

A First Map in QGIS: presentation notes

Mark Torrey

2024-05-30

(Title slide)

- *My name is Mark Torrey, I'm a GIS Specialist at DOHMH.*
- *You all have QGIS installed?*
- *And you got the data we're going to use? Did you unzip it?*
- *And the instructions to follow (hold up sheet)*
 - *A lot of this is going to be clicking-along with me, which will follow this instruction sheet — that way if you get lost, you can look up what you are supposed to do next.*

What is GIS?

- *What does GIS stand for?*
 - *Geographic Information System*
- *And does anyone know what it does?*
 - *Makes maps*

Why isn't it just called “mapping” software?

- *Why isn't it just called mapping software?*
 - *“Geographic Information” more accurately describes what the software does. It's not primarily about making maps, it's about poking your data. Maybe the simplest way to think of it is as an extension of your spreadsheet software like Excel. An extension that lets you look at your spreadsheet from a mappy kind of perspective (if you can). If all you wanted to do was make pretty maps, you want to be working in Illustrator and getting a lesson from a graphic designer (a useful, but different expertise). You can get pretty maps out of GIS, but that's kinda the secondary function after data analysis. Keep this in mind because it will help you understand some the of the more byzantine methods that come up in GIS sometimes.*

GIS software you need to know the names of:

- GUI:
 - ArcGIS: expensive commercial GIS software made by ESRI
 - QGIS: Open-source free ArcGIS work-alike software made by volunteers
- (R and Python: programming languages, also open-source)
- *If you're going to do anything in GIS you need to know that ArcGIS exists: it's the main commercial software that companies buy if they have tons of money to spend on their GIS work. It's made by a company called ESRI.*
- *QGIS is open-source and free, and designed to work pretty much just like ArcGIS. It does all the core things ArcGIS does, and is sometimes simpler to use.*
- *Arc and QGIS are both primarily graphical user interface GIS software.*

- You should also know that pretty much everything you can do in gui GIS you can also do in programming languages like R and Python, and a lot of people do. The big benefit to a GIS is you get a point-and-clicky interface, instead of having to write code.

NYC Environment & Health Data Portal — PM 2.5 Map

- This is a screenshot of the NYC Environment & Health Data Portal, which is exactly what it sounds like: a place to get environment and health data about NYC. There's lots of websites out there like this where you can download GIS data.
- Say we were doing research on fine particle air pollution. You can see in the screenshot they not only provide that data, but also give you a map of it.
- But this map is an answer to the question: "what neighborhoods in NYC have the worst fine particle pollution?" But what if what we want to know is: "What neighborhoods in The Bronx have the worst pollution?"
- You can see in their map of the whole city the Bronx comes out looking pretty much all the same color
- So you might decide we want to get the environmental health portal's data and make our own map. Which is how we're going to start today.

QGIS Main Window

- We'll use QGIS to make our map, and most of the rest of what I'll show you will be just a QGIS walk-through
- This is what the main windows looks like. Most of it should be pretty familiar from any other office software you've used.
- There's a bunch of tool buttons you can select from up top, a main window where you see your map, and a status bar on the bottom.
- Pay attention to this Layers list: each bit of data you add – the equivalent to a sheet in Excel – will show up here.
- The stack order really matters – what is on top of this layers list is what you will see in the map window. Make sense?

Let's get into QGIS

A new project

1. Start QGIS

- OK, fire up QGIS. Let me know if you have any trouble with that.
 - All these tools at the top can also be found from the menus. I like menus myself, but if you like tool buttons, you can mouseover them and the tooltip will tell you what they are.
 - Notice now we also have this Browser window where we can look for data sources. Today, we'll just be loading data from files, but this also lets you connect to databases and online servers.
- Everybody has QGIS up? And you know where our sample data is?

2. From the menu at the top, click on: Project → New (Or click the new button)

Adding your first data

1. In the "Browser" panel (on the left), open the "Home" folder and browse for the QGIS demo data I sent you.
2. Go into the "NYC-community-districts" folder and select the file that ends with .shp
 - .shp are "shapefiles" — a format ESRI developed for geographic information. It's not the only file that GIS data comes in, but it's a common one. Unfortunately the "shapefile format" requires a half-dozen or so other files with different endings/formats to go along with the .shp file, which you'll see if you

look at your data with the system file browser. You need to remember this sometimes when downloading and opening data — it can be annoying/tricky.

3. Now just drag that .shp file over to the map window and drop it there.
 - You should see a map of NYC

Adjust the Properties

- The first thing we're going to do is change the color
1. Right-click your "NYC CDs" layer in the layer window on the left again, and click on `Properties...`
 - (This will open the Properties window)
 2. Select `Symbology` and then click on the line that says `SimpleFill`. Click on the drop-down box for "Fill style" and select `NoBrush`
 - This "fill" and "stroke" stuff should be familiar if you have used PowerPoint or Illustrator. Those are "vector based" drawing programs. "Vector" data is one of the main formats of GIS data. So like PowerPoint and Illustrator, you can change the color of the fill, or the outline of the vector object. Does that make sense?
 - There's 3 kinds of vector data in GIS: points, lines, and polygons. Points and lines are just what they sound like, and polygons are basically an outline of any shape, which is what we're working with right now.
 - UNLIKE Powerpoint and Illustrator, in GIS you can't just grab the object and change it's size or shape. That's because it's not just a drawing, what you see is tied to an underlying spreadsheet of data, remember? We'll look at that spreadsheet data in a minute.
 - You CAN edit the data and the shapes, but it's more involved than Illustrator or Powerpoint and we won't really get to that today.
 - (BONUS): What is a vector? A picture, like a jpg, is made up of lots of tiny dots, right? In a vector image, each line and object is actually described by fancy math inside the software that tells the software how long each line is supposed to be, and it's curve, and color. Because everything is just a math equation, it can be scaled up and down without losing quality. In graphics that have a few shapes and colors, it can be very very small (because it's stored as just a bit of math). But if you wanted to make a photograph all vectors, it ends up being huge because you need a bit of math to describe every pixel.
 3. Click the `OK` button.
 - This will leave you with outlines of the community districts polygons with no fill.

Select The Bronx

- Since we just want to work with The Bronx, we need to select those community districts in The Bronx
1. From the menu at the top, click on `Edit` → `Select` → `SelectFeature(s)` (or click the tool button) to activate the selection tool.
 2. Hold down Ctrl and click on each of the community districts in The Bronx. They will light up yellow as you select them.
 - (While still holding down Ctrl, click again on anything you might have selected by accident to unselect it.)
 - This is a cool thing about GIS — you can select data by geography.

Let's look at the Attribute Table

- OK, now let's take a look at the spreadsheet view of the data
1. Without changing your selection, right-click on the `NYC-CDs_PM25` layer in the "Layers" panel on the left and And select `OpenAttributeTable`

- The Attribute Table will open in a new window — this is the spreadsheet table view of the data for that layer in the map window
 - Notice the selected rows: each row represents one of the community districts on the map, so you can hold down Ctrl and right-click the row number to select or unselect a community district there too. (Notice the selection changes on the map at the same time.)
2. Look at the column names at the top. Notice we have columns with the CD names, and information about the pollution levels in each CD as well as some other data.
 3. You can close the Attribute Table window.

Export The Bronx data

- *Now we'll save our selection of just The Bronx data*
1. Without changing your selection, right-click on the *NYC-CDs_PM25* layer in the “Layers” panel on the left again, and select **Export** → **SaveSelectedFeaturesAs...**
 2. From the “Format” drop-down menu, select **GeoJSON**
 - *GeoJSON is another commonly used GIS data format. You can't do as many things with it as you can with a Shapefile, but it conveniently only saves a single file, instead of that mess of files that the Shapefile format creates. So it's good for us right now.*
 - (Notice that “ESRI Shapefile” is also on that list)
 3. Click the **...** button and make sure you are saving in a directory that makes sense. Give the file a name like “Bronx-CDs_PM25”. Click **Save** to get back to the export data window.
 4. Everything else you can leave at the defaults. Click **OK** to export your data.
 - Your new layer of just Bronx CDs will be added to the map in a new color.
 - *Remember on the layers panel, things are stacked. You can turn a layer on or off by clicking the checkbox over here. If you turn off your Bronx layer, you'll see that our whole-city layer still has the Bronx selected.*
 - (You can click the **DeselectFeaturesfromAllLayers** button now to clear your selection.)

Classify CDs by fine particle pollution

- *A map of one color means nothing, so let's classify the Bronx by fine particle pollution.*
1. Right-click your “Bronx-CDs_PM25” layer and select **Properties...** to open the properties window again. Make sure you are on the “Symbolology” tab (on the left)
 2. From the drop-down menu at the very top of the window where it says “Single Symbol” select **Categorized**
 - *Categorized is going to split up our data based on categories within a column.*
 3. When you click the “Value” drop-down menu you can see the column names of all the data we saw in the Attribute Table spreadsheet view. Select **PM2.5_{Val}**
 - *“PM 2.5_Val” is the column in our Attribute Table that has the values for the fine particle pollution.*
 4. From the “Color ramp” drop-down, select “Reds”
 5. Click the **Classify** button
 - *The classify button will add all the different categories of the column you selected in “Value” to the Properties window.*
 - *Does anyone want to take a shot at explaining what that did? It's important...*
 - *For every value in that column in the attribute table, it created a color here. You can click **Apply** right now if you want to see those colors on your map immediately.*

6. In the “all other values” line, double-click on the color swatch. Left-click on the color bar where it says “Color”. Change the color to some shade of gray and click and again to close the two color choice boxes.
 - *We want to make these gray because the ‘all other values’ are parks in the Bronx.*
7. click the button to apply your classification to the map.
 - *And we can see our community districts where darker red means higher fine particle pollution*

Save your map

- *Now is a good time to save your map*
1. Click on → to open a file browser window.
 2. Browse to a location that makes sense, give your project a good name, and save it.
- *Be aware that the project file (.qgz) will not* include your data. So if you move the data your project depends on (or rename it) you will have to tell QGIS where it is when you open your project again later. For that reason, it often makes sense it keep your map file with the data that goes along with it.**

Next up

- Other things you can show them:
 - How to classify by “Graduated” instead of “Categorized”
 - * Why Categorized is a problem — people have trouble reading a map with more than 4 or 5 colors
 - * You can also get into accessibility choices for people with color disabilities
 - Setting the projection:
 - * If you load data and get a mostly blank with tiny data far apart, your projection is the problem — time to google it!
 - * Spot the chance in projection from WGS84 (EPSG: 4326) to NAD83.
 - * Know that the standard projection for NYC is NAD83 / New York Long Island (ftUS) (EPSG:2263)
 - Describe/show how a join works:
 - * You often have data that is just in CSV, but maybe still can be associated with a geography like community districts. Then you do a join, which is just like a join in Excel, except one of the things you are joining is the CDs shapes.
 - Where the plugins manager is
 - * Sometimes they just work, sometimes you need to install additional python libraries, sometimes they are experimental
 - Add labels (often a pain)
 - More on the relationship to Adobe Illustrator and vector programs:
 - * If you see a fancy map, the data was probably worked in GIS software, and then the layers exported so a graphic designer can clean them up in Adobe software.
 - How to play with line weights and colors to get a better map
 - * Add more layers like one to show borough boundaries and label the boroughs, and hydrology
 - How to use layout manager
 - * What does ARCGis do that QGIS doesn’t? (Not sure, have to look up good examples of this.)
- *Now I’m going to let you go at it on your own. If you want to follow the directions sheet, it will walk you through adding some points data of hospital locations and classifying those.*
- *But if you have specific questions, we could talk about those as well. Or, heck, we can go off-script and try something random if you like. (See other things you can show them list above.)*