Getting started with QGIS

Tutorial by Dante Castellanos-Acuna & Andreas Hamann

This lab comprises a curated mix of basic and advanced GIS functions to create maps of your study area and to obtain spatial data for your sampling or site locations. The lab is designed for beginners who have never used a GIS system before, and for those familiar with ArcGIS but who are are interested in a free and open-source alternative. The lab will reveal most of the important toolboxes and menu structures of QGIS, so that you likely find other things you need by yourself.

QGIS has now been around for 10 years and has become as powerful and polished as ArcGIS. In many ways, it exceeds ArcGIS in ease of use and in the logic of organizing tools and menus. QGIS now also has a critical mass of community support, and a Google search of "*QGIS … something*" will reliability give you the answer on how to do something via StackExchange or YouTube.

Getting good with a complex software package is quite a large investment in time and effort. Selecting well-established open-source offerings appears to be a reasonably good insurance that your skills don't become obsolete in a few years time. If you are interested, you can download QGIS here for your own computer: http://www.qgis.org. We recommend using the latest Long Term Release (LTR) version.

1. Opening and working with shapefiles

Shapefiles are a data format for polygons, lines or points. There are many formats, but we recommend working with ESRI shapefiles that consist of 4 files that have the same name but different extensions. The files with the .shp and .shx extension contains the coordinates and vectors required for drawing shapes, the .prj file contains projection information (e.g. geographic lat/long or projected systems like UTM), and the .dbf file is an attribute table that describes the properties of each point, line, or polygon.

Shapefiles for all kinds of things can be obtained from many websites, and the UofA library has an excellent guide to these resources: https://guides.library.ualberta.ca/geospatial-data-maps. However, for this lab, we have already prepared a data package that you can download here: https://tinyurl.com/QGIS-dataset. This tutorial works in conjunction with a PDF that has numbered screenshots for reference. The screenshots are downloadable here: https://tinyurl.com/QGIS-screenshots. The link for an electronic version of this tutorial is: https://tinyurl.com/QGIS-getting-started.

1.1. Open and subset shapefiles (Screenshots 1-4)

- Download and unzip the files from the link above to a convenient place (e.g. C:\QGIS Lab), and open QGIS from the start menu.
- Look for the Open Data Source button, or press Ctr+L, select the Vector tab, navigate to the directory where the lab files are, and look for the Outlines shapefile ("outlines.shp") ***(Screenshot 1)***.
- Click Add and then Close to see outlines for British Columbia, Alberta and Saskachewan appear on the main screen ***(Screenshot 2)***.
- Often, you want to work with only a portion of the downloaded shapefile, e.g. for an Alberta study area. Look for the select feature button and select Alberta ***(Screenshot 2)***.
- Next, right-click on the outline file on the layers manager (bottom left), and look for the Export and Save selected features as options ***(Screenshots 3 & 4)***.

1.2. Change the appearance of shapefiles (Screenshots 5-10)

- You can use the Layers box, bottom right, to turn layers on or off (checkbox), remove layers altogether (right click remove layer), and drag to change the order of layers ***(Screenshot 5)***.

- Also play with the zoom and pan tools at the top to modify the layout **(*Screenshot 5)***.
- To change the fill color of the Alberta shapefile, double click on it, select the Symbology tab, and pick a color you like. You can also give it a cross-hatching look or other classic effects ***(Screenshot 6).***
- You further have fine grained control over the color with the color wheel, or by using the pick color tool. This will give you a pipette to suck in a color from anywhere on your screen. Useful, for example if you want to emulate the exact same color scheme as already used on another document ***(Screenshot 7).***
- Next, import a line shapefile of Alberta roads (just like before ***(Screenshot 8)***
- Double click the road layer ***(Screenshot 9)*** and take a look at how customizing the visualizations changes from the polygon shapefiles to the lines shapefiles ***(Screenshot 10).***

2. Dealing with projections for spatial data

One of the more challenging parts when working with GIS data is that data files come in different projections that attempt to faithfully represent areas, distances and shapes (i.e. projecting the round surface of earth on a flat piece of paper). The bigger the area, the more difficult this is. Therefore there are hundreds of local projections that work well for a single country, province or state. Those projections get revised and updated, so there are literally hundreds of projections in active use or needed to import legacy datasets.

QGIS can re-project any layer on the fly as long as the projection is known and readable by the software. Normally, the first layer you import defines the projection of your data frame (although you can also change that later), and any other layers will be re-projected to match that first layer you opened.

Our first layer was projected in a Universal Trans Mercator Region 10 projection (UTM-10), and you can check it by opening the outline.prj file or the alberta.prj file with a text editor. Now, we want to import a geology layer (geol_srf_latlong.shp) that has no .prj file. This is quite common, and a first good guess is that it is a geographic lat long projection, given the file name. Let’s try to import it and manually define the projection.

2.1. Importing an undefined, WGS 84 projected file

- As before, go throught the import layer routine **(*Screenshot 11)***. Now you get a pop-up screen that asks for the projection since the .prj file was missing*.*
- Search for **4326** in the filter field, which finds the standard geographic lat long projection **WGS 84** that is widely used, e.g. for GPS coordinates, select it and hit OK **(*Screenshot 12)***.
- Success! We guessed correctly and it matches our previous layer. Now save a new copy to generate a .prj file, so that we don’t have that problem again in the future ***(Screenshot 13).***
- For the new output file, you can pick the original WGS 84 projection, or any other projection, including the current “Project” projection for consistency ***(Screenshot 14).***
- Note that when you import data without a defined projection, if you pick the wrong projection, your data will appear shifted relative to other layers, or will not appear at all because it is outside the frame. In that case you have to research the data documentation, talk to the people that generated the layers, or try common projections for an area from the long and exhaustive list of options that are built into QGIS.
- You can always view an undefined layer in a brand new QGIS project, but you can’t add anything to it that is in a different projection.

3. Coloring a layer by attributes

Once the geology shapefile has been imported, we are going to display the different types of soils with different colors, but first, it would be good a good idea to explore the attribute table of the shapefile and get an idea of the type of data we are dealing with.

- First, right-click on the file and select the Open attribute table option ***(Screenshot 15)***. Once open ***(Screenshot 16)***, we can see all the data associated with each one of the polygons of the shapefile, such as area, perimeter, and type of soil. Any of those variables can be used for color classification.
- Close the attribute table, double click on the GeofSurf layer, and select the Symbology tab. From the upper menu, select the **Categorized** option and then select the variable to be displayed ***(Screenshot 17)***. In this case, we are going to select the **SurfGeol** variable.
- You can chose a specific color ramp, or leave the random colors option (recommended for a class variable) and then hit the **Classify** button. The empty space in the middle should get filled with all the soil classes and their associated colors. By double clicking on the color or legend, you can make individual changes to any of the classes ***(Screenshot 18)***.
- Hit **Apply** or Ok and take a look at the new visualization ***(Screenshot 19)***.

4. Working with your own data: Part 1

When working with your own data in GIS, a common objective is to extract data from other spatial layers for specific points. For example, if you obtained coordinates with a GPS for your field sites (or if you are able to determined them accurately with Google Maps), you can use these coordinates to obtain additional information such as climate, elevation, or soil type, from GIS layers produced by other research groups. Subsequently, you can for example analyze if this additional information has any association with the measurements in your own study.

For this lab, we created a dataset similar to you might have yourself: a CSV file that contains coordinates for sample plots. Open the file “AB_Trees.csv” with tree species frequency information. We will use this example file to show you how to import this as a GIS layer and extract information for those coordinates. You should be able to do the same with your own data tables, which may contain anything between one study site and thousands of sample plots.

4.1. Import a layer of sample points

- Add the CSV file by going to the Open Data Source button, or press Ctr+L, look for the Delimited text tab, and navigate to the file ***(Screenshot 20)***.
- In the **Geometry definition** option, specify long for the X field, lat for the Y field, and look again for the WGS84 reference system (EPSG:**4326**) in the Geometry CRS filed by clicking on the small globe ***(Screenshot 20)***.
- Once the CSV has been loaded, you can change which layer to view by moving them up or down, or by turning them on or off ***(Screenshot 21).***

4.2. Extract information from a GIS layer to your own sample points

- If you moved the tree sample plots to the top of the visualization, you can see that plots fall on different soil types. To get this information into your CSV, go to the top menu, look for the **Vector** tab, **Data Management Tools**, and then the **Join Attributes by location** option ***(Screenshot 22)***.
- In the window that appears, select the layer with the tree sample plots as the Input layer, and the layer with the soils data as the Join layer. In the Joined layer option, specify where do you want to save the file that is going to be generated, and if you want a CSV file or a shapefile, among other formats ***(Screenshot 23)***. Click Ok.

- Once done, open the attribute table of the new layer*,* or open the generated file outside of QGIS (i.e. in Excel), and you will see that next to the tree frequency data, now there are columns that contain the data of the soils layer ***(Screenshot 24).***

5. Saving & opening projects, general navigating and query tools

At this point you have invested quite a bit of time to put your layers together, import them, order them, and color them. It's a good time to save your progress as a project. As a general habit, it is good to place all layers that belong to a project in the same folder, and save a QGIS project file in that same folder as well that recalls how to put this all together in the map window.

By default, QGIS will save relative paths to the layers, so from the place where the project is saved, QGIS will look in the same folder for your data layers. Generally it's a good idea to place all layers that belong to a project in the same folder (and in subfolders within that folder) where your project is saved, even if that means you have to duplicate some commonly used layers. The big advantage is that you can move your folder to another computer, or reorganize or rename your outside directories without the links to your layers being broken, as long as you don't change the file and folder structure within your project folder.

5.1. Save your project

- Go to the Project menu, and then select save, or click on the button resembling a floppy disk, type a name and hit save ***(Screenshot 25)***.
- You can now safely close QGIS.
- Open your project again by double-clicking your saved project file in Windows explorer. You should be back where you left off.

5.2. Zoom, pan, and query tools

- Since we are covering some general functionality, let's try out some basic zoom, pan and query operation. Locate the zoom and pan toolbar at the top and play with all the available options and explore what they do. Useful, and just like in ArcGIS if you are familiar with that ***(Screenshot 26)***.
- To the right of the zoom and pan toolbar, it's the identify or query button ***(Screenshot 27)***. After selecting it, highlight a layer, e.g. the geol_surf layer, on the bottom left panel, and click on any point in the map to view the attributes of the layer at that location. The results are shown in a new window on the right
- Try to highlight the tree sample layer (move it to the top so you can see it), and click on any of the dots to see the tree species frequency information.

6. Working with raster data

Another important type of spatial file is raster data, which is essentially the same as an image file. This image file is formed by a matrix, in which each cell represents a location and a value for a variable of interest, such as climate or elevation. Rasters are widely used in remote sensing by satellites, and many interesting layers are derived from those, such as land cover, primary productivity, vegetation types, etc.

6.1 importing a raster layer

- In the lab dataset there is a file named DEM.asc file that contains elevation information, typically called a Digital Elevation Model (DEM). Go to the Open Data Source button, or press Ctr+L, look for the Raster tab, browse to the file, and add it ***(Screenshot 28)***.

- Re-order your layers so that the sample plots are at the top, next the Alberta boundaries, and then the DEM layer that we just imported. Display the Alberta boundaries only with the outline by selecting an empty fill with the “no brush” option ***(Screenshot 29)***.

6.2. Coloring a raster layer

- Now that we have our raster loaded, it would be good to look at it with some nice colors. Double click on it, and go to the **symbology tab**. Then select the **Singleband pseudocolor** option in the **Render type** option. Select a color ramp, a classification mode, the number of classes, and then click **classify** to see the classes appear in the empty frame.
- If you are not happy with the appearance double click it again to change it. Click **apply** at the bottom to see the raster change ***(Screenshot 30)***.
- One great feature of QGIS is that you can re-create any color ramp that you can find in other software packages, e.g. in R, ArcGIS or a color scheme that you found in a book or on the web. Select “edit color ramp” instead of a built in color ramp ***(Screenshot 31)***.
- Then you need to define the colors where you want to start and end the color ramp. Next, create as many color steps as you want. In the example, we create an elevation ramp that starts at blue for sea level elevations, and ends at white for mountain pikes. In between there are green and brown steps ***(Screenshot 32)***. You can use the color wheel or the color picker to select a color from an outside source (e.g. take a photo with your phone of a color legend in a book, email it to yourself, then use the color picker tool to import the colors – very flexible)

6.3. Clipping a raster layer to the required extent

- As you can see, the DEM we imported also includes data for BC and Saskatchewan, but we only want to show Alberta. To do this, we are going to clip the DEM based on the outline of Alberta. Go to **Raster tab** on the top panel, then **Extraction**, and then Clip raster by Mask layer **(*Screenshot 34*)**.
- In the window that appears, select the DEM as the Input layer, the Alberta outline as the Mask layer, and make sure that the option “Match the extent of the clipped raster to the extent of the mask layer” is selected. Next, click on the button with the three dots and then the **Save to file (*Screenshot 35*)**.
- Navigate to the folder where you want the file to be saved, give it a name, and select a format **(*Screenshot 36*)**, The default TIF is an excellent format. You can try ASC, but it does not always work in this particular tool. Once all is selected, hit “Run”
- To add the right color scheme, we can copy the legend from the other raster. Right click on the raster already colorized in the layers manager, and find the **Styles** option, then select **Copy style (*Screenshot 37*)**. Now, right click on the clipped raster, find the **Styles** option, and then **Paste style (*Screenshot 38*)**.

6.4. Adding a hillshade for better visualization of terrain

- While elevation can be represented just fine with a colorized scale, another or additional way to represent terrain is a hillshade. This generates a 3D effect for valleys and hills.
- Start by duplicating the clipped DEM for Alberta by right-clicking on it and selecting **Duplicate layer (*Screenshot 39*)**. Double-click the top layer of the two, and in the **Symbology tab**, instead of the singleband pseudocolor render option, select the **hillshade** option **(*Screenshot 40*)**..
- You can now play with the position of the sun with the **Azimuth control**. The default of 315° usually gives good results, so we are going to leave that value **(*Screenshot 40*)**.
- Z-values depend on the latitude and the unit of measurements. Since this layer is in lat-long projections with degrees as unit, very small values are needed (Google “z values for hill shades). A value around 0.00005 gives good contrast for ridges and hills. Play with the brightness, saturation and contrast controls and hit “Apply” until you are satisfied with the appearance, then hit OK.

- Finally, switch from the **Symbology** tab to the **Transparecy** one, and set the **Global opacity** parameter to around 50% so that you can see the color layer underneath **(*Screenshot 41*)**.
- Perhaps move your points and roads to the top, so that they appear clearly above the background **(*Screenshot 42*).**

7. Working with your own data: Part 2

7.1. Extract raster values to points

- Similar to the extraction of soil types data, we can obtain data of elevation from the raster DEM for each one of the sample plots. At the top right of the interface, there should be the **Toolbox button**. Look for the **Raster analysis** option, and then the **Sample raster values (*Screenshot 43*)**.
- In the window that appears, select the sample plots file for the Input, and the clipped raster for the **Raster layer to sample.** In the **sampled points** option, navigate to the directory where you want the new file to be saved. Click Ok.

8. Final layout of maps for presentation and publication

Lastly, you might want to add labels, titles, scale bars, grids, or north arrows to your map. Plus, it is quite common to compose a final layout with multiple maps showing different aspects, topics, or time periods. Similar to ArcMap, QGIS offers you all a layout tool to compose your final map image without needing to do any final touches outside of the software.

- Go to the Main menu, click on **Project** and on **New Print Layout (*Screenshot 44*)**. A new small window should appear where we have to give a name to our layout. After that, we should end up in a new window **(*Screenshot 45*)**, which is where we are going to put our map as if it were a page on your report.
- In the left side of the screen, look for the Composer toolbar, and click the **Add new map** button. then and draw a frame on the blank sheet **(*Screenshot 45*)**. Our visualization of elevation and hillshade should appear.
- To zoom in and out or pan the imported map, look for the **Move item content** button near the Add new map button. Then click and drag to center and use CTRL-scrollwheel to zoom **(*Screenshot 46*)**. Alternatively, you can also go to the panel on the right, named **Item Properties,** then **Extents subsection**, and click on the **Set to map canvas extent** button, to adjust the extent. Finally, you can also specify a more accurate zoom by entering values on the extents fields just above the Set to map canvas extent button.
- Also near the Add new map button are the buttons to add scales, legends, text, and images (e.g. of a North arrow). Add a scale bar, right-click on it, and select the **item properties** option. You can play with some of the options, but for this example, we just changed the width of the segment to 100 units **(*Screenshot 47*)**.
- Once you are happy with the zoom and the scale bar, look for the **lock layers** option in the **item properties** section, also in the right panel **(*Screenshot 48*)**. This locks the frame, and prevents any change to it, even if we modify the main visualization.
- Go back to the Project screen and activate the layer showing the soil surfaces. Go back to the Layout view and draw a new map. By default the new frame should show the soil surfaces visualization. If not, look for the refresh button and that should do the trick **(*Screenshot 49*)**.
- Now you can export your maps either to Image or Pdf **(*Screenshot 50*)**. Save your Layout composer and your QGIS Project. Whenever you want to work again in the Layout Composer, just go to the Main menu and click on Project and then Layout Manager. This should open a new window with all the layouts for the Project.