Using NHGIS: An Introduction

August 2014

Overview

Selecting data for download with NHGIS requires **three** primary steps.

Apply Filters & Select Data

To browse available tables and GIS boundary files, apply one or more data filters. NHGIS provides four filter types, which can be used in any combination to hone in on desired data.

Geographic Levels - Counties, tracts, blocks...

Years - 1860, 2010, 2008-2012...

Topics - Total Population, Race, Educational Attainment....

Datasets - 2010 Summary File 1, 2012 ACS 1-year, 1870 Manufacturing Census....

When you find tables or GIS files you would like to download, add them to the Data Cart.

Select Data Options

Next, select which geographic levels, years, and breakdowns you would like to download data for. The available options will vary according to the sources and types of data in your Data Cart. You may also choose to remove tables and GIS files from the Data Cart.

Review & Submit

The final step before submitting your request is to choose the output format for any data tables you have selected. You may choose a **comma-delimited** format with or without a **descriptive header row**, or a **fixed-width** format.
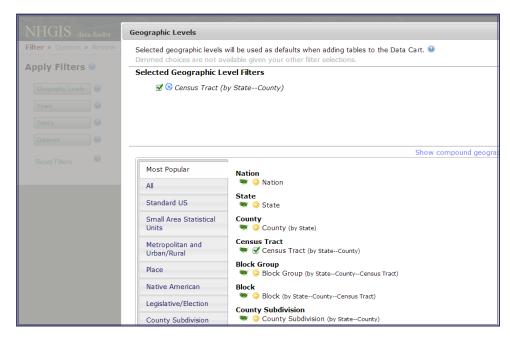
When requesting time series tables, you will also choose a file layout. The options are to arrange data from different times **by file**, **by column**, or **by row**.

When your request is ready to download, NHGIS will send you an email. Click through to your Extracts History page to download your request.
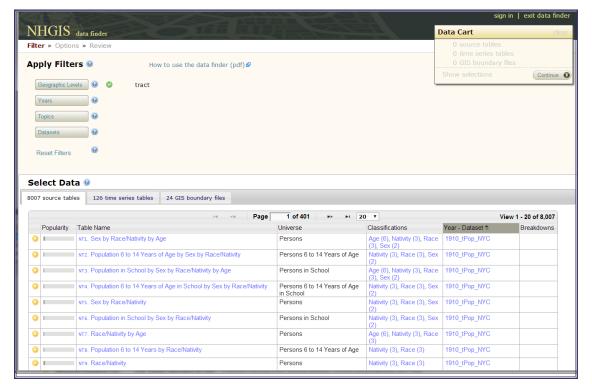
NHGIS

1 Apply Filters & Select Data

Browse for data by selecting any combination of filters

For example, selecting the Geographic Levels filter of 'Census Tract' results in the display of 8,000+ tables and GIS files that are available at the census tract level. If you select the '1970' Years filter as well, the list is reduced greatly to show only 1970 data available for census tracts.

Question mark icons link to additional information

Apply Filters & Select Data

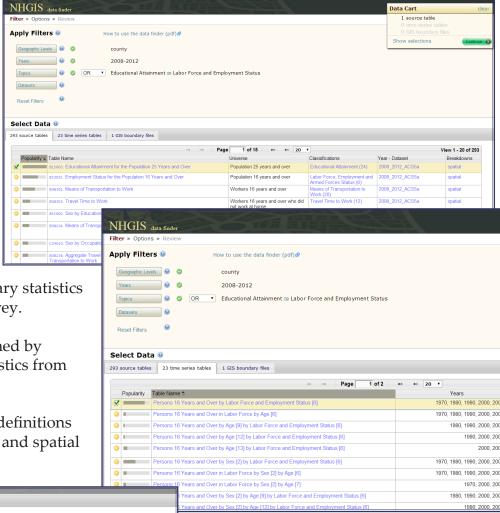Select any combination of filters you would like!

Sort by popularity, name & other features to find data of interest quickly.

Click on the tabs to find data of different types.

Source tables supply summary statistics from a source census or survey.

Time series tables are designed by NHGIS to link together statistics from multiple censuses.
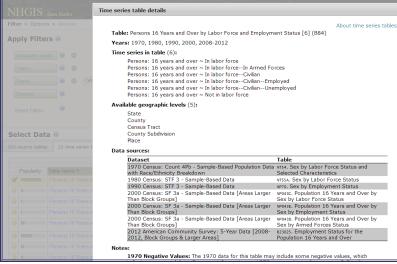
GIS boundary files provide definitions of census areas for mapping and spatial analysis applications.

Click any blue text in the Select Data grid to find additional information about available data.

Click the '+' next to any table or file to add it to your Data Cart. The Cart will reflect changes as you select and deselect data.

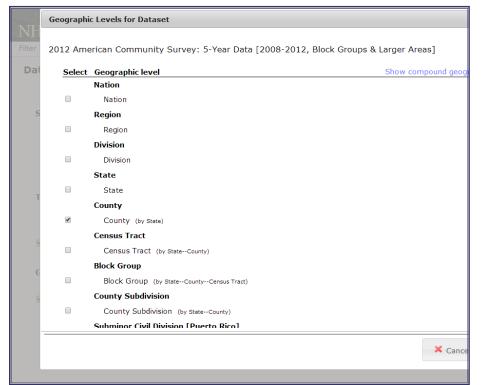When satisfied with your selections, click on **Continue**.

2 Select Data Options

Select geographic levels and confirm other options. Remove any data no longer wanted.
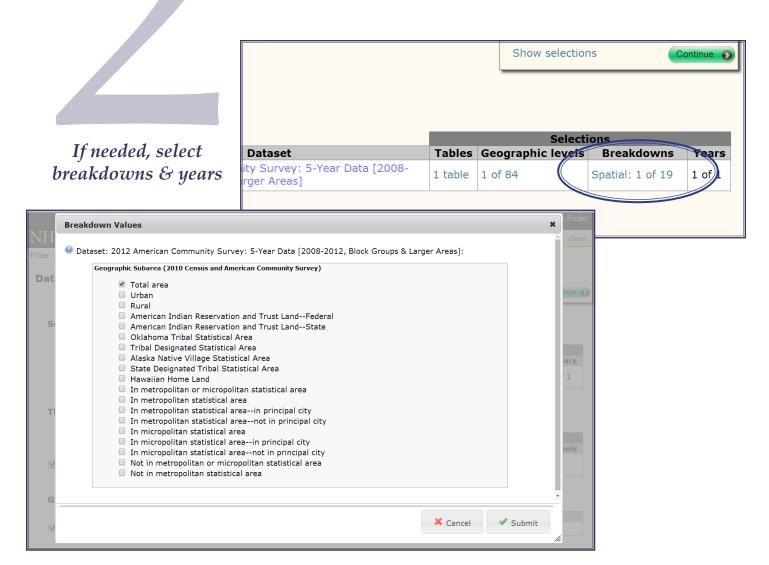

Depending on the selections made in Step 1, there may be some additional options to select or change in Step 2. If you have any tables in your Data Cart, you will need to select which geographic levels you would like the tables for. If you had any geographic level filters selected when you added the tables to your Cart, those levels will be preselected here, but you may also add to or change those selections.

Most NHGIS files cover *all* areas in the U.S., so there is no option here to select a subset of areas for download. For example, if you are interested only in Tennessee counties, then you may request a nationwide 'County' file and isolate the Tennessee counties after downloading the data.

NHGIS does, however, provide data for blocks and block groups by individual U.S. state. For these levels, you may select which states to download in Step 1 for GIS files or in Step 3 for tables .

2 Select Data Options

If needed, select breakdowns & years

| | | | Selections | |
Dataset	Tables	Geographic levels	Breakdowns	Years
...ity Survey: 5-Year Data [2008-...rger Areas]	1 table	1 of 84	Spatial: 1 of 19	1 of 1

Breakdown Values

Dataset: 2012 American Community Survey: 5-Year Data [2008-2012, Block Groups & Larger Areas]:

Geographic Subarea (2010 Census and American Community Survey)

- ☑ Total area
- ☐ Urban
- ☐ Rural
- ☐ American Indian Reservation and Trust Land--Federal
- ☐ American Indian Reservation and Trust Land--State
- ☐ Oklahoma Tribal Statistical Area
- ☐ Tribal Designated Statistical Area
- ☐ Alaska Native Village Statistical Area
- ☐ State Designated Tribal Statistical Area
- ☐ Hawaiian Home Land
- ☐ In metropolitan or micropolitan statistical area
- ☐ In metropolitan statistical area
- ☐ In metropolitan statistical area--in principal city
- ☐ In metropolitan statistical area--not in principal city
- ☐ In micropolitan statistical area
- ☐ In micropolitan statistical area--in principal city
- ☐ In micropolitan statistical area--not in principal city
- ☐ Not in metropolitan or micropolitan statistical area
- ☐ Not in metropolitan statistical area

✗ Cancel   ✓ Submit

You may also select *breakdowns* and *years* for source tables in datasets that include breakdowns or multiple years. NHGIS datasets that cover multiple years include County Business Patterns and FDIC Bank Deposit Data. By default, all years are selected. Breakdowns—described as *iterations* or *components* in Census documentation—are population subgroups for which the tables in a dataset are separately reported. For example, a table in a dataset with Spatial and Race/Ethnicity breakdowns is available not only for the total population (the default breakdown) but also for subgroups distinguished by geographic (e.g., urban or rural) or racial/ethnic (e.g., White or Hispanic) characteristics. In many datasets, breakdown data are reported only for areas where the identified subgroup surpasses a minimum population.

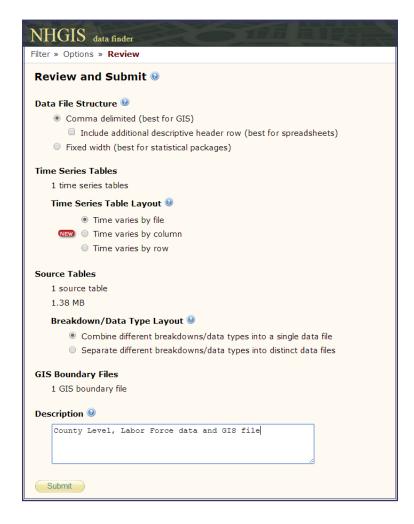When satisfied with your selections, click the **Continue** button.

3 Review & Submit

Select the output data format and layout

Data File Structure

You may download NHGIS tables in *comma-delimited* or *fixed-width* formats. If you plan to use the table in a spreadsheet or GIS application, the comma-delimited option is best. The option to include a descriptive header row is helpful for interpreting data in a spreadsheet but is not supported in standard GIS software. The fixed-width option is best for users of statistical software packages as it includes command files for SPSS, SAS and Stata.

Time Series Table Layout

Time series tables include data for multiple areas over multiple times. How the unique area|time combinations are arranged is up to you.

Time varies by file separates data from each time into a separate file. *E.g.,* data for Chicago 1970 and Chicago 1980 are in separate files.

Time varies by column provides a single file with one record for each area and separate columns for each time. *E.g.,* data for Chicago are in one row with both 1970 and 1980 columns.

Time varies by row provides a single file with separate records for each area|time combination. *E.g.,* one row contains Chicago 1970 data and another row contains Chicago 1980 data.

Note that NHGIS time series provide data for areas as they were defined *at each time* without adjustments for name changes or redrawn boundaries. *E.g.,* data for "Tract 3" in a particular county may describe different areas in 1970 and 1980.
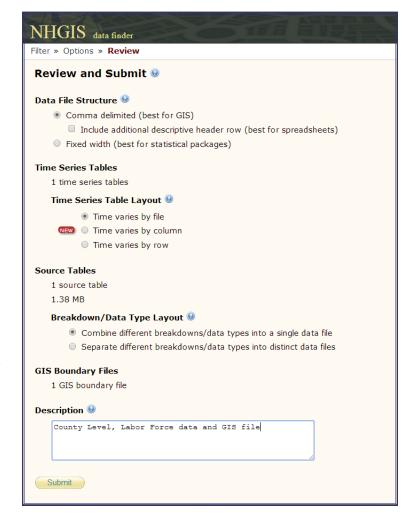
3 Review & Submit

Source Tables: Geographic Extents

If you have selected block or block group geographic levels for any tables, a 'Geographic extents' link will appear on this page, through which you must select which state(s) the block and block group tables will cover. You may select all states at once, but a nationwide extract for these levels (especially for blocks) may take several hours to complete, or it may time out. For a faster result, select only one or a few states at a time.

Breakdown / Data Type Layout

If you selected multiple breakdowns for any source tables in Step 2, or if your request includes any tables from the American Community Survey (ACS), you will be able to choose whether to combine all of the breakdowns (*e.g.*, urban and rural) and data types (the estimates and margins of error in ACS datasets) into a single data file or to separate them into distinct data files. The default is to combine them all in one file.

Description

You may optionally enter a brief description of what it is you are requesting. This description will appear in the Extracts History for your account and can be helpful if you ever return to revise or resubmit your request. The description is purely for your benefit, so use whatever format you would like. The example description mentions the geographic level and the basic topic, and indicates there are multiple data types (tables and a GIS file).

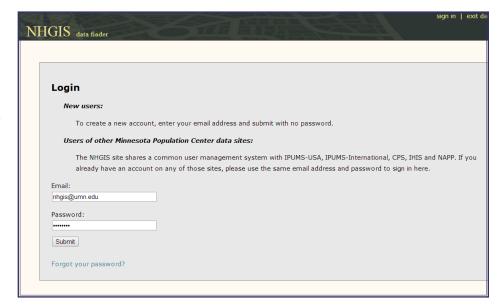Finally, when you are satisfied with your choices, **Submit** your extract.

Download

After submitting the extract, you will be prompted to login to NHGIS if you had not already done so. If you do not yet have an account, simply enter your email address and **Submit**. You will be taken to another page where you will create a password and provide some basic information. All NHGIS accounts are absolutely free.

Once you've registered and logged in, your extract will begin processing and your browser will open your Extracts History page. This page lists every extract you have requested. When the Extracts History page first opens, the extract status for your current request will be 'queued.' To see if the status has changed, use your browser controls to reload the page. Each extract will eventually be 'in progress' and finally, when it is ready to download, 'complete.' You can also leave the page, and NHGIS will email you when your extract is ready.

The record for a completed extract will include links for 'tables' and/or 'gis' data. Click the link(s) to begin downloading your data.
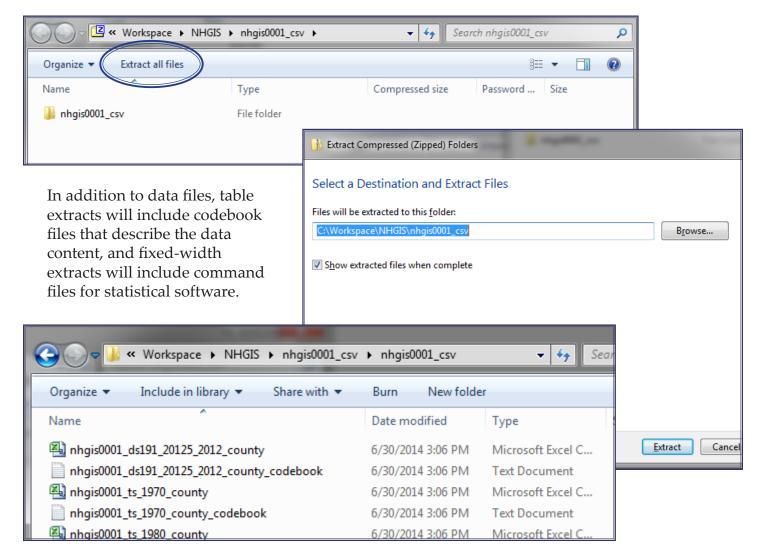
Unzip & Open

NHGIS delivers data extracts in zip files. To 'unzip' the extracts, you may use the default zip tool within Windows Explorer or another third-party program. When unzipped, your extract may consist of any number of files, depending on the data types, datasets, geographic levels, and layout options you selected.

In addition to data files, table extracts will include codebook files that describe the data content, and fixed-width extracts will include command files for statistical software.

Because the GIS files are in a shapefile format, and each single 'shapefile' is comprised of multiple data files, the GIS files are double zipped, with each shapefile individually zipped within the main extract zip file. Make sure to unzip all files before using them in a GIS.

You can now open your data and begin your analysis!

Please email **nhgis@umn.edu** with any questions you may have.