

Using the NHGIS Data Finder

October 2012

This guide steps users through **two scenarios** of differing complexity.

Scenario 1 will use NHGIS to select and download one data table from one census year, along with the corresponding GIS boundary files.

Let's assume we want to know how many farms are in each state across the United States in 1920, and we want to map that information in a GIS.

Scenario 2 will use NHGIS to select and download a time series table, which includes data from different years. In addition, GIS boundary files will be selected and downloaded.

Let's assume we want to compare owner occupied housing rates of Whites versus Asians & Pacific Islanders for the years 1980, 1990, 2000, and 2010 at the county and state geographic levels to see if they have shifted over time.



Funding provided by the National Science Foundation and National Institutes of Health.
Project support provided by the Minnesota Population Center.

Scenario 1: Basic Extract Request

Let's assume we want to know how many farms are in each state across the United States in 1920, and we want to map that information in a GIS.

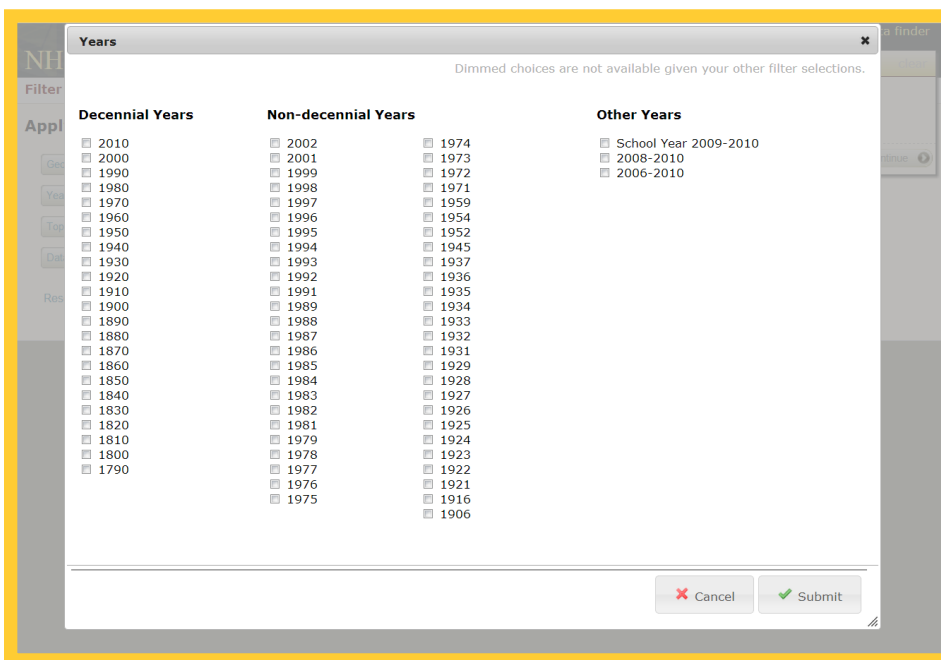
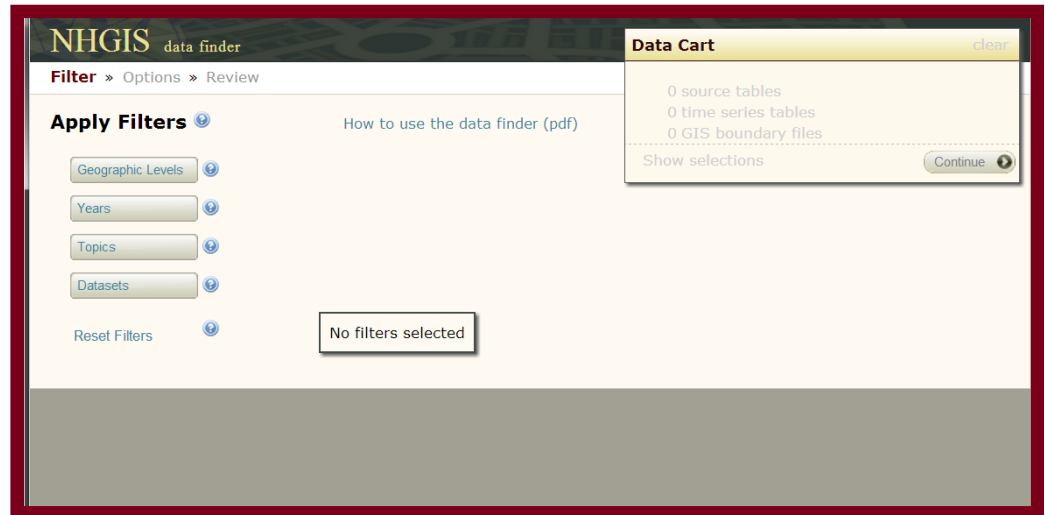
Step 1 - Filter

Scenario 1: Basic Extract Request

NHGIS utilizes four filter types to search for data. Using any combination of filters one chooses allows for maximum flexibility in honing in on the desired data.

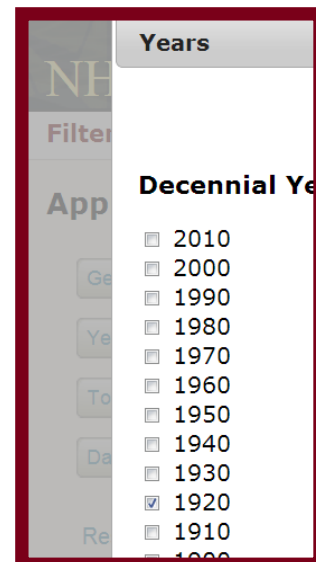
The four filters, as shown below, are:

- Geographic levels•
- Years•
- Topics•
- Datasets•



Clicking the **Years** filter button will open this popup window. From here, you can select any combination of available options. After you make your selection, and click **Submit**, you will see corresponding data tables.

For this example, let's select the year '1920'.



Step 1 - Filter

Scenario 1: Basic Extract Request

Clicking **Submit**, the resulting tables are displayed. Notice the '1920' visible to the right of the **Years** filter button. Clicking the green check mark will toggle the filter selection on and off, altering the tables displayed. With only one filter in use, however, turning it off will result in zero tables displayed.

The screenshot shows the NHGIS data finder interface. At the top, there is a navigation bar with 'Filter', 'Options', and 'Review'. Below this, the 'Apply Filters' section is visible, with a 'Years' filter set to '1920'. A 'Data Cart' on the right shows 0 source tables, 0 time series tables, and 0 GIS boundary files. The 'Select Data' section below shows a table with 190 source tables, 2 time series tables, and 3 GIS boundary files. The table has columns for Table Name, Universe, Classifications, Year - Dataset, and Breakdowns. The table is currently displaying 11 rows of data.

Table Name	Universe	Classifications	Year - Dataset	Breakdowns
NT001. Average Value Of Farmland And Buildings	Farms		1850_1959_cFV	
NT1. Percentage of Employed Population 10 Years of Age and Older by Occupation by Reporting Years	Employed Persons 10 Years and Over	Occupation (9), Year (2)	1920_sOccFarmer	
NT2. Employed Population 10 Years of Age and Over	Employed Persons 10 Years and Over		1920_sOccFarmer	
NT3. Employed Population 10 Years of Age and Over by Occupation	Employed Persons 10 Years and Over	Occupation (9)	1920_sOccFarmer	
NT4. Total Farms	Farms		1920_sOccFarmer	
NT5. Race of Farmer	Farms	Race (2)	1920_sOccFarmer	
NT6. Sex of Farmer	Farms	Sex (2)	1920_sOccFarmer	
NT7. Race of Farmer by Sex of Farmer	Farms	Race (2), Sex (2)	1920_sOccFarmer	
NT8. Race/Nativity of Farmer	Farms	Nativity (2), Race (2)	1920_sOccFarmer	
NT10. Farms by Size	Farms	Farm Size Class (10)	1920_sOccFarmer	
NT11. Farm Size by Race of Farmer	Farms	Farm Size Class (10), Race (2)	1920_sOccFarmer	

From this window, you are able to browse through the source tables, also referred to as the data tables. In addition, you can select the tabs that say time series tables and GIS boundary files which contain more files; time series data will be discussed in-depth in Scenario 2, while additional detail on selecting GIS boundary files is found at the end of Step 1.

As you begin to browse the Source tables, the first thing you may wish to change is the number of tables to display per page from the default of 20. Notice that changing this value results in changing the number of pages of results.

The five fields shown (Table name, Universe, Classifications, Year - Dataset, Breakdowns) provide detailed information on the tables' content. Each of field can be sort in ascending or descending order to make browsing easier. In addition, text in blue can be clicked on to reveal additional information about the cells' content.

To take full advantage of NHGIS, it is important to understand what these five fields are telling you.

Step 1 - Filter

Source Tables:

"Table name" is self-explanatory. The name itself, for later census years, matches the Census Bureau's name they gave the particular table. Table names found in the historical census years are a combination of actual names created by the Census Bureau and names created by those parties who converted the data from hard copy, published tables to a machine readable format.

"Universe" refers to the body of people or things (e.g. housing units) being counted. This is important as population totals, for example, will vary wildly depending on whether all persons are being counted or just those of a certain age or a particular race.

"Classifications" details how the data are categorized. The numbers in parentheses refer to the number of categories the data are broken down into. This can range from one value for the table, such as a *Total* field to well over 1000 values; for example, a table from the 2000 Census called Population by Race by Sex by Age has a classification schema of Race (7), Sex (2), Age (103) that creates 1442 fields in the table.

"Year - dataset" lists the dataset in which the table is located. Tables from older censuses, like 1790, are contained in one or two datasets each. The year 2000, however, yields 29 datasets.

"Breakdowns" describe ways the data's universe can be subdivided into more specific groups. Options that may be available include spatial, industry, and race-ethnicity among others. For example, a spatial breakdown could show the total population for counties, along with those values divided into the urban and rural population totals within each county. With a race-ethnicity breakdown, the total population for counties could be subdivided to return just the total population that identified themselves as Laotian or Apache American Indian. Breakdowns are not available for all data tables and are nonexistent prior to 1970. While you see which breakdown options exist here, the actual breakdown selections are made in the next step of the data extract process.

Time Series Tables:

"Table Name", **"Years"**, and **"Geographic Levels"** are discussed later in Scenario 2.

GIS Boundary Files:

"Year" is self-explanatory. Geographical boundaries change from year to year, so it is important to select the appropriate year's GIS files that will correspond to the tabular data you are using.

"Geography" refers to what spatial area is being delineated by the GIS file. Each GIS file only contains one geography (e.g. only states, only census tracts, etc.).

"Extent" details what area is being included in the GIS file. NHGIS provides data at the nationwide and state level extents. If you want a GIS file of census blocks for one specific county, you download the entire state's census blocks. Be aware that the extent does not imply full coverage for the area. Year 1920, for example, includes the Census Tract *geography* at the *extent* of the United States. This includes all census tracts in the United States, but only those areas which had census tracts in 1920 are included; in this case, only the major cities.

Step 1 - Filter

Scenario 1: Basic Extract Request

The screenshot shows the NHGIS data finder interface. On the left, the 'Apply Filters' section includes 'Geographic Levels', 'Years' (set to 1920), 'Topics', 'Datasets', and 'Reset Filters'. The 'Select Data' section shows a table of 190 source tables, 2 time series tables, and 3 GIS boundary files. The table lists various datasets, with 'Total Farms' (table ID 174) highlighted with a yellow circle and a checkmark. A 'Data Cart' window in the top right corner shows '1 source table', '0 time series tables', and '0 GIS boundary files', with a 'Continue' button highlighted in green.

Don't forget!

Click on
blue text to
reveal more
detailed
information.

Clicking the plus sign next to a table, such as Total Farms, adds that table to the Data Cart. Clicking the check mark would remove the table from the Data Cart. Switching over to the GIS boundary files tab, clicking the plus sign next to the 1920 State file adds it to the Data Cart.

Once you have finished filtering and making selections, you can click on *Show selections* to review the tables and GIS files selected, or simply click **Continue** to move on to the next step. Conversely, you can start over if you wish by clicking the *clear* on the Data Cart to remove everything you had selected. Choosing *Reset Filters* will remove filter selections but will not affect your Data Cart.

Remember!

You can apply
different filter
combinations
in the same
extract session
and download
any data in
one simple
extract .

The screenshot shows the NHGIS data finder interface. On the left, the 'Apply Filters' section includes 'Geographic Levels', 'Years' (set to 1920), 'Topics', 'Datasets', and 'Reset Filters'. The 'Select Data' section shows a table of 190 source tables, 2 time series tables, and 3 GIS boundary files. The table lists various datasets, with '1920' (table ID 1920) highlighted with a red circle and a checkmark. A 'Data Cart' window in the top right corner shows '1 source table', '0 time series tables', and '1 GIS boundary file', with a 'Continue' button highlighted in green.

Step 2 - Options

Scenario 1: Basic Extract Request

Once the data and/or GIS files have been chosen, the next step is to define which Geographic Level you are interested in. Notice that the text in red is indicating that a geographic level has not been selected. Clicking either on the **Select geographic levels** button, or the red *0 of 2* text will open a window where the geographic selection level can be made. Let's select the 'State' option, as it will match with the GIS file we selected.

Data Options

One or more tables lack a geographic level selection (see below).

Source Tables

Select geographic levels

Dataset	Tables	Selections		
		Geographic levels	Breakdowns	Years
1920 Census: Occupation Data & Farmer Characteristics [US & States]	1 table	0 of 2		1 of 1

GIS Boundary Files

Year	Geographic level	Extent
1920	State	United States

Data Cart

1 source table 1 dataset
0 time series tables 0 years
1 GIS boundary file 0 geographic levels

Show selections Continue

Geographic Levels for Dataset

1920 Census: Occupation Data & Farmer Characteristics [US & States]

Select Geographic level Show compound geographic levels

Nation

Nation

State

State

Cancel Submit

Click on **Submit** to confirm your selection, and then **Continue** in the Data Cart to proceed.

Step 3 - Review

Scenario 1: Basic Extract Request

Once you have selected the tables and GIS files you want along with the desired geographic level, the final step before submitting your request is to choose the format in which you want the data tables. GIS files only come in the shapefile format, so no options are available. The data tables, however, can be prepared in a **Comma delimited** format, either with or without **additional descriptive header rows**; or the data can be downloaded in a **Fixed width** format.

The screenshot shows the 'Review and Submit' page in the NHGIS data finder. The page is titled 'Filter » Options » Review'. Under 'Review and Submit', there are three main sections: 'Data File Structure' with radio buttons for 'Comma delimited (best for GIS)' (selected), 'Include additional descriptive header row (best for spreadsheets)' (checkbox), and 'Fixed width (best for statistical packages)'; 'Time Series Tables' showing '0 time series tables'; 'Source Tables' showing '1 source table' and '0.0 MB'; and 'GIS Boundary Files' showing '1 GIS boundary file'. There is a 'Description' text area and a 'Submit' button at the bottom.

Comma delimited is the ideal choice for using the data in a GIS. The field names consist of a unique ten character code created by NHGIS that does not contain spaces or other special characters. This allows for the .csv file (comma delimited) to be added to the GIS without errors. This format can also be used with a spreadsheet software like Excel. To decipher the nondescript field names, you will need to use the codebook file that is automatically included in your data extract. This is a .txt file that includes important information about your extract. In addition to the field name key, the codebook also lists the table and dataset names you downloaded, along with citation information.

If your intentions are only to use Excel or another spreadsheet software, you may be better served by clicking the **Include additional descriptive header row** box. This option will return the same .csv file, but it will include an extra row of field names that can be read and understood without using the codebook. The format of these descriptive field names, however, prevents the data from being used in a GIS without error.

Fixed width is the best format to select when you will be using the data in a statistical package like SPSS, Stata, or SAS. This download option will provide the .dat fixed width file in addition to three separate command files, one each for the aforementioned statistical packages. The command files are a .sps for SPSS, .do for Stata, and .sas for SAS. The .txt codebook file is also included in the download.

Step 3 - Review

Scenario 1: Basic Extract Request

After making your selection on the data file structure, it is recommended that a brief description of your extract request be typed into the Description text box. This description is a helpful way of keeping track of each extract. Ideally, the description should be no more than a few key words. In our example, let's call our extract "Total Farms - 1920 - all states - .csv - GIS." This way, you can easily remember what is included in the extract. Of course, you may need to get more creative or use more words when downloading multiple tables, geographies, and years.

Once you click **Submit**, you will be taken to the Extracts History page if you are already logged in to NHGIS. Otherwise, you will be prompted to login or set up a new account. Setting up an account is both free and easy.

Once on the Extracts History page, you can watch the progress of your extract request by refreshing your browser page. You'll notice that the status changes as it transitions from queued, in progress, and complete. You may also exit NHGIS and simply return to download your data at a later time. An email will be sent once your extract is ready.

NHGIS data finder

[Create another extract](#)

Extracts History

The download progress of your latest extract may be seen by refreshing your internet browser.

Extract Number	Date Created	Download Data	Download GIS	Status	Revise Extract	Resubmit	Description (click to edit)
1	2012-Oct-03	data (2 KB)	gis (26 MB)	complete	revise	resubmit	Total Farms - 1920 - all states - .csv - GIS

Documentation for NHGIS datasets is available [here](#).

Research using NHGIS data should cite it as:
Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011.

The data tables and GIS are downloaded separately by clicking on *data* and *gis*. Both are zipped folders that contain the data you requested.

In addition to simply downloading your extract, from the Extracts History page you can revise your extract, resubmit your extract, or change its description.

Revise Extract is very useful as it allows you to create a new extract with all the parameters preset from the original extract. If you want both the comma delimited and fixed width files, for example, revising the extract would be the fastest way of getting the two formats.

Resubmit allows you to request the same exact extract again. The downloadable data is cleared off the Extracts History page every two weeks to preserve storage space, while the parameters of the extract are maintained indefinitely. This method benefits NHGIS immensely while keeping users' inconvenience to a minimum.

Description can be altered by clicking on the existing text or blank space if no text was entered.

Download Your Data

After you have clicked to download your extracts, you'll be prompted to open or save the zipped folder(s). Inside each zipped folder, again, will be multiple files. In our 1920 example, the data folder will contain two files. The first is named *nhgis0001_ds44_1920_state.csv* and is the comma delimited file that can be brought into a GIS or opened in a spreadsheet software. The second file is named *nhgis0001_ds44_1920_state_codebook.txt* and is the codebook that includes the key to understanding the field names along with other information about the downloaded data. In both instances, your naming structure may vary slightly as the *0001* that follows *nhgis* is the sequential extract number. In our case, it is the very first extract we have ever done. If it is your second extract, you would expect a *0002*.

The zipped folder with the GIS data is typically much larger and contains zipped folders for each shapefile group, with each of those folders containing several files. In our example, the internal zipped folder is called *nhgis0001_shapefile_us_state_1920.zip* and in it is 21 separate files. These 21 files, however, are just three shapefiles that can be opened in a GIS. One is for the state of Alaska, another is Hawaii, and the final one is the contiguous United States; note that even though Alaska and Hawaii were territories at the time, the GIS naming structure still refers to them as states. Take care when moving or copying these files to keep the seven files for each shapefile together. If the files are separated from each other, the shapefile will not open in a GIS. Ideally, file management of shapefiles should be done within Esri ArcCatalog or a similar software.

The image shows two screenshots of Windows Explorer windows. The top window, titled 'GIS Download', shows a folder named 'nhgis0001_shape' with a size of 27,263,562 bytes and a packed size of 27,254,914 bytes. The bottom window, titled 'Data Tables Download', shows a folder named 'nhgis0001_csv' with a size of 5,403 bytes and a packed size of 1,975 bytes. A third window, partially visible in the background, shows a folder named 'nhgis0001_shapefile_us_state_1920.zip' with a size of 29,060,380 bytes and a packed size of 23,383,897 bytes. This folder contains 21 files, including shapefiles for the US, HI, and AK states, along with their respective metadata files (shp.xml, shp, shx, sbn, prj, dbf).

Name	Size	Packed Size
US_state_1920.shx	492	428
US_state_1920.shp.xml	59,594	12,191
US_state_1920.shp	29,060,380	23,383,897
US_state_1920.shx	164	90
US_state_1920.sbn	612	435
US_state_1920.prj	467	281
US_state_1920.dbf	4,031	1,766
HI_state_1920.shx	108	58
HI_state_1920.shp.xml	57,992	12,031
HI_state_1920.shp	720,624	590,917
HI_state_1920.shx	116	60
HI_state_1920.sbn	132	68
HI_state_1920.prj	459	271
HI_state_1920.dbf	335	158
AK_state_1920.shx	108	59
AK_state_1920.shp.xml	59,979	12,163
AK_state_1920.shp	3,812,620	3,245,060
AK_state_1920.shx	116	60
AK_state_1920.sbn	132	68
AK_state_1920.prj	460	272
AK_state_1920.dbf	335	159

Congratulations! You have now gone through all of the steps necessary to get the data needed to prepare a simple map of the number of farms by state for 1920. Remember to review the additional help documentation if assistance is needed in joining the data tables to the GIS shapefiles or with other issues.

Final Notes

Number of files in download - The number of .csv files you receive in your download is dependent on the number of datasets and geographic levels from which you requested tables. If you only extract tables from a single dataset and geographic level (like our 1920 example), you will only get one .csv. Selecting tables from the 1920 and 1930 Censuses, however, would result in two .csv files. If you added the county level geography, in addition to the state level geography, the two .csv files would become four .csv files.

Too many columns or rows for GIS - In addition, it is important to understand that the number of fields in the .csv can be very large if you select many tables or even a few large tables. Older versions of Excel cannot handle files with more than 255 columns in a worksheet and will simply cut off the remaining columns. Even more problematic is that ArcGIS (prior to the 10.1 release) also limits tables to 255 columns (fields), except in geodatabase tables, and will exclude all other columns beyond that range. Another unfortunate quirk when importing .csv files directly into ArcGIS is that when the .csv exceeds 2048 rows (shown as *2000 in the attribute table), it will not include the extra rows UNLESS in the attribute table you immediately click the arrow to go to the last record. Then it will process and include all of the records. As you would expect, exceeding 2048 rows is pretty common when working with most census geographies.

Join data in GIS - It is very important to realize that even when downloading data tables and GIS files in the same extract that the data tables are not attached automatically to the GIS files. This 'join' must be done by you, inside of the GIS. NHGIS has made it very simple, however, to successfully complete a 'join' operation. In both the data tables and the attribute table within each GIS shapefile, you will find a field called *GISJOIN* that is the key to linking the separate files together. For detailed information on completing the 'join' operation, please review the *Using NHGIS with GIS* user's guide.

Scenario 2: Time Series Extract Request

Let's assume we want to compare owner occupied housing rates of Whites versus Asians & Pacific Islanders for the years 1980, 1990, 2000, and 2010 at the county and state geographic levels to see if they have shifted over time.

Step 1 - Filter

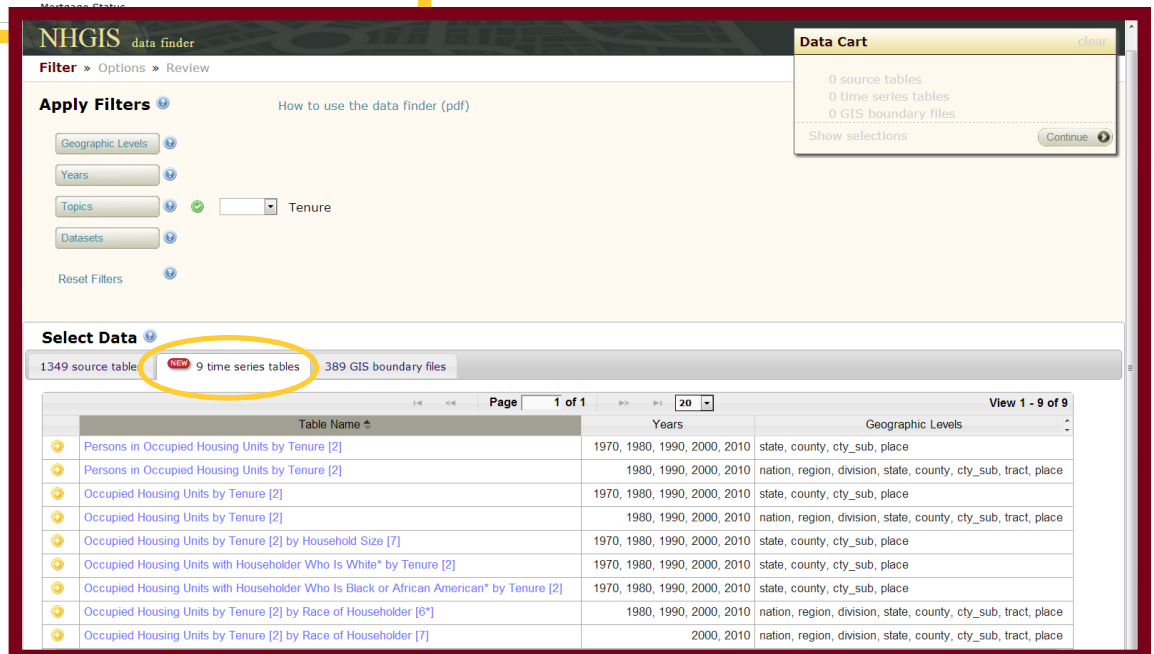
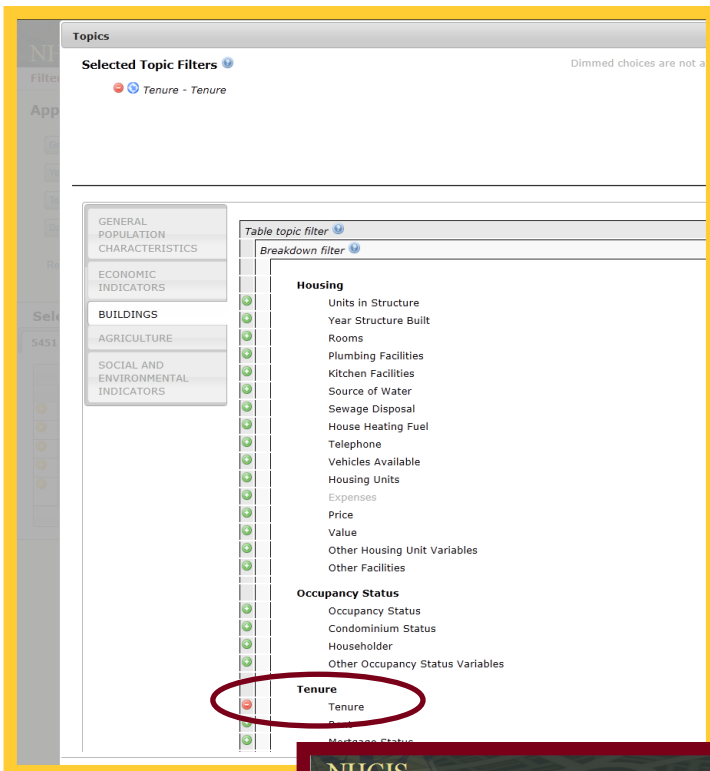
Scenario 2: Time Series Extract Request

This advanced request is more in line with the type of questions NHGIS users are commonly seeking answers to. It is a complex extract because it touches on data from different years, geographies, and multiple topics. In addition, trying to find comparable data from year to year frequently poses challenges as data is not always collected and reported in the same fashion from one census to the next. However, NHGIS Time Series Tables alleviate much of this difficulty as the task of selecting equivalent tables from each decade and harmonizing the changing variables from each census has been done for you!

On the Filter page, we can narrow the data down in a number of ways as we have particular years, topics, and geographic levels that we need. While the **Year** filter or **Geographic Levels** filter will help narrow down the potential tables, it is the

Topics filter that will help most as the topic of owner occupied housing encompasses far fewer tables than our other options. Open the **Topics** filter and select the **Buildings** tab. Since we are interested in homeownership, let's scroll down and select the 'Tenure' table topic filter that is listed under the Tenure subcategory. Tenure, to the US Census Bureau, describes whether the housing unit you reside in is owned or rented.

If you click 'Submit', you'll see a staggering 1349 resulting data tables spanning the years 1880 to 2010. Rather than apply more filters to narrow down the results, we can save time and effort by using a time series table. Notice the time series tab, listing 9 available tables.



Step 1 - Filter

Scenario 2: Time Series Extract Request

The time series tab, as shown below, looks much like the source table and GIS boundary file tabs. Users can sort the available tables in descending or ascending order of any of the three columns in the Select Data grid. But what are those three columns of data telling you?

Time Series Tables:

"Table Name" indicates the name of the time series table. Such names do not always reflect actual U.S. Census table names because these were created by NHGIS combining data from many different tables.

"Years" lists those decades for which data will be provided. At this time, users do not pick and choose which of those years to include in an extract, rather receiving all years by default.

"Geographic Levels" identifies the geographies available to the user for download. Time series tables are available for a number of geographies ranging in size from the census tract level to the state level, though availability does vary by table. Users can choose any number of the available geographic levels to include in the extract request.

The screenshot shows the NHGIS data finder interface. The top navigation bar includes "Filter" > "Options" > "Review". The "Apply Filters" section has buttons for "Geographic Levels", "Years", "Topics" (with a green checkmark), "Datasets", and "Reset Filters". A "Data Cart" on the right shows 0 source tables, 0 time series tables, and 0 GIS boundary files. The "Select Data" section shows 1349 source tables, 9 time series tables (marked as NEW), and 389 GIS boundary files. Below is a table with columns for Table Name, Years, and Geographic Levels.

Table Name	Years	Geographic Levels
Persons in Occupied Housing Units by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Persons in Occupied Housing Units by Tenure [2]	1980, 1990, 2000, 2010	nation, region, division, state, county, cty_sub, tract, place
Occupied Housing Units by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units by Tenure [2]	1980, 1990, 2000, 2010	nation, region, division, state, county, cty_sub, tract, place
Occupied Housing Units by Tenure [2] by Household Size [7]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units with Householder Who Is White* by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units with Householder Who Is Black or African American* by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units by Tenure [2] by Race of Householder [6*]	1980, 1990, 2000, 2010	nation, region, division, state, county, cty_sub, tract, place
Occupied Housing Units by Tenure [2] by Race of Householder [7]	2000, 2010	nation, region, division, state, county, cty_sub, tract, place

Looking at the nine tenure tables, a few seem like possible options for us as they include a race component. However, with another look, you'll notice in the Years column that only one of the tables goes back to 1980 that would possibly include Whites and Asians & Pacific Islanders. Let's take a closer look at that particular table. Click the table name to bring up additional information.

Blue Text can (and should!) be clicked on to reveal more detailed information.

Step 1 - Filter

Scenario 2: Time Series Extract Request

Table Name	Years	Geographic Levels
Persons in Occupied Housing Units by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Persons in Occupied Housing Units by Tenure [2]	1980, 1990, 2000, 2010	nation, region, division, state, county, cty_sub, tract, plac
Occupied Housing Units by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units by Tenure [2]	1980, 1990, 2000, 2010	nation, region, division, state, county, cty_sub, tract, plac
Occupied Housing Units by Tenure [2] by Household Size [7]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units with Householder Who is White* by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units with Householder Who is Black or African American* by Tenure [2]	1970, 1980, 1990, 2000, 2010	state, county, cty_sub, place
Occupied Housing Units by Tenure [2] by Race of Householder [6*]	1980, 1990, 2000, 2010	nation, region, division, state, county, cty_sub, tract, plac
Occupied Housing Units by Tenure [2] by Race of Householder [7]	2000, 2010	nation, region, division, state, county, cty_sub, tract, plac

Table: Occupied Housing Units by Tenure [2] by Race of Householder [6*] (AE2) [More documentation on time series tables](#)

Years: 1980, 1990, 2000, 2010

Time series in table (12):

- Housing units: Owner occupied ~ Householder is White (single race)
- Housing units: Owner occupied ~ Householder is Black or African American (single race)
- Housing units: Owner occupied ~ Householder is American Indian or Alaska Native (single race)
- Housing units: Owner occupied ~ Householder is Asian and Pacific Islander (single race)
- Housing units: Owner occupied ~ Householder is Some Other Race (single race)
- Housing units: Owner occupied ~ Householder is Two or More Races
- Housing units: Renter occupied ~ Householder is White (single race)
- Housing units: Renter occupied ~ Householder is Black or African American (single race)
- Housing units: Renter occupied ~ Householder is American Indian or Alaska Native (single race)
- Housing units: Renter occupied ~ Householder is Asian and Pacific Islander (single race)
- Housing units: Renter occupied ~ Householder is Some Other Race (single race)
- Housing units: Renter occupied ~ Householder is Two or More Races

Available geographic levels (8):

- Nation
- Region
- Division
- State
- County
- County Subdivision
- Census Tract
- Place

Data sources:

Dataset	Table
1980 Census: STF 3 - Sample-Based Data	RT988. Tenure by Race of Householder
1990 Census: STF 1 - 100% Data	NR9. Tenure by Race of Householder
2000 Census: SF 1a - 100% Data [Areas Larger Than Block Groups]	NR014A. Occupied Housing Units by Tenure by Race of Householder
2010 Census: SF 1a - P & H Tables [Blocks & Larger Areas]	H14. Tenure by Race of Householder

Notes:

Measured features:

- Housing Units

Classification dimensions:

- Race

- In 1980, people not belonging to one of the nine Asian or Pacific Islander groups, and who wrote in a different Asian race in the "Other" category, were included as "Other" in the 100% count, but were reclassified as "Other Asian and Pacific Islander" in sample tabulations. The 1980 data included in this table is from a 100% count table.
- **"Single race" as used in time series compares the race categories as used in 1990 to the race "alone" categories of 2000 and 2010. Persons of two-or-more races in 2000 and 2010 are counted separately and not combined with the race "alone" categories, nor compared to any 1990 and earlier categories. Persons of two-or-more races in 1990 and earlier were forced to choose a single race to report; 2000 was the first Census in which multi-race reporting was allowed.

In the table details popup window, a more complete picture of this particular table is presented. We see that for the four decades, we'll be able to get clearly defined values for the number of single race White and Asian and Pacific Islander owner occupied housing units, as well as the number of renter occupied units. Those values will be useful when creating rates of homeownership.

In addition to simply seeing the different categories, critical information such as a listing of the original census tables used to create this unique NHGIS time series table and other important information are provided. In the upper right corner of this window, a link to even more time series information is provided. Given the complexity of creating these tables, yet the ease and simplicity of using them, the amount of time series documentation each user wishes to review will vary greatly.

Step 1 - Filter

Scenario 2: Time Series Extract Request

Comfortable with the contents of this particular time series table, let's go ahead and add it to the Data Cart. Click the plus sign next to table name to add it to the cart. The plus sign will be replaced with a green check mark, indicating its presence in the Data Cart.

The screenshot shows the NHGIS data finder interface. The 'Apply Filters' section includes 'Geographic Levels', 'Years', 'Topics' (set to 'Tenure'), and 'Datasets'. The 'Select Data' section shows 1349 source tables, 5 time series tables, and 389 GIS boundary files. A table with 5 rows is displayed, with the second row selected and a plus sign next to its name. The 'Data Cart' summary shows 0 source tables, 1 time series table, and 0 GIS boundary files.

Table Name	Years	Geographic Levels
Persons in Occupied Housing Units by Tenure [2]	1980, 1990, 2000, 2010	nation, region, division, state, county, ct_y_sub, tract, place
Occupied Housing Units by Tenure [2]	1980, 1990, 2000, 2010	nation, region, division, state, county, ct_y_sub, tract, place
Occupied Housing Units by Tenure [2] by Race of Householder [6*]	1980, 1990, 2000, 2010	nation, region, division, state, county, ct_y_sub, tract, place
Occupied Housing Units by Tenure [2] by Race of Householder [7]	2000, 2010	nation, region, division, state, county, ct_y_sub, tract, place
Occupied Housing Units by Tenure [2] by Household Size [7]	1970, 1980, 1990, 2000, 2010	state, county, ct_y_sub, place

With our time series table selected, we now need to choose the GIS files that will complement our analysis by allowing us to create some maps using a GIS. If we select the GIS boundary file tab, we see 389 files from which we can choose. That is too many to sort through, so let's add an additional filter or two to quickly narrow down the list to the pertinent files.

The screenshot shows the 'Geographic Levels' filter section. It includes a 'Selected Geographic Level Filters' section with 'State' and 'County (by State)' selected. Below this is a list of geographic levels: Nation, State, County, Census Tract, Block Group, Block, County Subdivision, and Place. Each level has a 'Geographic Levels' icon and a 'Selected' status.

Returning to the Apply Filters portion of the screen, use the **Geographic Levels** filter to choose the State and County geographic levels, as that is all we are interested in. This reduces the available GIS files to 26. At this point, you can easily find the files of interest or you can filter one more time.

The screenshot shows the NHGIS data finder interface. The 'Apply Filters' section includes 'Geographic Levels' (set to 'county on state'), 'Years', 'Topics' (set to 'Tenure'), and 'Datasets'. The 'Select Data' section shows 1228 source tables, 9 time series tables, and 26 GIS boundary files. A table with 4 rows is displayed, showing 'Year' and 'Geography' columns.

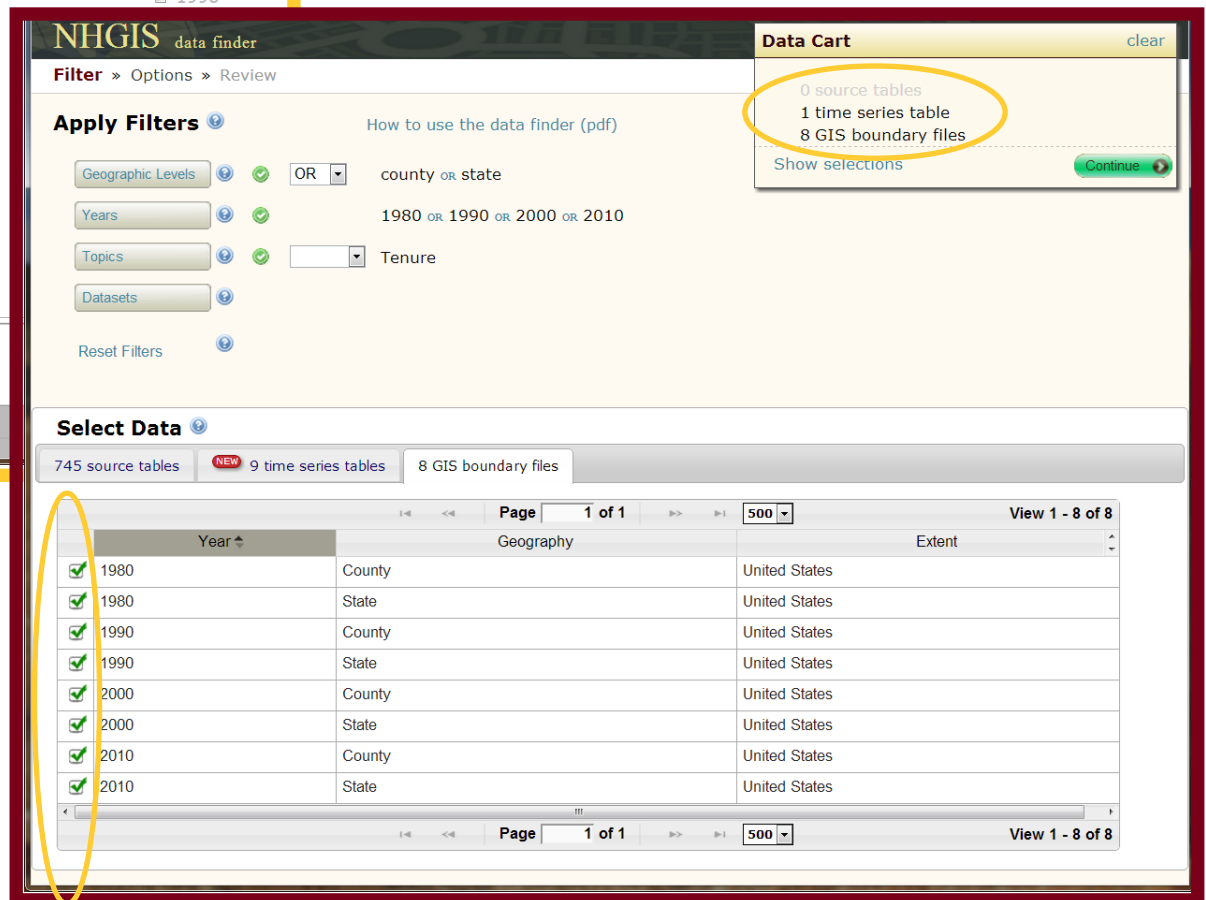
Year	Geography	Extent
1880	County	United States
1880	State	United States
1890	County	United States
1890	State	United States

Step 1 - Filter

Scenario 2: Time Series Extract Request



If you wish to reduce the number of GIS files being displayed to only those of interest, use the **Years** filter to select 1980, 1990, 2000, and 2010. This reduces the selection set to the eight desired files--a State and County shapefile for each of the four decades. Add them to your Data Cart. Once you have made your selections, you can review your cart, if you wish, by clicking on *Show selections*. When you are satisfied that you have selected the needed tables and GIS files, click **Continue** to move on to the Options page.



Please note that NHGIS does not allow for downloading data tables or GIS files for individual geographies (e.g. one particular state or county). Notice that the Extent for all of the files is the United States. This means that when we select 'State' and 'County', we will receive all the States and Counties in the country. Information on using NHGIS data within ArcGIS can be found in additional help documentation on the NHGIS website.

Step 2 - Options

Scenario 2: Time Series Extract Request

The Options page is where we actually select which geographic levels we want for our data tables. We want our data at the state level and the county level. We can select these details for both tables in the same window by clicking the 'Select geographic levels' button. For this time series table, eight different geographies are available. This entire time, we have only been talking about doing an analysis with state and county level data, but we could easily change our minds. If we decided to prepare a much more detailed census tract analysis, we can make that switch here. However, let's stick to our original plan and select 'State' and 'County'. Click **Continue** to move onto the next page.

Remember!

You must select at least one Geographic Level for each data table you have put in your Data Cart.

You cannot proceed to the final steps without doing so.

Data Cart clear

0 source tables 0 datasets
1 time series table 4 years
8 GIS boundary files 0 geographic levels

Show selections Continue

Filter » Options » Review

Data Options

One or more tables lack a geographic level selection (see below).

Time Series Tables

Select geographic levels

	Table	Selections	
		Years	Geographic levels
✓	Occupied Housing Units by Tenure [2] by Race of Householder [6*]	1980, 1990, 2000, 2010	0 of 8

GIS Boundary Files

	Year	Geographic level	Extent
✓	1980	County	United States
✓	1980	State	United States
✓	1990	County	United States
✓	1990	State	United States
✓	2000	County	United States
✓	2000	State	United States
✓	2010	County	United States
✓	2010	State	United States

Geographic Levels for Time Series Table

Filter » Occupied Housing Units by Tenure [2] by Race of Householder [6*]

Data

Select	Geographic level
<input type="checkbox"/>	Nation
<input type="checkbox"/>	Region
<input type="checkbox"/>	Division
<input checked="" type="checkbox"/>	State
<input checked="" type="checkbox"/>	County
<input type="checkbox"/>	County Subdivision
<input type="checkbox"/>	Census Tract
<input type="checkbox"/>	Place

Data Cart clear

0 source tables 0 datasets
1 time series table 4 years
8 GIS boundary files 2 of 8 geographic levels

Show selections Continue

Filter » Options » Review

Data Options

Time Series Tables

Select geographic levels

	Table	Selections	
		Years	Geographic levels
✓	Occupied Housing Units by Tenure [2] by Race of Householder [6*]	1980, 1990, 2000, 2010	2 of 8

GIS Boundary Files

	Year	Geographic level	Extent
✓	1980	County	United States
✓	1980	State	United States
✓	1990	County	United States
✓	1990	State	United States
✓	2000	County	United States
✓	2000	State	United States
✓	2010	County	United States
✓	2010	State	United States

Step 3 - Review

Scenario 2: Time Series Extract Request

The final step of the extract request is to select your data file structure and extract layout, and write a brief description to personally identify it. We will label our extract "AsianPacIsland/White Tenure - State/County - 1980/2010."

Detailed information about the data file structure options can be found under Step 3 of Scenario 1. This example, again, is assuming you aim to map the data in a GIS, so the option to choose is **Comma delimited** with the descriptive header row box left unchecked.

The Time Series Table Layout options appear when an extract contains time series data. We must choose between two Time options-vary by file or vary by row. Time varies by file will result in a separate file for each decade of data. Our 1980-2010 table will result in four separate files. Time varies by row will combine the data from each year into a single file, with one row per year-location combination. Thus, a state geographic level file would contain four records for California-a 1980, 1990, 2000, and 2010 record. Select the default, **Time varies by file**.

After clicking **Submit**, we are taken to the Extracts History page where we can download the data. If you have not yet logged into NHGIS, you will instead see the Login screen. Login, or if you have not yet created an account, simply enter your email address to proceed to the account registration page. Once finished, your extract request will continue. You will NOT need to reenter your request.

NHGIS data finder

Filter » Options » **Review**

Review and Submit

Data File Structure

- Comma delimited (best for GIS)
 - Include additional descriptive header row (best for spreadsheets)
- Fixed width (best for statistical packages)

Time Series Tables

1 time series tables

Time Series Table Layout

- Time varies by file
- Time varies by row

Source Tables

None selected.

GIS Boundary Files

8 GIS boundary files

Description

AsianPacIsland/White Tenure - State/County - 1980/2010

NHGIS data finder sign in | exit data finder

Login

New users:
To create a new account, enter your email address and submit with no password.

Users of the legacy NHGIS site:
Your account has been preserved. You will need to [reset your password](#) before using the new site.

Users of other Minnesota Population Center data sites:
The new NHGIS site shares a common user management system with IPUMS-USA, IPUMS-International, CPS, IHIS and NAPP. If you already have an account on any of those sites, please use the same email address and password to sign in here.

Email:

Password:

[Forgot your password?](#)

Final Steps

Scenario 2: Time Series Extract Request

Hint!

Refresh your browser to watch the progression of your extract request.

NHGIS data finder
Create another extract

Extracts History

The download progress of your latest extract may be seen by refreshing your internet browser.

Extract Number	Date Created	Download Data	Download GIS	Status	Revise Extract	Resubmit	Description (click to edit)
2	2012-Oct-05	data (382 KB)	gis (494 MB)	complete	revise	resubmit	AsianPaclsland/White Tenure - State/County - 1980/2010
1	2012-Oct-03	data (2 KB)	gis (26 MB)	complete	revise	resubmit	Total Farms - 1920 - all states - .csv - GIS

Documentation for NHGIS datasets is available [here](#).

Research using NHGIS data should cite it as:
Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011.

Research using data on school attendance areas (provided by the [SABINS project](#)) should cite the data as:
The College of William and Mary and the Minnesota Population Center. School Attendance Boundary Information System (SABINS): Version 1.0. Minneapolis, MN: University of Minnesota 2011.

The next step is to click the blue *data* text to download the data tables and the *gis* text for the GIS shapefiles. Both downloads consist of a zipped folder, while the GIS folder contains additional zipped folders for each shapefile within. When we click *data*, we are prompted to open the zipped folder. When we click **OK** we see a folder called *nhgis0002_csv*. The *0002* corresponds to this being our second extract ever made from NHGIS. If you have made a different number of extract requests, the number will be different. If we open this folder, 16 separate files are there. Half are the .csv files and half are the codebook files. Remember that we downloaded data for two separate geographic levels (state and county) for four years (1980, 1990, 2000 and 2010) which equals eight unique data files

File Edit View Favorites Tools Help

Add Extract Test Copy Move Delete Info

C:\Downloads\nhgis0002_shape.zip\nhgis0002_shape\

Name	Size	Packed Size
nhgis0002_shapefile_us_county_1980.zip	65 341 862	65 327 284
nhgis0002_shapefile_us_county_1990.zip	65 389 334	65 374 598
nhgis0002_shapefile_us_county_2000.zip	65 389 760	65 374 724
nhgis0002_shapefile_us_county_2010.zip	157 099 780	156 253 185
nhgis0002_shapefile_us_state_1980.zip	27 264 835	27 256 374
nhgis0002_shapefile_us_state_1990.zip	27 264 825	27 256 387
nhgis0002_shapefile_us_state_2000.zip	27 265 956	27 257 585
nhgis0002_shapefile_us_state_2010.zip	84 049 997	83 534 480

Data Tables Download

GIS Download

when choosing the Time varies by file option. Had we selected the Time varies by row, our extract would have contained four files—a data file and codebook file each for the state and county geographic levels.

File Edit View Favorites Tools Help

Add Extract Test Copy Move Delete Info

C:\Downloads\nhgis0002_csv.zip\nhgis0002_csv\

Name	Size	Packed Size
nhgis0002_ts_1980_county.csv	240 263	78 593
nhgis0002_ts_1980_county_codebook.txt	7 938	1 398
nhgis0002_ts_1980_state.csv	4 919	2 649
nhgis0002_ts_1980_state_codebook.txt	7 853	1 370
nhgis0002_ts_1990_county.csv	266 806	88 244
nhgis0002_ts_1990_county_codebook.txt	7 844	1 399
nhgis0002_ts_1990_state.csv	4 990	2 409
nhgis0002_ts_1990_state_codebook.txt	7 759	1 370
nhgis0002_ts_2000_county.csv	272 062	97 072
nhgis0002_ts_2000_county_codebook.txt	9 362	1 482
nhgis0002_ts_2000_state.csv	5 320	2 642
nhgis0002_ts_2000_state_codebook.txt	9 277	1 454
nhgis0002_ts_2010_county.csv	284 152	102 644
nhgis0002_ts_2010_county_codebook.txt	9 180	1 471
nhgis0002_ts_2010_state.csv	5 499	2 755
nhgis0002_ts_2010_state_codebook.txt	9 095	1 442

Congratulations! You have now gone through all of the steps necessary to get the data needed to compare homeownership rate changes among White and Asians/Pacific Islanders in each state and county in the country for the time period of 1980-2010. Remember to review the additional help documentation if assistance is needed in joining the data tables to the GIS shapefiles or selecting a specific location of interest from the entire United States to map separately.