Welcome to the GIJN 2025 mapping class: Using QGIS

Andy Lehren, Mongabay investigations editor
Kuek Ser Kuang Keng, Pulitzer Center data editor
The Global Investigative Journalism Conference 2025
Kuala Lumpur, Malaysia

A quick introduction

Mapping is used by investigative reporters for a wide range of stories -- Revealing crime hot spots. Highlighting environmental destruction. Uncovering demographic patterns. Documenting war crimes.

Goals for this class

In this class you will quickly go from learning the basics to analyzing data in ways you can only do with mapping software like QGIS.

- In this class you will learn how to use QGIS to analyze maps and data
- You will learn the key parts of the program for everyday use
- You will learn how to import in data and maps
- You will use several different maps, so you can learn different tools for different scenarios that are common
- Among the tools you will learn is how to join data using intersects and more.

You do not need to know anything about mapping software for this class. But if you do, that's great, and we'll work with you to build those skills.

Let's get set up!

First, we hope that you have already downloaded and installed QGIS, a free, widely-used mapping program. If not, please download and install it on your laptop using this link: https://www.qgis.org/download/

You should see a screen like this. Accept all the standard defaults for your laptop's operating system. It should be quick.

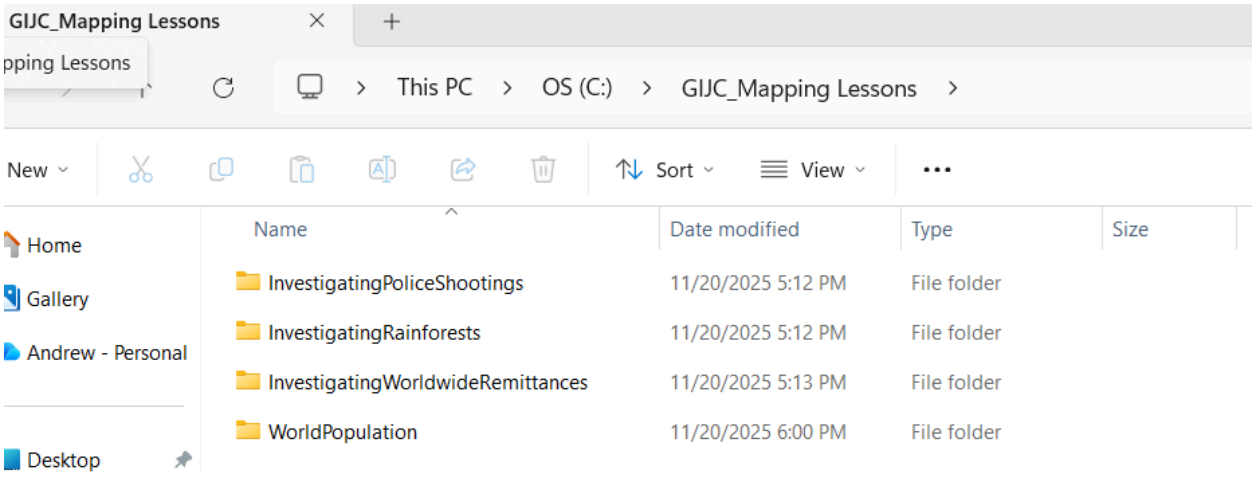
***

The data we are using for this class is located here:

https://tinyurl.com/GIJCMapping

Please make sure that the data is also downloaded on your laptop, and you know where you saved it.

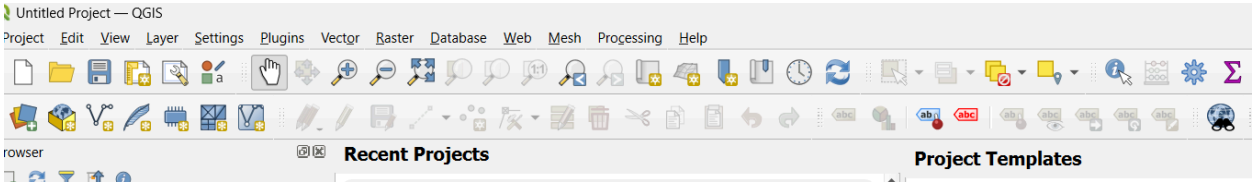
I've saved my data in a folder like this (so that I know where it's at)



Starting QGIS

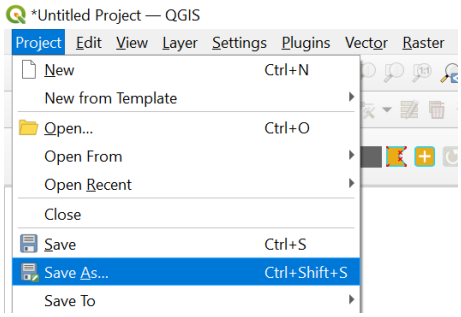
When you start up QGIS, your screens may look slightly different from the person next to you, depending on your laptop and your version of the software. No worries. Let's make sure that all our screens look the same.

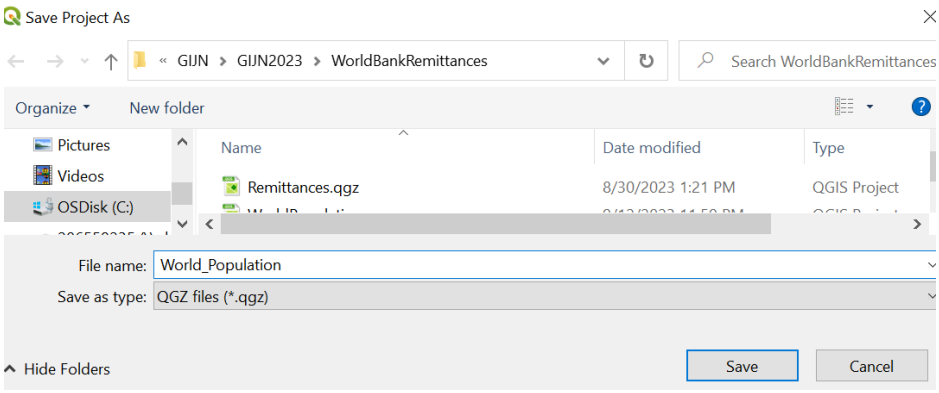
Let's say your first look is a menu bar and a blank page



There are a lot of buttons and options -- but here is the secret -- you really just need to know a dozen or so of these to do most investigative work in mapping.

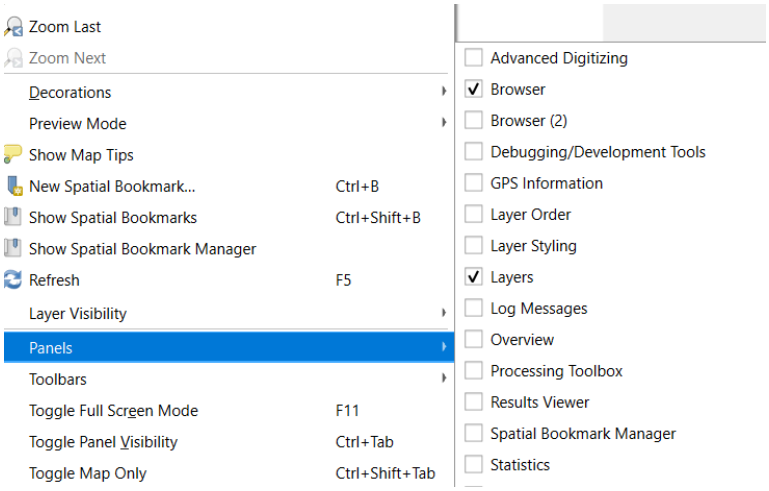
First, we're going to create and save a project -- this will be the place where we will create our first map in this class. Go to the left-most part of the menu bar, Project, and save a project. We'll call it World_Population.





Now let's do one last thing before we bring in our first map -- let's make sure we can see a legend to go along with our map. It will help us better navigate our maps, especially as they get more complex.

Go to View … Panels … Layers … and check the layers box



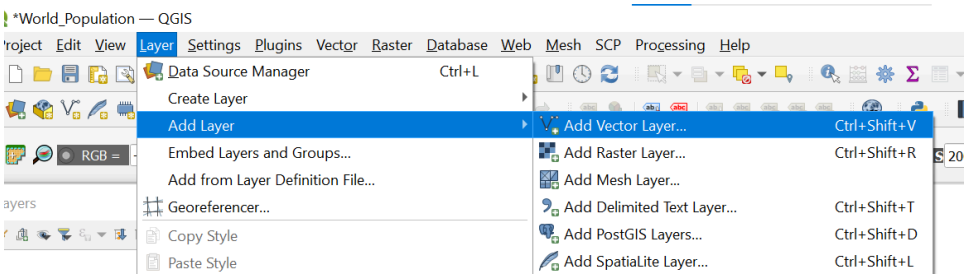
For some of you, this may be checked in your default settings.
We just want to make sure that you have this set up.

<u>Bringing in our first map</u>

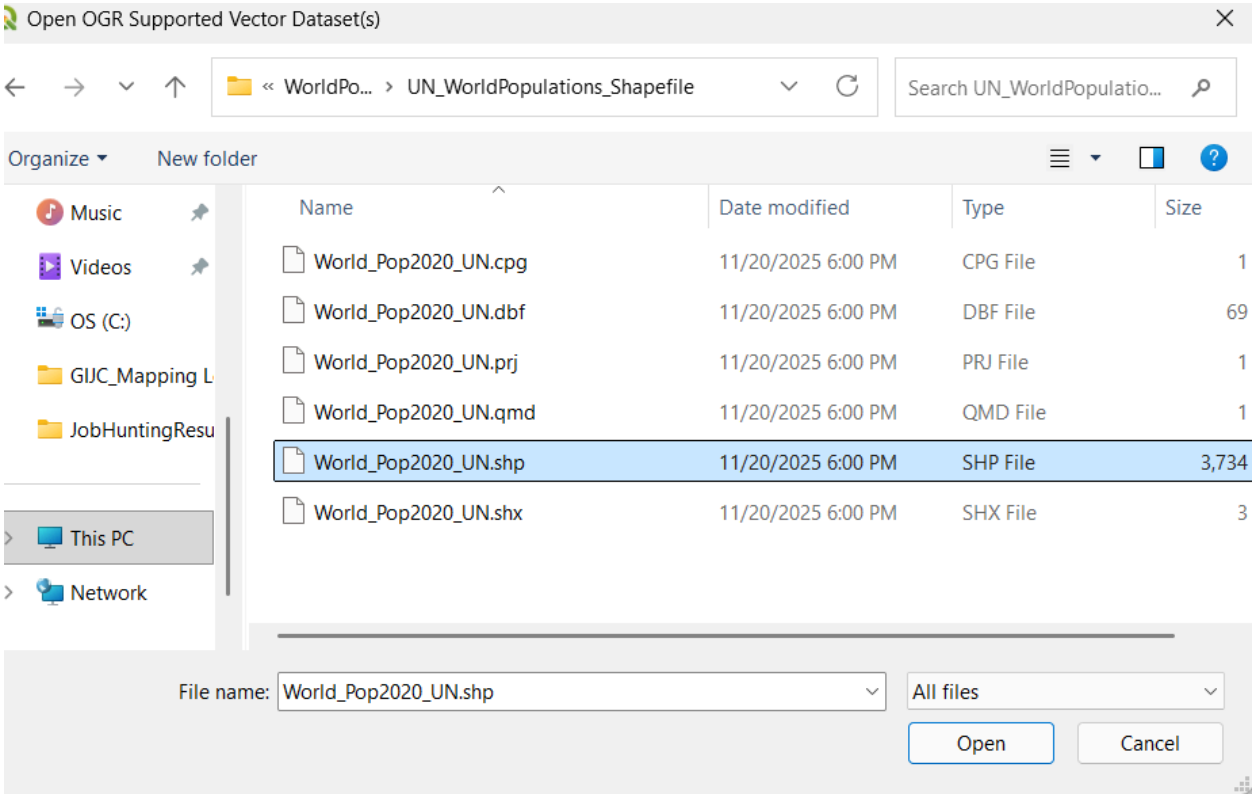
Whew!

Now that we've done a little housekeeping, it is time to bring in our first map -- and in the process, learn some more of the key buttons that will help us dive deeper.

Maps in QGIS are called layers. And there are a bunch of different types. For starters, we will use some of the most common kinds of files. QGIS calls these 'vector layers.' (We will talk later and explore the other options. But that is later).



Once you've picked this, you'll need to go to where you downloaded our training data files. The file we are looking for is called World_Pop2020_UN.shp. It will be in the folder labeled WorldPopulation.

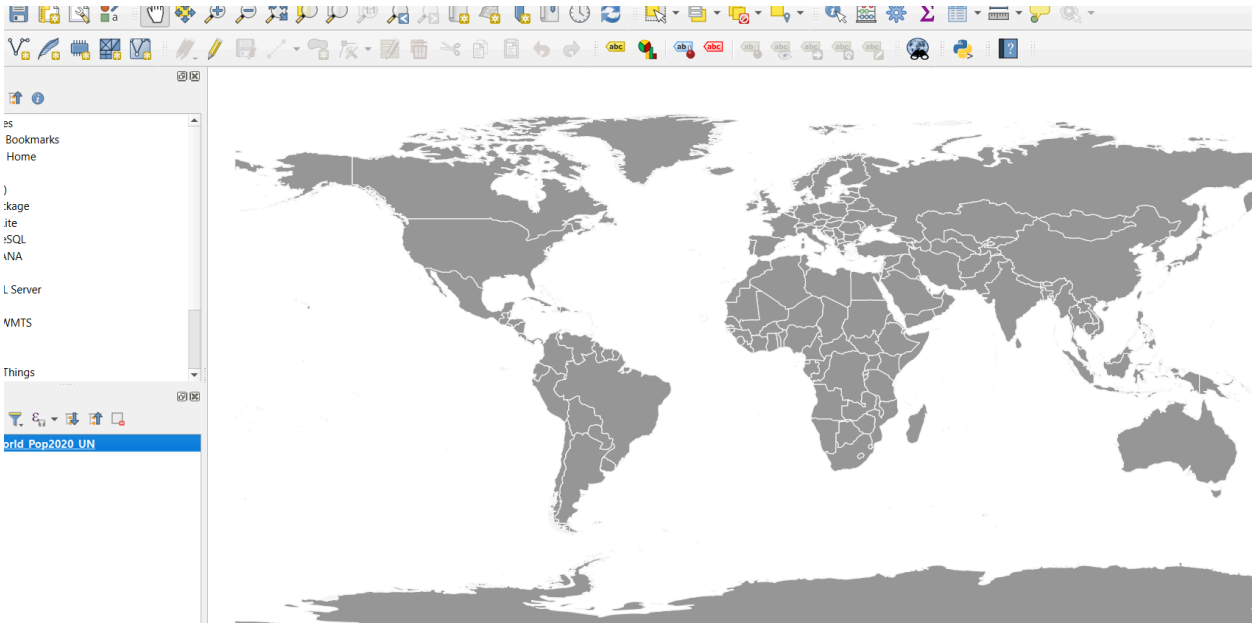


Let's find that file and load it into our project.

Once you have selected it, click Open … and then Add.

This filetype -- a shapefile with a .shp extension -- is among the most common mapping files out there. There are a lot of free shapefiles available for anywhere in the world. In this case, the file is from the United Nations, with population data.

Your screen should now look something like this:

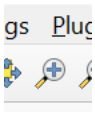


(The overall color may be different. We will play more with this later on)

Let's learn several buttons that will help us better navigate this map -- and the deeper analysis we'll do later on in the class. Some of these may be intuitive.

Zoom in

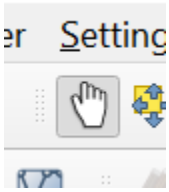
First, you may want to zoom into a country or region.
There is a zoom button along the top menu -- the icon that looks like a magnifying glass with a plus sign:



Just click and drag a square around the region you wish to zoom into.

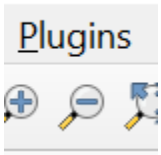
Move your map around

You may want to move around your map. In this case, go for the white gloved hand along the top menu:



Zoom out

You want to zoom out. That is the magnifying glass with the minus sign:



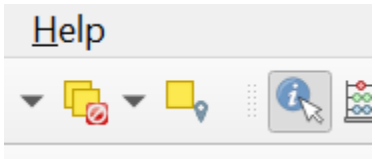
Just click and drag a square around the region you wish to zoom out from.

Going back to the way things looked before

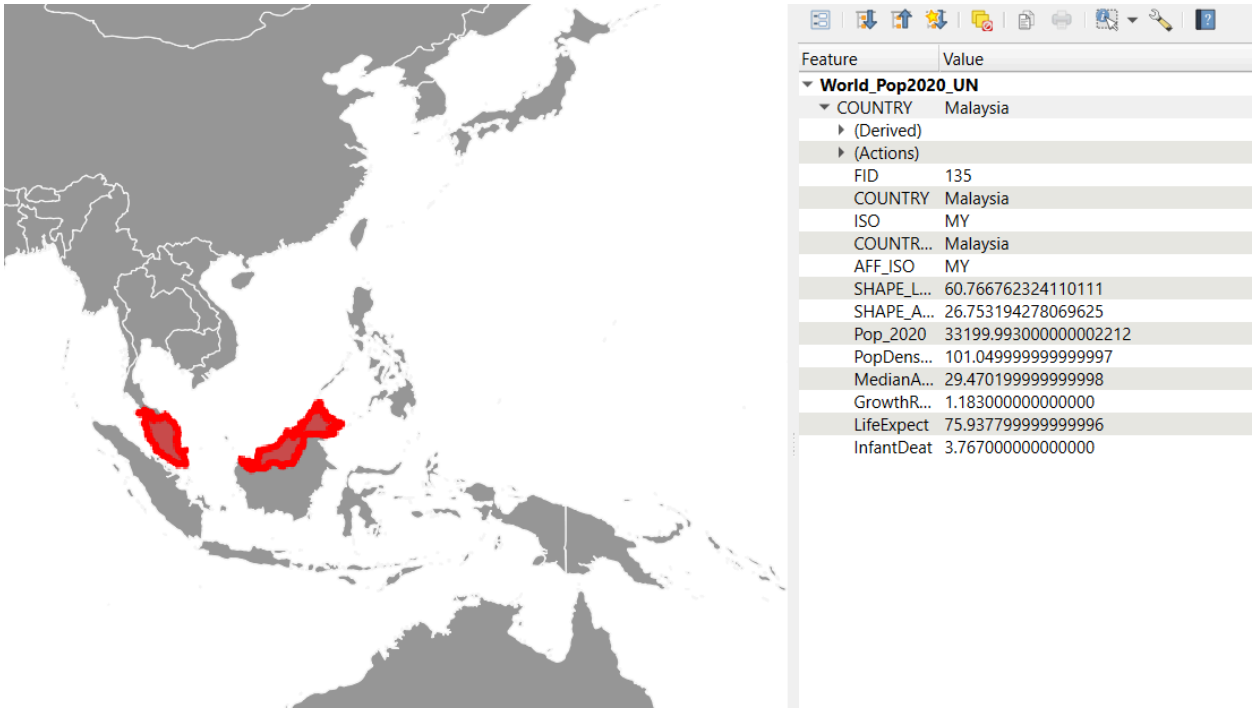
Another handy button is to go back to how your map looked one step back. Like you zoomed in too far and now want to go back one step. It's the magnifying glass with the back arrow.

Get more information

You may also want to know more about an area on your map -- use the icon with the little i (for information) :



When you use this and click on a place, you should get a popup box that shows you the underlying data that comes with this map:



Every digital map has underlying data. And you can quickly add more data. We will do that as this lesson progresses.

For now, let's play around for a few minutes to make sure everyone is comfortable with these tools before we dig deeper.

And let's take a moment to appreciate what we've learned so far.
At this point you've learned how to create and save a project, bring in maps from a governmental source, learned some QGIS buttons and begin to navigate your way around the map.

Now let's do more analysis.

Doing analysis

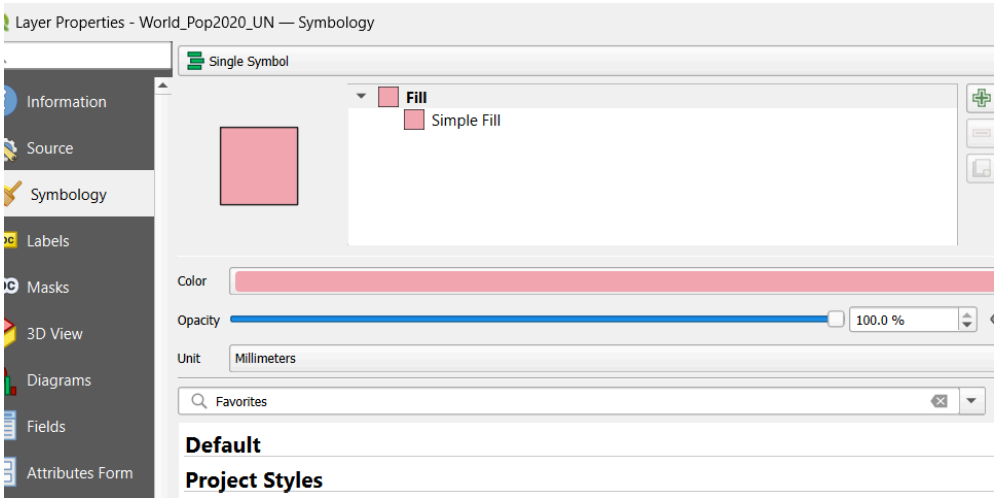
We're going to do different kinds of analysis.

Let's start with working on just a single map.
And don't worry -- Once we get this down, we'll tackle more complicated tasks.

The basic UN map we have is ok, but it could tell us much more. Let's get to work.

First, it's a single color. We are now going to learn more about symbology. How things look on a map -- and how it can help you spot patterns.

Double click in the legend on the name of the map

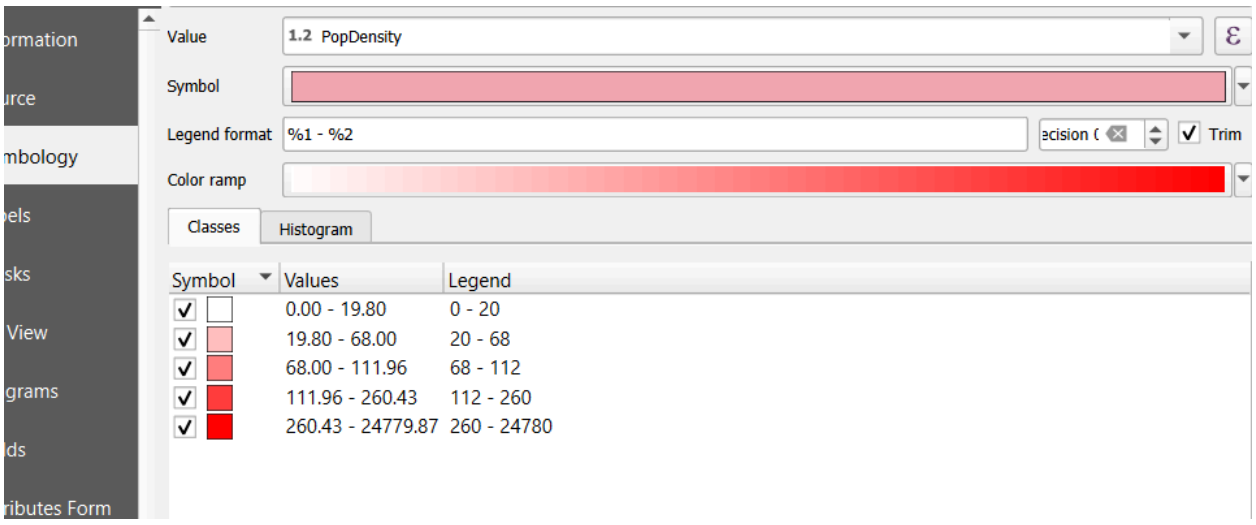


You should get a pop-up window that gives you a range of options:

Make sure “Symbology” is highlighted.

Notice how on the top it says “Single symbol” ?
That’s why our map is just one color.

A great way to begin to see patterns is by changing “Single symbol” to “Graduated colors.”
But there are decisions to make now.



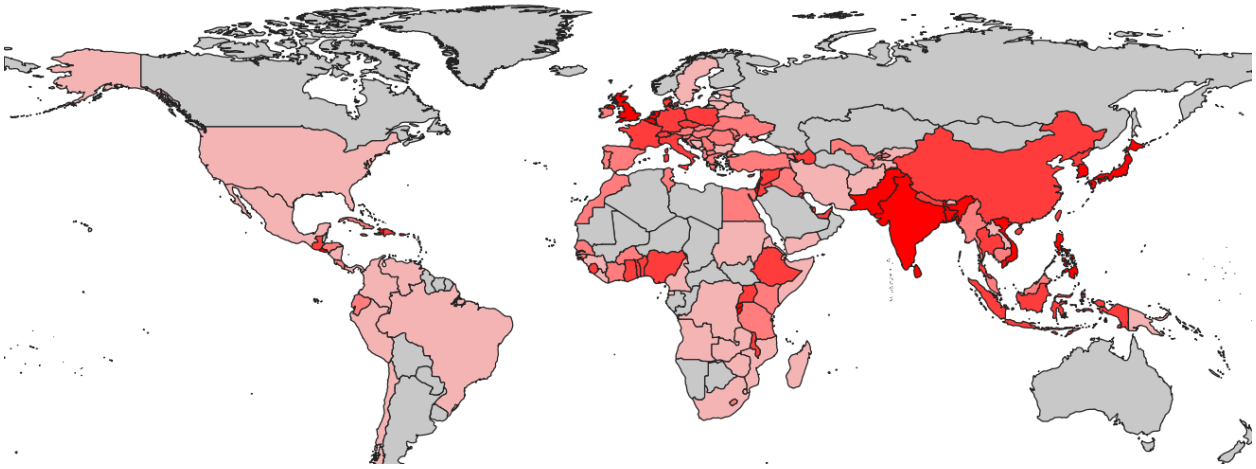
In this example, there are a series of decisions.

-- Which bit of data should we analyze? Here we chose population density -- how many people per square kilometer does a country pack on its land?

-- Then at the bottom, how will we break down the data? To show you clear differences, I chose equal count. Others can also be helpful, like natural breaks and standard deviation. We'll talk about that in class.

--How many different colors will we use? Here, we'll stick with the default setting of 5. You may want to choose fewer. So we can play with the color ramp.

The result tells us a lot:

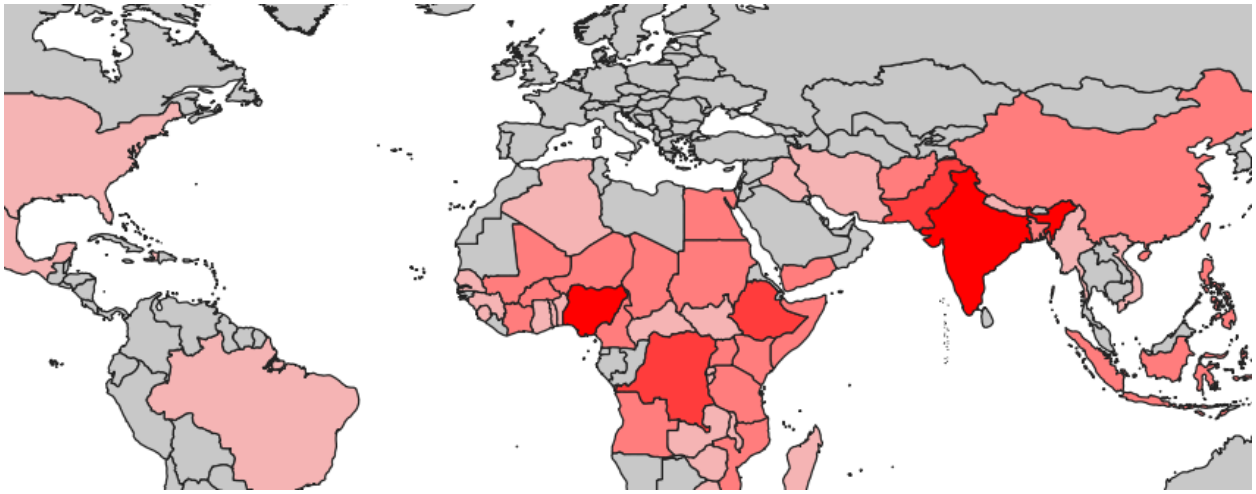


Parts of Europe and Southeast Asia are among the most densely populated areas of the planet -- including where we are at for this conference.

But there is a lot more data underneath this map. You can try these.

What happens when you look at life expectancy? At the median age of a population? At population growth rates?

At infant death rates (this is a good one to analyze using the natural breaks method, to reduce the signal to noise)?



You can begin to see regional demographic differences.

Notice in these examples that, to make things clearer at least for analysis, we've used a brighter color for the places we would likely report on for a story. (If we were publishing this map, we would more likely use gradations of a single color).

With a map like this, especially for a first analysis, using the right colors can make you spot things quicker. And having red show troubled spots can make a lot of sense to your eye.

Now you have begun to learn how to do basic map analysis.

Congratulations !!

Let's take it up a notch.

First, save this project again. And it is always a good idea to save a project.

==========
==========
==========
==========

Mapping different kinds of data: Spreadsheets. Shapefiles and combining them. Using police shootings data.

This time we're going to map different kinds of data.

One will be a shapefile, just like what we just used in the last lesson.
The other will be an ordinary spreadsheet that happens to have latitude and longitude or address -- we can map these.

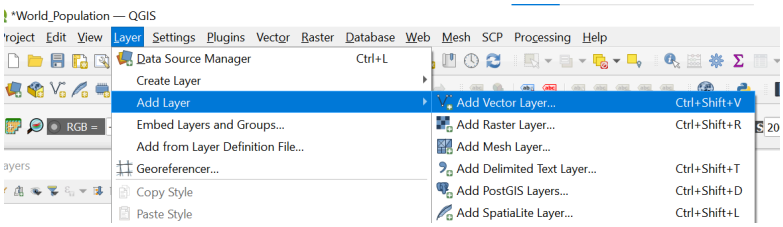
We will close our last project, and now create a new project. We can call it “Police Shootings.”

This data is part of an actual award-winning investigation into New York police shootings.

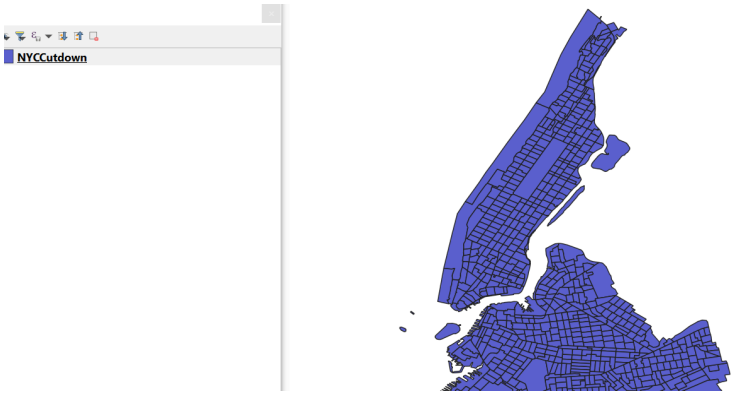
Let’s go the folder labeled NYCPoliceShootings
All the data we need for this lesson is in a folder called InvestigatingPoliceShootings

First let’s bring in a map for part of New York. It is labeled NYCCutdown, and covers Manhattan and Brooklyn. It’s in the folder labeled NewYorkCity_DemographicShapefile.

Bring it in the same way we brought in our countries map, going along our menu bar to Layer … Add layer …

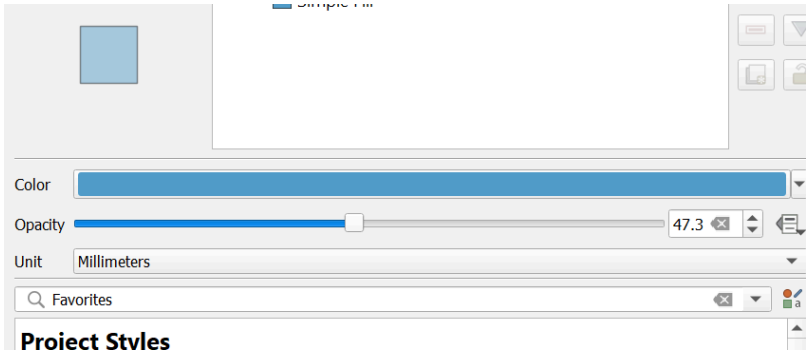


Your screen may look something like this:



That dark color may be a bit much for when we overlay police shootings, so let’s lighten it up.

Let’s click on the symbology, and lighten the color along with lowering the opacity.
(Opacity is like how thick the color is put down; whether it is a bit translucent and you can see what might be under it.)



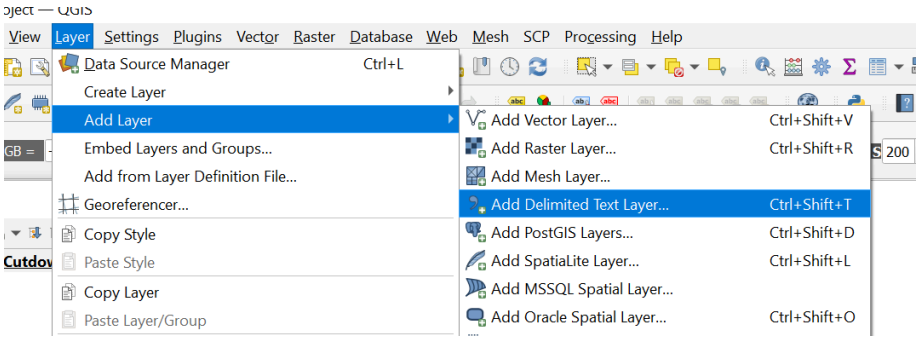
Now we are going to add in New York police shootings. This is from a reporter's spreadsheet, where she recorded each time an innocent bystander was shot by New York police, and was either injured or killed. It is part of a much larger data set. She built this spreadsheet through reporting out every New York City police officer shooting.

How can we map a spreadsheet when it is not a mapping file?
This is a common problem.
There are several ways to do this.

In this case, we have the street address and also identified the latitude and longitude. You might find similar files for this for crimes, chemical spills, tornadoes and more.

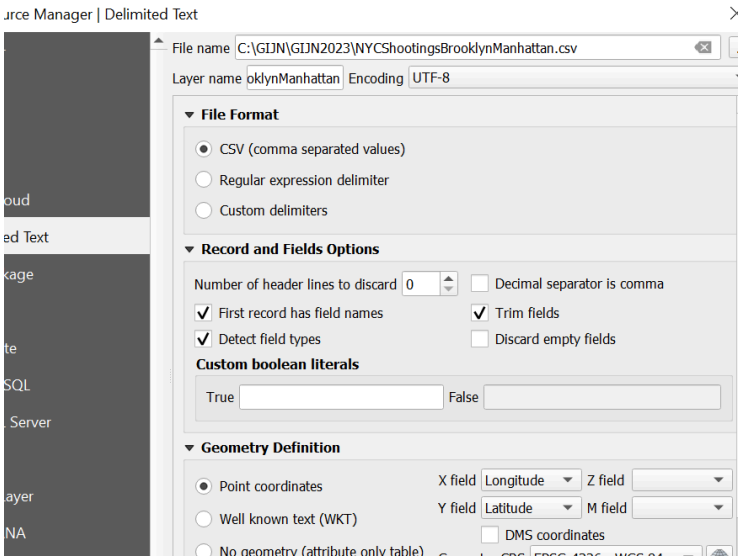
We'll go to Layer … Add layer

But this time we'll choose a delimited text layer



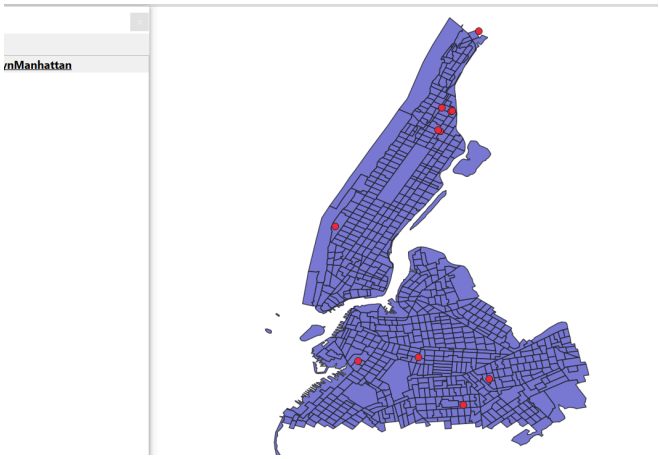
Now let's choose NYCShootingsManhattanBrookly.csv

It is a spreadsheet saved as a csv file.



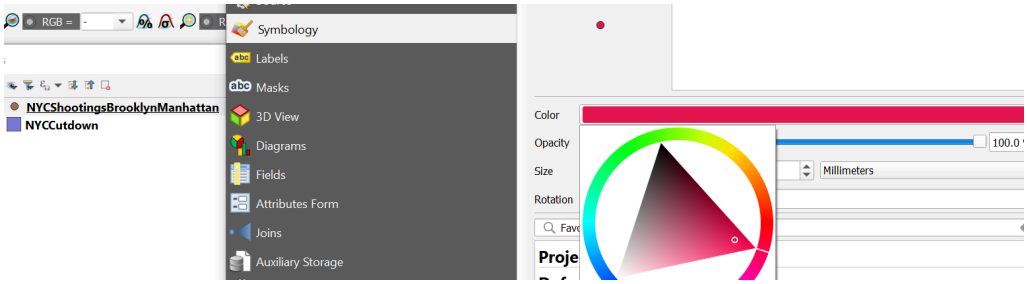
It happens to have columns labeled Latitude and Longitude, to make it easier for QGIS to know the column names.

Your result may look like this:



If the colors coming in are close together and hard to see, let's take a few minutes to make them clearer for our eyes.

To make the dots or points for each shooting appear red, make sure the shootings file is selected in the legend, double click the file, and let's change its symbology to red by double clicking the color.



Now we have a map of shootings in Manhattan and Brooklyn.
Let's dig deeper.

What are the communities like where these shootings happen?

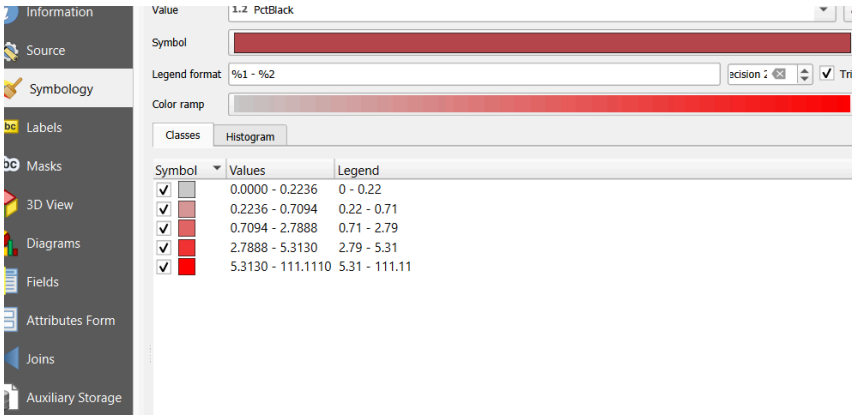
Let's look closer at our map of Manhattan and Brooklyn.
We'll take a similar approach like we did with the world population map.

In this case, we'll use US Census data that tells us a little about the demographics of each small neighborhood.

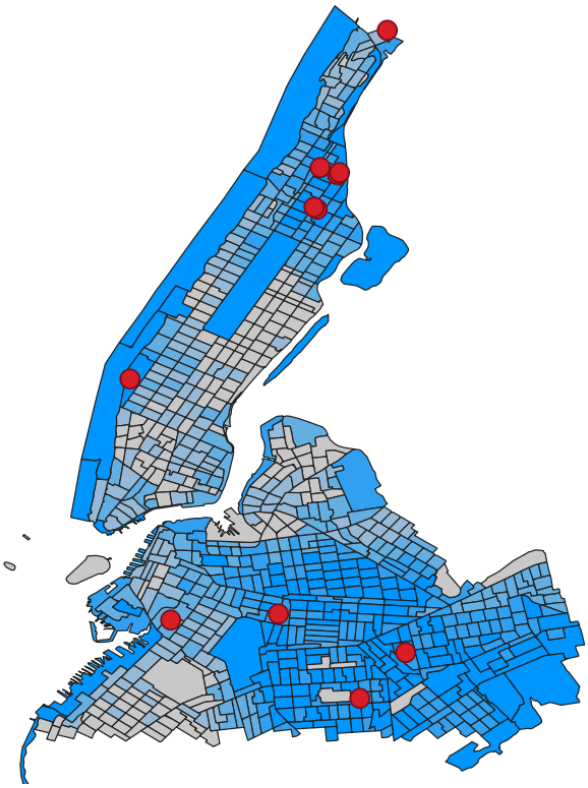
We'll double click on the file name, call up symbology, and notice it has demographic data.

Let's choose PctBlack -- the proportion of the population that is African American.

We also want to choose Equal count … and click the Classify button to do the math.



The result now tells us a lot more about where police shootings of innocent bystanders are happening.



See any pattern?

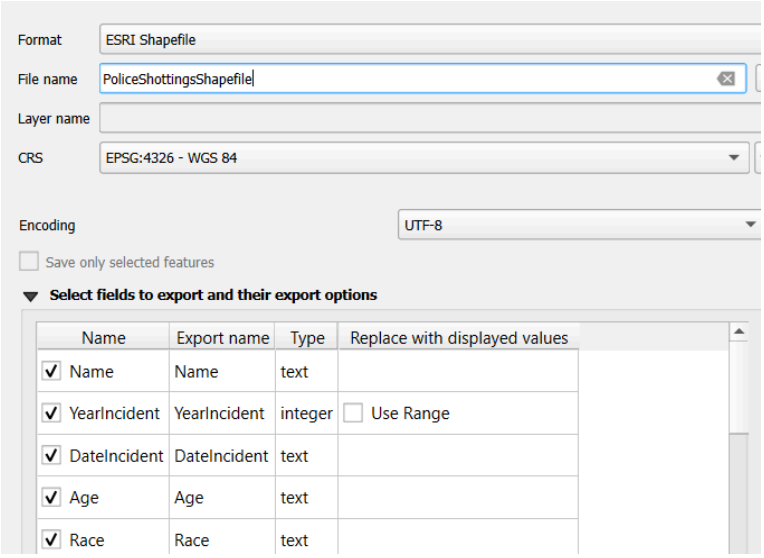
The shootings are happening in neighborhoods with a larger proportion of African Americans.

We'll want to document our work and save it.

==========

For a bonus lesson, if time allows we may do this:

If we'll want to improve our shooting data to now include neighborhood demographic data.



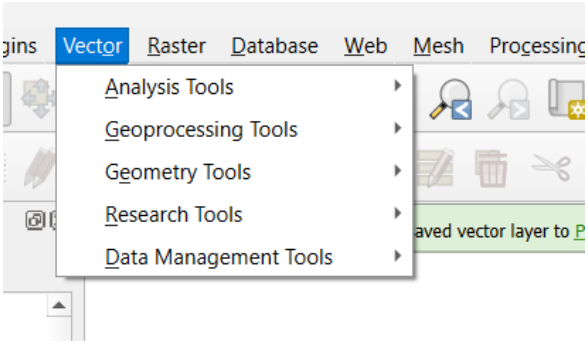
What does this do? It takes your spreadsheet and now has officially created a mapping file, known as a shapefile. This can work in multiple other programs.

Now we can take an extra step and actually join the census data with our shootings data -- so we can later scroll through the data of our new shootings file and be able to write quickly, and with precision, the demographics of the neighborhood where the shooting occurred.

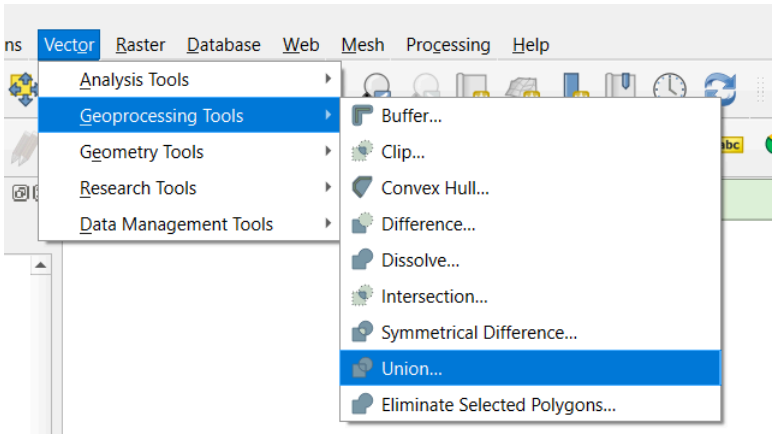
We're going to do what is called a union -- this will physically join our two maps.

Be patient -- there are a couple of default settings we'll need to quickly change.

First, we navigate to the vector part of our QGIS toolbar menu

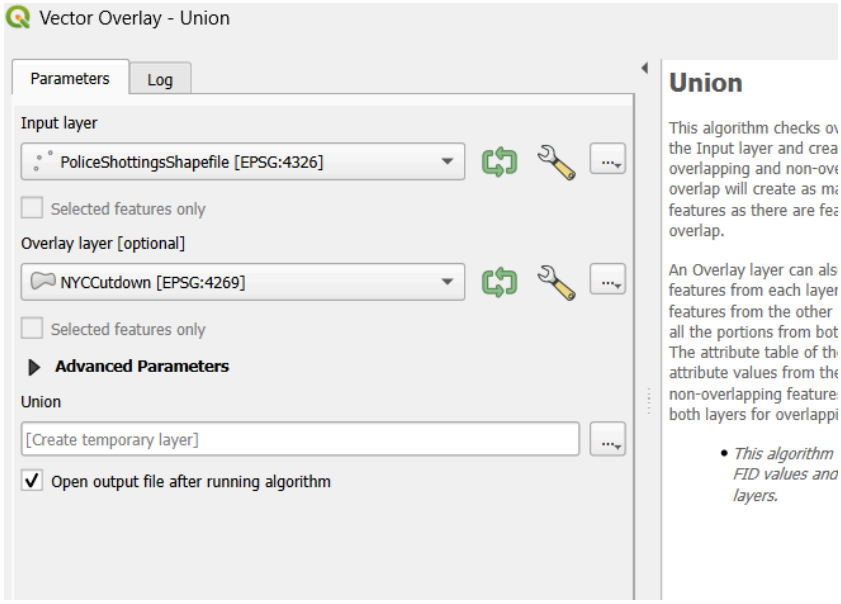


Select Geoprocessing tools … choose Union.



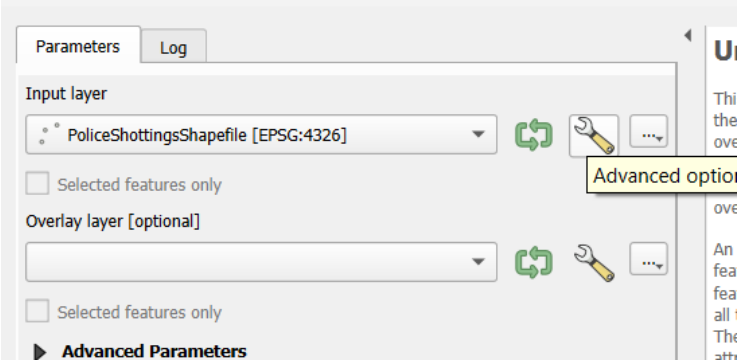
You then need to choose how you want your union to work.
We want to see each shooting -- but now know the demography of the neighborhood where the shooting happened. So we will choose this:

- The input layer is the shooting file
- The overlay layer is the demography file

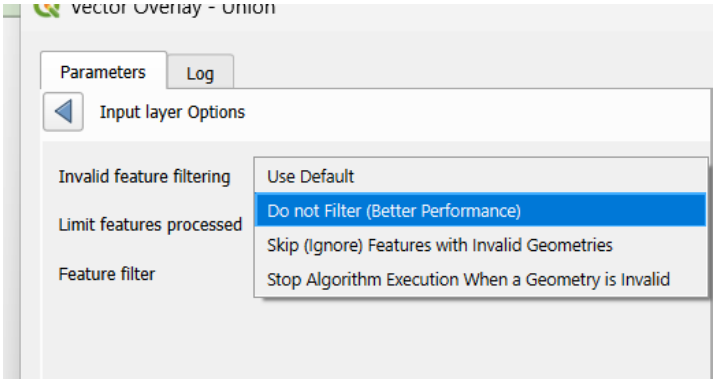


<u>But … The tricky part here are settings</u>

1. **Click on the little wrench icon for the demographic file -- the one labeled NYCCutdown.**

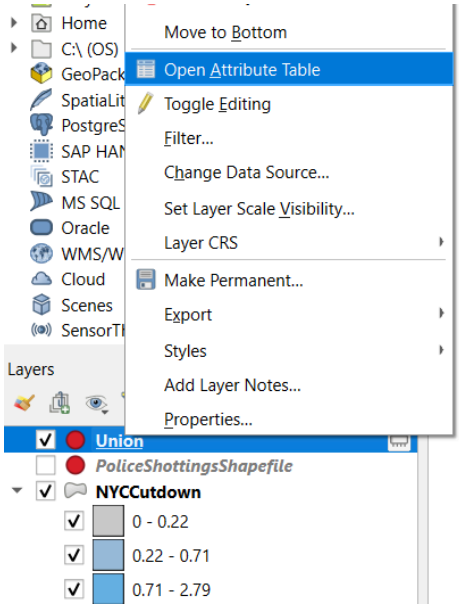


Then select Do not Filter (Better Performance)

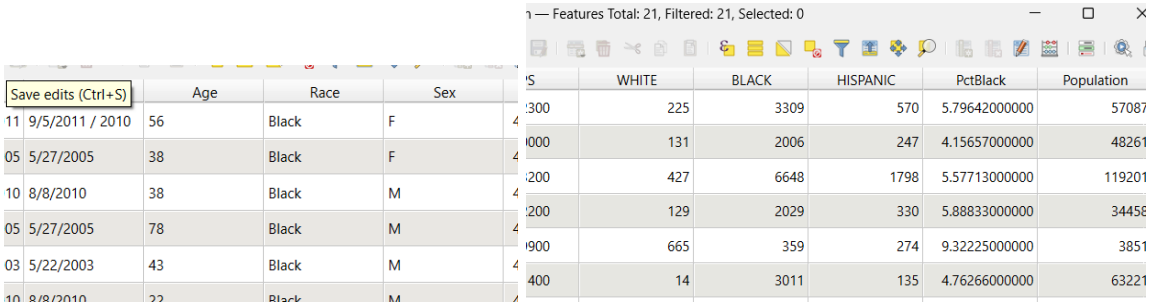


Now … if you want to see how we've combined the datasets, let's check out the union shapefile's underlying data table -- known in mapping software as an attributes table

RIght click on the Union file … navigate to open the attributes table



And you should see this as you scroll across:



This file, now saved on your harddrive, can be easily brought into a spreadsheet program. That can be handy.

==========

==========

==========

==========

==========

Investigating a rainforest proposed to be cut down for mining

Doing intersects and unions, calculating area and creating a new shapefile

Now let's start a new project, using data from an investigation published earlier this year. All of this lesson involves real data from government agencies.

This data comes from a Pulitzer Center investigation into proposed mines in Colombia, involving protected rainforests, lands for indigenous people and a large multinational mining company.

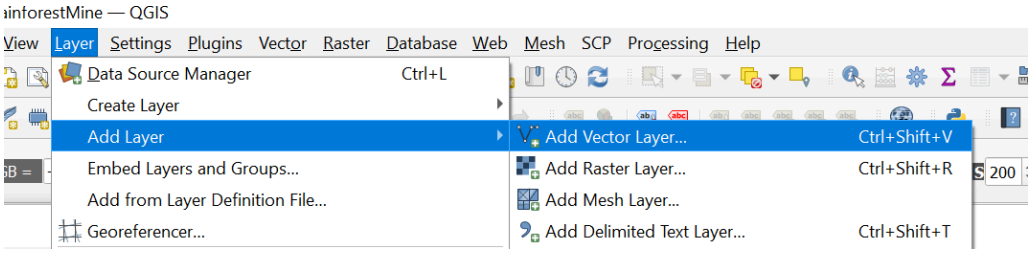
Our examples so far have used one or two maps.

This story involves six maps.

All of them are in the InvestigatingRainforests folder.

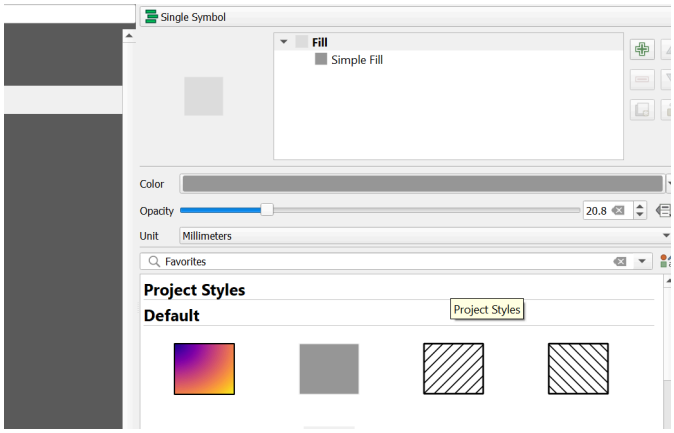
Let's first add a base layer we are familiar with -- the world population map we used earlier.

Remember, we go along the menu bar to: Layer … Add vector layer … and choose the World_Pop2020_UN.shp file. … then click Add and Close.



When the map comes in, QGIS may pick a bright color. Let's tone this down. Why? We're bringing in six maps, and we want to reduce any confusion. This is just a base map for this exercise after all.

To tone it down, let's click on the file name and go back to symbology.
Let's choose a gray color, and reduce it's opacity to about 20 percent.



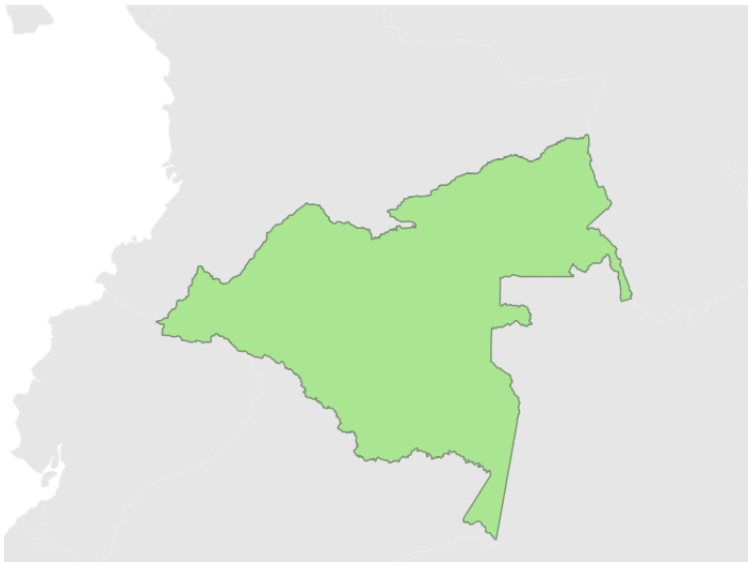
This base layer will just help keep us oriented as we add in more maps.

This is a rainforest story, so let's next add in a map of rainforests in Colombia.

To add more maps, we go along the menu bar to: Layer … Add vector layer … go to the rainforest folder and choose the shapefile. … then click Add and Close.

Then let's zoom to this layer.
Also it may have come in as some unusual color. Like what we just did with the first map, for this one, let's change that color by clicking on the file name, and choosing a green, and reduce its opacity to like 40 percent.

Your map should look something like this:



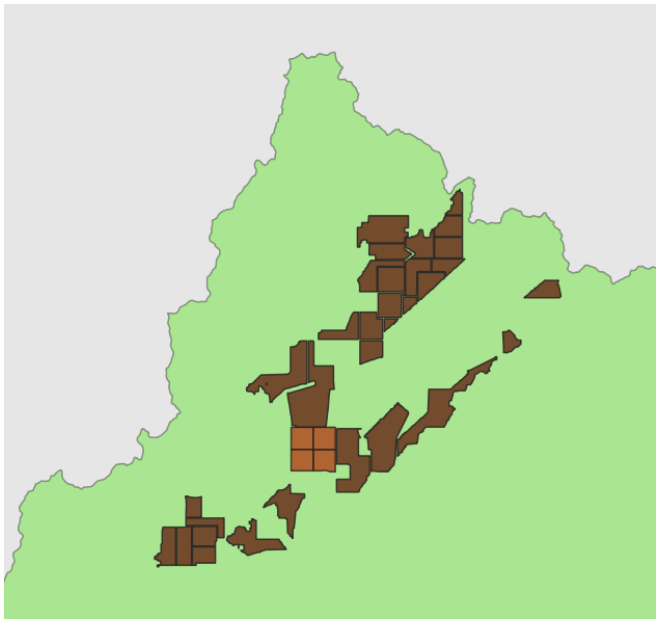
A mining company plans to start digging in this region. It has applied for permits for two phases of work. This would involve cutting down rainforest. Let's see where they plan to work.

Go to the proposed mine folder. You should see two folders, one for each phase. Each contains one shapefile. Let's bring them both in.

Again, to add more maps, we go along the menu bar to: Layer … Add vector layer … go to the proposed mine phase 1 and phase 2 folders and choose the shapefiles. … then click Add and Close.

Let's zoom to those proposed mines. And let's play with the colors and opacity so we can distinguish between the two phases of mine development.

At this point you should be getting comfortable with importing in shapefiles and changing how they look.



At this point, one part of the story should be clear: This proposed mine would cut into rainforest.

Might there be more to the story?

Let's learn more about this region.
From reporting, we know there are lands that have certain government protections.

- One is for the rainforest.
- The other is for indigenous communities.

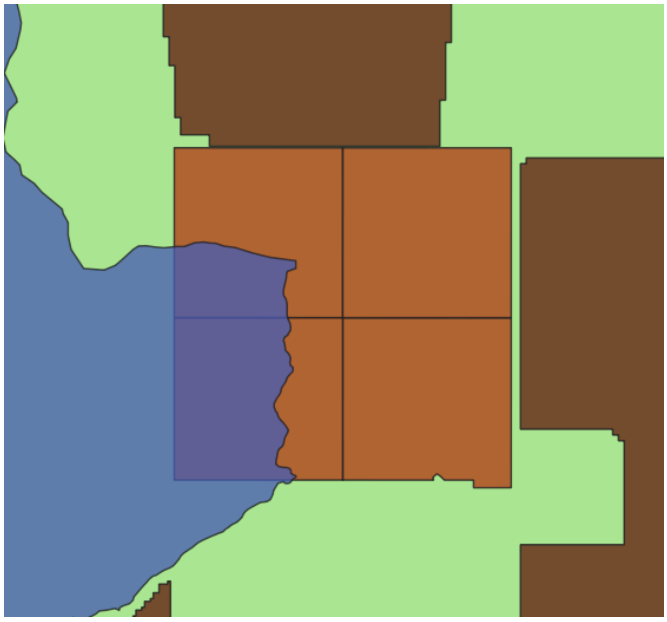
Let's look at those maps.

Let’s bring them into this project.

Let’s look at land that the government has deemed environmentally protected -- where development is not supposed to happen.

Now you should see that the mine -- in its first phase of development -- would cut not just into rainforest, but into environmentally protected land.

Let’s zoom into that. Your map may look something like this, depending on the colors and opacity you selected.

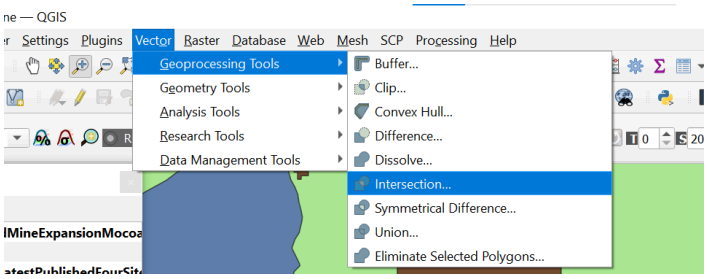


You can see where there is an overlap.

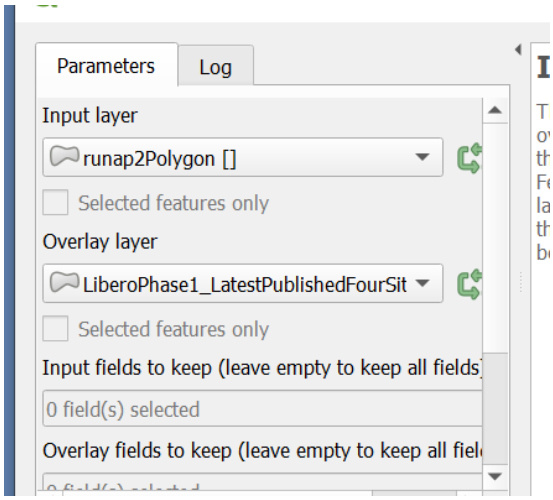
An intersect

Now we are going to take advantage of one of the most-used tools for doing mapping analysis: We are going to create our own map of this area where the phase 1 one development and the protected lands meet.

We are going to go to a different part of the menu bar. Go to Vector … Geoprocessing tools … Intersection



We'll choose to intersect our phase 1 map with the environmental protection map.
Leave all the other default settings.



Now you have created a new map just for the parts where the two maps intersect, or lay on top of one another.

Let's save our work. We've done a lot.

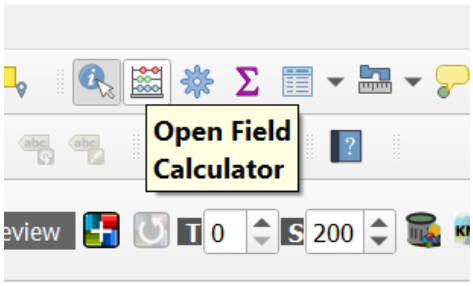
Calculating the area

We may want to not just include a visual, but also calculate the area for the overlap.

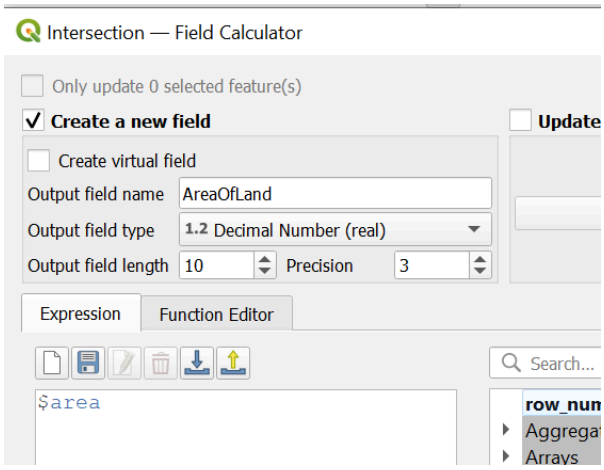
Like everything in QGIS, there are a few ways to do this.
And this gives you a glimpse at all the different kinds of math you can do in QGIS.

Let's open the Field Calculator -- the part that lets us do math and save it in our map.

Let's go the icon along toward the right that looks like a miniature abacus. (Please don't blame us if these are not too intuitive. We promise this is among the last buttons we will show you.)



It asks a few questions.
We will want to create a new field. We named ours AreaOfLand.
We want the numbers in decimals, so we have a little more precision.
And in expression -- to do the math -- type $area



Whew.

Now run it.

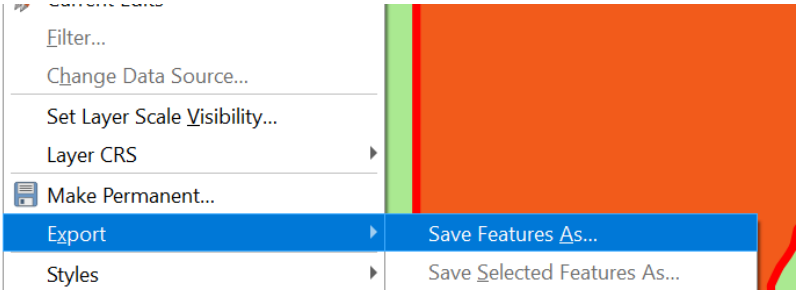
Now we know the area.

When you use your information button, you can scroll down and see the resulting number.

The units are in square meters. We can talk about how you find that out.

Saving your new file

QGIS at this point is saving all this information in its virtual memory.
Let's hard code this file and save it, so we have a more permanent copy.
We are going to create our own shapefile.



I right clicked, or apple clicked, on the file name, chose Export, and Save features as …
Then named the file.

We now have more detailed information to ask both the government and the mining company about.

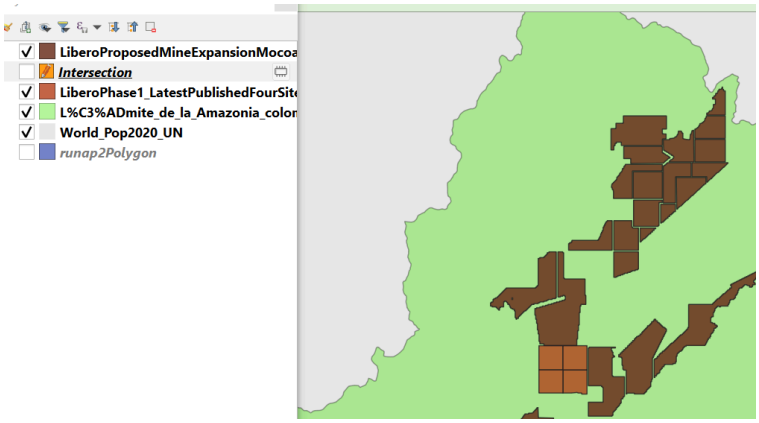
Doing another intersect

Now you have sorted out the area of environmentally protected land that the mine would damage, let's look at the land of indigenous peoples who live in the region.

It's a way to practice our new skill of doing intersects.

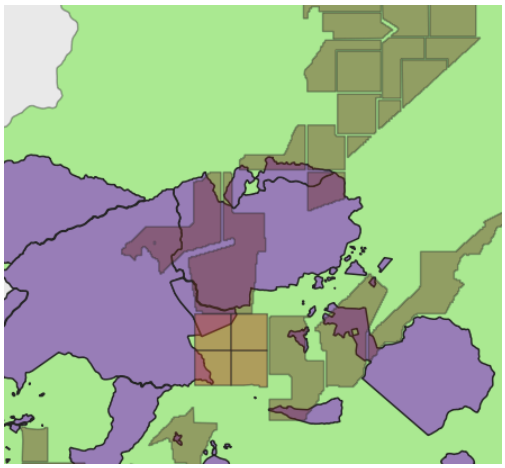
To make things simpler, let's turn off our protected lands map and out newly created intersection map. Just use the check boxes in your legend.

We should be back to a map that looks like this:

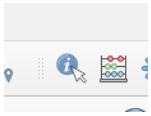


Let's bring in the shapefile stored in the IndigenousLands folder.

The result, after you play with colors, might look like this:



As reporters, this gives us a few options.
First, we could take the information button



Just be sure on the legend that the indigenous peoples file is the active or selected layer.

Now click on the indigenous peoples' lands that seem to overlap the mine -- to find out which communities we will want to call, to ask them what they think about the proposed mine, and whether they plan to oppose the development.

We can also figure out how much land in the proposed mine is covered by territory officially designated as belonging to indigenous peoples.

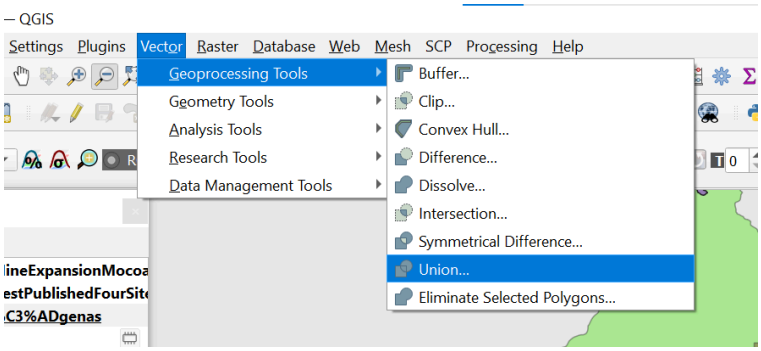
We will do an intersect -- the same analysis that we did just before this.

A union
If there is time in this class ….

… You may also notice that environmentally protected areas overlap a bit with indigenous communities. What if you wanted to know how much land the mine might span, including either environmentally protected land AND indigenous communities. But also avoid double counting any territory.

First we would start as if we would be doing an intersect, going to Vector … Geoprocessing tools ….

Then we would select Union, and pick the two files -- the ones covering indigenous peoples and the other for environmentally protected lands.



This will merge the two files -- and avoid any duplication between the two.

For bonus points, you could also use a union to join the mining files -- both phase 1 and phase 2 -- into a single overall mining development proposal shapefile.

Once we have this union file, then we can, once again, do an intersect with the mining file.

==========
==========
==========
==========

More joining data with maps

Sometimes we may have data that is not in a map, and does not have an address or latitude and longitude, and we want to map it.

Here is how to do that in QGIS.

We get data -- in this case from the World Bank, about the money that workers who go overseas send back home to their families, and other kinds of remittances; where people send money overseas, often back home.

It is not in a map.

But it does include the name of each country, and it includes the two-letter ISO abbreviation for the country name.

Here is how we can map the data to better see the overall pattern.

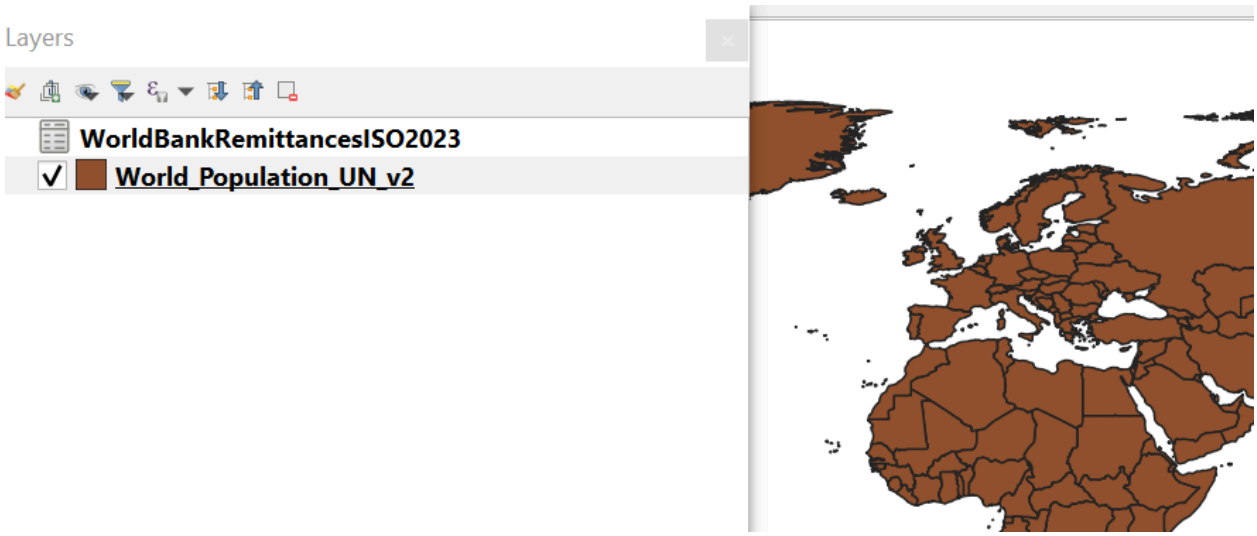
First let's have our standard world map with UN population data.

Now let's add the spreadsheet -- actually a .csv file -- to our map.

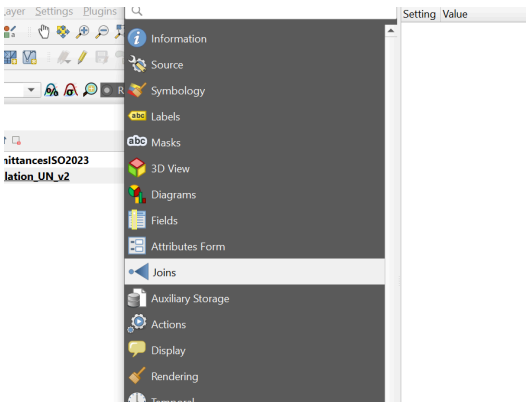
The next part is also a join, but a little tricky.

The order of things matter.

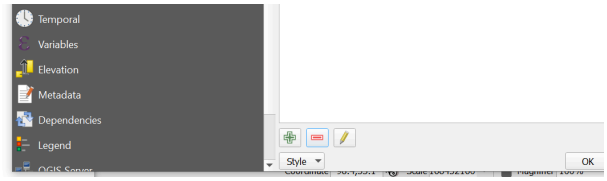
1) Make sure your shapefile -- the population map -- is the active layer. Click once with your cursor to be sure. Notice here how the file name is now highlighted? This is how you know it is the active layer.



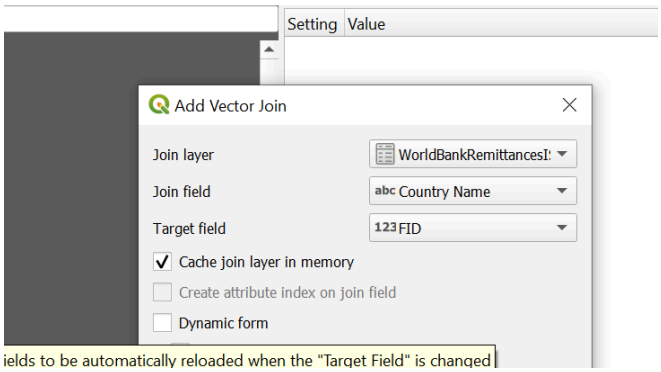
2) Now double click on the file name, and you should again get the properties popup window. Like this:



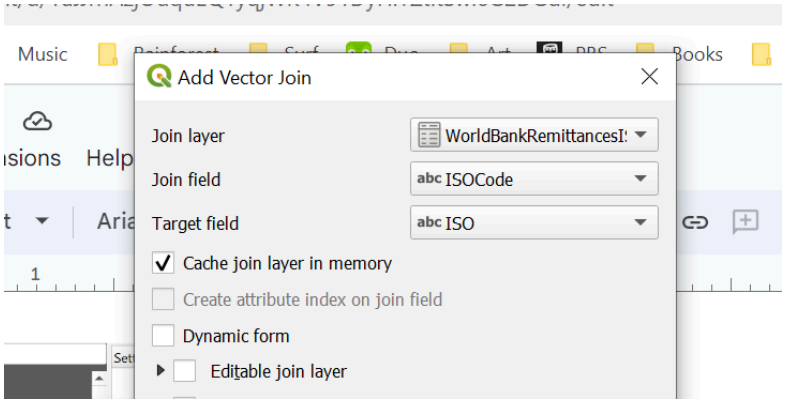
3) Now click Joins …we are about to do a join between the map and the data … and click the green plus sign you see lower down.



4) You should get a popup box. It's going to make guesses on how you want to do this join. But it's probably guessing wrong, and we need to fix this.

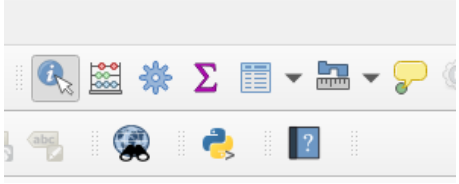


5) We want the join to be: Joining with the remittances data, using the ISO codes.



6) Now hit OK … and in the next window Apply and OK. Just to be sure.

7) Now open the attribute table. You could right click on the file name. Or hit the icon along the top of the menu bar, on the right side, that looks like a mini spreadsheet.

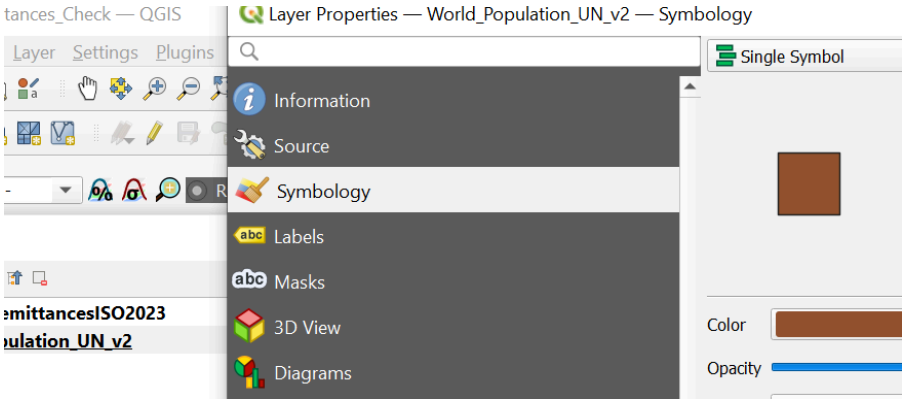


8) Notice now when you scroll through the data you now have connected country-by-country remittances with the map.

ty	MedianAge	GrowthRate	LifeExpect	InfantDeat	nittancesISO2023_	nittancesISO2023_	kRemittancesISO2	kR
000...	36.9044000000...	-0.368	76.9893000000...	0.26	Albania	ALB	1586574271	
000...	27.6050000000...	1.66000000000...	74.4527999999...	18.7330000000...	Algeria	DZA	2599027039	
999...	26.8625000000...	-2.4470000000...	72.4432000000...	0.006	NULL	NULL	NULL	
999...	42.1032000000...	1.75000000000...	79.0233999999...	0.003	Andorra	AND	NULL	
000...	16.1488000000...	3.16800000000...	62.2612000000...	64.8550000000...	Angola	AGO	17972037	
000...	36.1952000000...	1.15500000000...	76.9020000000...	0.001	NULL	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	

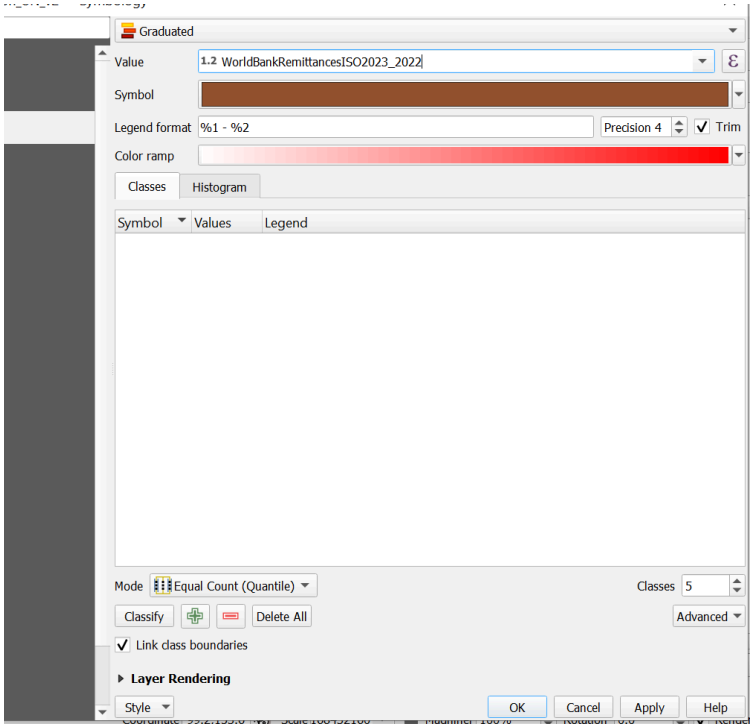
9) Now let's play around mapping remittances.

Go back to the map name … double click the name and get the popup properties window .. and go to symbology.

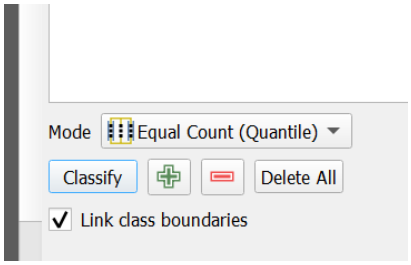


Choose the field name to do the mapping -- let's pick the most recent year of data.

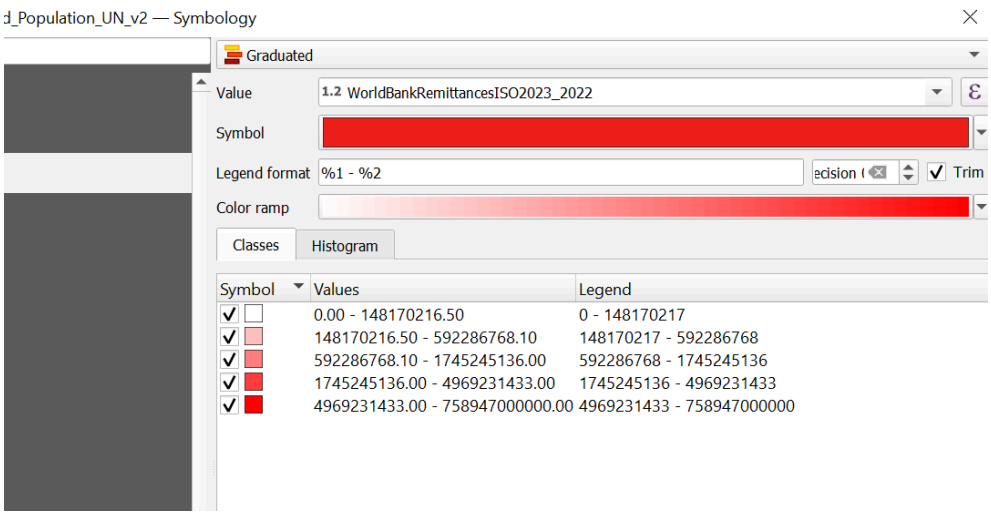
And equal count to classify the data range into five buckets.



Now be sure to hit the classify button lower down in the popup box. This does the math for us.

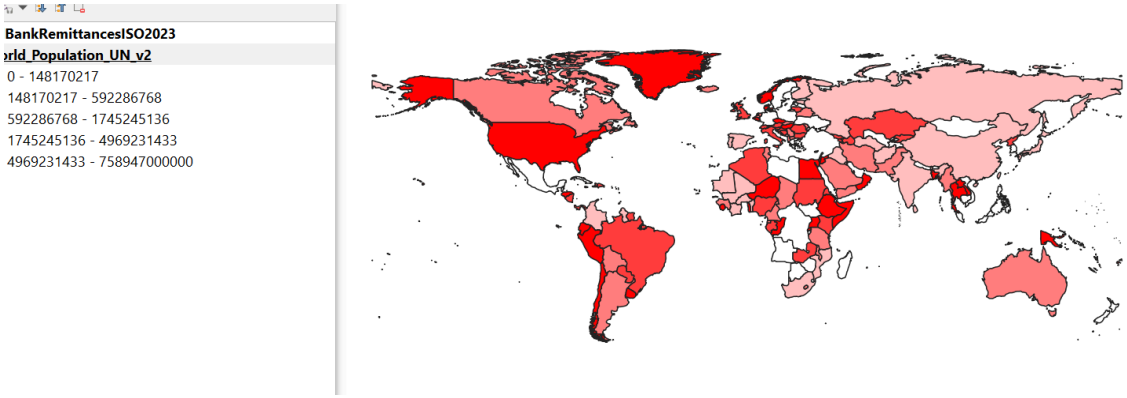


Your popup box should now look like this:



Hit apply and OK.

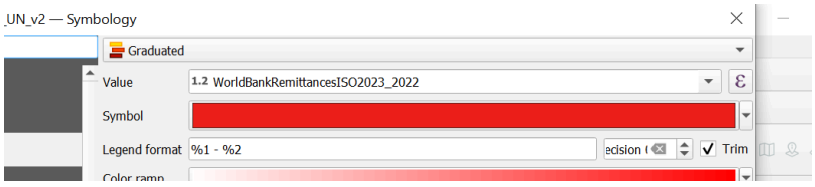
Your map may look something like this:



And you can now begin to hunt for interesting patterns. And also this may help you spot missing data you may wish to inquire more about.

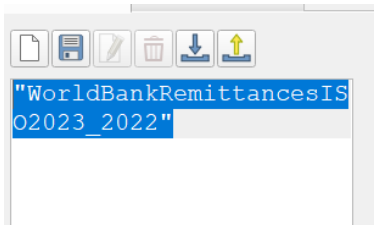
Calculating rates

You can do this.
It is tricky.

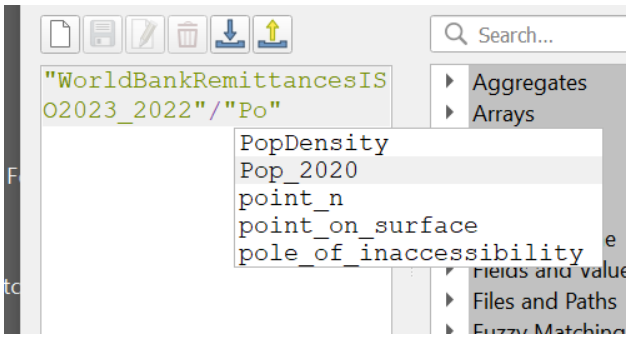


Look for the icon in the upper right that looks like the Greek letter Epsilon. (Don't ask)

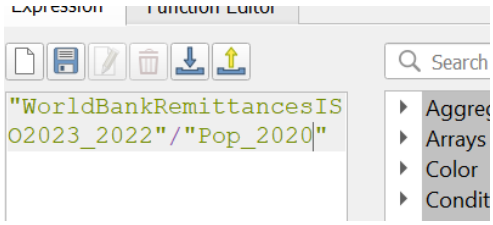
You will see the menu now ready to calculate a rate. To make things easy, copy the field name "WorldBankRemittancesISO2023_2022"



Type a dividing sign, since we're doing math. A rate.
Paste in the field name -- this is to ensure you get the right double quotes.



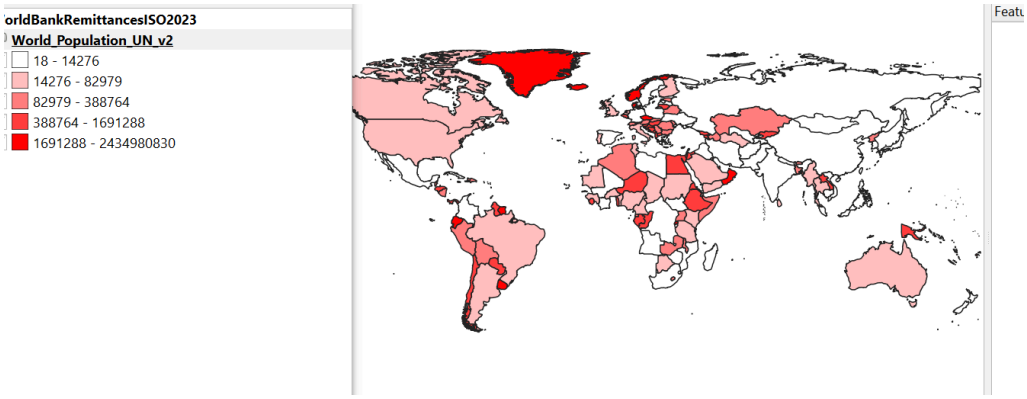
And divide remittances by recent population.



Hit OK …

Hit Classify … Apply … OK

And now -- aside from places with missing values that raise questions -- you see some interesting patterns



The Horn of Africa, Southeast Asia and parts of Latin America remain of interest by either standard.