

GIS for Economists 5

Giorgio Chiovelli Sebastian Hohmann Tanner Regan

26/05/2023

Table of contents

Overview

The plan for the session

Introduction to QGIS

QGIS Installation

Working with QGIS

Georeferencing and Digitizing an Old Map

Georeferencing

Digitizing

Geocoding

Introduction and Geopandas approach

Data Visualization

Making a Map

Overview

The plan for today

Introduction to QGIS

- QGIS Installation
- Working with QGIS

Old maps

- Georeferencing
- Digitizing

Geocoding

- Using **GeoPandas**

Making a map

- Using **Matplotlib**

Introduction to QGIS

QGIS – What is it good for?

QGIS is a programme for analysing, modifying, and creating geo-spatial data.

Some of the things it can do:

- Visualise and troubleshoot your data
- Make maps
- Georeference maps and images
- Digitize vector data from scratch
- Point and click all the functions from geopandas
- ...

Introduction to QGIS

QGIS – Installing QGIS

Go here:

<https://qgis.org/en/site/forusers/download.html>

- Go to the tab for your operating system
- Choose the appropriate standalone installer for your system
- We recommend the “Latest release (richest on features):”
- Execute the installer, choose default options, don’t install the datasets

Mac users may face an “unidentified developer” problem. See here for solutions:

[https://www.macworld.co.uk/how-to/
mac-app-unidentified-developer-3669596/](https://www.macworld.co.uk/how-to/mac-app-unidentified-developer-3669596/)

Mac users may also encounter an issue with GDAL geoprocessing tools (we will need this later). See here for a solution:

[https://gis.stackexchange.com/questions/276853/
gdal-scripts-not-found-in-qgis-3-on-osx](https://gis.stackexchange.com/questions/276853/gdal-scripts-not-found-in-qgis-3-on-osx)

Introduction to QGIS

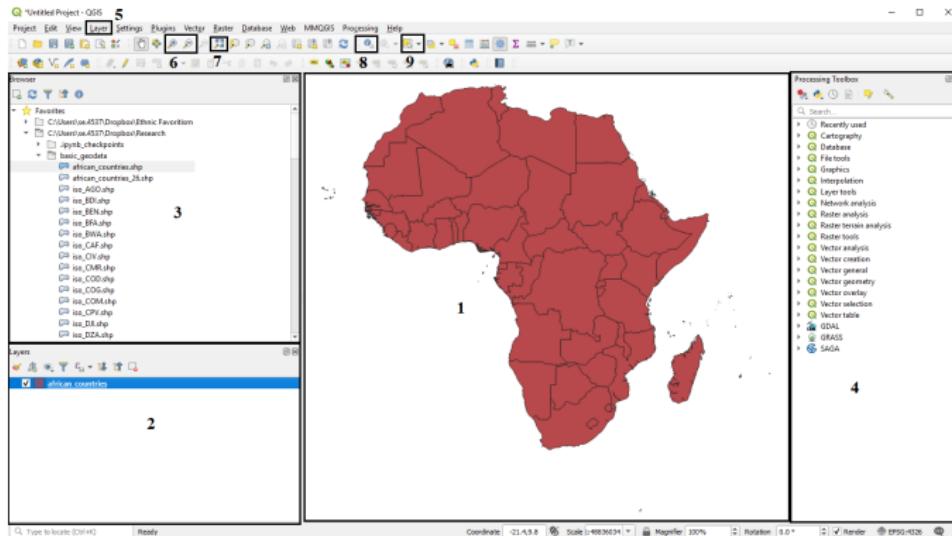
QGIS – Working with QGIS

Installing QGIS will put a bunch of programmes on your computer.

- We will introduce you to QGIS – the software which allows you to work with geospatial data.
- QGIS has a built in “Browser” – a type of file viewer.
- As you see quickly when using QGIS to do geoprocessing, doing anything in QGIS can produce a lot of intermediate files at every step.
- Every shapefile, for example, comes with an *.shp* file (which stores the feature geometry itself), a *.dbf* file (which stores the attribute table, see below), a *.prj* file (which stores the projection), and often with an *.shx* file (which stores a positional index of the feature geometry to allow seeking forwards and backwards quickly).
- The Browser can make managing this soup of files easier as you always just see one file.
- In practice, it is better to just automate as much as you can in python and access the Browser from within QGIS if you absolutely have to.

Introduction to QGIS

QGIS – What the buttons mean



- 1 "Canvas": data is visualized here
 - 2 "Layers": which datasets are loaded?
 - 3 "Browser": quickly load datasets from disk locations, delete data
 - 4 Processing toolbox
- Note: showing 2-4 is a good default configuration, but other windows can be shown (see "View" → "Panels")
- 5 Among other things, add layers (data) to canvas
 - 6 Zoom in and out
 - 7 Zoom to full extent (very useful)
 - 8 Get information on geographic features
 - 9 Select elements

Introduction to QGIS

QGIS – Adding data

Download some data

- from the google [drive](#) or from
<http://www.naturalearthdata.com/downloads/10m-cultural-vectors/>
- download “Admin 0 – Countries”, save to some directory and unzip

Add directory to “Favorites”

- facilitates loading data
- navigate to the folder containing the file you downloaded and add it to favorites
- sometimes may have to refresh () to see new data

Add the data

- In the folder you just added, right-click, double-click (or do the mac thing) to “Add Layer to Project” or single click and click on “add layer” ()
- Can also go more complicated: “Layer” → “Add Layer” → “Add Vector Layer” → “Browser” → navigate to your folder and add the data.

Introduction to QGIS

QGIS – Inspecting the data

You should see a political map of the world in the main window and an entry in “Layers” listing the dataset.

We have loaded a file with **polygon features**. Features come with “attributes”.

Attribute table

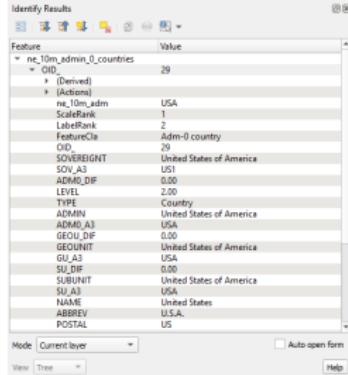
- right click on *ne_10m_admin_0_countries* in Layers → Open Attribute Table
- a table opens, listing a bunch of variables
- each row in the table corresponds to one polygon feature on the map
- pick a country and click on the small grey square at the start of its row with the number to select the row and close the attribute table
- the country you picked is highlighted
- click on Zoom to Selection (
- click on  to zoom back out
- click on  to de-select all features
- in select mode () , you can also control-click on a feature on the canvas to un-select it
- play around with zooming, panning, selecting, until you are comfortable

Introduction to QGIS

QGIS – Inspecting the data

We can also use **Identify Features** to view the information in the attribute table for one or more features

- click on  and click on the US
- you should see the window on the right
- this lists all the variables in the attribute table for this particular feature
- we can select more than one feature this way (hold down the mouse and drag it over several features)
- then expand and collapse individual feature attribute lists to find the one(s) we are interested in



Feature	Value
nc_10m_admin_0_countries	
OID	29
# (Derived)	
nc_10m_adm	USA
ScaleRank	1
LabelRank	2
Featurecla	Adm-0 country
OID	29
SOVEREIGNT	United States of America
SOV_A3	USA
ADM0_DIF	0.00
LEVEL	2.00
TYPE	Country
ADMIN	United States of America
ADM0_A3	USA
GEOID_0	00000
GEOUNIT	United States of America
GU_A3	USA
SU_DIF	0.00
SUBUNIT	United States of America
SU_A3	USA
NAME	United States
ABBRREV	U.S.
POSTAL	US

Introduction to QGIS

QGIS – Inspecting the data

Extent and coordinate system

- back in Layers, right click on the layer and “Properties” at the bottom
- select the “Information” tab and check “Information from provider”
- select the “Source” tab

The screenshot shows two stacked QGIS Layer Properties dialogs for a layer named "ne_10m_admin_0_countries".

Information Tab:

- Name:** ne_10m_admin_0_countries
- Path:** C:\Users\asus.457\Dropbox\Pulkit\GeogIS\QGIS\lecture_1\095\database\10m\ne_10m_admin_0_countries\ne_10m_admin_0_countries.shp
- Storage:** ESRI Shapefile
- Comment:**
- Encoding:** UTF-8
- Geometry:** Polygon (MultiPolygon)
- CRS:** EPSG:4326 - WGS 84 - Geographic
- Extent:** 79.999783491996149, -89.9998233894376527 ; 180.0000000000001137, 83.6338109340297393
- Unit:** degrees
- Feature count:** 253

Source Tab:

- Layer name:** ne_10m_admin_0_countries displayed as ne_10m_admin_0_countries
- Data source encoding:** UTF-8
- Geometry and Coordinate Reference System:** Set source coordinate reference system: EPSG:4326 - WGS 84
- Buttons:** Create Spatial Index, Update Extents

Introduction to QGIS

QGIS – Changing color, outline width, labelling features

Change color of countries and outline width

- Back under Layers, right click the *ne_10m_admin_0_countries* layer → Properties → Symbology
- Change Fill color, outline (“Stroke”) color, and outline width
- Click “OK” and “Apply” and see how the canvas display changes

Label features

- Back to Layer Properties window, select the “Labels” tab
- From the drop-down menu at the top, choose “Single Labels”
- From the “Label with” drop-down menu, select “ADMIN” (this is one of the variables from the attribute table)
- Click “OK” and “Apply” and see how the canvas display changes

Hide a dataset

- in Layers, uncheck *ne_10m_admin_0_countries*

Introduction to QGIS

QGIS – Adding and inspecting some raster data

Download some data

- from the google [drive](#)
- download the data, save to some directory and unzip

Add the data

- exactly as with the feature data (the data are under *suit/hdr.adf*)
- ignore the warning message (we will deal with this in a later lecture)

Inspect the data

- zoom in closely so you can make out individual cells
- use  to look up individual pixel values
- look in the source tab of the Properties menu (note the absence of a CRS)

Change the color scheme

- Layer → Properties → Symbology
- Band Rendering → Render Type → Singleband pseudocolor → choose Linear Interpolation and your favorite Color Ramp → OK + Apply

Old Maps and QGIS

How to Georeference in QGIS?

- Georeferencing means to associate something with locations in physical space.
- Basically you assign latitude and longitude information to the map

Let us try to Georeference and Digitize

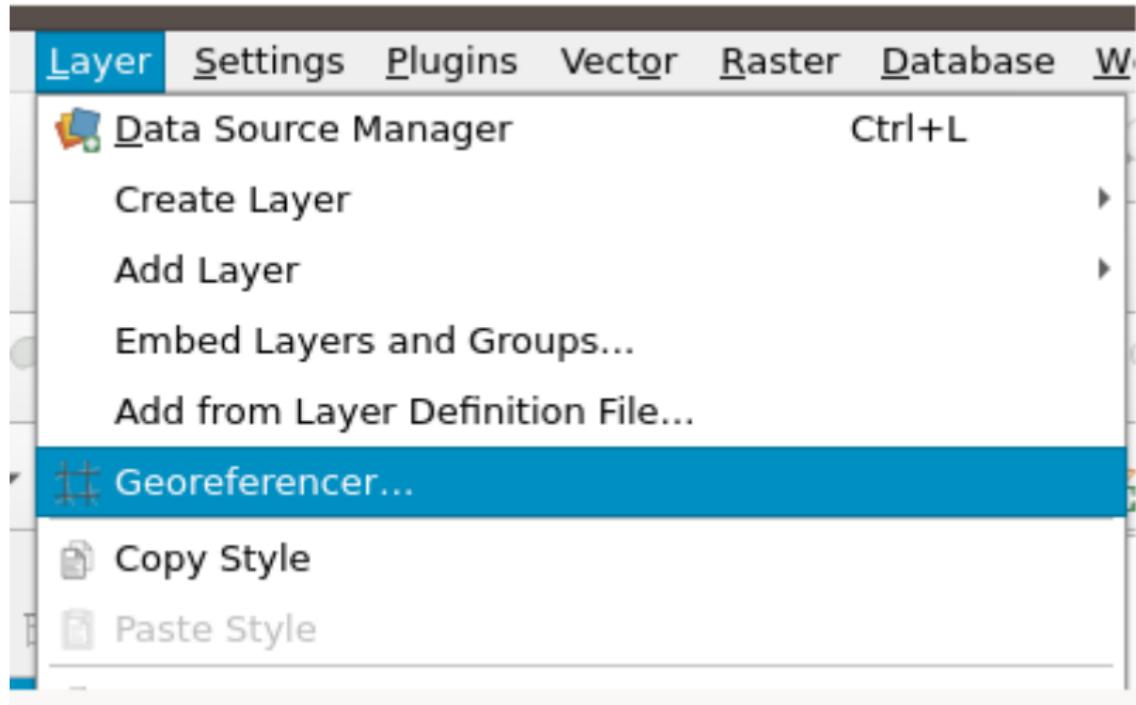
- **Niassa Company - Mozambique**

https://drive.google.com/file/d/1kxQWAbLDsG_Hpu-EzGP6bmWIRR1j7zix/view?usp=sharing

Georeference in QGIS

Georeference is going to be your best friend

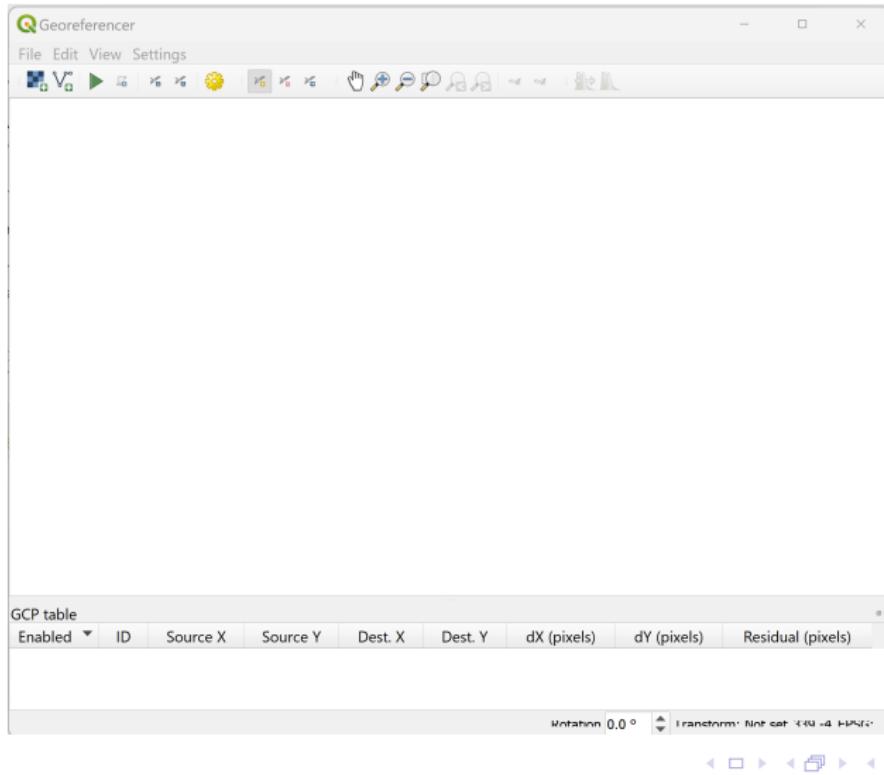
- Click on Layer → Click on Georeferencer.



Georeferencer

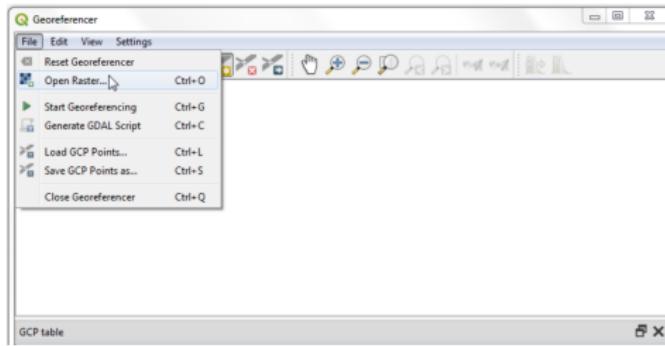
How the Georeferencer looks like?

- The plugin window is divided into 2 sections.



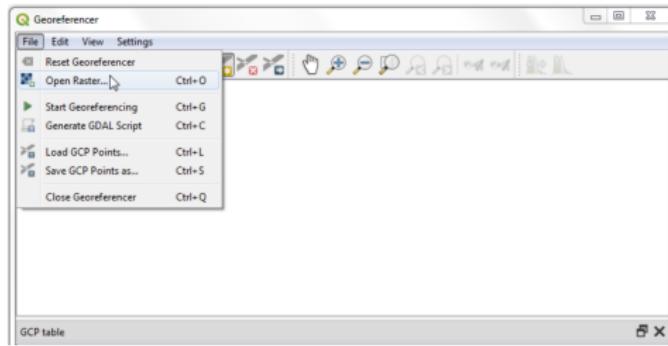
Georeference

- To open our image, go to File → Open Raster. Browse to the folder storing the scanned map and click Open.



Georeference

- To open our image, go to File → Open Raster. Browse to the folder storing the scanned map and click Open.



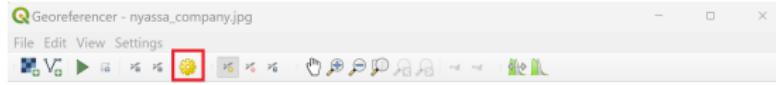
- you will be asked to choose the coordinate reference system (CRS). Click Cancel as we will assign this later.

Georeference

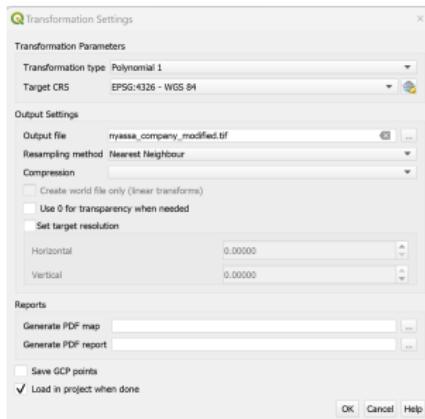
- We need to assign coordinates to some points on this map.
- Almost all the maps of the project will have coordinate grid with markings.
- These are Latitude and Longitude grid lines.
- you will be asked to choose the coordinate reference system (CRS). Click Cancel as we will assign this later.

Georeference

- We start adding Ground Control Points (GCP)
- We need to define the Transformation Settings. Go to Settings → Transformation settings.



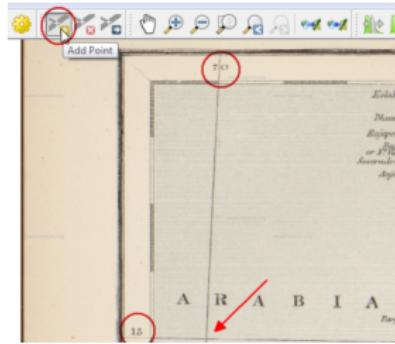
Georeference



- Choose the Transformation type as Polynomial 1 (we will see later the implications of this choice).
- Name your output raster as the **name of the map**
- Choose LZW as the Compression.
- Make sure the Load in QGIS when done option is checked.
- Click OK.

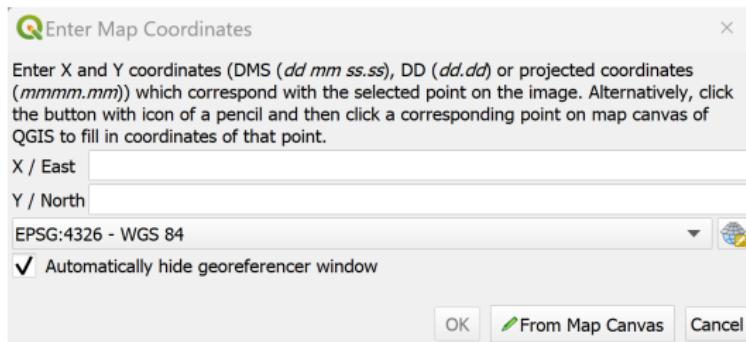
Georeference

- Now we can start adding the Ground Control Points (GCP).
- The intersections of the grid lines will serve as the ground-truth in our case.
- As the grid lines are labeled, we can determine the X and Y coordinates of the points using them.
- Click Add Point.



Georeference

- In the pop-up window, enter the coordinates. Remember that X=longitude and Y=latitude.



- Latitude may be positive (above the Equator) or negative (below)
- Longitude may be positive (right of Greenwich) or negative (left)

Georeference

Latitude and Longitude Africa



Negative Lon



Positive Lat

Positive Lon



Negative Lat

Georeference

- You will notice the GCP table now has a row with details of your first GCP.

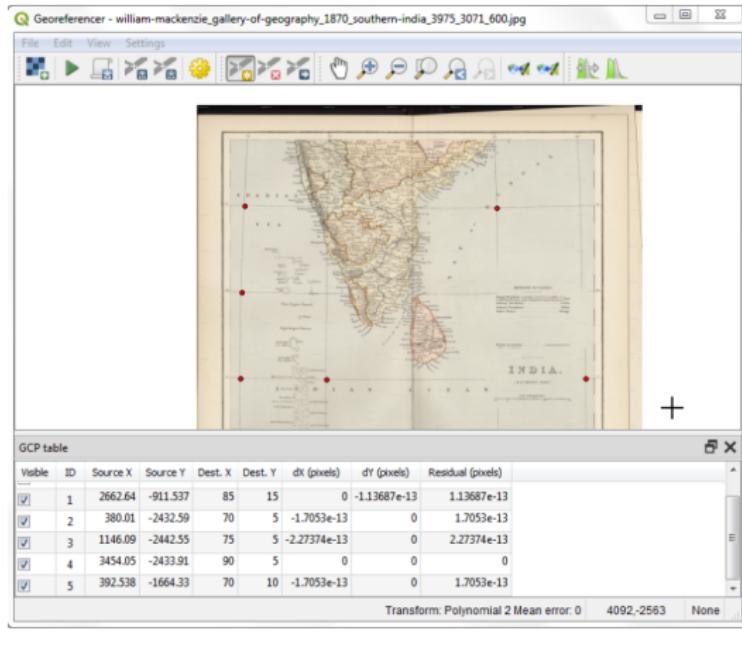
GCP table								
Enabled ▾	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
<input checked="" type="checkbox"/>	0	927.582859	-557.577802	35.000000	-11.000000	n/a	n/a	n/a

- Add at least more GCPs covering the entire image.
- The more points you have, the more accurate your image is registered to the target coordinates.

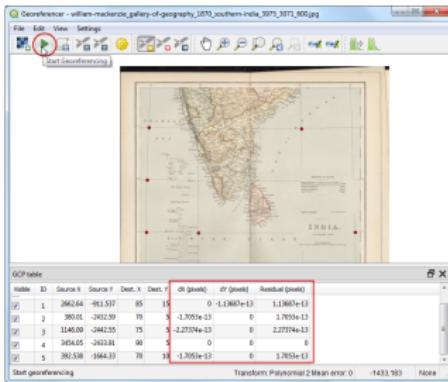
Georeference

- Different Transformation Types:

- Linear: most simple. Rotation and rescaling of the image.
- The higher the polynomial, the larger the number of GPCs you need
- For example, the Polynomial 2 transform requires at least 6 GCPs.

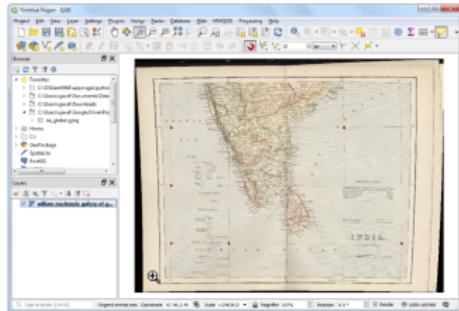


Georeference



- As you keep adding points, you will notice that the GCPs now have a non-zero dX, dY and Residual error values.
- If a particular GCP has unusually high error values, that usually means a human-error in entering the coordinate values.
- If this happens, delete that GCP and capture it again.
- Once you are satisfied with the GCPs, go to File → Start georeferencing.
- This will start the process of warping the image using the GCPs and creating the target raster.

Georeference

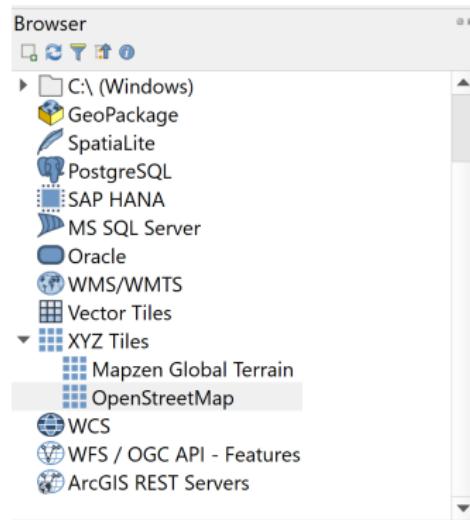


- Once the process finishes, you will see the georeferenced layer loaded in QGIS.
- The georeferencing is now complete.

How to Georeference when Latitude and Longitude are not available?

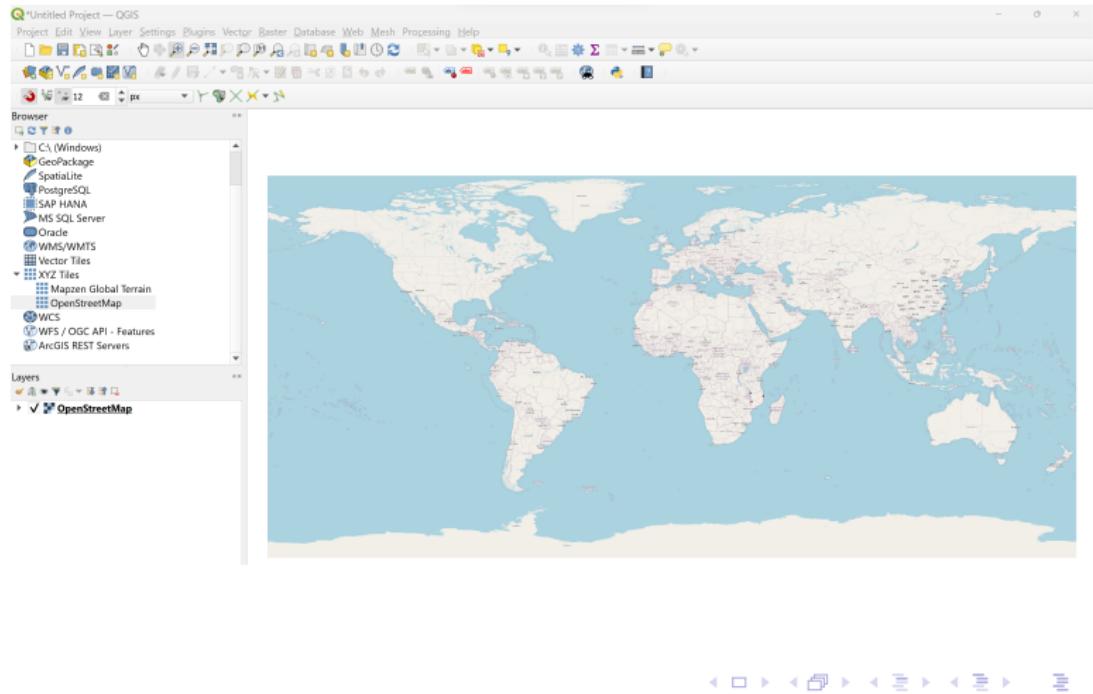
You can use a Basemap!

- In the left-hand browser menu, locate XYZ Tiles → OpenStreetMap



How to Georeference when Latitude and Longitude are not available?

OpenStreet Map



How to Georeference when Latitude and Longitude are not available?

Georeferencing from the Map Canvas

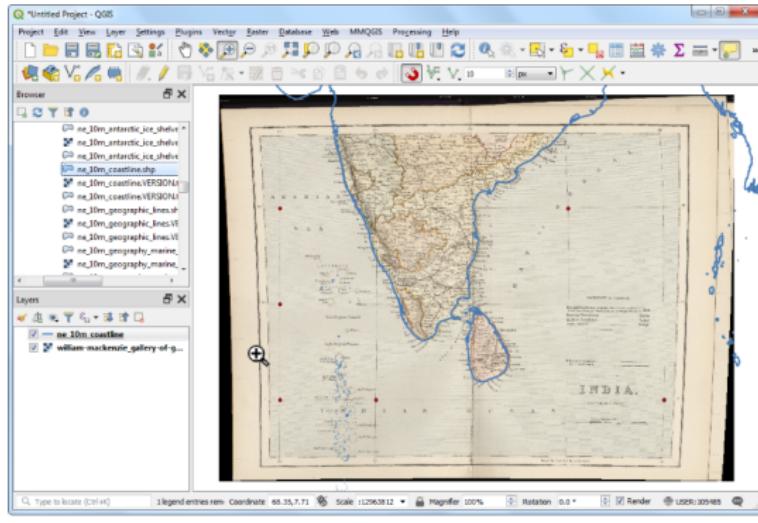
- Repeat the steps listed above to load the map into the Georeference
- Click add a Point
- Choose a spot on the historical map where you can find a corresponding modern location.
 - Example: city/town location!



- Find the location on the modern map.
- Click OK on the Enter Map Coordinates dialog box. The GDAL Georeferencer's map will not automatically snap to the correct location.

Georeference's Accuracy

- Last check: How do we check if our georeferencing is accurate?
 - You can load the boundary shapefile from a trusted source like the Natural Earth dataset and compare them.



Digitization

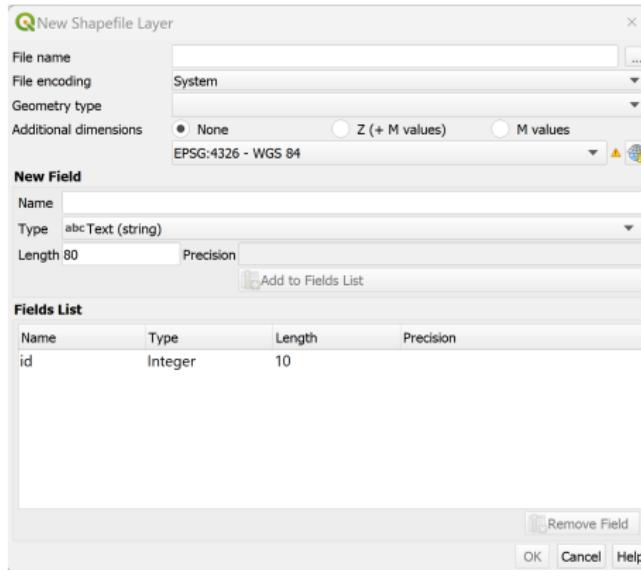
Now we can start Digitizing the map

- Digitizing means creating a shapefile
- It allows us to create new data!

Digitizing

Create a New Shapefile

- Let us create an empty shapefile
 - Select Layer → Create Layer → New Shapefile Layer



Digitizing

Create a New Shapefile II

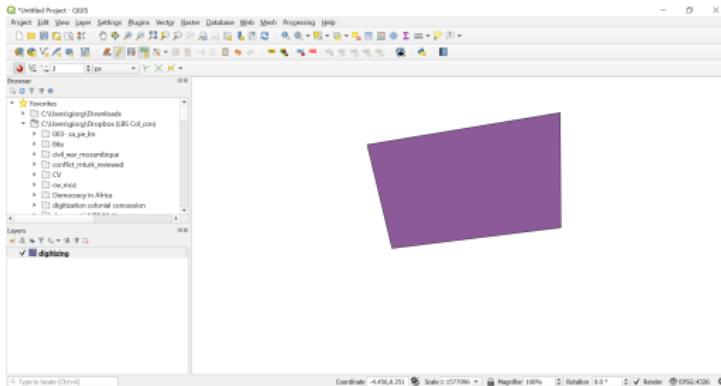
- Select the Geometry type: Points, Line, Polygons
- Structure your Attribute Table by creating the Fields you need to be in the shapefile

Digitizing Toolbar

- Right-click on the upper bars above the canvas and check:
 - Digitizing Toolbar



- To enter digitization mode, click on the “Toggle Editing”
- To start drawing, click on the “Add Polygon Feature”

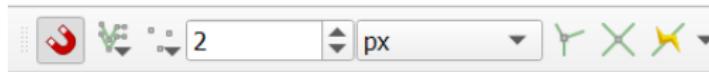


- To finish your sketch, right-click and insert the Field Values

Snapping Toolbar

Use existing shapefiles to improve our digitization.

- We can snap existing line/polylines to improve the precision of our digitization (e.g. coast, country borders, etc)
- Right-click on the upper bars above the canvas and check:
 - Snapping Toolbar → Enable Tracing



- To enter snapping mode, click on the “Enable Snapping”
- To trace an existing feature, click on the “Enable Tracing”
- **TRACE** tool to automatically capture the vertices, instead of manually clicking to snap.



Saving your Digitization

- Good practice is to save the digitized elements as you create them
 - Click on “Save Layer Edits”
- Once you finish the digitization, click again on “Toggle Editing”
- Now you are done! You can use your new data for your analysis!

Geocoding

Introduction and Geopandas approach

Suppose we have address data

country	city	post code	street name	number
UK	London	NW1 4SA	Sussex Place	26
UK	London	NW1 5PT	Marylebone Rd	117

- How do you go from this to latitude/longitude?
- Answer: **Geocoding**. Like a phone book for co-ordinates. Many ways. One way, google maps: Goole the address → find the place you are looking for → right-click → “What’s here?”
- Not very scalable. Want to automate → API providers: Google maps, Geonames, Open Street Map, ...
- We can geocode in GeoPandas using `geopandas.tools.geocode`. → open `geocoding.ipynb` in jupyter.

Making a Map

Introduction and Matplotlib approach

- Visualizations are an important part of good papers / policy arguments.
- Getting them right can be time-consuming.
- We have already seen some “on-the-fly” plots as part of data exploration.
- Now we want to make a high-quality plot.
- We will make use of the **Matplotlib** library – the go-to tool for plotting in python – together with some additional tools for geospatial visualization.
→ open `mapmaking.ipynb`