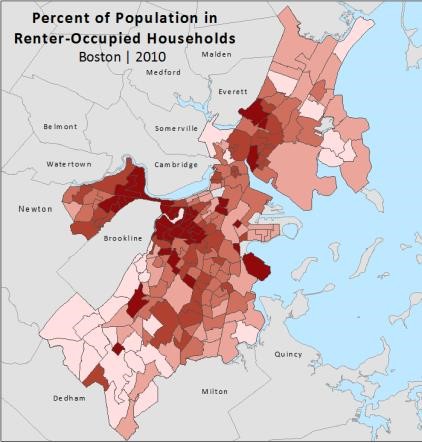
**USA Census Tutorial:** Downloading &



Mapping American Factfinder Census Data on

Renter Demographics for use in ArcMap

*Written by Barbara Parmenter, revised by Alyssa Kogan on January 26, 2018 Tufts Data Lab*

– Revised again by M Gottfried October 2019

## Contents

[Obtaining GIS Files from Census Geography 1](#_Toc20068)

[Obtaining Data from American FactFinder (AFF) 4](#_Toc20069)

[Preparing American FactFinder Data for Use in ArcMap 7](#_Toc20070)

[Joining the AFF table to your Census Tract polygons in ArcMap 11](#_Toc20071)

[Removing Water Only Census Tracts 14](#_Toc20072)

[Setting a Projected Coordinate System for your Map 15](#_Toc20073)

[Using Symbology to Map a Demographic Variable 16](#_Toc20074)

**Skills covered in this tutorial include:**

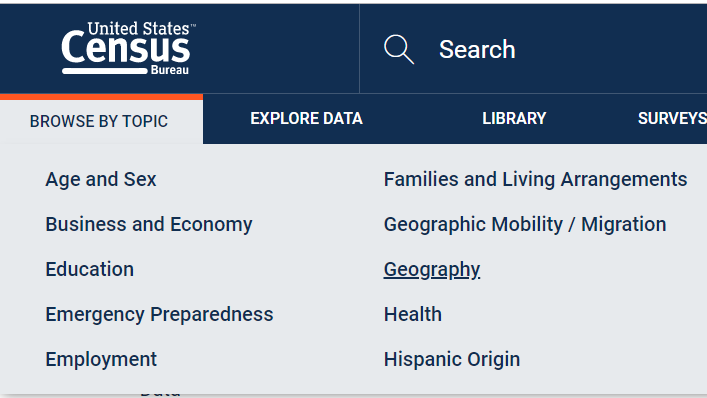
* Obtaining excel data from the American Fact Finder on Census.Gov
* Obtaining Geography Data from census.gov
* Cleaning Excel Data for Joins
* Joining Excel Data in ArcMap

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 GIS Files from Census Geography

The first step is to download the administrative geography spatial data (shapefiles) from Census.gov. This is the GIS data and contains only the boundary data – there is no demographic information included.

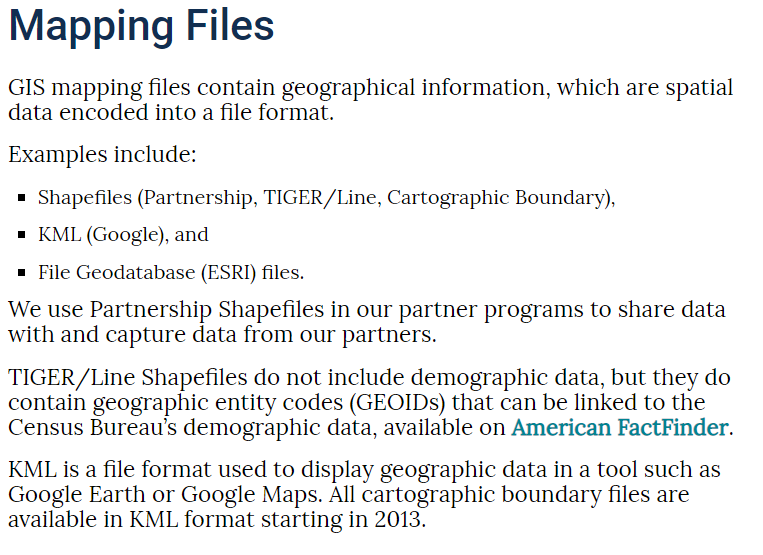
1. Data management is critical when dealing with the multiple tables of the Census. Before beginning this tutorial:
   1. Create a “**Census\_2010”** folder in your personal workspace
   2. Create *two* subfolders: **AFF\_Data** and **Census\_Geography**
      * The AFF folder will hold the excel tables you download from the Census
      * The Geography folder will hold the actual GIS Data shapefiles
2. Go to the Census web site ([http://census.gov)](http://census.gov/) and click on Browse by topic tab at the top and then Geography.



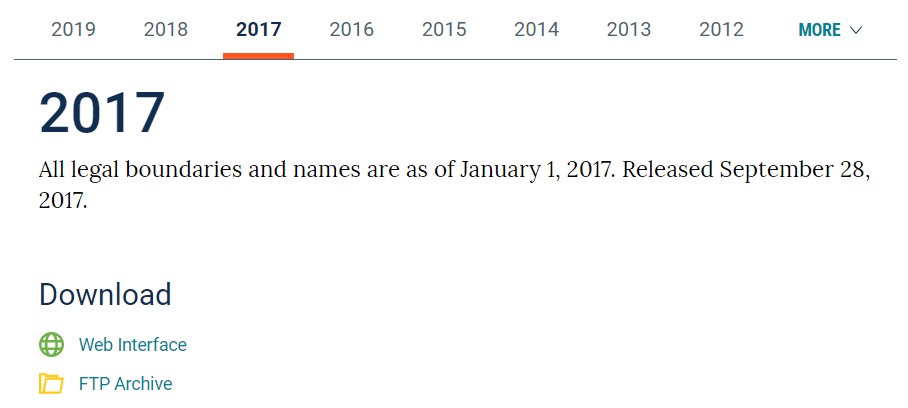
1. Under **Geography Program page**, select **TIGER/Line shapefiles** (bottom).



1. Then click on **Tiger/Line Shapefiles** in the **Table** as shown:



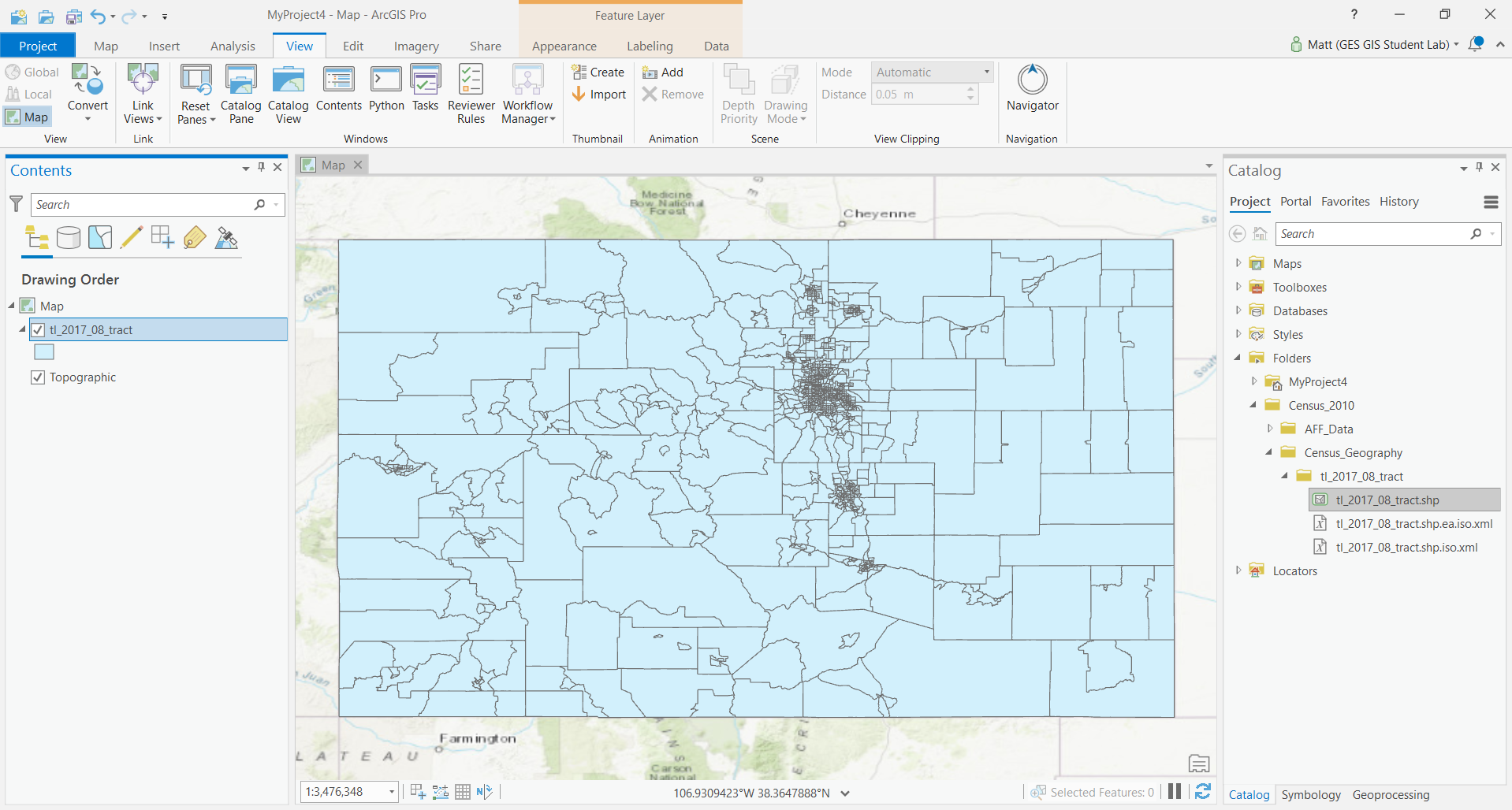
1. Click on **2017** and expand **Download** tab. Then click on **Web Interface (or manually download thru ftp)**. This will mainly depend on the browser you are using. You may need to switch.



1. Under **Select Year**, choose 2017 and under **Select a Layer Type**, notice how many different types of GIS data you can download from here! Remember this site!! Choose **Census Tracts** then **Submit**.



1. Select your **state of interest** (this tutorial will use Colorado data, but feel free to try it with your own area of interest if you wish!) and click **download**. It downloads the data as a zipped file.
2. Save the zip file into the **Census\_Geography** folder in your H drive. Navigate to the folder (in windows) and right click on the zipped file. Select *extract here*.
3. Open a **blank ArcGIS Pro session** and navigate to that folder in Catalog (Make a new folder connection). Drag in this new shapefile from your workspace. If you already had ArcGIS Pro open, you’ll likely have to refresh your workspace.



1. Change the name of the layer in the Table of Contents from tl\_2017\_25\_tract to Census Tracts. The census naming system for this particular files works as such: tl =tiger, 2017 is the year the data represents, 08 is the state number –aka Colorado, and Tract is the unit of the administrative boundary.

Remember: Census tracts are created to have approximately 4000 people per census tract – that is why they are varying sizes across the state.

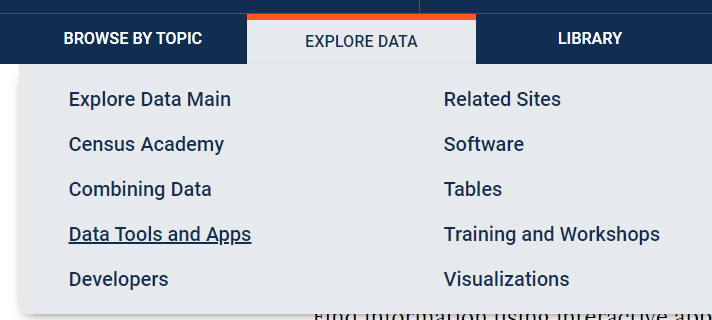
# 

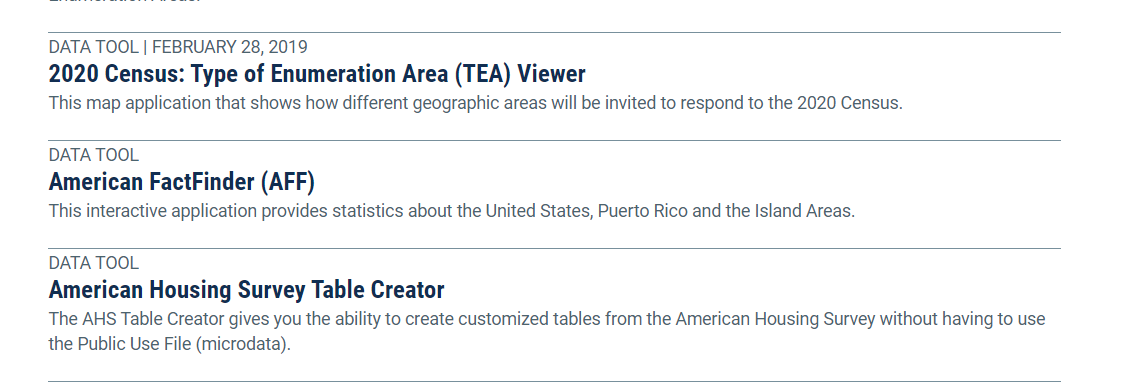
# Obtaining Data from American FactFinder (AFF)

Now we need to go and get the excel data containing all the demographic data per census tract.

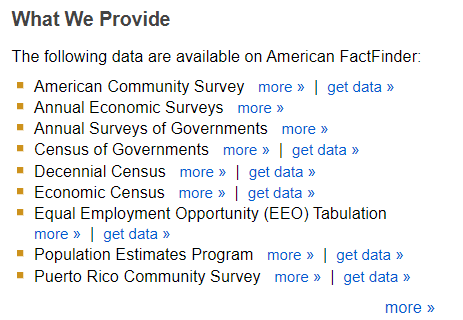
1. Go to the US Census web site [– http://census.gov](http://census.gov/)

1. Click on the **Explore *Data***tab ***Data Tools & Apps***select **American FactFinder.** This is the web interface to access census **excel/tabular data**.

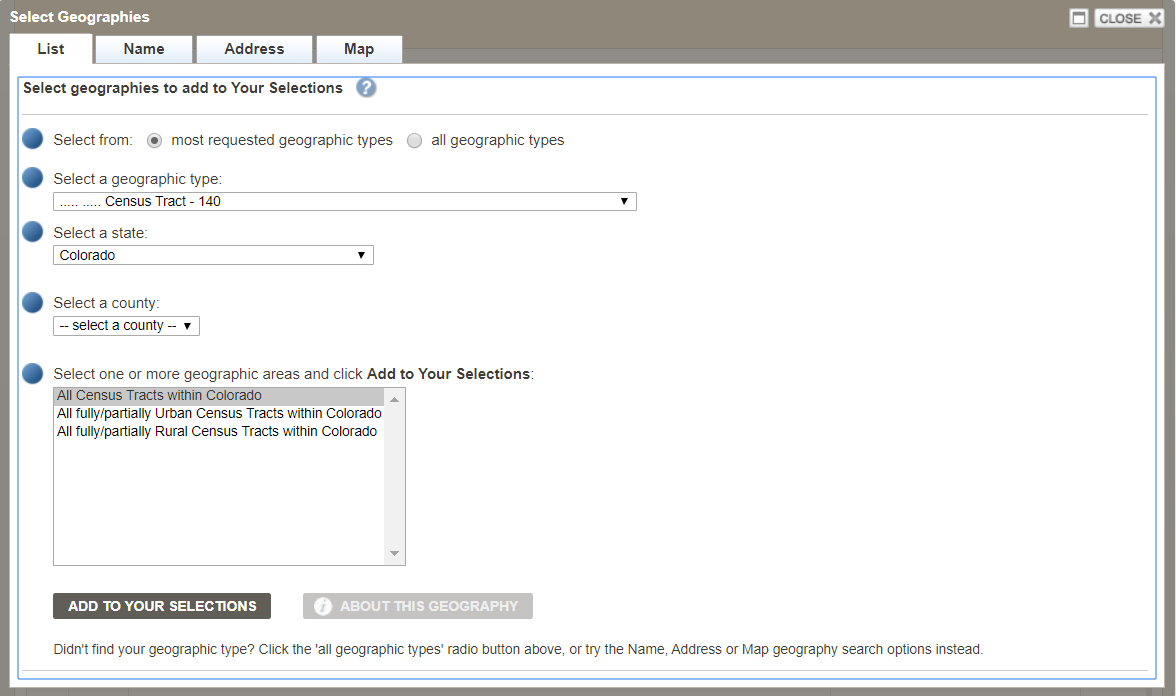




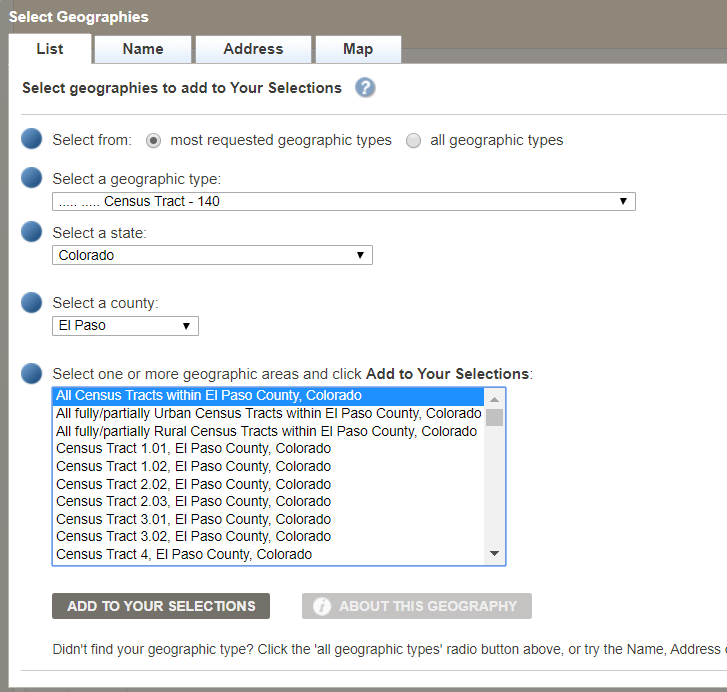
Find the Decennial Census at the bottom of the page and select get data



1. Click on **Geographies** in the left column – this brings up the *Select Geographies overlay*. This is where we tell it to get Colorado data ONLY by census tract! Once we set this search selection, the census website will only give us data that is available for this location and *scale*.



1. Fill out the box so that you are selecting **Census Tracts** for a specific state and a county in that state. You can follow the example below if you want to select all census tracts in El Paso County, Colorado. Alternatively, you could pick a state and county of your choosing.

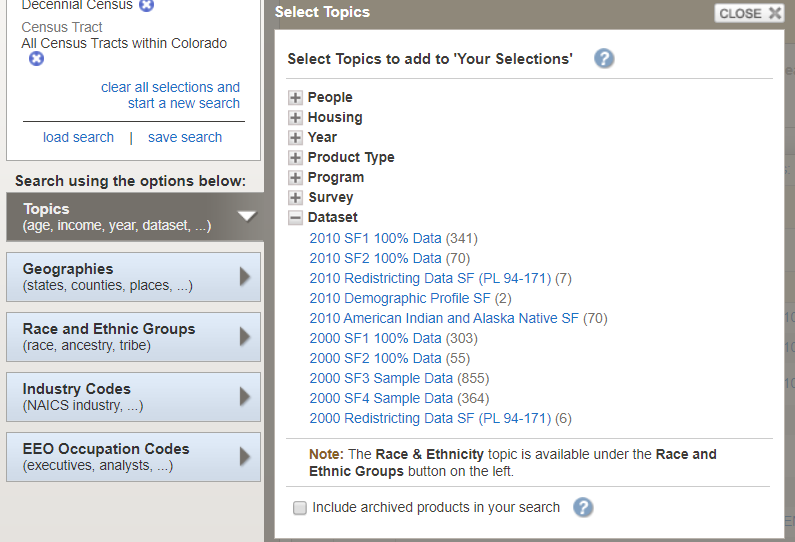


1. **Be sure to click on ADD TO YOUR SELECTIONS.**



1. Close the Select Geographies overlay by clicking in the top right corner of the box.

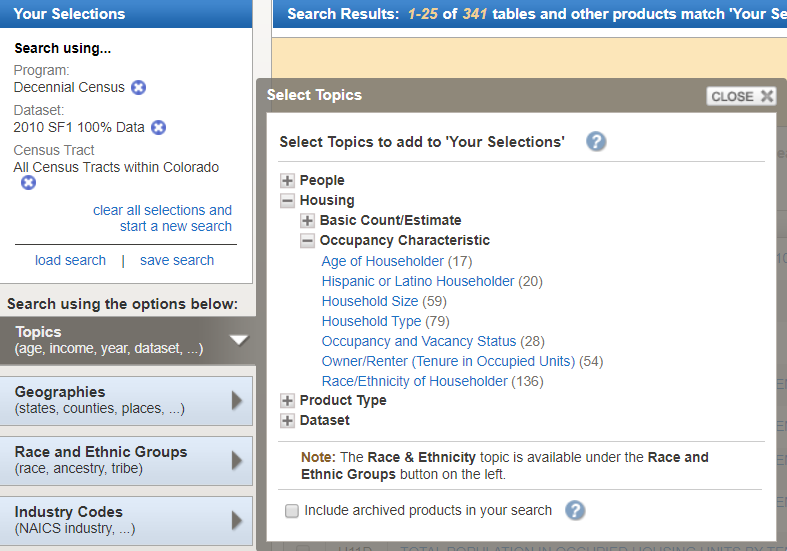
1. Now that you told it where you want the data, you need to tell it WHAT data you want to download. Today, you’ll be downloading information on Housing and Tenure.
2. Click on **Topics** in the left column and expand **Dataset.**



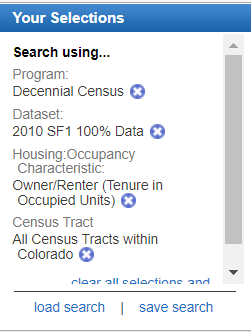
1. **S**croll down to click on **2010 SF1 100% Data** – this will send this criteria to your Selection box in the upper left of the site:



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



1. Close the **Topics** box by clicking in the top right corner of the box.
2. 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 4 above.



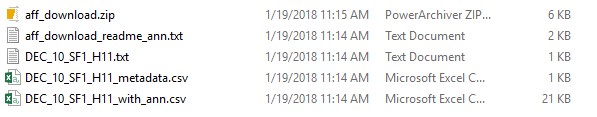
1. You should now have a list of available datasets about housing characteristics. **Checkmark** one of interest and see what variables it contains by clicking on the *Information* icon for that table.  This pulls up the information about this table. You’ll see what data will be included when you download it. This is very helpful for previewing datasets so you don’t have to go through all the steps of downloading it first. For this exercise, we highly recommend a table with just a few variables. In this exercise, we have used the **H11 table for total population in occupied housing**. Make sure this is checked.
2. After checking a table, click on **Download (** **)** and follow the instructions. This creates a zip

file. Save it in your *Census\_2010* *AFF Data* folder.

1. Navigate to your AFF folder. Right click on the zipped folder and select *extract here* or open with 7zip and extract to AFF folder.

# Preparing American FactFinder Data for Use in ArcMap

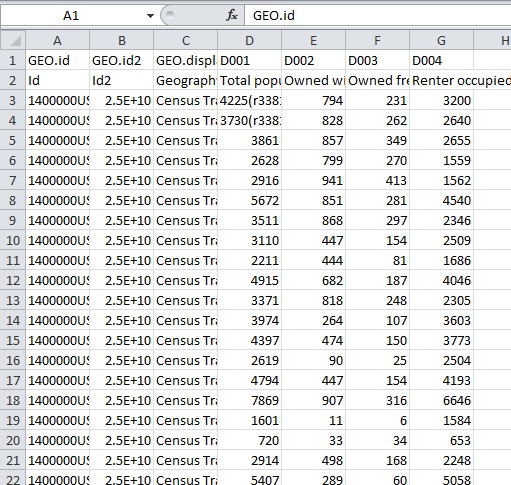
1. You’ll notice that 2 csv files have downloaded, along with two txt files. Double-click on both downloaded **CSV** files to open them:



Several files will be downloaded. The “with\_ann” file contains the data, while the “metadata” file contains the descriptions of the table headings. Open both excel files. Note: If you are opening the file from within Excel, you will need to set the option to look for all file types

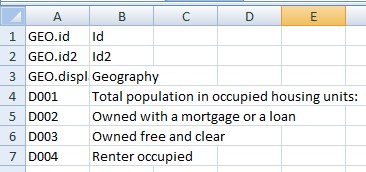
**Note**: If you are opening the file from within Excel, you will need to set the option to look for *all file types*:

1. The “*DEC\_10…with\_ann”* file should look something like this.



This file contains the **data** and the **headings**. Notice how there are essentially two headings. The first row contains “Data Speak” and the one below contains written out explanations.

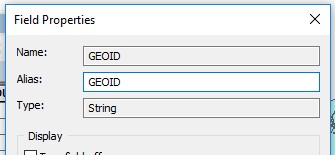
1. Now look at the “*DEC\_10… \_metadata”* file. This file explains the column header codes in the data file - it should look something like what you see below. This is a very important file!!!



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. 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. This process is called “normalizing.” **A few important steps left.**

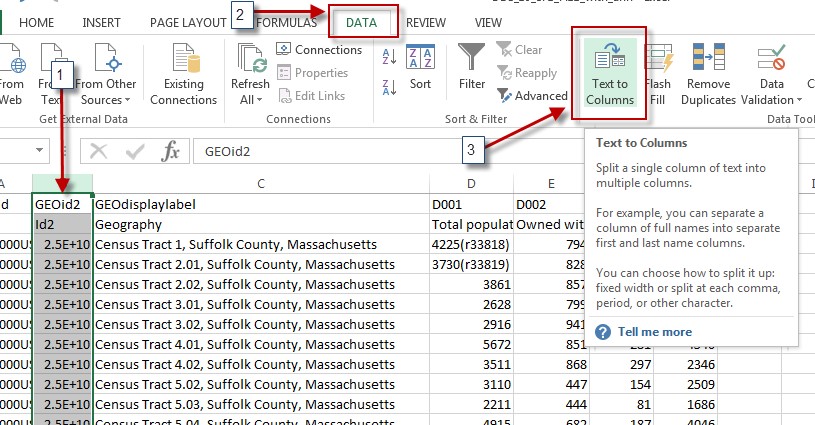
Go back to the excel file containing the data. Now we need to clean it EXTENSIVELY so that we can join it in ArcGIS Pro. ArcGIS Pro is VERY picky about how data is formatted!!! Make sure to follow all rules below:

1. The very top row (with all the data speak headings) will be our Attribute table headings. ArcGIS does not like ANY extra characters in the column names. Delete all periods (.) and extra characters (-) in all the **column names**/**heading**. The only acceptable character is underscores (\_).
2. In ArcGIS Pro, open the census tracts attribute table. Find the GeoID field heading. Right click on it and open the properties. You will notice that it says it’s a string. String means that it is formatted as text (not a number). It needs to be formatted as text so GeoIDs that start or end with a 0 don’t get adjusted to drop the 0 (for example, zip codes like 02144 don’t get automatically changed to 2144).

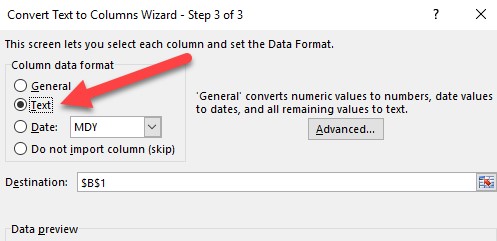


1. Since our joins need to have fields that match **EXACTLY**, both the excel data field and the attribute table field that will be matched (joined) need to both be strings/text.

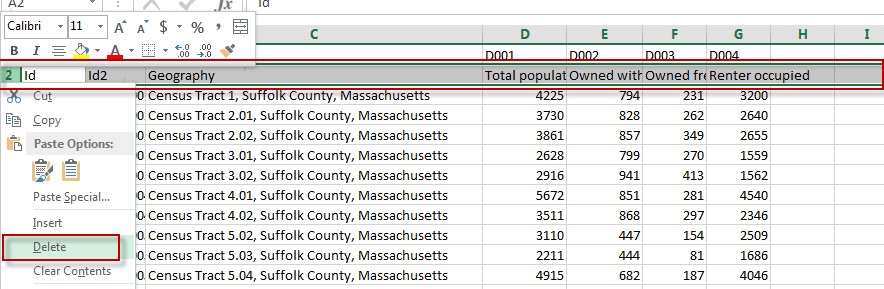
1. Back in excel, notice the **GEOid2 field. This is the field that exactly matches what is in the census tracts attribute table. However, we can’t really tell right now because it’s formatted as a number and therefore changes the number to read as 2.5E+10.** As a result, this field must be changed to *text* for the join to work properly. Follow the image below…
   1. Highlight the GEOid2 column.
   2. On the Excel toolbar at the top, click on the **Data tab**
   3. Click on *Text to Columns*:



* 1. Click **Next** to leave the first setting at Delimited.
  2. Click **Next** to leave the second setting at Tab.
  3. In Step 3, change the column data format to **TEXT**, and then hit **Finish**. This will format ONLY THIS ONE COLUMN as text (because we still want all the data itself to be a “number”).



1. Notice now how all the numbers in GeoID2 now look identical to how the GEOID field is formatted in ArcMap. We’re getting closer.
2. Census.gov now includes the description of the column under the column heading in the excel sheet (e.g. Under D001 it says Total Population). However, ArcMap does not like this extra row and the text causes the software to read it as a “string” (e.g. text) instead of “double” (e.g. numbers). Therefore, it is necessary to delete this row so that ArcMap realizes that this is a number field and not a text field.



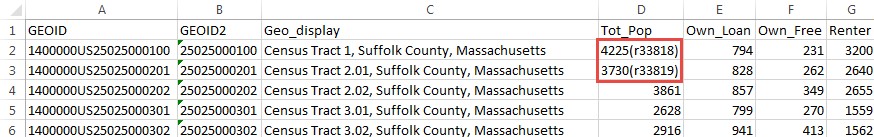
**Optional Tip –** Although you need to delete the 2nd row of text, you can change the column headings to the descriptions if it makes it easier (e.g. Change D001 to Tot\_Pop). However, there can be no spaces or periods and the heading needs to be under 9 characters. For excel sheets containing several fields, it’s probably easier to refer to the codes later rather than changing all the column headings.

1. To make things easier later, rename the worksheet to something comprehensible, e.g., *Housing\_Tenure -* the worksheet name will be the identifier in ArcCatalog.



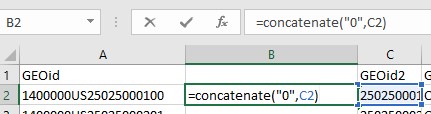
1. Very important step – **save your modified CSV file as an Excel Workbook (.xlsx)** – give it a comprehensible name, e.g., *2010CensusH11\_populationbyhousingtenure.xlsx.* Save it in your workspace Aff\_Data Folder

Note: Your table may have columns where the data has text values in it. For example, column D001 (which is total population) has data like 4225(r33818) or 3730(r33819). Since ArcMap uses the first eight rows to determine what the data type is for each column (ex: string, float, double, integer, etc.), it is necessary to delete the information in parentheses for it to be mappable as a number.



1. **Extra step for Alaska, Alabama, Arkansas, Arizona, California, Colorado, and Connecticut.** (*Ignore this section if you are not working in these states)*

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, create a new column next to GEOid2, in Row2 type in: **=concatenate(“0”,C2)**.



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. Save the file!

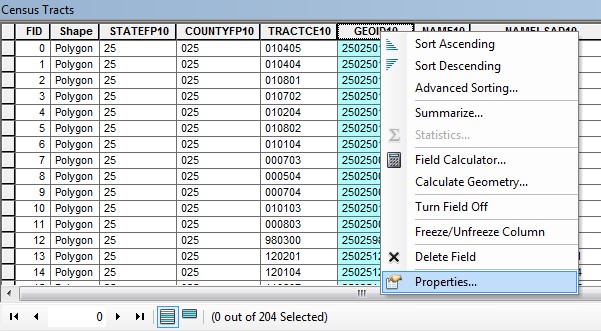
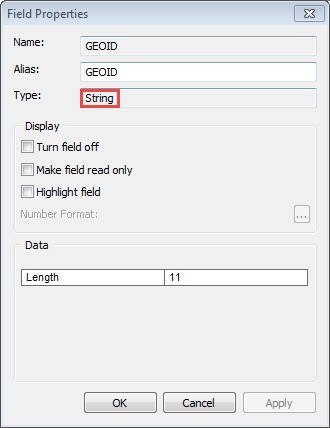
|  |  |
| --- | --- |
| **Final STEP: Save your file and EXIT out of Excel – you CANNOT have Excel open still when** | |
| **you work with this data in ArcGIS!** | **If you cannot open the excel file, save as .csv instead** |

Now you’re ready for mapping!

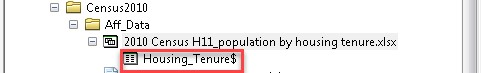
# 

# Joining the AFF table to your Census Tract polygons in ArcMap

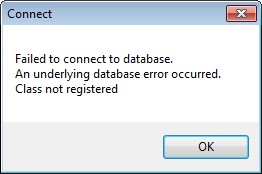
1. In ArcMap, open the Census Tracts polygon attribute table and take a look at it.
2. Which column matches the excel document EXACTLY? The GeoID in the attribute table matches the GeoID2 from the excel table exactly! **Therefore, we will use the *GeoID* column to join our AFF data.** Remember from class, that GeoID is a concatenation of the State FIPS Code (25 for Mass) plus the County FIPS Code plus the Tract FIPS code. Every single tract in the US has a unique ID! This unique ID is the best way to identify counties, especially since many different states contain counties with the same names.
3. Right-click on **GEOID** field name and choose **Properties**. You’ll see it is a **string** type attribute field – that’s important to know. Strings are the same as “text”, which is why we changed GeoID2 to “text” in excel! Close the table.

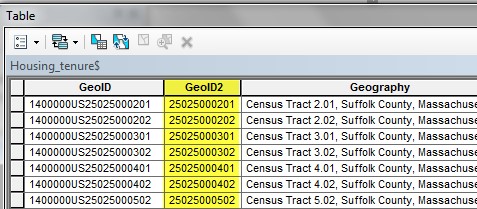
1. In Catalog, add your census excel sheet (or csv) to the map. If you don’t see it in Catalog, refresh the folder by rightclicking it and selecting “Refresh”. You will need to drill down to the sheet level, as shown below. If there is still a plus sign, you haven’t gone deep enough! Drag the “Housing\_Tenure$” sheet into your map like you would a shapefile. Or use add data. This step often will cause problems.



1. If you get the following error, it means that your version of ArcGIS and Excel are having connectivity issues. A solution may be to save your Housing Tenure excel sheet as an Excel 97-2003 Workbook (\*.xls) or CSV.



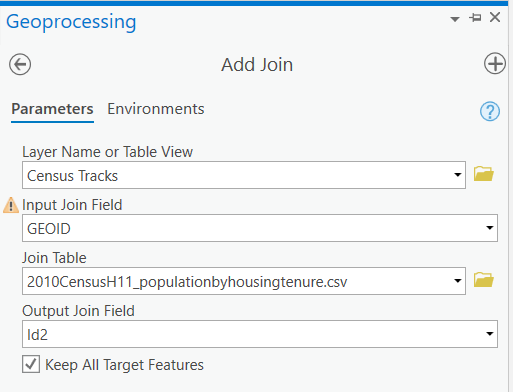
1. In the Table of Contents, open *Housing\_Tenure* Tableby right-clicking on it and choosing *Open-* very similar to how you would open an attribute table.



**GEOID2** will be used to join this AFF data to the 2017 Census Geography. These two fields match exactly. It’s OK that they have different names. Check its properties to ensure that it is also a STRING type, close the table when done.

**Take a second to realize that you will not being joining the GeoID to GeoID. You will instead be joining GeoID2 (from the Housing\_Tenure$ table) to GeoID in the Census Tracts shapefile.**

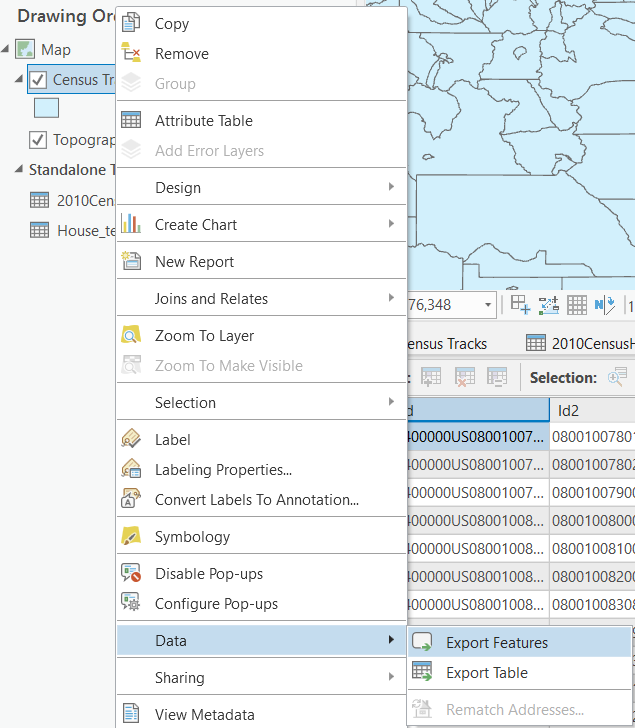
1. Right click on the GeoID2 field → Properties, and double check that the type is indeed a **String**. If it says double, you did not convert it to text in excel and must go back and redo that step (Page 8, step 4). That’s why I had you check twice. If it says string, you are good to proceed! Close the attribute table.
2. Right click on your **Census Tracts**and choose **Join & Relates**, then select **Join…**
3. Fill in the dialog box as follows. You are joining attributes from a table (the spreadsheet, in this case Housing\_Tenure$) using *GEOID* in your Census Tracts layer and *GEOid2* in your Excel table. Click OK when done:

****

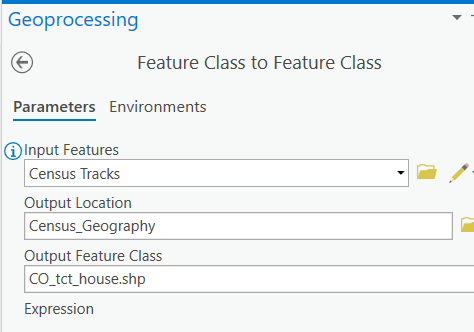
Note: By clicking “Keep only matching records”, only Suffolk County will remain visible in the shapefile because we only downloaded census tabular data for this one county in MA. If we select “Keep all records”, MA would remain whole, however, the attribute table would only have housing information for the census tracts within Suffolk County. The rest of the census tracts would read “Null” in those joined fields.

1. Open *Census Tracts* attribute table to ensure that the join was made correctly. If so, you should see your housing tenure AFF data when you scroll to the right in the table. Close the table.
2. **It is important to know that when you make a join it is not permanent until you EXPORT THE DATA**. Until you export, the join is only temporary. If you run any tools on this layer without exporting the data, it will drop the joined data. Exporting the data saves this shapefile as a NEW shapefile, where the join is now permanently part of the attribute table.

Export the data by right clicking on the census tracts, selecting **Data** and then **Export Feature.**



1. Save the shapefile with an appropriate name inside the Geography folder (include Census tracts so you know the boundaries and include the location). It can be good to acknowledge exactly what has been joined, especially if you end up having multiple joined layers. Also, make sure to save as a shapefile, otherwise you will encounter an error. Change location to Folder to save as shapefile

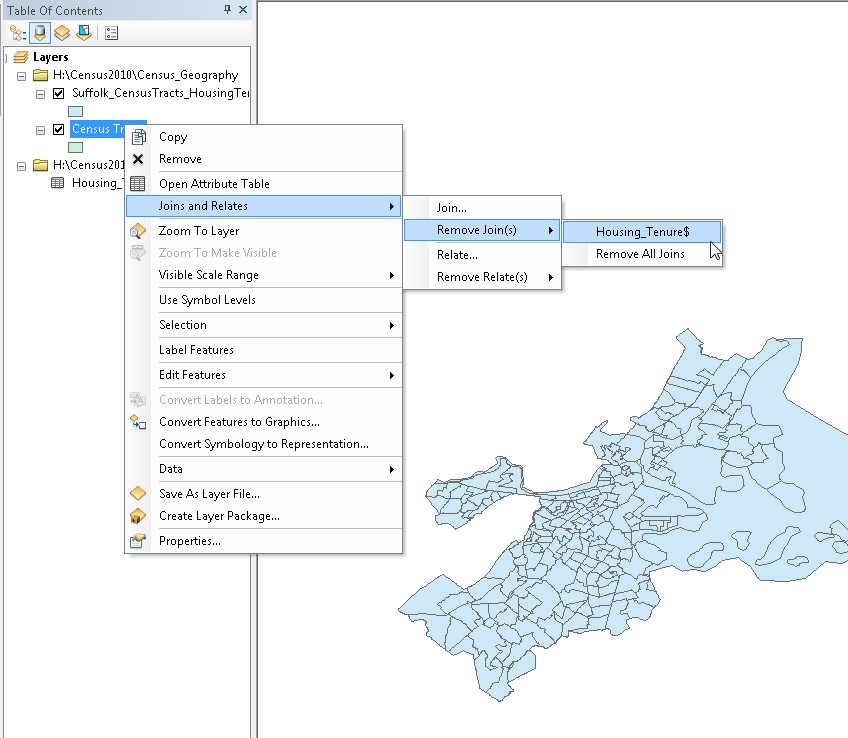


1. Exporting data is always good practice to ensure something is **permanent**. It also sometimes resolves minor ArcGIS glitches, such as layers not drawing on the map.
2. Now before hitting OK you will need to decide if you want to save this new shapefile to the coordinate system it came with (2016 Census Data uses GCS\_North\_American\_1983), or if you have already put the data frame into a certain projection you could select data frame (this map ultimately uses

NAD\_1983\_StatePlane\_Colorado\_central\_FIPS). By setting your data frame projection first, this then saves you the step of projecting your data later! You can do this on the fly in the Environments tab

1. Press ok and click yes when asked if you want to add the exported the data to the map as a layer. Now, you should see a new layer in the Table of Contents with your given name. It has also been added to the Catalog!

1. Since you have saved this join as new shapefile, you can remove the join from the original Census Tracts Shapefile. Right click on Census Tracts, and click **Joins and Relates** and then **Remove Join(s)**  **Housing\_Tenure$.**



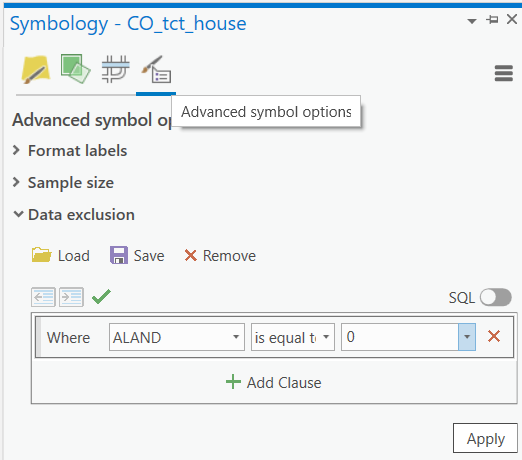
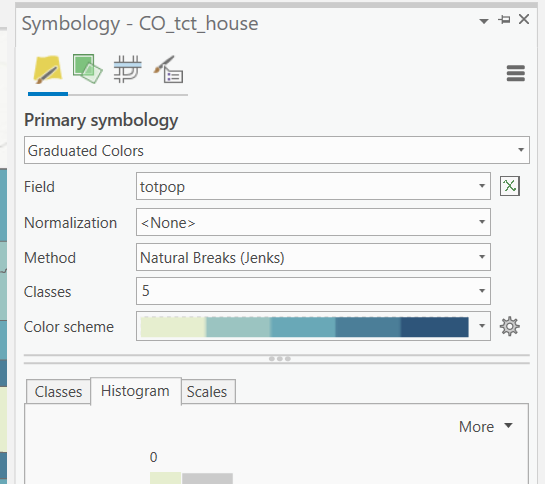
1. When you do this it removes the excel data in the Census Tracts attribute table, and the shapefile reverts back to all of Massachusetts. The reason why you would want to remove the join is so you can join a new or different tabular dataset to this shapefile without having to download it again from the census.
2. After you remove the join from the original Census Tracts layer, turn the layer off. You are done working with this one.

# Removing Water Only Census Tracts

If your area of interest is near water or has water features in it (like Boston), your census tract data will extend into the water. Why? Because people live on islands and on boats, and census tracts includes those areas! However, we don’t always want to map them, so let’s remove them.



1. To get rid of the water tracts from your new shapefile, right-click it and open the layer Properties, and click on the **Symbology** tab.
2. Click on *Quantities.*
3. Under *Value,* choose the variable **D001** (or Tot\_Pop if you renamed it in excel first).



1. Now you can map your data without having to include water-based census tracts.

**Note**: Do not just accept the default colors. Play around with the different color schemes!

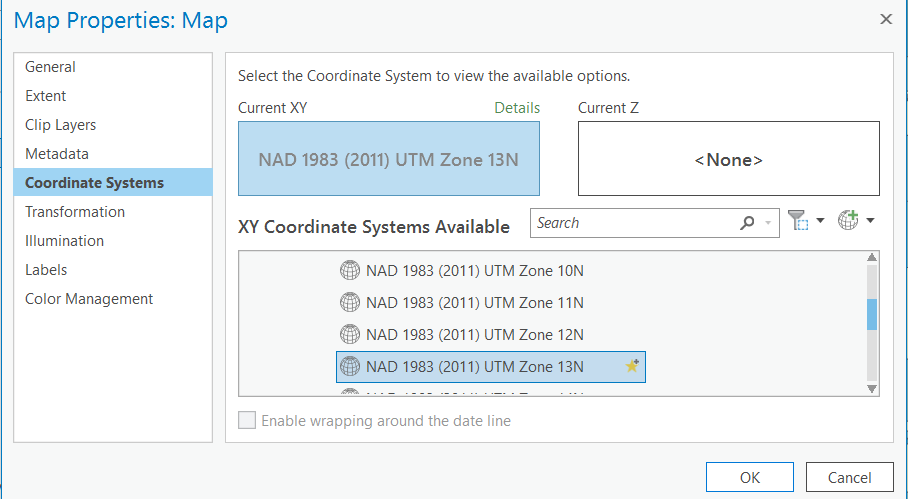
1. Your data is ready to map. Make sure to save often!

# Setting a Projected Coordinate System for your Map

It is always necessary to put your map into an appropriate projected coordinate system. The TIGER data is in a geographic coordinate system (GCS) and can appear distorted 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 of interest. In the case of Colorado, we will use the UTM zone 13 (NAD83) – meters coordinate system. If you don’t know which coordinate system to use, use [this resource.](http://www.geo.hunter.cuny.edu/~jochen/gtech201/lectures/lec6concepts/map%20coordinate%20systems/how%20to%20choose%20a%20projection.htm) It is a quick read that gives you an understanding of what factors to consider when selecting a projection, and provides a helpful table at the end.

1. In Contents Pane : Right Click on **Map**  **Properties.**
2. Click on the **Coordinate System tab.**
3. Scroll down till you find the **Projected Coordinate Systems** folder. Make sure you are not still in the “*Geographic Coordinate System*” folder.
4. Scroll down to the **UTM** – open that folder and select **NAD 1983 UTM Zone 13N (Meters)** from the list:



1. Click **OK** and click **Yes** when warned that the coordinate system is different from the data in your maps.

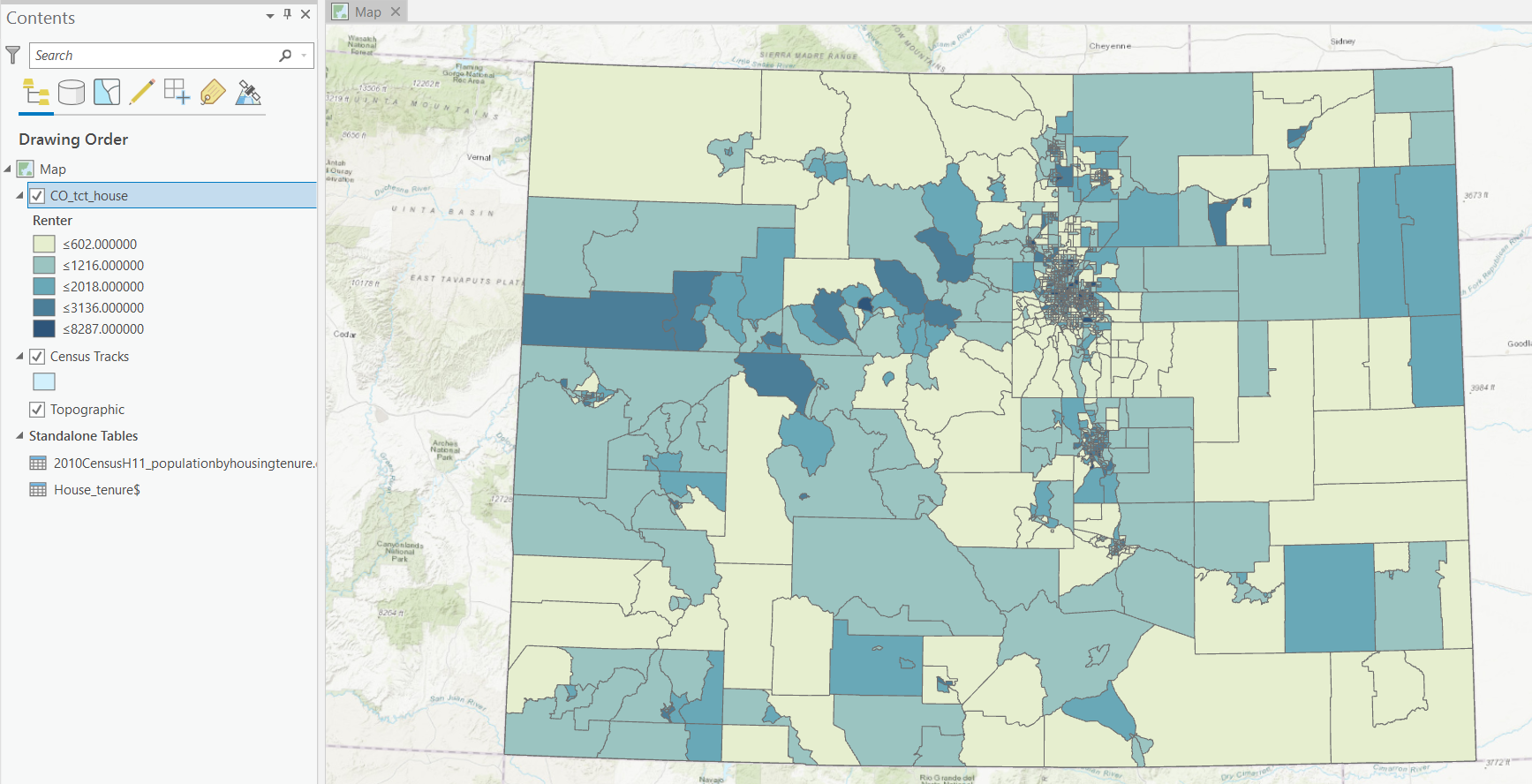
You’re done! You have successfully downloaded American Fact Finder data (AFF) and joined it to Census Geography

data in ArcMap. The process of finding and preparing data can be time consuming, but doing it correctly is important to ensure that your maps and the information they are conveying is accurate.

# Using Symbology to Map a Demographic Variable

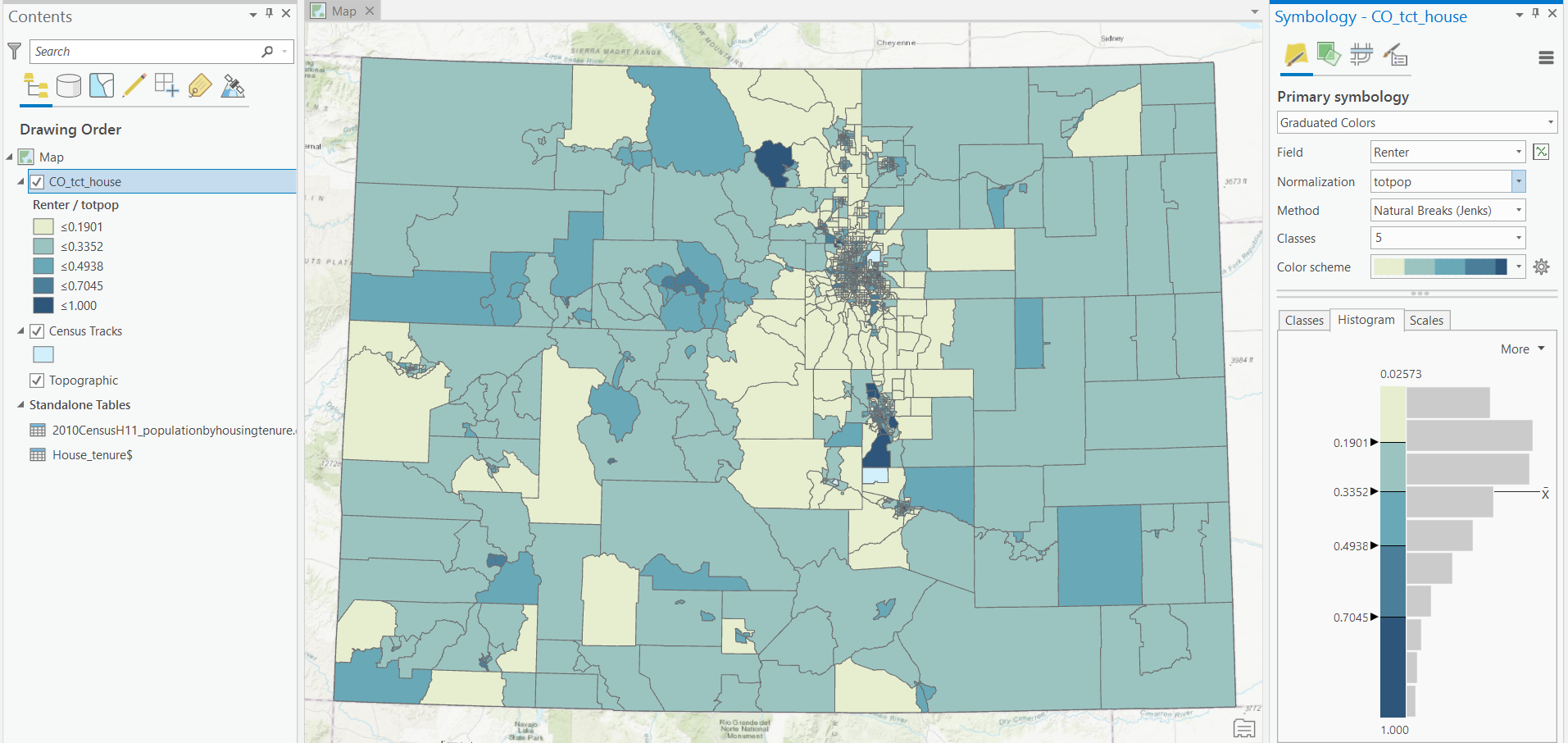
Now that our join is successful, we want to use symbology to show one of our demographic variables by census tract. Let’s open up the attribute table of the **Colorado\_CensusTracts\_HousingTenure** layer to remember what our options are for the demographic variables on tenure.

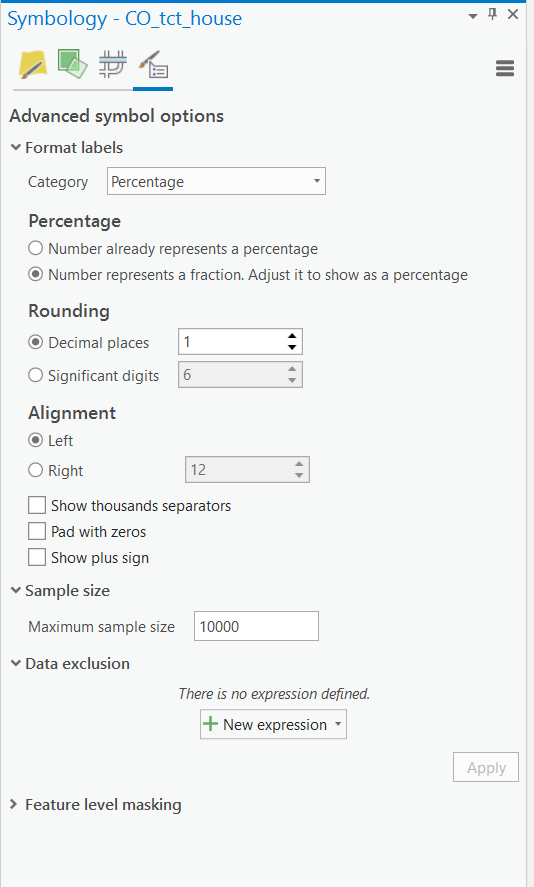
1. Let’s map the D004 variable (Renter Occupied). If we were to symbolize this variable, it would just be a *raw count* of the population living in renter-occupied housing units. Instead, we want to show a percent so that we can better compare between census tracts and better understand what the data means. This is called *normalizing*. It is very important when analyzing any dataset.
2. In order to make this number a percent, we must divide the **population of** **renter-occupied housing units (D004)** by the **Total population in occupied housing units (D001).**
3. Double click on your new shapefile **Colorado\_CensusTracts\_HousingTenure** to get to the properties. Then click on the **Symbology tab.**
4. Since we want to map numbers select **Quantities** on the left hand side of the symbology box.
5. The **Value** field dropdown contains all the **Attribute Table** headings. Let’s select **D004** or the variable that represents population of renter-occupied housing units (below it has been relabeled to **Renter\_Occ**)**.** Hit Apply – the map updates, even though the property box remains open. You can move the Property box out of the way to see the map! See below for an example of how to change the symbology.

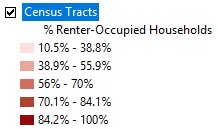


**Note**: As it stands, the map is NOT illustrating a percentage of the population living in renter-occupied housing units. This is just the **raw count of the population living in renter occupied households.** It is not useful for comparisons, since census tracts are varying sizes.

1. In order to compare values between census tracts, we must turn this number into a percent. This means taking the total population living in renter occupied households and **dividing it by the TOTAL POPULATION in occupied housing units**. This process is called **Normalizing** and it is an important statistical operation that allows comparison of two different sized areas.
2. To do this, we select the **Normalization** field in symbology. We know that **D001** is the field containing data for the total population in occupied housing units. If you changed field labels, you may have a field titled **TotalPop** or something similar instead**.** Change the normalization field to **D001** or **TotalPop**.



1. Click **Apply**. Notice how the map changes. Also, the numbers changed under the Label column. Now they are fractions representing population of renter-occupied housing units divided by total population of occupied housing units. To format the labels to show percentage, we must multiply them by 100.
2. ArcMap makes it easy to format numbers. Click on **Advanced Symbols tab in symbology**  **Format Labels…**
3. There are already a bunch of different number categories. Since we want to convert the fractions to percentage, click on **Percentage** on the left. Now, two options appear. Since the numbers currently represent a fraction, and we want to adjust it to show a percentage, select the 2nd option. Now, ArcMap will automatically multiply our numbers by 100 without us having to do anything!
4. If you hit OK now, our numbers will be in percentage, but they will have a lot of decimal places that make it hard to read. If you have hit OK, go back to the Number Format properties (click on **Advanced Symbols tab in symbology** > Format Labels > Percentage). In the **Number Format** box under Percentage, Set rounding to 1 decimal place.
5. Now **Apply**. Now our numbers are MUCH more simplified and easier to read.
6. Pick a gradual color scheme (light to dark) that you like in the **Color Ramp.** Press apply to see how the different colors look. Find one you like, and close the Layer Properties window by clicking OK.
7. Lastly, let’s change the layer name and heading in the table of contents so that when we add our legend it looks good! Change the name of Colorado\_CensusTracts\_HousingTenureto **Census Tracts**. Then change the **heading name** to **% Renter-Occupied Households.**



1. Save your map.

# Creating a Map Layout

1. Reference previous labs to revisit layouts.
2. Switch to Layout View (Insert> New Layout) so we’re setting up our piece of paper.
3. You’ll likely want to switch your paper to landscape, which is done in the File  page and print setup.
4. Set up the Map Frame (Insert > Map Frame) on the piece of paper how you would like it arranged. Then, zoom into Area of interest so it fills out the whole data frame area really well.