Downloading Census Data from American Factfinder for use in ArcGIS

Written by Barbara Parmenter, revised September 24 2013

In this tutorial, we will be obtaining information about housing tenure at the **Census Tract** level from the **2010 Census** for a single county using American Factfinder. You can then use a similar process to download any other Census 2010, American Community Survey, or Census 2000 data for other geography levels and/or for whole states or multiple counties. You have many, many options in American Factfinder – this shows one possible path.

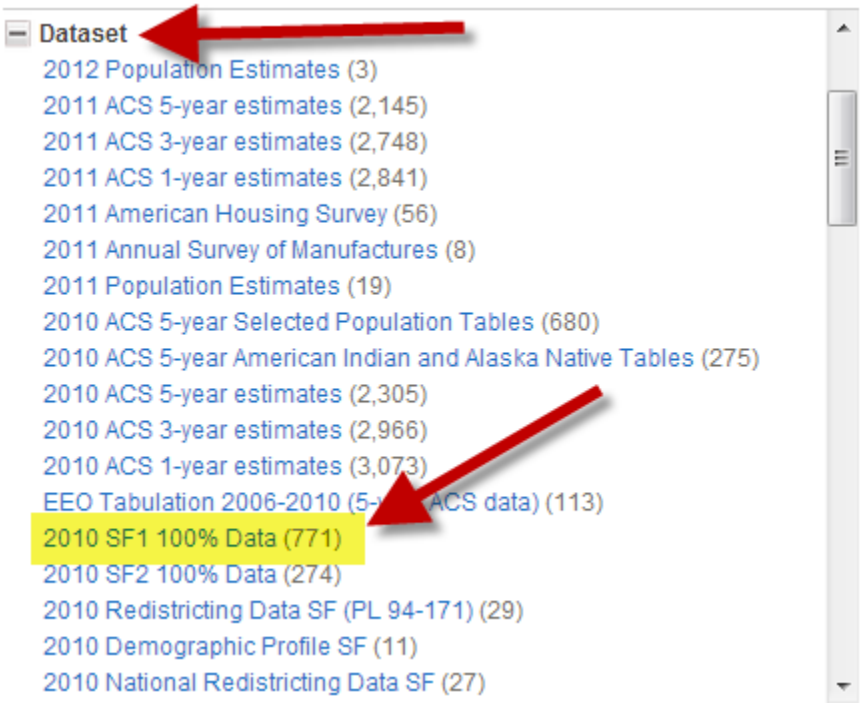
Obtaining Data from American FactFinder (AFF)

1. Data management is critical when dealing with the multiple tables of the Census. Before beginning this tutorial:
    a. Create a **Census 2010** folder in your personal workspace
    b. Create *two* subfolders: **AFF Data** and **Census Geography**

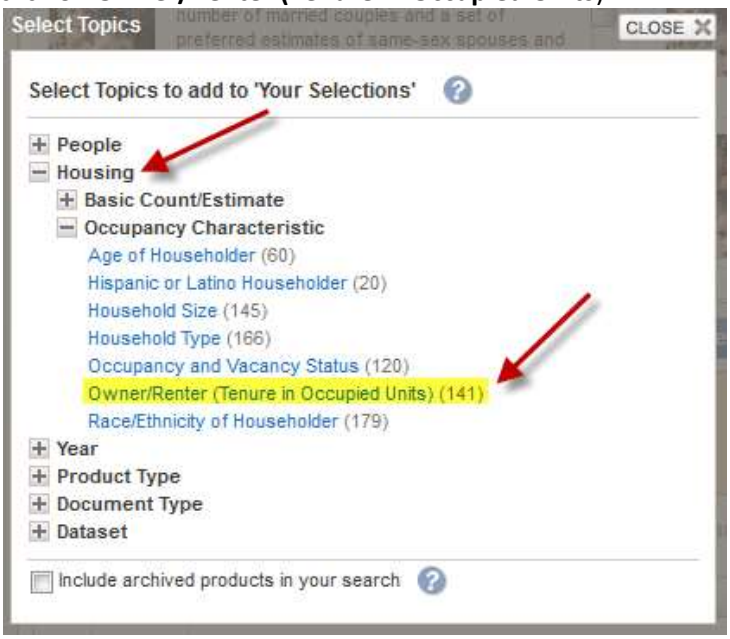
2. Go to the US Census web site – http://census.gov and under the *Data tab* select **American FactFinder**

3. Click on **Advanced Search** (and if you see it, **Show Me All**)
4. Click on **Topics** in the left column:

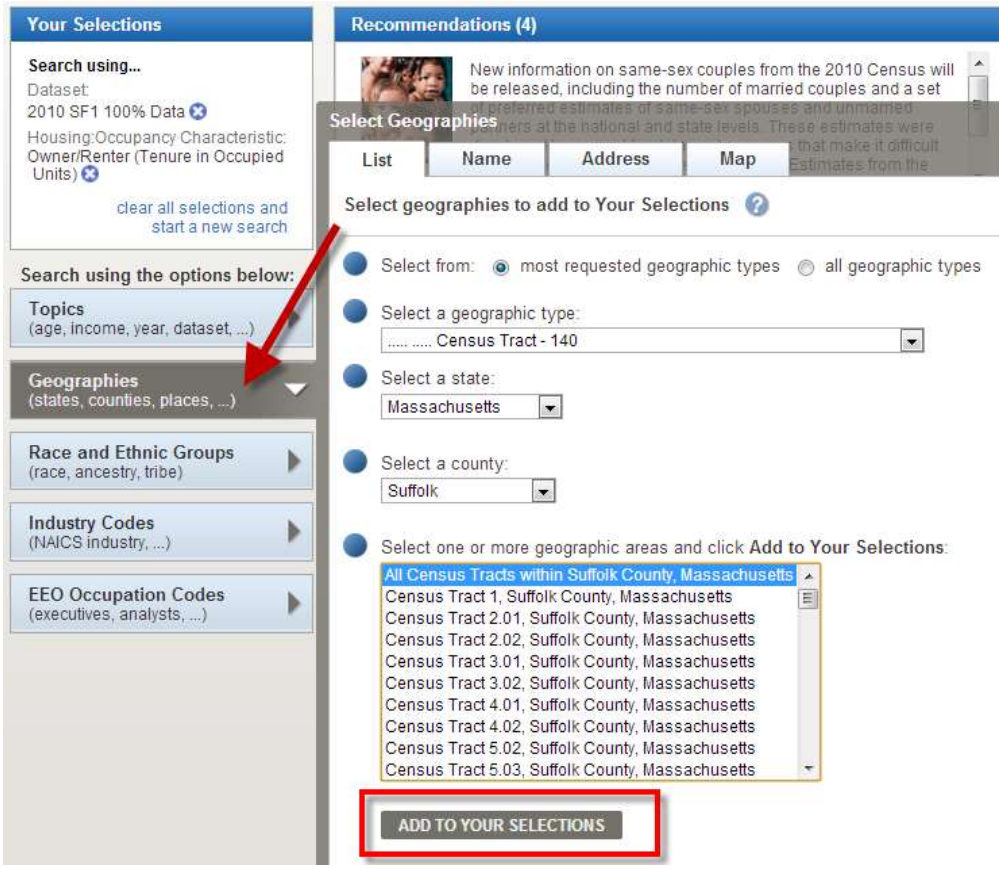
5. Click on **Dataset** and scroll down to click on **2010 SF1 100% Data** – this will send this criteria to your Selection box in the upper left of the site:



6. Scroll up in the **Topics** list and click on **Housing** – then under *Occupancy Characteristic*, click **on Owner/Renter (Tenure in Occupied Units**)

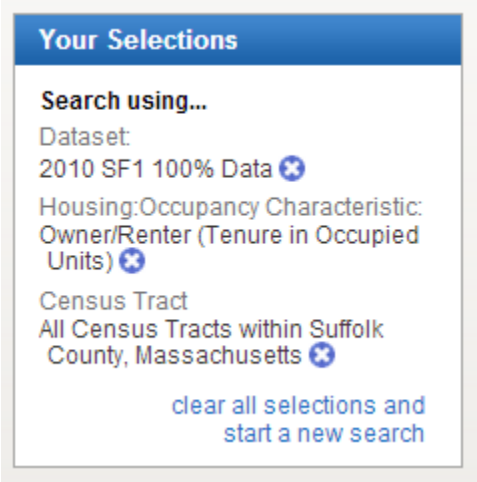


7. Close the **Topics** box (see above graphic)
8. Click on **Geographies** on the left column – this brings up *the Select Geographies overlay*
9. Fill out the box so that you are selecting all Census Tracts for a specific state and a county in that state – below we are selecting all census tracts in Suffolk County, Massachusetts



10. Be sure to click on ADD TO YOUR SELECTIONS
11. **Close** the *Select Geographies* overlay

12. Be sure that the **Your Selections** box in the upper left corner contains what you want – the data set, the general topic, and the census geography level for the specific location you want (all tracts, not just one tract). If it does not say this, clear your selections and start over from Step 3 above.

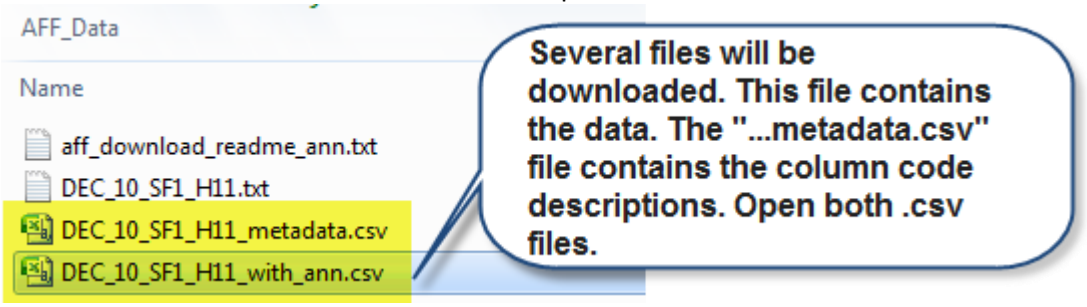


13. Checkmark a table of interest – to see what variables a table contains, click on the *Information* icon for that table. For this exercise, we highly recommend a table with just a few variables, for example**, H11. TOTAL POPULATION IN OCCUPIED HOUSING UNITS BY TENURE** – this will show you how many people live in rented units, units owned through a mortgage, and units owned free and clear or without payment. This is an easy table to process in Excel.

14. Click on **Download (**  **)** and follow the instructions (your file will be built, and then you will click on another *Download* option)
15. The file is in compressed format – extract it to your *Census 2010/AFF Data* folder

Preparing American Factfinder Data for Use in ArcMap

Double-click on both downloaded **.csv** files to open them in Excel:



Note: If you are opening the file from within Excel, you will need to set the option to look for *all file types*:

Your *DEC_10…with_ann.csv* file should look something like this – this file contains the data:

	A	B	C	D	E	F	G	H
1	GEO.id	GEO.id2	GEO.displ	D001	D002	D003	D004	
2	1400000US250	2.5E+10	Census Tr	4225	794	231	3200	
3	1400000US250	2.5E+10	Census Tr	3730	828	262	2640	
4	1400000US250	2.5E+10	Census Tr	3861	857	349	2655	
5	1400000US250	2.5E+10	Census Tr	2628	799	270	1559	
6	1400000US250	2.5E+10	Census Tr	2916	941	413	1562	
7	1400000US250	2.5E+10	Census Tr	5672	851	281	4540	
8	1400000US250	2.5E+10	Census Tr	3511	868	297	2346	
9	1400000US250	2.5E+10	Census Tr	3110	447	154	2509	
10	1400000US250	2.5E+10	Census Tr	2211	444	81	1686	
11	1400000US250	2.5E+10	Census Tr	4915	682	187	4046	

Now open the file - *DEC_10… _metadata.csv* file. This file explains the column header codes in the data file - it should look something like what you see below. This is a very important file!!! Typically the first data column (D001 here) is the **Universe** of things counted in this table. This table is counting people in occupied housing units. In the case of this table if you wanted to show the % of the population that is in rented housing units, you would divide D004 by D001 and multiply by 100:

	A	B	C	D	E
1	GEO.id	Id			
2	GEO.id2	Id2			
3	GEO.displ	Geography			
4	D001	Total population in occupied housing units:			
5	D002	Owned with a mortgage or a loan			
6	D003	Owned free and clear			
7	D004	Renter occupied			

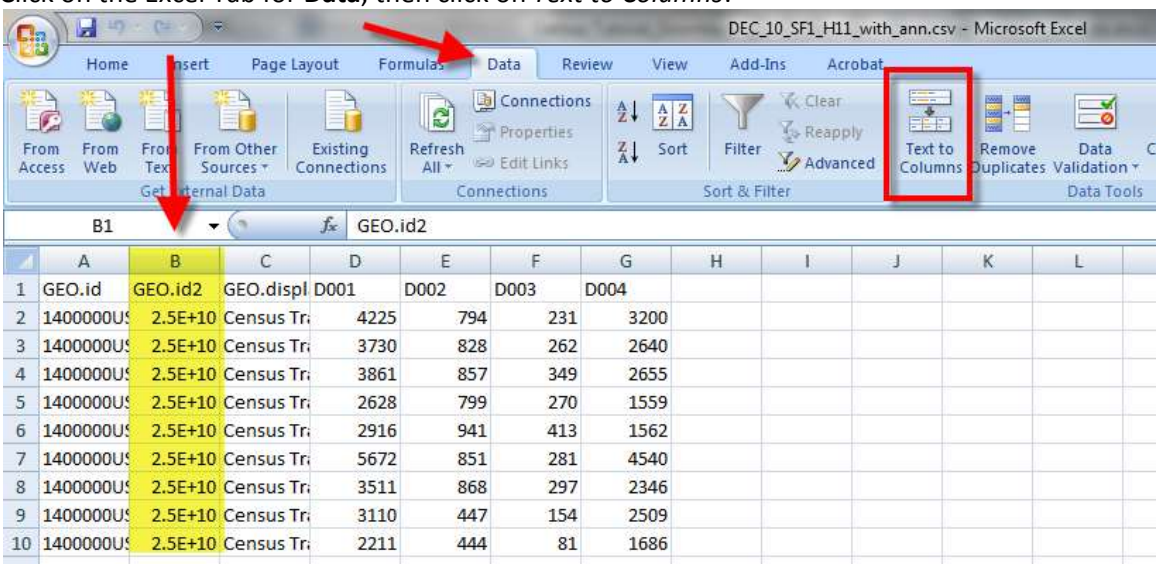
A few important steps left. First, ArcGIS does not like periods in the column names:

1. Delete all periods (.) in all the column names

Second, the Geo ID in the *Census Tracts polygon attribute table* to which you will be joining this data table is in a text format. **GEOID2** in this *CSV* file must also be *text* for the join to work properly.

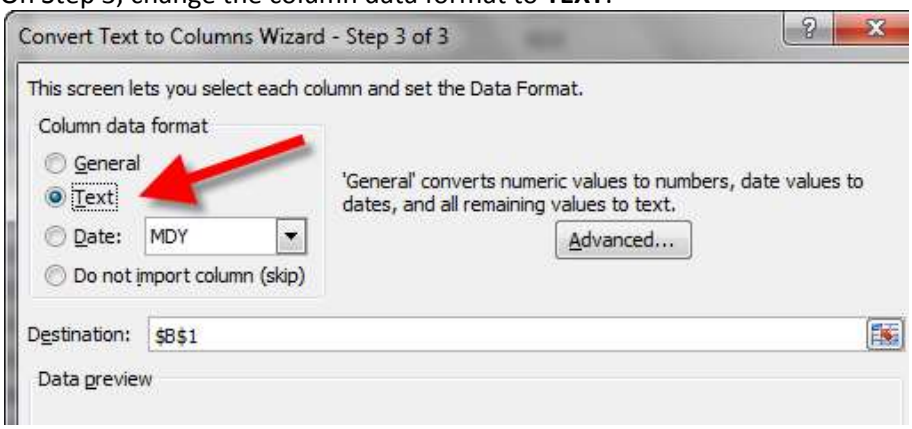
2. Click on the tab (B) above **GEO.ID2** to highlight the entire column

3. Click on the Excel Tab for **Data**, then click on *Text to Columns*:

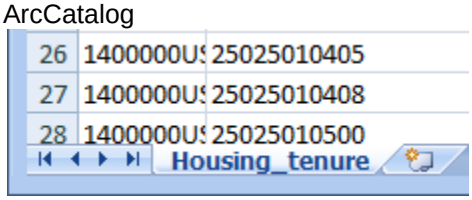


	A	B	C	D	E	F	G
1	GEO.id	GEO.id2	GEO.displ	D001	D002	D003	D004
2	1400000US	2.5E+10	Census Tr	4225	794	231	3200
3	1400000US	2.5E+10	Census Tr	3730	828	262	2640
4	1400000US	2.5E+10	Census Tr	3861	857	349	2655
5	1400000US	2.5E+10	Census Tr	2628	799	270	1559
6	1400000US	2.5E+10	Census Tr	2916	941	413	1562
7	1400000US	2.5E+10	Census Tr	5672	851	281	4540
8	1400000US	2.5E+10	Census Tr	3511	868	297	2346
9	1400000US	2.5E+10	Census Tr	3110	447	154	2509
10	1400000US	2.5E+10	Census Tr	2211	444	81	1686

4. Click **Next** to leave the first setting at Delimited
5. Click **Next** to leave the second setting at Tab
6. On Step 3, change the column data format to **TEXT**:



7. To make things easier later, rename the worksheet to something comprehensible, e.g., *Housing_Tenure* - the worksheet name will be the identifier (what is displayed) in ArcCatalog



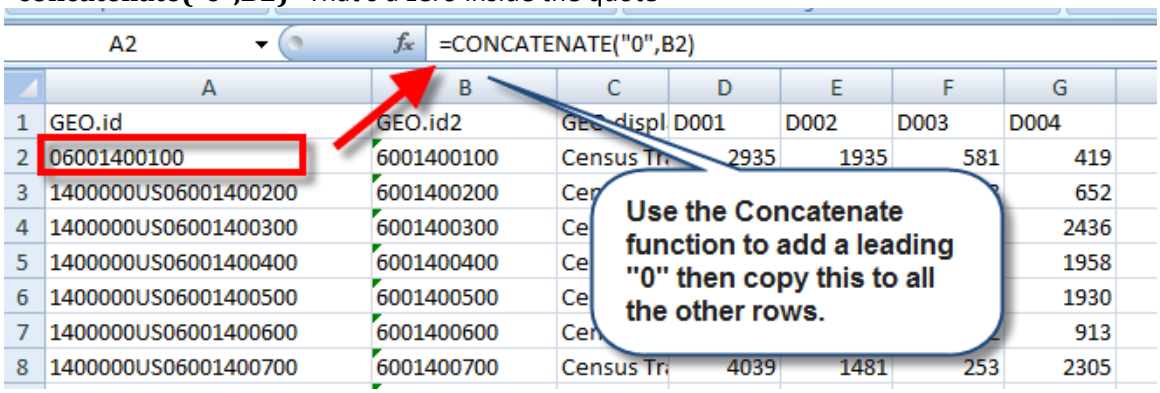
8. Very important step – **save your modified CSV file as an Excel file (.xlsx)** – give it a comprehensible name, e.g., *2010 Census H11_population by housing tenure.xlsx*

Optional Tip – you can change your column names to match what the metadata .csv file is telling you, or you can keep the metadata file handy so that you know what the codes stand

for when you are ready to use this data in ArcGIS. It's probably easier to refer to the codes later.

Extra step for Alaska, Alabama, Arkansas, Arizona, California, Colorado, and Connecticut
(if you're not getting data for one of these states, ignore this section)
Some states have FIPS codes that start with a zero, and because Excel removes that leading zero, the table won't join properly unless we put it back on. If you're working in Alaska, Alabama, Arkansas, Arizona, California, Colorado, or Connecticut, you'll have to add that zero back on manually.

To add the zero back on, at the beginning of the table, under Column A, at Row2 type in: **=concatenate("0",B2)** That's a zero inside the quote



	A	B	C	D	E	F	G
1	GEO.id	GEO.id2	GEO.displ	D001	D002	D003	D004
2	06001400100	6001400100	Census Tr	2935	1935	581	419
3	1400000US06001400200	6001400200	Cen				652
4	1400000US06001400300	6001400300	Ce				2436
5	1400000US06001400400	6001400400	Ce				1958
6	1400000US06001400500	6001400500	Ce				1930
7	1400000US06001400600	6001400600	Cen				913
8	1400000US06001400700	6001400700	Census Tr	4039	1481	253	2305

Hit **Enter**.

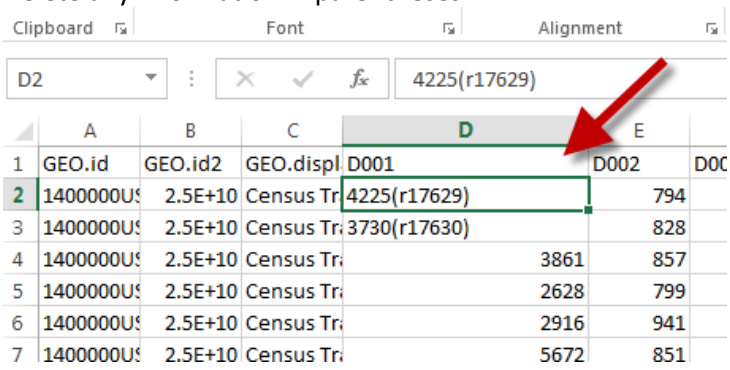
If the result of that formula looks right (i.e. it has a leading zero), copy that cell's formula to the rest of the column.

You're almost done! To keep this compatible with the rest of the directions, copy all of the cells in this new column, and right click on the GEOid2 column. Click the options below "**Paste Special**", choose **Values**, and your leading zeroes should be all set. Ensure your column is still named GeoID2.

Delete the column you added but be sure you still have the fixed GeoID2 column.

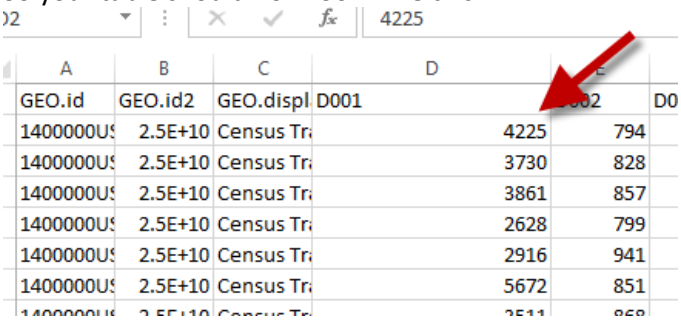
9. Note: your table may have columns, like D001, where the data has text values in it. This will cause this information to not be mappable in ArcMap. Follow the directions below if this is the case:

a. Delete any information in parentheses



	A	B	C	D	E
1	GEO.id	GEO.id2	GEO.displ	D001	D002
2	1400000US	2.5E+10	Census Tr	4225(r17629)	794
3	1400000US	2.5E+10	Census Tr	3730(r17630)	828
4	1400000US	2.5E+10	Census Tr	3861	857
5	1400000US	2.5E+10	Census Tr	2628	799
6	1400000US	2.5E+10	Census Tr	2916	941
7	1400000US	2.5E+10	Census Tr	5672	851

b. So your table should now look l like this:



	A	B	C	D	E
	GEO.id	GEO.id2	GEO.displ	D001	D002
	1400000US	2.5E+10	Census Tr	4225	794
	1400000US	2.5E+10	Census Tr	3730	828
	1400000US	2.5E+10	Census Tr	3861	857
	1400000US	2.5E+10	Census Tr	2628	799
	1400000US	2.5E+10	Census Tr	2916	941
	1400000US	2.5E+10	Census Tr	5672	851

Save your Excel file again!

Final STEP: Save your file EXIT out of Excel – you CANNOT have Excel open still when you work with this data in ArcGIS!

Obtaining GIS files for Census Geography

Now you need to get your Census tract polygons…

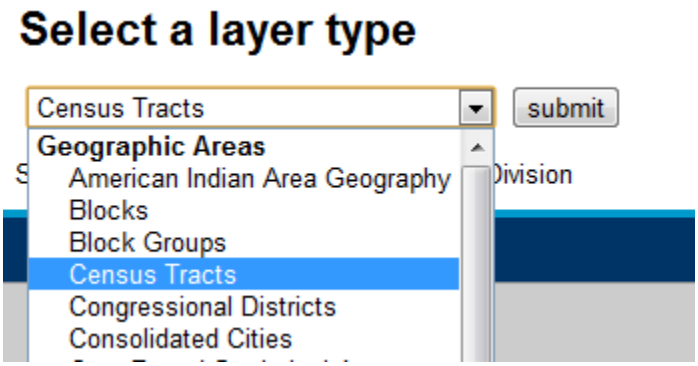
1. Go to the Census web site (http://census.gov) and click on the **Geography** tab and select the option *TIGER*

2. Click on **Tiger/Line Shapefiles and Files** in the **TABLE** as shown below:

3. Click on the link for the **2013 Tiger/Line Shapefile Main Page**

4. Click on **Download Shapefiles** then on **Web Interface**

5. Under *Select a Layer Type* use the pulldown menu to choose **Census Tracts** then press *Submit*

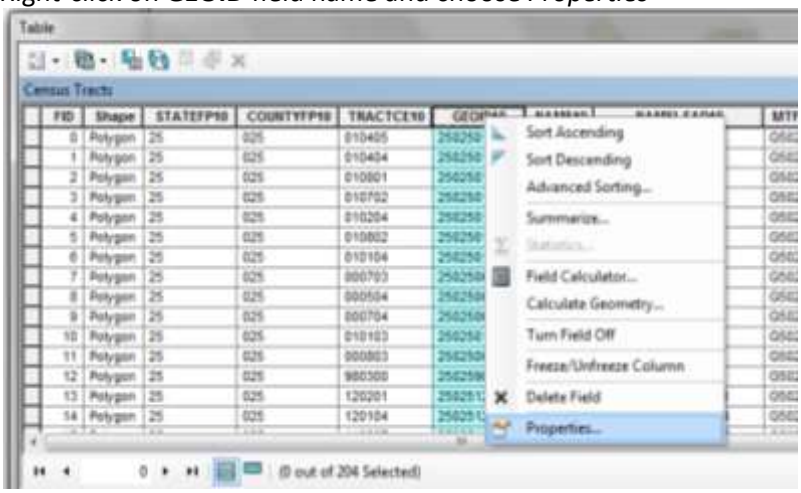


6. **Select your State of interest and download** the data set – it is compressed in a .zip file
7. Unzip the downloaded file to your **Census Geography** folder

Joining the AFF table to your Census Tract polygons in ArcMap

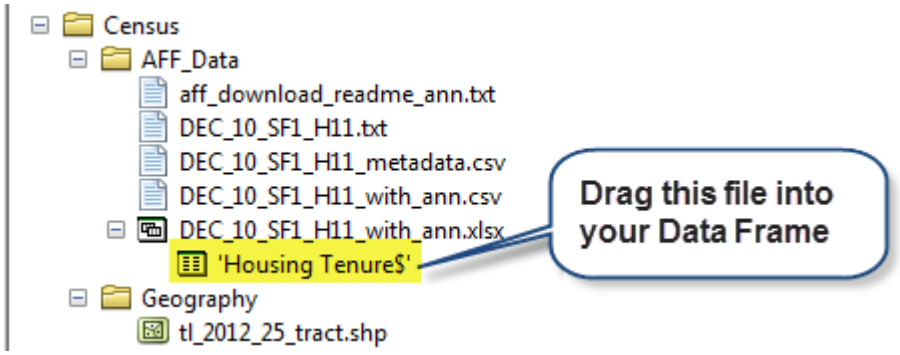
Now you're ready for mapping!

1. Start a session of ArcMap with a blank map
2. Add your **Census Tracts** geography data set to the map (e.g., *tl_2012_25_tract.shp*)
3. Rename your Census Tract geography layer to *Census Tracts*
4. Open the *Census Tracts* polygon attribute table and take a look at it
5. The *GeoID* column is what we will be using for joining our AFF data
6. Right-click on **GEOID** field name and choose *Properties*

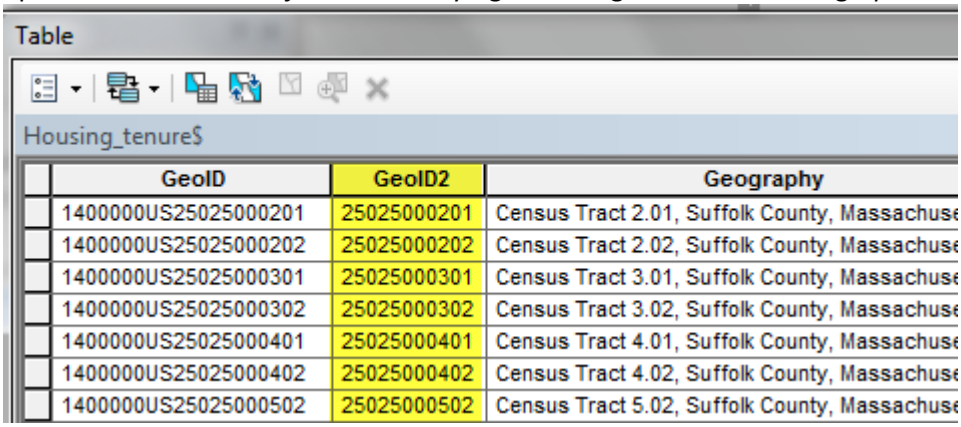


You'll see it is a STRING type attribute field – that's important to know. Close the table.

7. Add your American Factfinder table to the map - you need to drill down to the worksheet level:



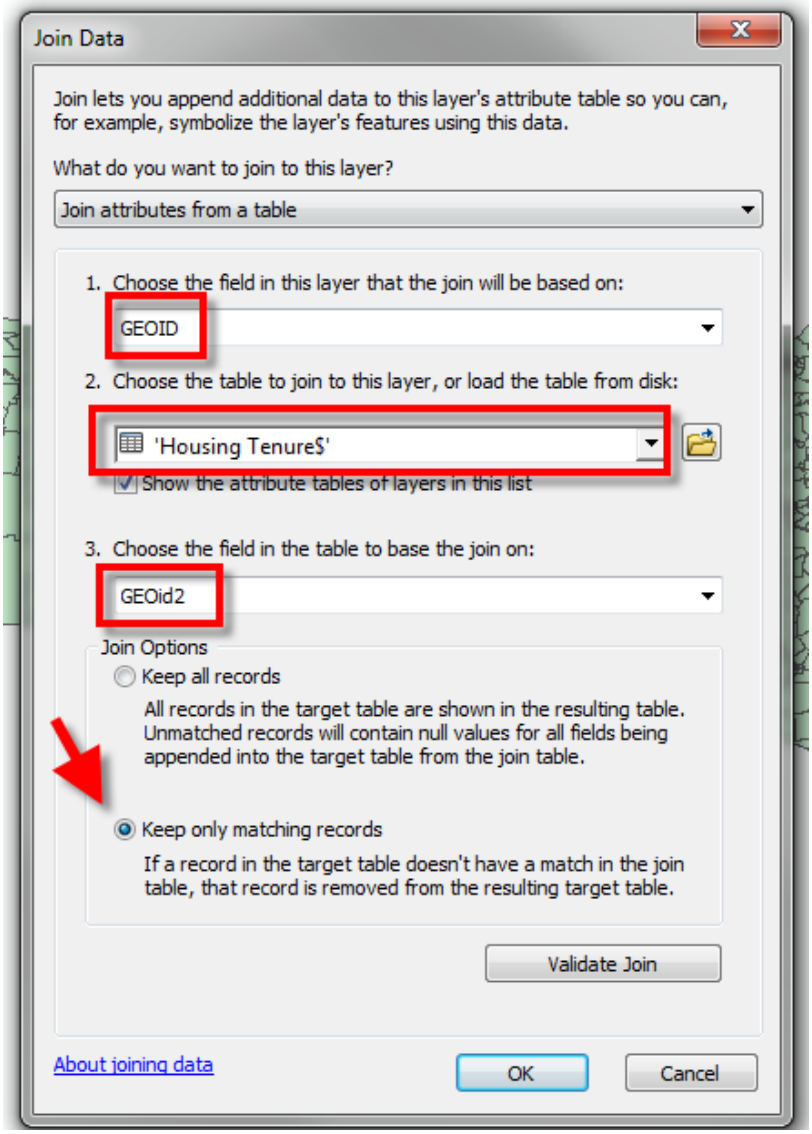
8. Open the *American Factfinder Table* by right-clicking on it and choosing *Open*



GeoID	GeoID2	Geography
1400000US25025000201	25025000201	Census Tract 2.01, Suffolk County, Massachuse
1400000US25025000202	25025000202	Census Tract 2.02, Suffolk County, Massachuse
1400000US25025000301	25025000301	Census Tract 3.01, Suffolk County, Massachuse
1400000US25025000302	25025000302	Census Tract 3.02, Suffolk County, Massachuse
1400000US25025000401	25025000401	Census Tract 4.01, Suffolk County, Massachuse
1400000US25025000402	25025000402	Census Tract 4.02, Suffolk County, Massachuse
1400000US25025000502	25025000502	Census Tract 5.02, Suffolk County, Massachuse

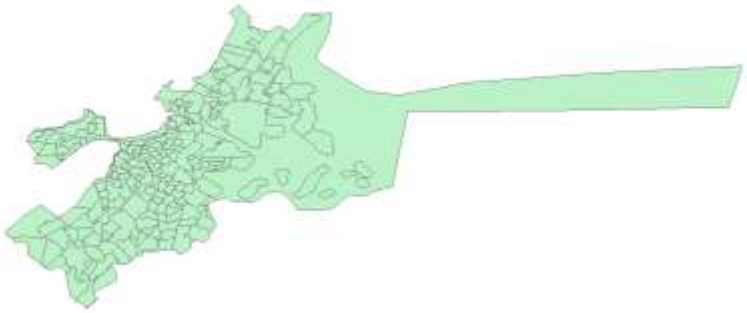
9. **GEOid2** will be your join field – check its properties to ensure that it is also a **STRING** type and that all the census data appears correctly
10. Close the table
11. Right click on your *Census Tracts* and choose **Join and Relates – Join…**

12. Fill in the dialog box as follows – you are joining attributes from a table, using *GEOID* in your Census Tracts layer and *GEOid2* in your AFF table – click OK when done:



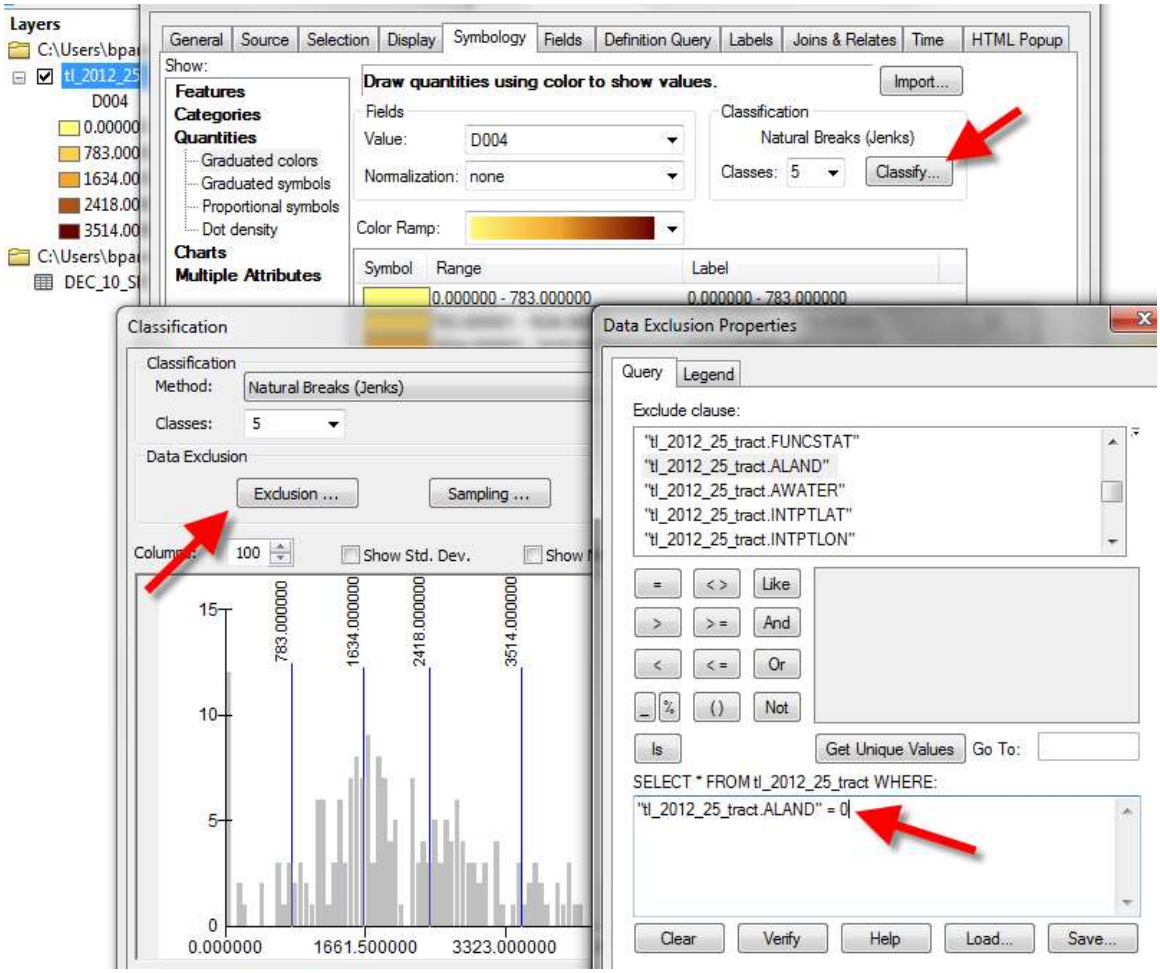
13. Open the *Census Tracts* attribute table to ensure that the join was made correctly. If it was, you should see your AFF data when you scroll to the right in the table.
14. Close the table

One tip – if your area of interest is near water or has water features in it (like Boston), your tract data set may look like this – it includes a lot of water:



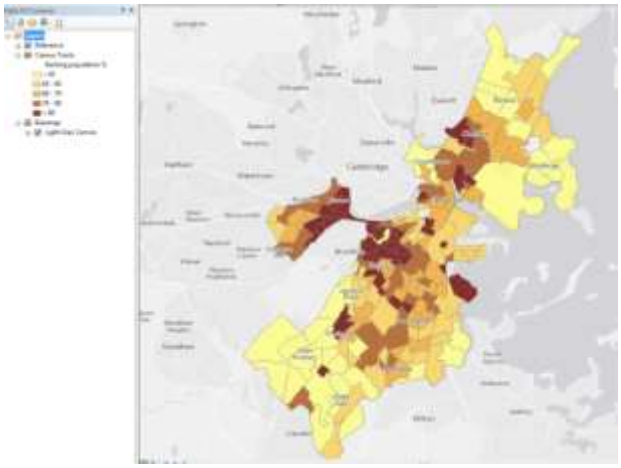
To get rid of the water tracts when you make a map, in the Symbology properties, click on *Quantities* and choose your variable, but then also click on *Classify*.

In the *Classify* dialog box, click on **Exclusion**. You can exclude all census tracts where the land area field = 0 (no land), as follows:

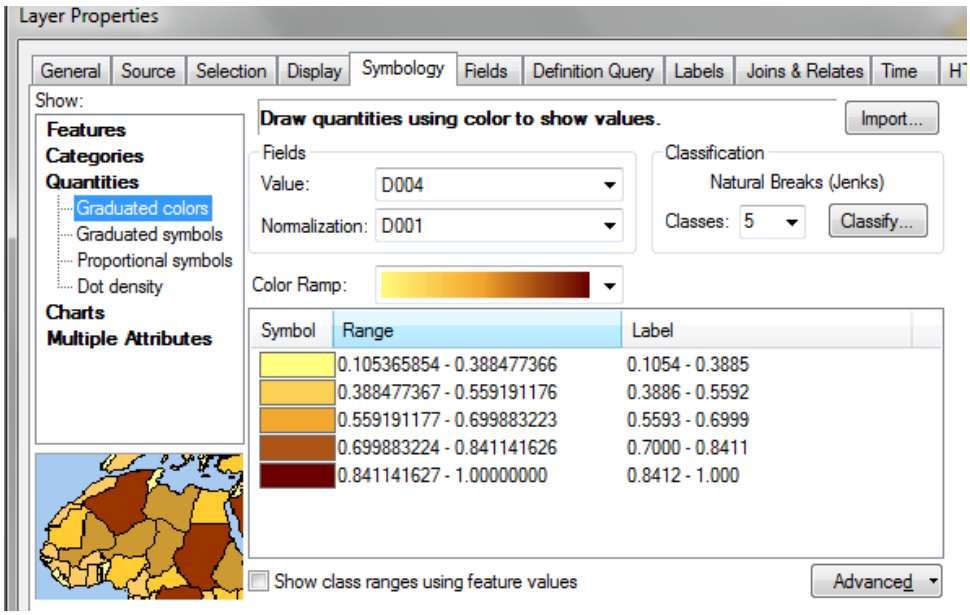


Now you can make a map of your data following the usual methods. If you are unfamiliar with mapping numeric values, see the ArcGIS 10.1 online help – [About Symbolizing Layers to Represent Quantity]

The next page shows an example of a map showing the % of people in rental housing units for each tract in Suffolk County (population renting normalized by total population in housing units). The map is using the "Light Gray Canvas" option from ESRI's basemap choices (Click on **File – Add Data – Add Basemap** to get this option):



Here is the symbology properties for the map above:

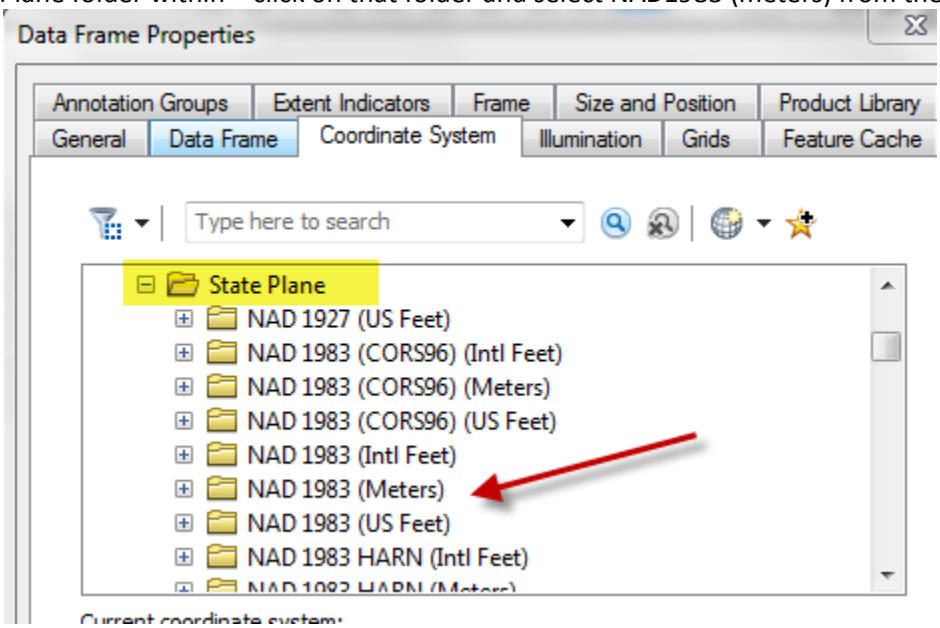


Setting a Projected Coordinate System for your Map

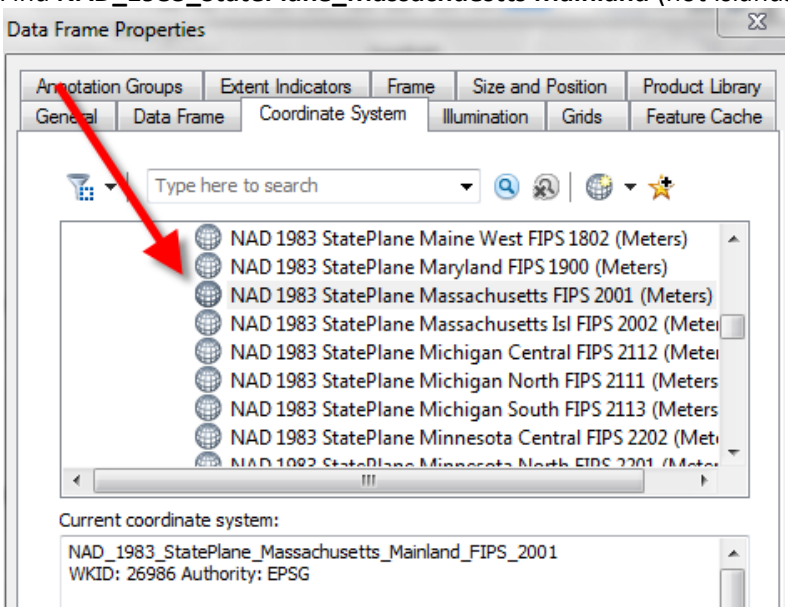
It is good cartographic practice to put your map into a projected coordinate system. The TIGER data is in a geographic coordinate system and can appear stretched in an odd way on your map. You can fix this problem by setting a projected coordinate system appropriate for your region.

You will need to know the best coordinate system to use for your area. In the case of Massachusetts, we will use the Massachusetts State Plane (NAD83) – meters coordinate system. If you don't know what coordinate system to use, you can leave your map as is or ask a lab assistant.

1. Click on **View – Data Frame Properties**
2. Click on the **Coordinate System tab**
3. In the dialog box, scroll to find the Projected Coordinate Systems folder and the State Plane folder within – click on that folder and select NAD1983 (meters) from the list:



4. Find **NAD_1983_StatePlane_Massachuestts Mainland** (not Islands!) and click on it



5. Click **OK**
6. Click **Yes** when warned that the coordinate system is different from the data in your maps.

You're done!