Video 4: IPUMS-National Historical Geographic Information System (NHGIS)

Jylisa Doney
University of Idaho, Social Sciences Librarian
jylisadoney@uidaho.edu

Transcript

00:00 Hello again, welcome to Video 4.
00:02 In this video, we're going to practice using the
00:04 IPUMS-National Historical Geographic
00:07 Information System repository.
00:09 This repository is unique in that it provides access
00:12 to spatially aggregated Census data,
00:14 which is data that summarizes individuals
00:16 within particular areas.
00:18 To protect anonymity, IPUMS-NHGIS
00:21 only provides access to raw aggregate data.
00:24 No individual-level records
00:26 are included within this repository.
00:28 Let's go ahead and navigate to the IPUMS main page.

[IPUMS: https://www.ipums.org]

00:31 All of the IPUMS data repositories are free to search,
00:35 but unlike data.census.gov, you must create an account
00:38 within each specific repository to download data.
00:42 To create an account with IPUMS-NHGIS,
00:45 first navigate to that repository,

[IPUMS-NHGIS: https://www.nhgis.org]

00:48 click-on "Log In,"
00:50 then click on "Create an Account" to register.
00:53 Remember, this registration will only allow us
00:55 to download data within IPUMS-NHGIS.
00:59 If you want to download data
01:00 from other IPUMS repositories,
01:02 you will need to visit their pages,
01:04 and log-in to your IPUMS account.
01:06 Your previous account information
01:08 will auto-populate the form
01:09 for that specific repository
01:11 and you will just need to agree
01:12 to the terms and conditions.
01:14 Let's navigate back
01:15 to the IPUMS-NHGIS homepage and click on
01:18 "Get Data" to be directed to a guided search.
01:21 As with data.census.gov, we see various filters.
01:24 If you knew the "Dataset" set you were interested in,
01:26 you could initially select that filter

01:28 and only see those data outputs.
01:30 But let's focus on browsing across data sets.
01:34 First let's choose a "Geographic Level."
01:37 Click the "plus sign" next to the geographic level
01:39 to add it to your search query.
01:41 You'll notice that we aren't able to select
01:44 specific subsets of geographic levels.
01:47 That's because most NHGIS files
01:50 cover all areas of the U.S.
01:52 and they do not create separate data files
01:54 for each of those subsets.
01:56 But once we download the data,
01:58 we can easily edit it so that it only displays
02:00 the information we are interested in,
02:02 such as county level data for Idaho.
02:04 After clicking "Submit,"
02:06 we can see that there are over 33,000 source tables,
02:10 379 time series tables,
02:12 and 56 GIS, or geographic information system files
02:17 that meet this geographic filter.
02:19 Source tables provide summary statistics,
02:22 time series tables link statistics
02:24 from multiple Censuses,
02:26 and GIS boundary files
02:27 provide definitions of Census areas
02:29 for mapping and spatial analysis applications.
02:33 Next we can select the data "Year"
02:34 that we're interested in.
02:36 Options in gray are unavailable
02:38 due to previous filter selections.
02:40 Click the years you're interested in then click "Submit."
02:43 Again we can see that the number of
02:45 available source tables and GIS files
02:48 changed based on this filter,
02:50 and the time series tables are no longer available
02:52 since we didn't select a decennial year.
02:55 Also, next to years, we see a drop-down box with "OR."
02:59 This means that we are interested in data outputs
03:02 from either 2012 to 2016 "OR" 2013 to 2017.
03:08 If we change this box to "AND,"
03:10 no data outputs will be available
03:12 as a single data file does not include data
03:15 from both of these date ranges.
03:17 So let's change the drop-down box back to "OR."
03:21 After selecting a geographic level and year
03:23 we can move on to "Topics."
03:25 We have two options within this filter.
03:28 The "Table Topic Filter" pertains to the
03:30 types of variables in each table.
03:33 For example, if we selected age,
03:35 IPUMS would provide tables such as median age.

03:38 In comparison, the "Breakdown Filter"
03:41 would show us tables that are broken down
03:43 based on certain segments of the population.
03:45 For example, if we selected
03:47 the breakdown filter for race,
03:49 IPUMS would show us data sets such as age by sex,
03:52 with a race/ethnicity breakdown option.
03:55 Let's scroll down to and check the "plus sign"
03:57 in the "Table Topic Filter" area
03:59 for "educational attainment,"
04:02 and "poverty,"
04:03 and then click "Submit."
04:05 Now we can see the source tables and GIS files
04:09 that meet our filter specifications.
04:11 If you want to remove any of the filters you selected,
04:14 hover over the "green check mark" and click.
04:16 You can also click "Reset Filters"
04:18 to start a new guided search.
04:20 To download data click the "plus sign"
04:22 next to the tables you are interested in
04:24 to add them to your "Data Cart."
04:26 Click on the different tabs, such as "GIS Files,"
04:29 to add those to your cart as well.
04:31 Now, within the "Data Cart" box,
04:33 let's click on "Continue" to move to the next screen.
04:36 First, we have an opportunity to confirm our selections
04:39 and either edit them or click continue.
04:42 Next we can select our download options.
04:44 I'm going to use the default "Table File Structure"
04:47 and "Breakdown/Data Type Layout" options,
04:50 but you could change these based on your data needs.
04:52 Finally, click "Submit" to generate your data extract.
04:56 If you aren't logged-in when you click "Submit"
04:59 you'll be prompted to log-in at that time.
05:01 After logging-in, you're directed
05:03 to the "Extracts History" page,
05:05 which also includes a suggested citation
05:07 for the NHGIS data you downloaded.
05:10 IPUMS will send you an email
05:12 when your data extract is ready to download,
05:14 which could take a few minutes or longer
05:15 depending on the file size.
05:17 Once your data extract is ready,
05:19 either navigate back to the "Extracts History" page
05:22 or follow the link and instructions in the email.
05:25 When we download the "Tables" file,
05:27 we can see that it includes the raw aggregate data.
05:30 If we were to scroll down within this spreadsheet,
05:33 we could identify counties in Idaho
05:35 and only analyze data from those areas.
05:38 The "Tables" file also includes the codebook,

05:41 which is like the data dictionary
05:43 we downloaded in Video 3.
05:45 From the Extracts History" page,
05:47 we can also download the GIS file.
05:50 If you're interested in learning more
05:52 about this repository,
05:53 IPUMS-NHGIS provides additional help
05:56 via their "User's Guide."
[IPUMS-NHGIS – User's Guide: https://www.nhgis.org/user-resources/users-guide]

05:58 IPUMS also provides a variety of tutorials
06:01 for their different programs.
[IPUMS – Tutorials: https://www.ipums.org/tutorials.shtml]

06:03 If you have any questions about using GIS data
06:06 as well as accessing Esri software,
06:08 ArcGIS, or Story Maps,
06:10 please reach out to Bruce Godfrey, our GIS Librarian.

[Bruce Godfrey (GIS Librarian), bgodfrey@uidaho.edu]

06:14 Coming up in Video 5, we will discuss
06:17 best practices for evaluating and citing data.