FINDING HISTORICAL CENSUS DATA

STEP 1: USING NHGIS

Open a browser and navigate to the NHGIS website: www.nhgis.org.
Familiarize yourself with this main page. NHGIS stands for National Historical Geographic Information Systems, and it is a great online resource for free census datasets from 1790 to 2014. In order to use NHGIS, you must first be a registered member. On the right hand side of the main page, click "Sign in". You will be navigated to Log in page where you can choose "Create an Account". Follow the necessary steps to create your free account.

Once you have logged in, return to the home page (www.nhgis.org). Notice the variety of options available to you on the left hand side of the page, including "Select Data" and "Data Documentation". These are links that will provide you with the resources to download different census datasets as well as to understand what the data means and how to use it.

NHGIS is specifically useful because it has a set of time-series tables already created that allow you to look at different variables such as race, or population density, across time. Instead of having to download a dataset for each year you are interested in, NHGIS Time Series Datasets provide groupings of years.

To download a time-series dataset, navigate to "Select Data" by clicking on the link on the left hand side of the main page.
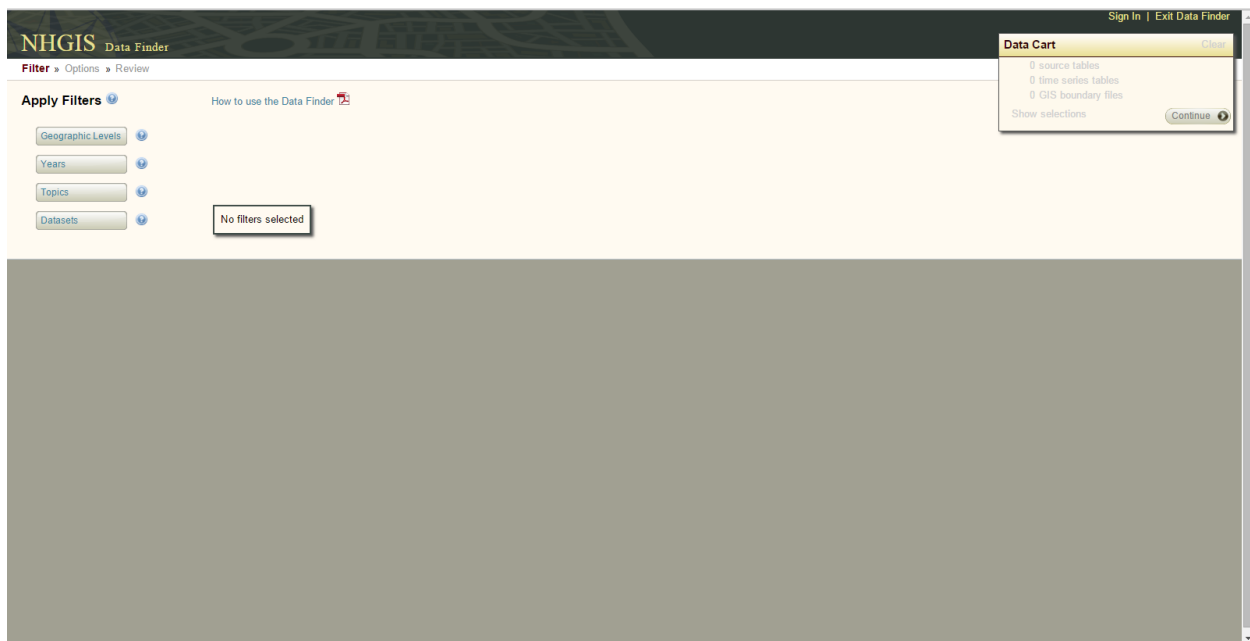
You will see the page above, which on the left hand side provides you with different selections including "**Geographic Levels**," "**Years**," "**Topics**," and "**Datasets**". Depending on what you're looking for, you may want to explore each of these selection tabs.

For example, if you were interested in a time-series dataset on race that gave you the variable Race from 1970 to 2010, you would first want to specify your Topic, which is....Race!

So, click on the "**Topic**" button. A pop-up will appear that allows you to specify your topic.
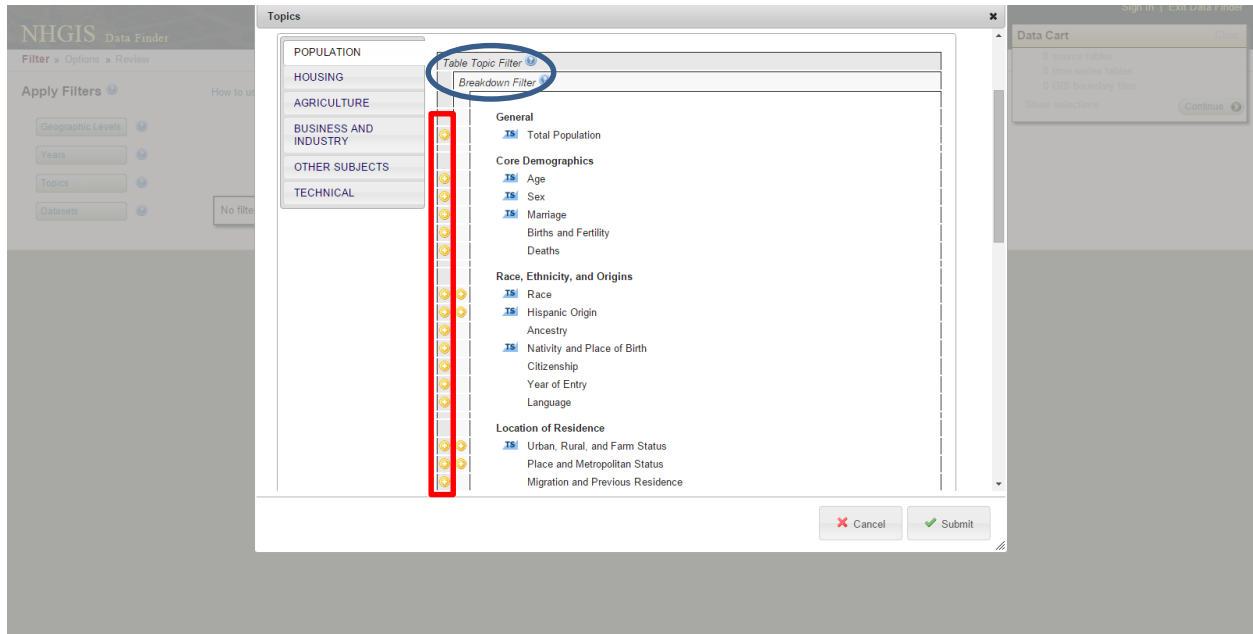
You will see different variables listed with little yellow plus-sign icons appearing next to them. Some topics have two yellow plus-sign icons, one for "Breakdown filter" and one for "Table Topic Filter". This is because some tables provide multiple breakdowns of variables such as Housing Characteristics by Age, or Income by Homeownership. In this case though, for a time series dataset, you must click the yellow plus-sign in the "Table Topic" column.

Once you have clicked on your variable's yellow plus-sign, you will see a checkmark appear next to it. Click "**Submit**" to finish this selection.

Your page should now list a number of tables available based on your selection. At the top of the listing, there are three tabs, "**Source Tables**," "**Time Series Tables**," and "**GIS Boundary Files**". Right now, we are interested in Time Series Tables. Click on the tab. You will now see basic breakdowns of the tables including their Popularity, Table Name, Geographic Integration, Years, and Geographic Levels.

You can also specify your geographic level of interest. If you want data at the level of the county or census tract, for example. Back at the top left, click on the "Geographic Levels" Filter button. Another pop-up will appear, this time with specific geographic levels to choose from.

Decide what unit you are interested in, such as the county, state, or census tract, and click on the yellow plus-sign next to that unit. Click "Submit".

You may notice the number of available tables has diminished. Make sure you are again looking through the Time Series Tables Tab, and choose the table best suited to your work. In the Race example, we may choose the first table, also the most popular, named "**Persons by Race**". Click the yellow plus-sign next to that table's name and popularity level.

Notice the pop-up in the top right hand side of your page, "**Data Cart**" which now shows that you have selected "1 time series table". Now that you have a table chosen, click "**Continue**" in the "**Data Cart**". You will then navigate to a page showing the table that you have selected. If this all looks right, once more click "**Continue**," in the "**Data Cart**".

You will now see a page titled "**Review and Submit**". Because each dataset is specific to the variables you have chosen, NHGIS takes a few minutes to about an hour to prepare you tables for you. The website sends you and email and a link when your dataset is ready.

The Review and Submit page allows you to choose how you want your data to be formatted. Under "**Table File Structure**" You have a few options. "Comma delimited is already chosen, and this is the format you want. Make sure to also choose "**Include additional descriptive header row**". This will give you definitions for each column in your data.

You also have a few options for your "**Time Series Table Layout**". For GIS and the analysis you will be working on, make sure to choose "**Time varies by column**".

Once you have those options selected, double check and then submit your data request.

Within a few minutes or about an hour, you will receive an email to the address you provided when you registered your account. The email will be from nhgis@umn.edu. The email subject heading will be "NHGIS Data Extract". It can sometimes take up to about a day if you request the data on a weekend or off-hours. Be patient!

This email will include a link which when clicked will prompt you to sign into your account. Once logged on, you will see your "**Extract History**" which shows when you requested the data, the status of your request. This table also provides a download link for the dataset, under "**Donwload Table Data**", click the "**tables**" link.

A zipped file will download onto your computer. Open and extract the folder onto your working drive.

This folder will include your data table, and a text file called the "codebook" which is your data dictionary, and explains what each column name means. If you did not specify a geography, your download may include two or more datasets, labeled by their geography (based on the last word in the table name such as nhgis0001_ts_nominal_**county**).

Notice that you did not specify which county, census tract, or state you want to look at. This is because NHGIS only allows you to download data for the entire country, divided up by your geographic level.

Open your dataset, and find the county or state you are interested in by looking at the column labeled "**STATE**," or if you know the name of the county you are interested in, "**COUNTY**".

Select every row labeled with your specific state or county, and copy and paste that into a new sheet in Excel.

Make sure to insert a row above the pasted rows, so you can also copy and paste the Column names (first row in your original dataset) and include them in this new, cleaned up sheet. This way, you can focus only on your region or place of interest, and also be able to tell which column corresponds to which variable.

The codebook is very useful to keep track of this as well. Because all of the column names are confusing combinations of letters and numbers, you may want to rename them so that you know which column is which.

Use the codebook to determine what each column name means, then rename it with your own title. For example. If you have a column called "B18AA2010" that corresponds to 2010 estimated number of White residents, you may want to rename that column, "White_Pop2010," and so on.

Save your dataset. When making graphs in your file, make sure to save it as an Excel not a CSV file!