Basic Census Cartography with QGIS

Prof. Hachadoorian

Summer II 2016

Overview

- Become familiar with the QGIS interface.
- Add data from the PostGIS server to QGIS.
- Create a basic thematic map in QGIS.

This tutorial assumes that you are already familiar with working with GIS software, but may not have used QGIS specifically. You should be familiar with concepts such as attribute joins, data normalization, and projections and coordinate reference systems.

Downloading Data from NHGIS

Go to http://nhgis.org. On the welcome page will be a link to a short YouTube video showing how to select a "data extract", which is a set of linked tables and shapefiles. Watch the video. Then click the Select Data link.

On the Apply Filters page, begin by setting Geographic Levels to State, and then set Years to a recent year of interest. Note that if you select a year from the Decennial Years column, you will get Decennial Census data, whereas a year or year range from the other columns will get American Community Survey data. The list of available files at the bottom of the page (Select Data) will update as you apply these filters.

For this tutorial I am going use data on race. From Census 2010, this is table B02001. You may use this year and topic to follow along, or you may pick another topic. Pick something that adds represents a break down of total population or housing. For example, you could choose sex, employment, housing occupancy, etc. Do not *only* pick a statistical aggregate, such as median age or median income.

You can apply a topical filter, or just select the data table of interest in the Select Data area at the bottom of the page. Make sure to select at least one

Source Table (these are the tables which hold the demographic data), and on GIS Boundary File (there should be only one if you apply a single Year and Geographic Level filter).

Once the files are selected, you should see the count of selected files in the Data Cart on the upper right. Click continue. Create an account for the NHGIS website (or log in if you already have one). Depending on the download size, creating the extract may take a while. Since we are using a gross geography (states) and only selecting one or a small number of source tables, the extract should run pretty quickly. Refresh the page to see the extracts, and click on "tables" and "gis" to download the files. They will be in ZIP archives. I suggest you unzip them and place the unzipped GIS and attribute data in the same folder.

Adding and Joining Data in QGIS

The attribute data that we downloaded from NHGIS is in a CSV (comma-separated values) text file. Because the file uses text, QGIS needs a hint in order to correctly recognize numeric data. We do this by creating a CSVT (for "CSV Template") file which contains a list of data types. You may use the data types String, Real (meaning decimal data), Integer, Time, and DateTime. [1]

Open the CSV file that you downloaded with a text editor such as Notepad++. [2] **Before you make any changes to the file, immediately Save As to create a new file with the *exact* same name, but the extension csvt.** Now delete the header row and all of the lines except for **one** data row. Replace all text, such as "codes" (like the state FIPS code) and names, as well as blanks (commas separated by nothing), with the word `String`. At the end of the line should be all of the actual demographic data. For count data, replace all the numbers with the word `Integer`. If you have also downloaded data such as median age, replace the number with the word `Real`. Save and close the file.

Open QGIS Desktop (*not* QGIS Browser) by double-clicking the icon on your desktop, or find it in the Start Menu in the OSGeo4W folder.

The central pane is where you will see the geographic data that you add. The left and right panes are dockable panels that can be torn off (grab the title bar of the panel) if you prefer floating panels in your workspace. We won't be using the Processing Toolbox today, so if it is open (by default in a right-hand panel) you can close it in order to increase the area avaialable for the map canvas. The left panel is the Layers pane. This will be an index of layers you have added.

---

[1]See http://www.gdal.org/drv_csv.html for additional documentation. Note that more recent versions of the OGR library that QGIS uses for data import can determine data type automatically, making this step unnecessary.

[2]Do **NOT** use Excel, which is what will probably launch if you double-click the file without specifying what software you want to use to open it.
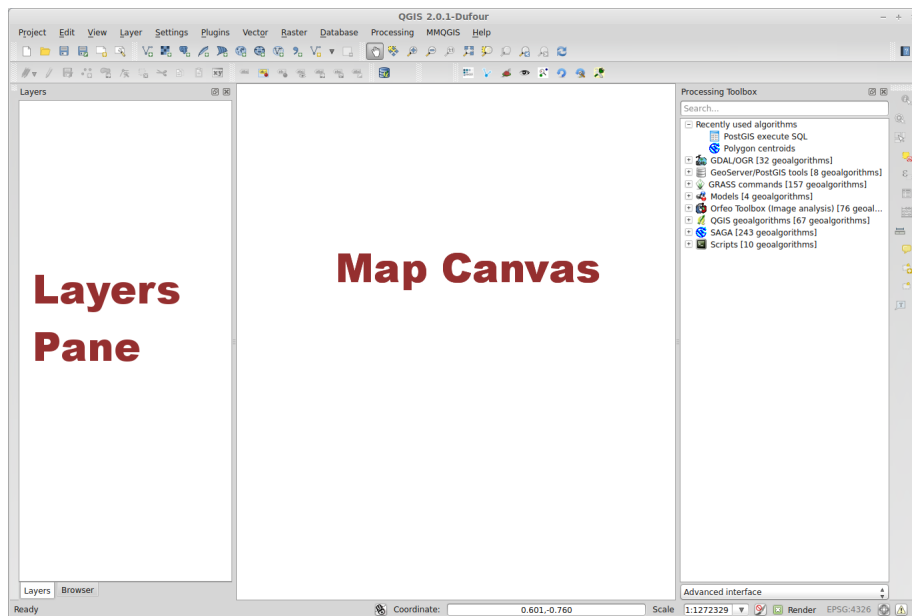
Figure 1: The main QGIS window

Now let's add some data. QGIS allows you to pretty easily add file-based data (such as shapefiles) as well as data from web map services (WMS, WCS, WFS) and spatial databases (such as PostGIS). As with many software, there are multiple ways to accomplish the same thing. I find that the easiest way to add data is to drag and drop from the Browser Panel on the left. Navigate to your unzipped data files, and drag the stages shapefile (might be named something like US_state_2010.shp) to the map canvas. The state layer should appear. Then drag the CSV file (which I suggested you store in the same folder as the shapefile) to the map canvas. Because this file lacks geographic data, nothing should change. Switch back to the Layers Panel in the left-hand pane.

In order to map the demographic data, we need to join it to the spatial layer. Double-click the states layer in the Layers Panel. In the left column, choose the Joins tab. At the bottom of the window hit the Green plug sign to open the Add vector join dialog. Choose a Join layer. (Since we are only using one attribute table in this exercise, the correct table should be selected.) Then choose GISJOIN as both the Join field (matching field in the Join layer) and Target field (matching field in the base table, which is usually your spatial layer).

Cartographic Display of Demographic Data

This section is incomplete. Prof. Hachadoorian will demonstrate in class.

To understand which columns have the data you are interested in, use the codebook that you downloaded with your data.

Create a Choropleth Map

A ***choropleth*** map is a map that color the geographic entity (states in this tutorial) based on the value of underlying attribute data. Choropleths should virtually never be used to represent raw count data! They should be sued for density (population per unit area), rates (cases or incidents, such as a disease or crime, per population), or proportions (such as percent of the population in a racial or ethnic group). Prof. Hachadoorian will demonstrate how to use the Expression dialog to turn the population count into a population density.

Create a Proportional Symbol Map

Raw count data can be represented by a symbol, such as a circle, of varying size.

Create a Proportional Pie Chart Map

For data that adds up to a total, we may display the result using a pie chart. This may be a proportional symbol map where the pie chart size is controlled by total population, and the pie chart itself shows the components of the total. For example, for housing, you could display a pie chart showing renter-occupied, owner-occupied, and vacant housing.