ArcMap will look for your shape file and the .dbf file and re-join them.

You can also create a permanent shape file with the census data attached. To do this, make sure the data is joined as described above. Then **right-click** on the Tract geographic data layer and choose **Data - Export Data**. In the dialog box that appears, leave the default settings as they are but give the new shape file an appropriate name, like *tracts\_sfl.shp* and make sure you use Browse to put it in your personal folder. Next time you add this data layer to a map, the census data will come with it. It will also be easier to map because you can read the field names a bit more easily.

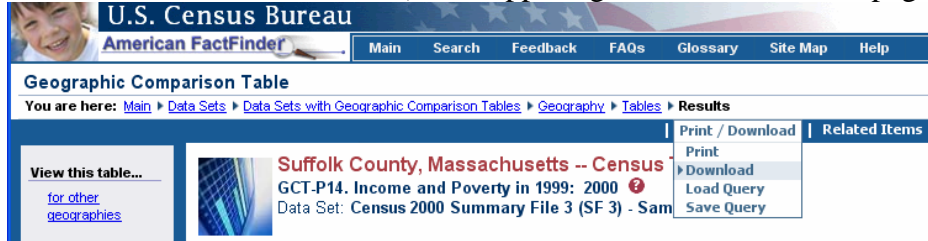
## Getting Census Data from the Census Bureau

You can get more detailed census data from the Census Bureau and from many state, regional, and local agencies (or you can buy it). In this example, you are going to get tabular data from the Census Bureau itself. Note that this data will not be as easy to use as the data that you previously downloaded, but going directly to the source gives you access to more detailed data for free.

In this section of the tutorial, you will be accessing census tract-level data again.

1. Go to the Census Bureau on the web at (<http://www.census.gov/> ). Then, click on *American FactFinder* (left column).
2. Click on the link to *Data Sets* (on the left column)
3. Click on the button for Census 2000 Summary File 3 (SF3) - Sample Data. *2000 Summary File 3* contains data from the 2000 sample census form and is useful for information about education, income, family status, housing costs, employment status, etc.
4. Click on "*Geographic Comparison Tables*" to the right side of the SF3 section
5. Once you are in the "geographic comparison tables page," for **Select geographic type**, choose *County*. Wait for the screen to change.
6. When you see **Select State**, choose the state for which you downloaded Tract data in the first part of this tutorial. Press Enter and wait for the screen to change.
7. For **Select Geographic Area**, choose the county for which you have tract boundaries, and then for **Select Table Format**, choose *Census Tract*, then press **Next**
8. Pick a table from the list (only one) - e.g., you could start with *Income and Poverty in 1999:2000*, and press **Show Result**
9. The resulting table shows you the requested information summarized for the entire county, then the data for each individual census tract within the county. Look it over to understand what it is showing.

- To download the information for use in a spreadsheet or GIS, choose **Print/Download - Download** (in the upper right corner of the web page):



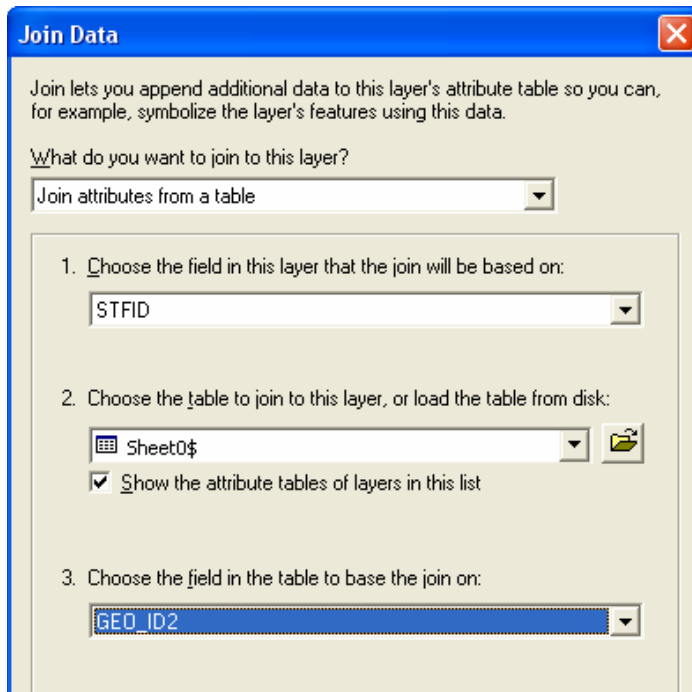
- Choose to download a *database compatible* version in *Excel*, and check mark the option to *include descriptive data element names*
- When prompted, choose *Save*. The downloading file will be named *output.zip* by default. Give it a new name to match the data (e.g., *income\_poverty.zip*), and save it to your personal folder.
- Open the windows file manager and unzip this file - it will open to a file named something like *gct\_dec\_2000\_sf3\_u\_data0.xls*
- When the Excel file has unzipped, double-click on it to open it in Excel
- You now have the data in an Excel spreadsheet:
  - The first row contains the attribute field names in abbreviated format (e.g., *GCTP14\_CO1.C0*)
  - The second row contains the descriptive data element names you asked for (e.g., *Median income in 1999 (dollars); Households*)
  - Look carefully at the third row and note that the SUMLEVEL or Geographic Summary Level says *050*. This is the data for the county as a whole. Everything below it (*SUMLEVEL=140*) is for each individual tract.
- Next, to avoid problems later, save your Excel file to a new name (e.g., *IncomePoverty\_SF3.xls*)
- In the first row, rename the field headings for the data columns to something that makes more sense. **BUT** - for it to be compatible with with ArcMap, it must follow dBase table rules - no field names longer than 10 characters, no spaces, no funny stuff. E.g., instead of *GCTP14CO1.C8*, you might name the field *PctPovFam*.
- Delete any fields (columns) you're not interested in
- Delete the second row (field descriptors)
- Make sure there are no (x) data values in the spreadsheet (delete them if there are) - if there are, field will some into ArcGIS as a text field and you need it to be numeric (you can automatically replace all (x) values with nothing if you use the Edit – Replace option on the menu – replace all.
- Choose **File – Save**
- Exit out of Excel

## Joining Census Bureau data to the Tract Shapefile

Now it is time to join the tract SF3 data to the tract shape file.

- If you don't have your tracts boundary layer up in ArcMap, open ArcMap and add it now.

2. **Add** the *income\_poverty.xls* file (or whatever the Excel file you created in the last step) to your dataframe. **Note** you cannot add the file .xls that you first see in your folder list – instead, you need to double-click on the Excel file to get the worksheet (e.g., *sheet0\$*)
3. Right-click on the table and choose **Properties** and then the **Fields** tab – make sure they numeric fields (e.g., *medfaminc*) came in as double or some other type of numeric type. If they came in as text, you need to trouble-shoot this. See the Tip Sheet for *Working with Excel Files*.
4. Assuming the field types are correct, open the table and examine its contents. You will see that it has a *geo\_id* field, a *geo\_id2* field, a *sumlev* field (summary level, in this case 140 = tracts), and other columns with the actual data.
5. You will see in the data table that there is no *STFID* to join to the Tracts shapefile, only a *GEO\_ID* and *GEO\_ID2*. But in the table for the *tracts* shapefile, there is no *GEO\_ID*. But note that the two ID fields are very similar. The *GEO\_ID* field simply has a beginning part (14000US) and then has the same numbers as the *STFID*. The *GEO\_ID2* field is exactly the same as the *STFID* field. We are going to use this information join the two tables together based on the *STFID* field in the GIS layer and the *GEO\_ID2* field in the census data table.
6. Right-click on the census tracts GIS layer in ArcMap and remove any existing joins. Then right-click on tracts again and choose **Joins and Relates – Join**. Fill in the table as follows:



7. Click OK
8. Check your work by opening the attribute file of the census tracts GIS layer. You should see all the joined attribute data from your .dbf table. Check for any <null> values - this would indicate an improper join. If you see this, examine the two join fields carefully to see why they wouldn't join.

9. Try creating a map (if you need to, consult the original Excel file for reference as to what the field names mean). Make sure there are no blank polygons that didn't join up properly.
10. If the join worked properly, try creating some maps of the census data. Use the Normalization option where appropriate.

You can repeat this for any table you download from American Factfinder. Remember that in ArcMap, you can also export joined layers to create a permanent shape file with all the information. I recommend this. For example, you could export the joined table in this tutorial and call it *2000\_tract\_income\_poverty*. If you then download another table, I recommend that you join it to the original tract shape file, and export that to a different shape file (e.g., *2000\_tract\_financial\_housing*).

### **More Census Data Downloads – including SF3 at the blockgroup level**

Once you have the hang of how all this works – downloading census tables and joining them to census geography layers, you can move on and try out the **Census Download Center** - [http://factfinder.census.gov/servlet/DownloadDatasetServlet?\\_lang=en](http://factfinder.census.gov/servlet/DownloadDatasetServlet?_lang=en) – where you can download whole tables at a time for many levels of geography. This is especially useful if you are looking for SF3 data at the block group level. There are more data processing tangles to deal with here another tip sheet is needed.