



American Planning Association *Making Great Communities Happen*

Planning — October 2012

Demographic Data for a Changing Nation

NHGIS provides current and historical U.S. census tables and GIS data.

By Jason Borah and Jonathan Schroeder

Take a peek inside any planning office, and chances are good that you will find planners working with census data. Population and housing counts, income, educational attainment, racial composition, and many other census statistics can play a part in creating the right plan for the right place. Despite this ubiquity, finding appropriate census data and processing it to fulfill project needs is often a struggle that eats into a planner's time and budget. So, where do *you* turn for census data?

In many cases, going straight to the source really is the best choice. The U.S. Census Bureau delivers most of its latest data online through the American FactFinder, a robust data access system that offers several search and filtering options to help users find the data they need. This site, however, has several limitations in data format and scope that make it unhelpful for some common applications.

Want data for any census before 2000? Or the latest American Community Survey data on median incomes among block groups? Or tract-level data for multiple states in a GIS-ready format? In these cases and others like them, the data are either unavailable on American FactFinder or, as in the last case, difficult to piece together.

Several other data providers have been working to fill the gaps in American FactFinder's offerings. One is the National Historical Geographic Information System, a website that provides U.S. census data from 1790 to the present, aiming to be as comprehensive as possible. It is a "historical GIS" site because it also provides shapefiles for geographic areas as they were defined in each census, enabling NHGIS users to map and compare distributions of population and housing characteristics across time.

The screenshot shows the NHGIS data finder interface. At the top, it says "NHGIS data finder" and "Filter > Options > Review". There are three main sections: "Apply Filters", "Data Cart", and "Select Data".

Apply Filters: Includes buttons for "Geographic Levels" (checked), "Years", "Topics" (checked), and "Datasets". A search box contains "tract" and a dropdown menu is set to "Tenure". There is also a "Reset Filters" button.

Data Cart: Shows "0 source tables", "1 time-series table", and "3 GIS boundary files". It has "Show selections" and "Continue" buttons.

Select Data: Shows a summary of "643 source tables", "7 time series tables", and "9 GIS boundary files". Below this 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, tract |
| Occupied Housing Units by Tenure [2] | 1970, 1980, 1990, 2000, 2010 | state, county, cty_sub, tract, place |

Why use NHGIS?

The NHGIS project began in 2001 at the Minnesota Population Center, a unit of the University of Minnesota, with funding from the National Science Foundation. The project had an ambitious goal to collect, format, and freely disseminate all available aggregate census information for the entire country from the past 220 years. The completed site (www.NHGIS.org) launched in 2005 to a primarily academic audience. Now urban planners are also discovering NHGIS and its ever-expanding cache of 15,000 tables and 430 shapefiles when they are seeking data from the past or present.

NHGIS data are derived from the U.S. Census, but much of the information comes from historical printed reports that were hand entered by several prior projects. You can also get many of these reports online from the Census in pdf format. While planners may find these reports useful, they are unfortunately of little use in GIS or spreadsheets.

In addition, since 1960, the Census has released some form of "machine-readable" data, meaning that it was tabulated using computers and stored on magnetic tape or in other digital formats. But that doesn't mean it's easy to access or use. Would your laptop read the raw 1960 data? Not a chance, unless your laptop has a vintage UNIVAC II-A tape drive installed. For your sake, we hope it doesn't. Changing technologies have made accessing raw census data from the early digital era nearly impossible, forcing a reliance on other sources for the converted data.

Not only does NHGIS provide all of this tabular data in one place in a standardized format, it also provides U.S. Census Bureau TIGER/Line-derived shapefiles, which are also much in demand by planners. If you just need a quick map, there are numerous sites that offer basic census web mapping. Often, however, planners need data they can use inside their own GIS to prepare publication-quality maps. For this reason, many will find the robust collection of historical shapefiles the most valuable product within NHGIS. Whether you need 1840 county boundaries, 1980 places, or 2010 blocks, NHGIS has you covered.

All shapefiles come "clean": data projected, coastlines clipped, metadata provided, and geographic identifiers standardized for table joins. In short, NHGIS did the dirty work so you don't have to.

NHGIS is not the only alternative census data source. Planners can also use Social Explorer, Esri Community Analyst, Geolytics, the US2010 project, and others. Each site has its own benefits in addition to making at least some historical data available. NHGIS, however, is the only site to provide data tables and shapefiles for all censuses from 1790 to 2010 while being free of charge to all users — public, private, academic.

Data in harmony

Easy access to historical census data in a GIS-ready format is a big step forward, but for many change-mapping projects, there remain two hurdles to clear: linking variables and linking geographies across censuses. After all, it's not just the people that change from one census to another; the sets of measured characteristics and areas keep changing, too. How do you find out which statistics you can map over time? And how do you measure population changes in areas where the boundaries have changed, too?

That's where NHGIS's newest product — time series data — comes in. With a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the Minnesota Population Center has begun harmonizing NHGIS data across time and packaging linked data in time series tables. Because the site already delivers more than 300,000 variables, each potentially available for any number of geographic levels, complete integration of NHGIS data across all census years is a slightly daunting task. Therefore time series data will be rolled out in phases over several years, iteratively providing a broader coverage of topics, years, and geographies.

The first release provides linked data for 1970 through 2010. Because the Census tabulated and delivered data digitally throughout this period, the data from these censuses cover a broader range of topics and areas than earlier data. Still, the set of measured characteristics and categories has continued to change. Many cross-tabulations, for example Persons by Race by Sex by Age, are provided in some years and not in others, and even when cross-tabulation types recur, the categories used may vary in number, names, or definitions, making it difficult to determine what's really comparable.

With NHGIS time series, however, even a novice user can quickly identify and download any of the comparable statistics for a given period of time. For example, for the 1970–2010 period, NHGIS provides separate tables for the three possible subdivisions of Hispanic or Latino population: one table of 25 time series spanning 1990–2010, one table of four time series for 1980–2010, and one table with just one time series spanning 1970–2010. NHGIS also assigns a single name to each time series (no confusion of "Spanish Origin" in one year versus "Hispanic" in another), sums subgroup counts as needed to align with counts of larger groups from other years, and documents all known changes in definitions.

Time series tables are now available for eight geographic levels, including tracts, places, county subdivisions, and counties, but the initial version has only tackled one of the two major integration challenges — the tables link variables across time, but they do not yet standardize geographic units. Thus, a tract-level time series provides data for tracts as they were defined in each census — 1970 data for 1970 tracts, 1980 data for 1980 tracts, and so on. This simplifies the task of mapping each year's tracts separately, but it remains difficult to compute changes between censuses because it is impossible to match all of one census's tracts directly with another's.

NHGIS is working to address this problem in two ways. First, as two other data providers — Geolytics and the US2010 project — have also done, NHGIS will provide interpolated tract data: estimates from multiple censuses for one year's tract extents (1990 and 2000 estimates for 2010 tracts). Second, both tracts and counties will be aggregated into "least common denominator" superunits that have fixed boundaries across multiple censuses. This aggregation strategy reduces spatial resolution but maintains accuracy, unlike interpolated estimates.

As the NHGIS time series rollout continues, planners and other GIS users will find they no longer need to battle with changing census codes and boundaries when preparing a temporal analysis.

Chicago Home Ownership Rates

Among Census Tracts within 2010 city limits

PERCENT OF OCCUPIED HOUSING UNITS THAT ARE OWNER OCCUPIED

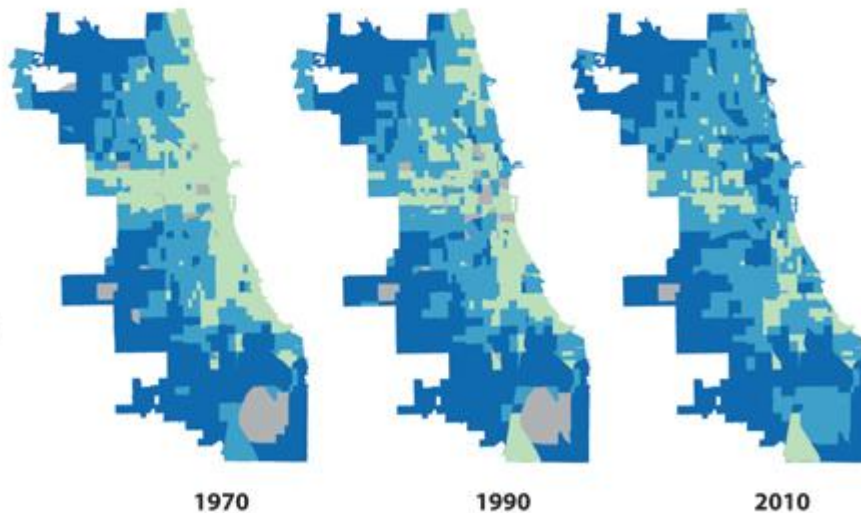
UNDER 25%

25% to 50%

OVER 50%

Fewer than 50 occupied housing units

Home ownership rates increased sharply in many areas, especially in neighborhoods immediately north and west of downtown. Almost no census tracts showed declines.



Source: Data www.NHGIS.org. O'Hare Airport extension excluded.

How to use it

Using the NHGIS website is simple. You can start on the NHGIS home page and click the "Select Data" link to enter the Data Finder. Next, click on any of the four filter types in the "Apply Filters" portion of the page: Geographic Levels, Years, Topics, or Datasets. For each type, a pop-up window will appear that presents the complete set of options for NHGIS data. Users can select any combination of specific filter values — for example, one topic, two geographic levels, and several years — and the Data Finder will list all tables and GIS boundary files that correspond to the selection.

Among the four filter types, the Geographic Levels type offers the largest array of options, with 343 different levels. While the pop-up initially lists only the most popular geographic levels, like county and census tract, clicking on different tabs reveals many more choices, like tribal subdivisions and state legislative districts, or compound levels like places by county (which divides multicounty cities like New York into separate county parts).

With the Year filter, you can quickly dial in on the time frame of interest. While the obvious 1790–2010 decennial choices are present, many other year options are available. These include the 2006–2010 ACS, the 2009–2010 School Attendance Areas from the School Attendance Boundary Information System, and several others.

The Topics filter is a great starting point for exploring what types of characteristics are covered in census data and for quickly narrowing down the range of tables presented. If you've determined that you need public assistance income data, for example, utilizing that topic filter will locate the applicable tables faster than any other options.

The Dataset filter is useful for those who want to see the full range of NHGIS source datasets and identify what types of data are provided by each. It also allows users to limit their selections to short-form, 100 percent-count sources (like 2010 Summary File 1) or to long-form, sample-based sources (like the ACS or 2000 Summary File 3).

Once you've selected whichever filter options are of interest, you can then browse through all of the corresponding tables and shapefiles in the selection grid that appears at the bottom of the page. Clicking on any hyperlinked labels brings up details pages with additional information. When you've identified a desired table or file, simply add it to your Data Cart. You may then continue browsing and adjusting filters to select as many tables and shapefiles as you like.

After making selections, proceed to the Options page, where you may finalize the geographic levels and "breakdowns" for which you would like to extract data. NHGIS breakdowns correspond to the "geographic components" and "characteristic iterations" available in recent census datasets. They generally provide detailed spatial or race-ethnicity group subtotals. For example, using a spatial breakdown on a 1970 "Aggregate Family Income" table would result in a separate total value for families in "Rural — farm" areas, "In MSA — not in MSA central city," or in other defined areas.

Finally, on the "Review and Submit" page you may select an output file structure for your data and, when applicable, the geographic extent. The data output options include three formats, one each for GIS, spreadsheet, and statistical software, which means you don't have to spend time

getting it into a useable format. With geographic extent selection, users can choose which state(s) to limit the data tables to. Currently, this feature is available only for the block and block group geographic levels, and data files for other levels are always nationwide.

Once you submit your data request, you will be directed to your Extracts History page, where you will be able to download the zipped files containing your data. Every extract you make is saved indefinitely, making it easy to reacquire data that has been misplaced or unintentionally altered.

Time saver

To plan effectively for long-term changes in regional and local populations requires first that we have effective tools and data for understanding how we have already changed. Until recently, the data needed for assessing demographic change have often been impossible to find, difficult to use, or too expensive to obtain.

With NHGIS and others making ever increasing amounts of demographic data available through integrated systems that let users assemble multiple years of data quickly, you will now be able to spend more of your time analyzing data and putting it to use. Spending valuable time aimlessly searching for needed census or GIS data shouldn't be your job.

Jason Borah is a GIS analyst at the Minnesota Population Center. Jonathan Schroeder is a research associate there.

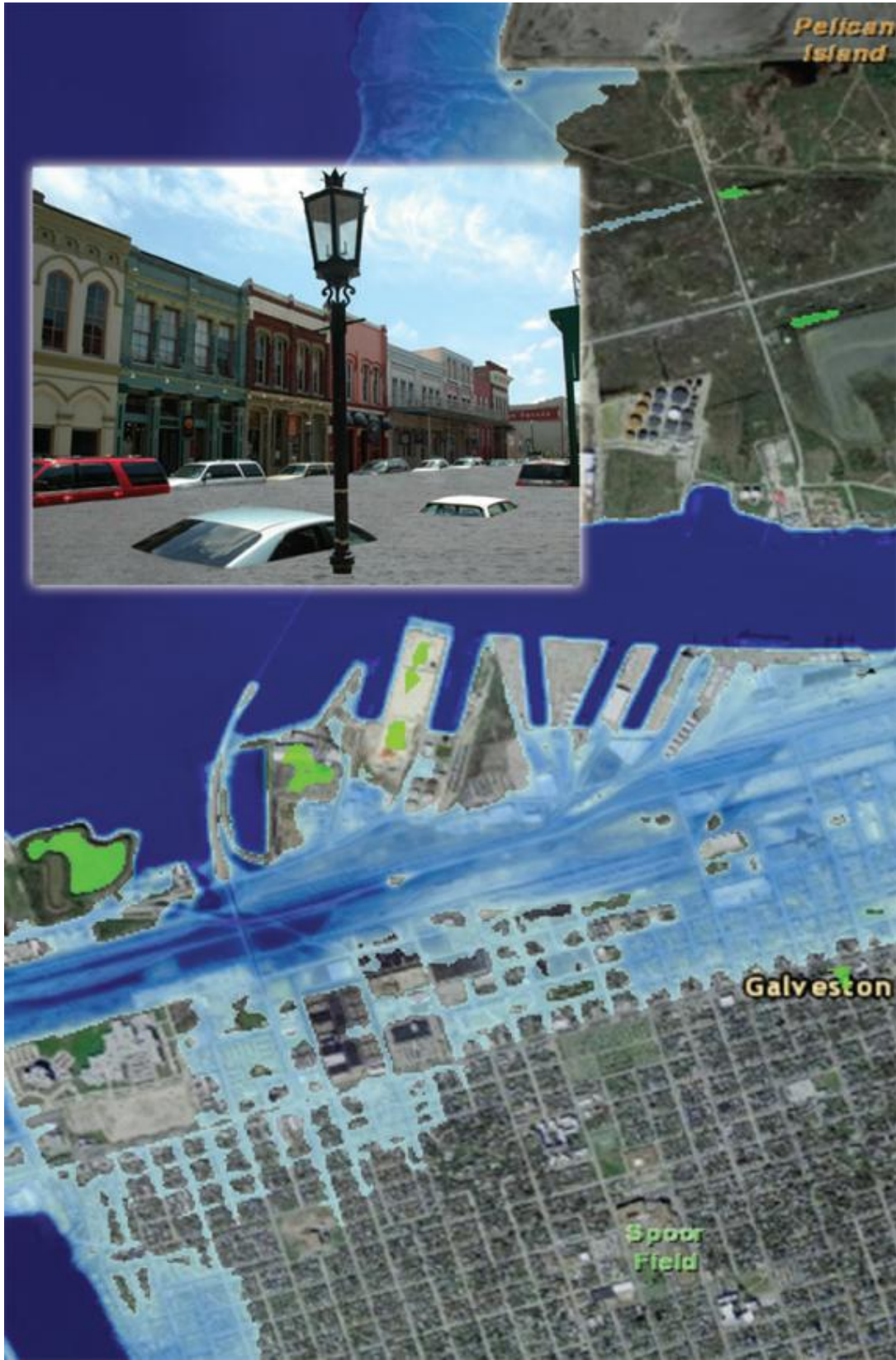
Coastal Data, Visualized

By Kitty Fahey

Planners need only look at communities along the U.S. oceans and Great Lakes to see the changing state of the nation writ large. According to the U.S. Census Bureau, the population living in coastal watershed counties grew by 50.9 million between 1970 and 2010. Other coastal sectors that have grown apace include development and infrastructure, the coastal economy, and — for a variety of reasons — hazards such as intensified storms, floods, and sea-level rise.

Coastal community planners have their hands full trying to understand and weigh these factors in their decisions, let alone clearly explaining the data that help justify decisions. In such cases a picture truly is worth 1,000 words. Planners using the Digital Coast have at their disposal an online suite of data, tools, and training that makes data and information easy to picture and use when considering potential plans and decisions.

"We develop our products to be accessible and visual, because we want Digital Coast to be used by all of the coastal players, not just those of us who are technologists," says Nicholas Schmidt, acting deputy director of the National Oceanic and Atmospheric Administration's Coastal Services Center. The center leads the Digital Coast Partnership effort.



The American Planning Association is one of seven national partners behind the Digital Coast. Many training sessions and webinars have been approved for AICP certification maintenance credits. Resources include the following:

CanVis. With this tool, existing photos provide the backdrop and more than 700 objects from the CanVis library can be added, including water, docks, marsh grass, and buildings. For instance, users have illustrated the likely visual impacts of coastal flooding and shown how offshore wind farms could affect shoreline vistas. Web-based CanVis training is available.

Coastal County Snapshots. This tool provides users with an instant summary of their county's natural resources, coastal economy, and vulnerability to hazards. Downloadable fact sheets and graphics show the number of critical facilities and percentage of vulnerable population in the flood zone, the percentage of ocean-related jobs and businesses, and practical steps to help communities become more resilient.

Data Access Viewer. Geospatial data users can view and download a wide variety of coastal data sets, simply by typing the location name or "drawing" a box around the area on a map. For example, a data search for "San Francisco Bay" brought up 18 different data sets detailing land cover, imagery, and elevation data.

Economics: National Ocean Watch. Seeing the facts about a coastal community's economic activity and characteristics can help officials protect local businesses while addressing natural hazards and the impacts of climate change. ENOW spotlights six economic sectors that depend on the oceans and the Great Lakes. The website features county, state, and national data about jobs, wages, and the U.S. gross domestic product.

Land Cover Atlas. Atlas users get instant, county-specific maps and charts that enable them to understand the effects of past land-use decisions, document trends, and consider future land-use planning needs. The data cover development and impervious surfaces, forests and fragmentation, and wetlands.

Sea-Level Rise and Coastal Flooding Impacts Viewer. An easy slider bar and "zoom" feature enable users to view local models of sea-level rise plus simulations illustrating the possible effects of various sea levels on community landmarks. Users can overlay social and economic data on mapped areas illustrating different sea levels. They also learn tips about flood frequency and communicating sea-level rise uncertainties.

The Digital Coast continually adds to product offerings. "We know that dealing with the 'people' part of coastal issues is a big aspect of planners' jobs, so we will be adding more social science data and tools over time," says Schmidt. "Giving coastal communities as complete a picture as possible is what we aim for."

Visit APA's Digital Coast Survey pages at www.planning.org/research/digitalcoast.

Kitty Fahey is a writer with I.M. Systems Group at the NOAA Coastal Services Center. The center supports the environmental, social, and economic well-being of the coast by linking

people, information, and technology. More information on the Digital Coast resources may be found at www.csc.noaa.

Resources

Image: How might a segment of Galveston, Texas, be affected by six feet of sea-level rise? The mapped model and visualization shown here paint a picture. Image courtesy of the Sea Level Rise and Coastal Flooding Impacts Viewer.

Primary Data Holdings in NHGIS

| Tabular Data | Source Years Available |
|---|--|
| Selected tables from printed decennial census reports | 1790–1960* |
| Decennial Census of Population and Housing | 1970–2010 |
| American Community Survey (ACS) | 2010 1-Year Summary Data 2008–2010 3-Year Summary Data 2006–2010 5-Year Summary Data |
| County Business Patterns | 1970–2002 |
| GIS Shapefiles | Source Years Available |
| State | 1790–2010 |
| County | 1790–2010 |
| County Subdivision | 1980–2010 |
| Tract | 1910–2010 |
| Block Group | 1990–2010 |
| Block | 1990–2010 |
| Places | 1980–2010 |
| Urbanized Areas | 1990–2010 |
| Metropolitan Statistical Areas | 1950–2010 |
| New England County Metropolitan Areas | 1990–2000 |
| New England City and Town Areas | 2010 |
| Congressional Districts | 1990–2010 |
| State Legislative Districts | 2010 |
| Voting Districts | 1990–2010 |
| School Districts | 2010 |
| School Attendance Boundaries | 2009–2010 Academic Year |
| ZIP Code Tabulation Areas | 2000–2010 |

* Data from the Interuniversity Consortium for Political and Social Research (ICPSR) & Andrew Beveridge

Resources

Image: Using NHGIS: an example — let's say you'd like to map home ownership rates in Chicago since 1970.

©Copyright 2012 American Planning Association All Rights Reserved