

GIS for Economists 4

Giorgio Chiovelli Sebastian Hohmann Tanner Regan

16/06/2021

Table of contents

Overview

- The plan for the session

Introduction to QGIS

- QGIS Installation

- Working with QGIS

Geocoding

- Introduction and Geopandas approach

- Optional: Geocoding example via MMQGIS

Data Visualization

- Making a Map

- Heatmaps

Georeferencing and Digitizing an Old Map

- Georeferencing

- Digitizing

Overview

The plan for today

Introduction to QGIS

- QGIS Installation
- Working with QGIS

Geocoding

- Using **GeoPandas**
- Optional: using **QGIS** via MMQGIS Geocoder

Making a map

- Using **Matplotlib**
- Optional: using **QGIS**

Creating a heatmap

- Passenger numbers in London tube stations

Old maps

- Georeferencing
- Digitizing

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/>

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>

Introduction to QGIS

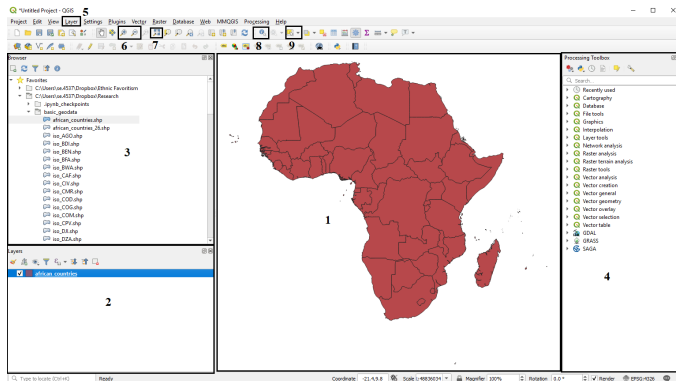
QGIS – Working with QGIS

Intalling 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

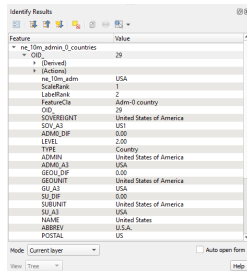
- right click on *ne_10m_admin_0_countries* in Layers → Open Attribute Table
- a table opens, listing a bunch of variables
- each row in the table corresponds to one polygon feature on the map
- pick a country and click on the small grey square at the start of its row with the number to select the row 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

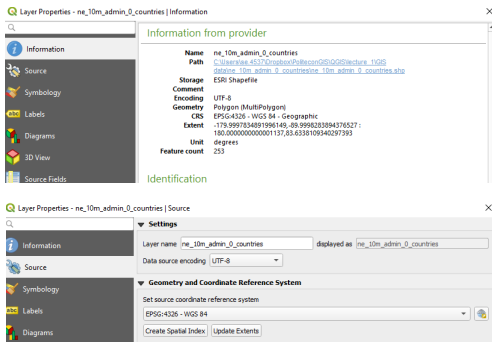


Introduction to QGIS

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



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 → Symbolology
- 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 or from
`https://nelson.wisc.edu/sage/data-and-models/atlas/data.php?incdataset=Suitability%20for%20Agriculture`
- 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

Introduction to QGIS

QGIS – Creating latitude/longitude data

So far we have added data downloaded from the web. There is one type of spatial data that we can easily create ourselves: **Point features**.

Create point features in text editor

- open a text editor
- in the first line, type: *point_name, latitude, longitude*
- in the second line, type: *some name, -12.3, 117.2*
- in the third line, type: *some other name, 65.2, -56.3*
- keep adding as many points as you like, until you're bored.
- make sure latitude $\in [-90, 90]$, longitude $\in [-180, 180]$
- save the data under *my_points.csv* in the directory where you saved the files you downloaded


Introduction to QGIS

QGIS – Adding and inspecting latitude/longitude data

Add and display data

- Layer → Add Layer → Add Delimited Text Layer → Under File name, browse to the point layer; Pick a layer name
- File Format → CSV
- Geometry Definition → Point coordinates
 - X field: longitude
 - Y field: latitude
 - Geometry CRS: Project CRS: EPSG: 4326 – WGS 84
- Click “Add”
- dots appear on the map

Inspect the data

- look at the attribute table
- use Identify Features () to get information on individual points

Geocoding

Introduction

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: Google 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.
- Alternatively (we won’t cover this here, but look at the slides) is to do it in QGIS using the MMQGIS plugin. (if don’t have MMQGIS plugin (in the top menu bar), go to Plugins → Manage and Install Plugins → Search for MMQGIS → Install)

Geocoding

Optional: Geocoding example 1 via MMQGIS

Geocoding simple example: cities

- locate the file *geocode_example_mmqgis.csv* (part of *GIS data* for this session on the google drive)
- Click on MMQGIS → Geocode → Geocode CSV with Google / OpenStreetMap
- Browse to the cities file
- Select fields for Address, City, State, and Country (leave blank what you don't have)
- Use the OSM geocoder (you can sign up for a google API key for bigger applications and possibly better results)
- specify shapefile and list of addresses not found and run it
- it may take a moment depending on the size of the .csv (there are query limits)
- inspect the .csv with the addresses not found (it should have found all)

Geocoding

Optional: Geocoding exercise

Geocoding exercise: Costa Coffeeshops

- Locate *costa_locations.csv* (scraped from costa coffee website), inspect the data (look at the “city” column, may be not 100% clean)
- repeat the above steps (make call which variables should go where)
- are you able to geocode all addresses?

Making a Map

Optional: Matplotlib

- 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.
- Alternative approach: use **QGIS** to make the map. Following slides (we will not cover) show you how if you are interested.

→ open `mapmaking.ipynb`

Making a Map

Optional: QGIS

- Load your data on the Canvas
- Rename the layer with the appropriate name you want to be shown in the Map Legend
- From the Project menu select New Print Layout
- Type a Title for your Print Layout

Making a Map

Optional: QGIS

- Inside the New Print Layout screenshot, use the Add Item to add:
 - Maps
 - Legend
- You have to drag the arrow to select the area where you would like each map element to appear
- Let us produce on Map of London and the UK and export it to pdf

How to visualize Spatial Pattern in your data

- Heat Map

Advantages

- Help you detecting spatial patterns: cluster, spatial correlations
- Appreciated by policy-makers and practitioners

Data Visualization Tools

Heat Map

Heat Map

- Represents the geographic density of features on a map
- Useful for layers with a large number of features

Example: Spatial pattern of crime in London

Data

- You can download data on street level crime from all UK
- <https://data.police.uk/data/>
- Monthly frequency (2016 to present)
- Let us focus on Central London

Example: Spatial pattern of crime in London

- Load the OpenStreetMap layer under XYZ Tiles
- Import the crime data CSV
 - use Add Layer Delimited Text
 - or use MMQGIS to automatically a shapefile out of the CSV
- Now let us play around with the Layer Styling Panel (shortcut F7)

Example: Spatial pattern of crime in London

- Select the option Heatmap from the menu
- We can now set different parameters altering the visualization
 - Colour
 - Opacity
 - Radius: assume the crime event will impact 500 meters from its gps location
 - Kernel shape: regulates the distance decay from the centroid

Example: Spatial pattern of crime in London

- We can now set the visualization to take be weighted by events importance
- Let us give more weights to Violence and Sexual Assault and Criminal Damages
- Open the Attribute Table and Select the Open Field Calculator
- Select the function CASE (Show Help will help you understanding the syntax)
- Type the following statement

```
CASE  
WHEN "Crime type" LIKE 'Violence%' THEN 10  
WHEN "Crime type" LIKE 'Criminal%' THEN 5  
ELSE 1  
END
```

- Let us now takes the newly created weights into account

Example: Spatial pattern of crime in London

- To further modify the heatmap (different kernels or dynamic radius) we need to save it as permanent raster
- Reproject the raster in projected coordinate system
- And now run the Heatmap (Kernel Density Estimation)

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

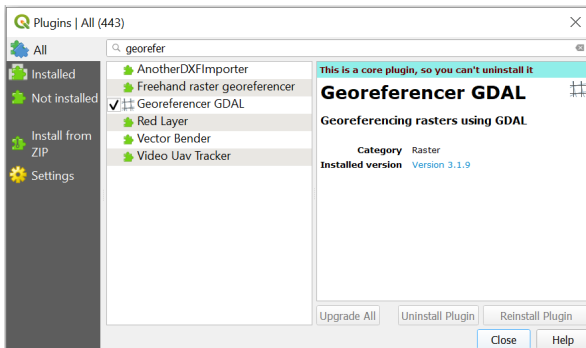
- **Niassa Company - Mozambique**

<https://drive.google.com/file/d/10cJjXKKr4FHg5w0uMLYBdQ7CrX3vzYnJ/view?usp=sharing>

Georeference and QGIS

GDAL is going to be your best friend

- Georeferencing in QGIS is done via the **Georeferencer GDAL** plugin.

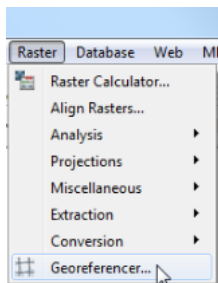


- Go to “Plugins” → “Manage and Install Plugins” and enable the Georeferencer GDAL plugin

How to Access GDAL

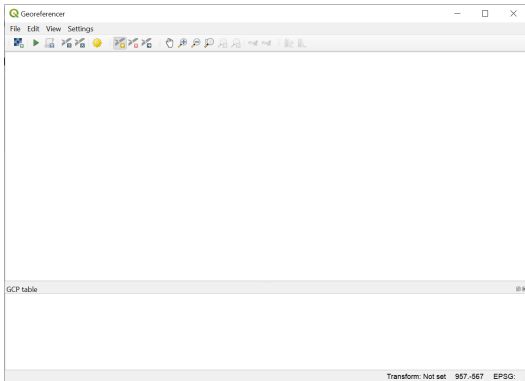
Access GDAL from Raster menu

- Click on Raster → Georeferencer to open the plugin.



How the Georeferencer looks like?

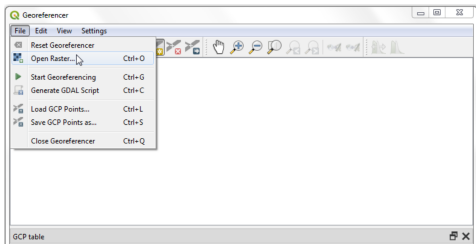
- The plugin window is divided into 2 sections.



- The top section where the image will be displayed
- The bottom section where a table showing your GCPs will appear.

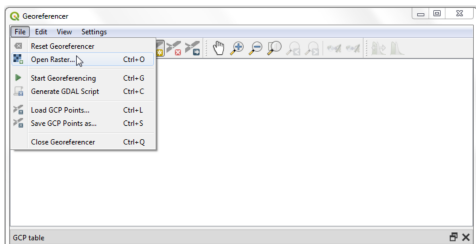
GDAL Georeferencer

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



GDAL Georeferencer

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

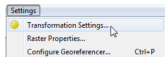


- you will be asked to choose the coordinate reference system (CRS). Click Cancel as we will assign this later.

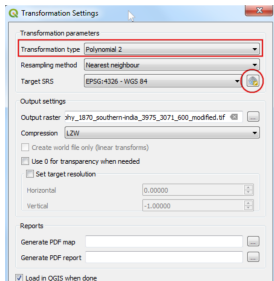
- We need to assign coordinates to some points on this map.
- Almost all the maps of the project will have coordinate grid with markings.
- These are Latitude and Longitude grid lines.
- you will be asked to choose the coordinate reference system (CRS). Click Cancel as we will assign this later.

GDAL Georeferencer

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



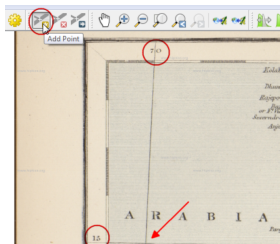
GDAL Georeferencer



- 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.

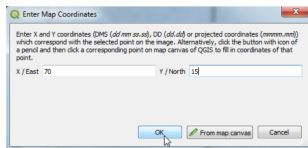
GDAL Georeferencer

- 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.



GDAL Georeferencer

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



- Remember we will be working with African countries
- Latitude may be positive (above the Equator) or negative (below)
- Longitude may be positive (right of Greenwich) or negative (left)

Latitude and Longitude Africa



Negative Lon



Positive Lat



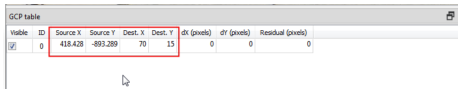
Positive Lon



Negative Lat

GDAL Georeferencer

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

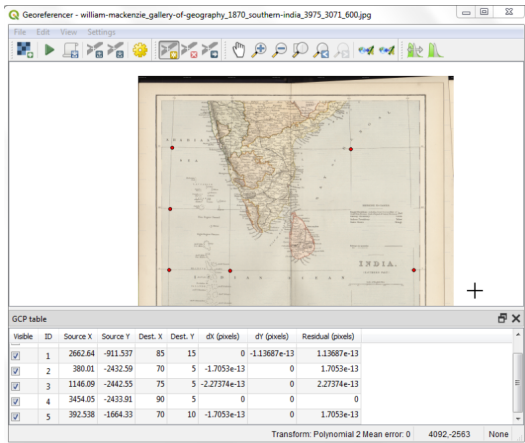


Visible	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
<input checked="" type="checkbox"/>	0	418.428	-893.289	70	15	0	0	0

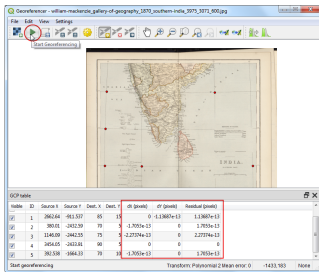
- Add at least more GCPs covering the entire image.
- The more points you have, the more accurate your image is registered to the target coordinates.

GDAL Georeferencer

- 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.

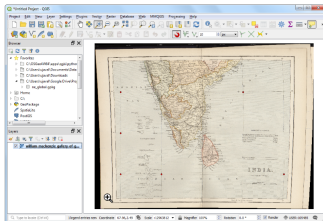


GDAL Georeferencer



- 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.

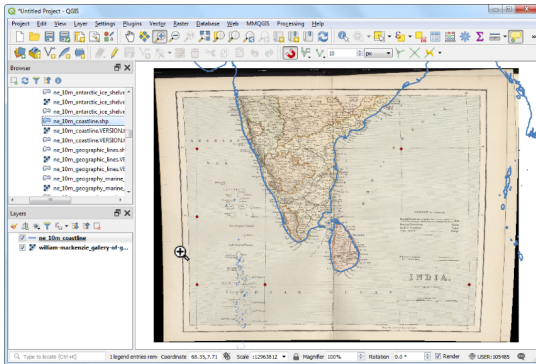
GDAL Georeferencer



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

GDAL Georeferencer

- 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 → New → New Shapefile Layer

The screenshot shows the 'New Shapefile Layer' dialog box in QGIS. The dialog has several sections:

- File name:** A text input field.
- File encoding:** A dropdown menu set to 'System'.
- Geometry type:** A dropdown menu set to 'Point'.
- Additional dimensions:** Three radio buttons: 'None' (selected), 'Z (+ M values)', and 'M values'.
- CRS:** A dropdown menu set to 'EPSG:4326 - WGS 84'.
- New Field:** A section for adding a new field with fields for 'Name', 'Type' (set to 'abc Text data'), 'Length' (set to 80), and 'Precision'.
- Fields List:** A table listing existing fields.

Name	Type	Length	Precision
id	Integer	10	

Buttons: 'Add to Fields List' and 'Remove Field'.

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

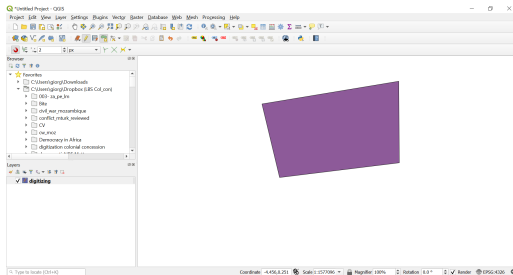
Digitizing Toolbar

- 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”

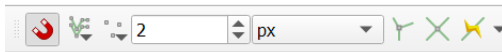


- To finish your sketch, right-click and insert the Field Values

Snapping Toolbar

Use existing shapefiles to improve our digitization.

- We can snap existing line/polygons 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!