SunoikisisDC Digital Approaches to Cultural Heritage, Spring 2024 – Session 4

Geographic Information Systems: From Maps to Analysis

Justin Colson (University of London)
Rebecca Seifried (University of Massachusetts Amherst)

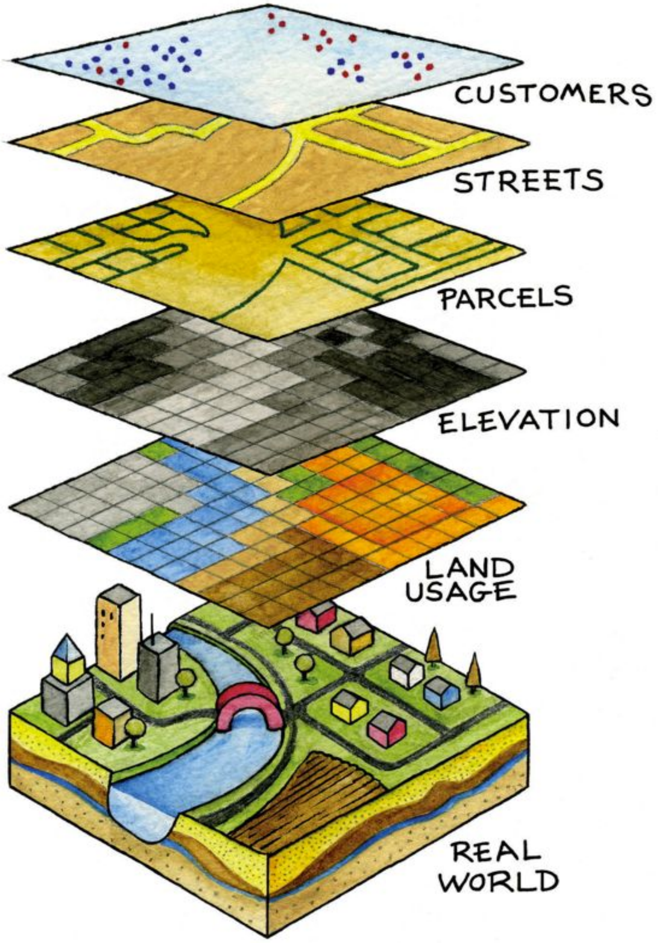
Outline

- How GIS data works
- Coordinate systems
- Vectors and rasters
- Symbology
- **Demo**: spatial queries
- **Demo**: georeferencing

How GIS data works

GIS as layers

- GIS data are essentially an abstraction of the real world
- Each *kind of thing* in the real world is usually represented by its own GIS data layer
- Layers at the top of the stack are rendered above lower layers

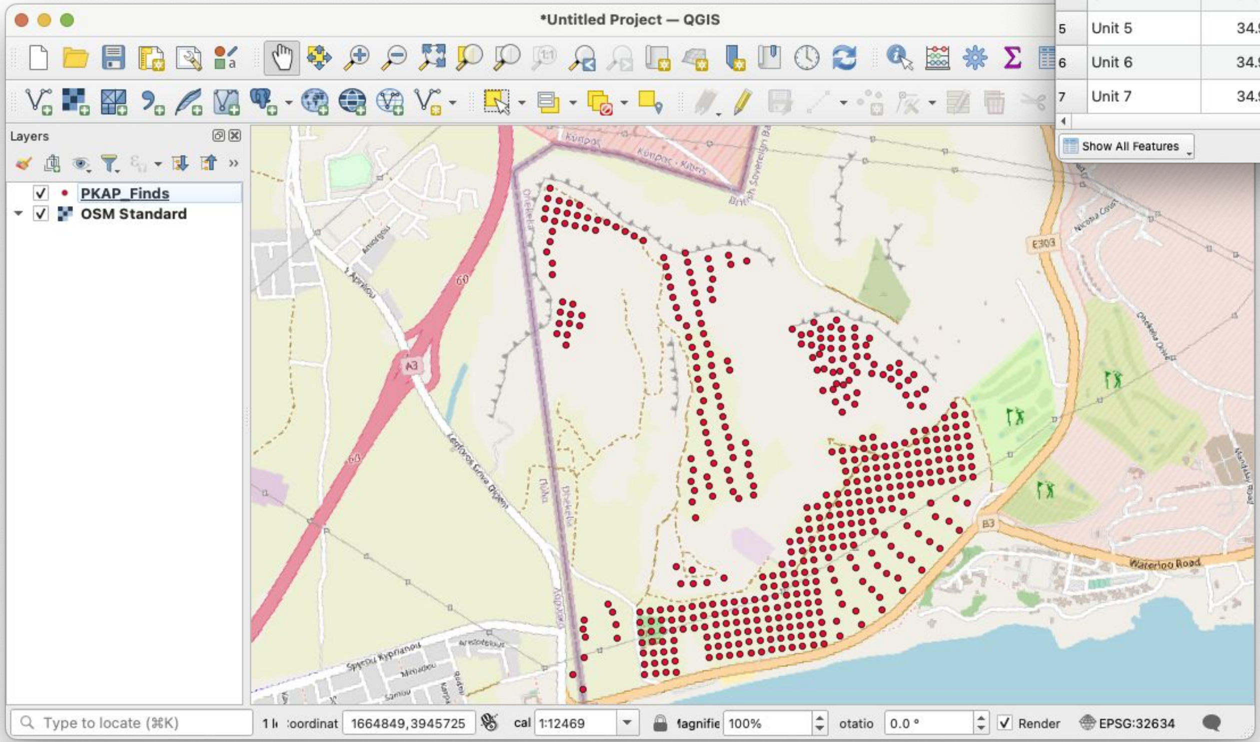


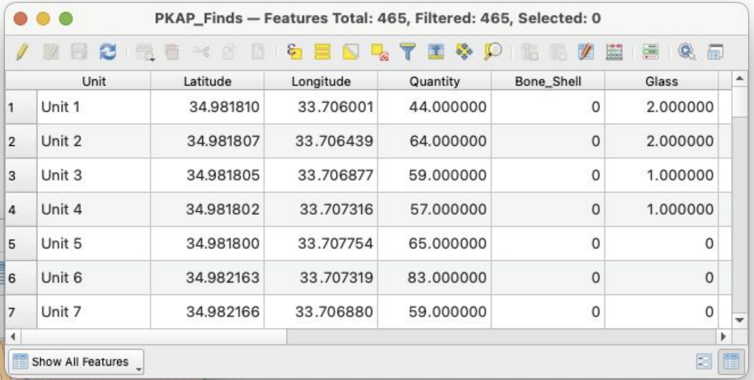
https://saylordotorg.github.io/text_essentials-of-geographic-information-systems/s05-03-geographic-information-systems.html

GIS as a database

Most GIS data is made up of:

- Location(s) in space, e.g. XY coordinate pair
- Some kind of additional data, e.g. a table containing attribute information (id, name, size, date) or a value (elevation, RGB color)

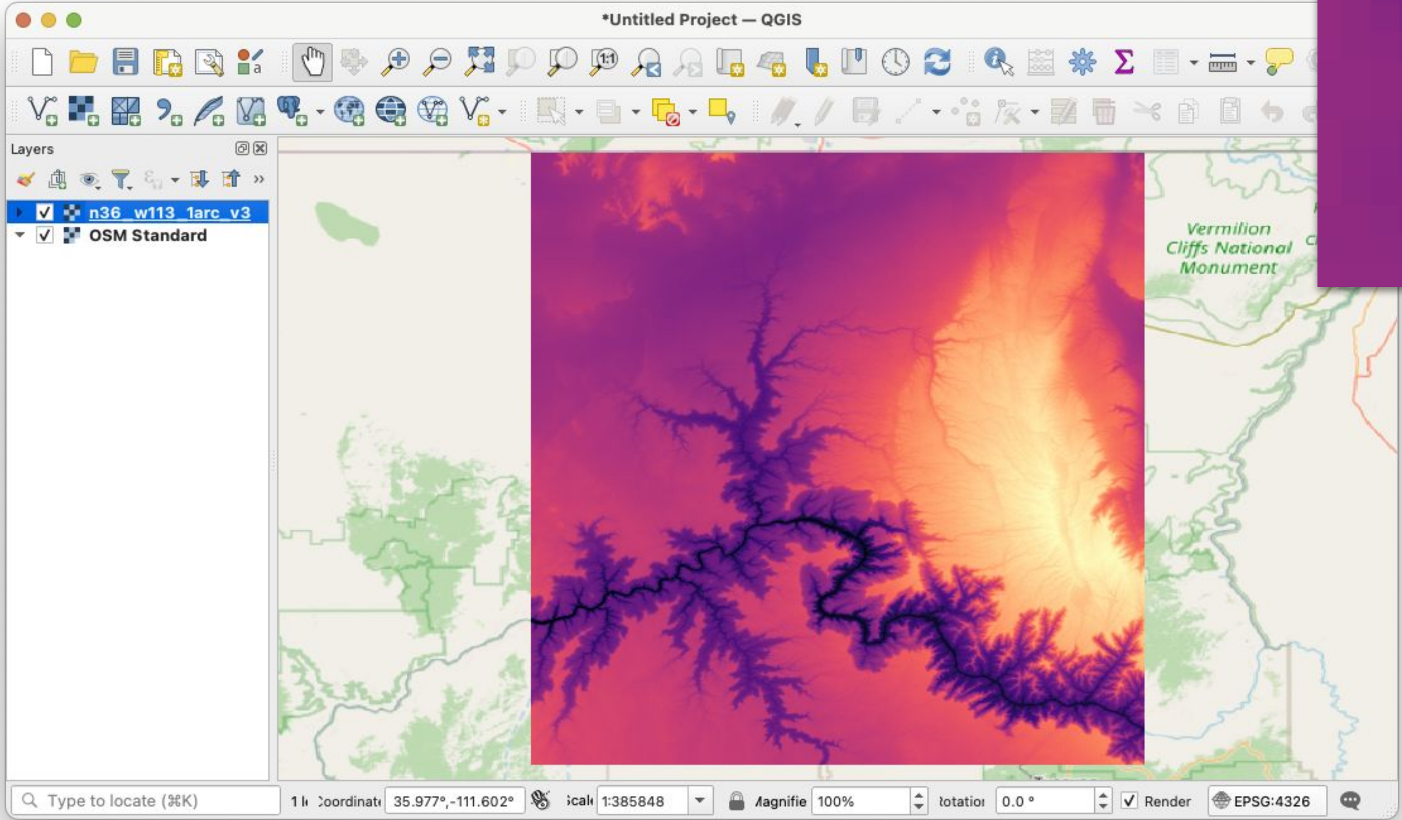




PKAP_Finds — Features Total: 465, Filtered: 465, Selected: 0

	Unit	Latitude	Longitude	Quantity	Bone_Shell	Glass
1	Unit 1	34.981810	33.706001	44.000000	0	2.000000
2	Unit 2	34.981807	33.706439	64.000000	0	2.000000
3	Unit 3	34.981805	33.706877	59.000000	0	1.000000
4	Unit 4	34.981802	33.707316	57.000000	0	1.000000
5	Unit 5	34.981800	33.707754	65.000000	0	0
6	Unit 6	34.982163	33.707319	83.000000	0	0
7	Unit 7	34.982166	33.706880	59.000000	0	0

Show All Features


*Untitled Project — QGIS
Layers
n36_w113_1arc_v3
OSM Standard
Vermilion
Cliffs National
Monument
Type to locate (⌘K)
35.977°,-111.602°
1:385848
100%
0.0 °
Render
EPSG:4326

Coordinate reference systems (CRS's)

Why CRS's are important

If I told you it was 40 degrees today, what would you imagine?

Every CRS has a datum

Datum = a system of reference points for measuring spatial positions on the Earth's surface

SATAN 2, MZ1900, 2, 20131102

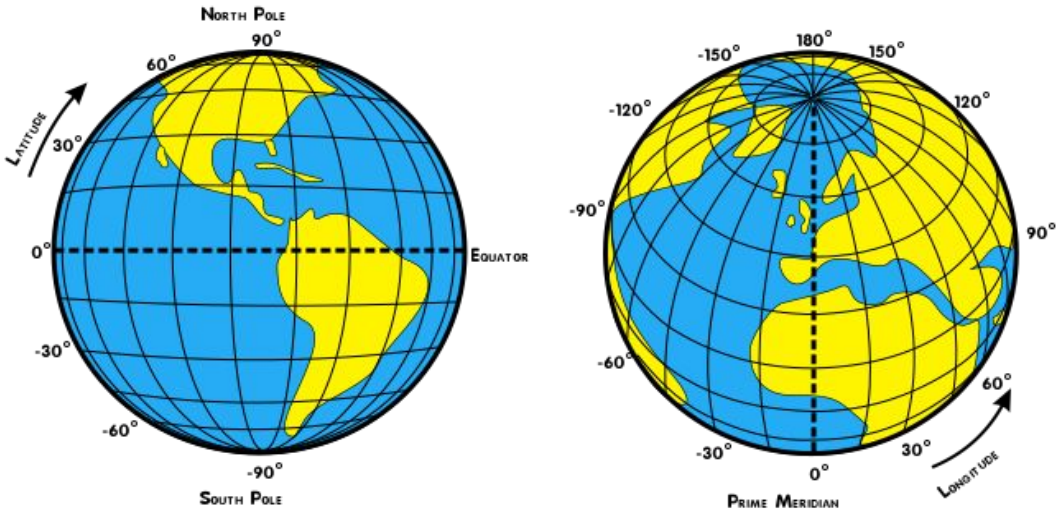
SATAN 2, MZ1900, 1, 20150425

Geodetic point in Greenfield, MA:
"SATAN 2. Mass. Dept. of Public Works
Triangulation Station. 1937"

Find a marker near you:
https://geodesy.noaa.gov/datasheets/ngs_map/

Geographic CRS's

Location is represented by the intersection of latitude and longitude lines.

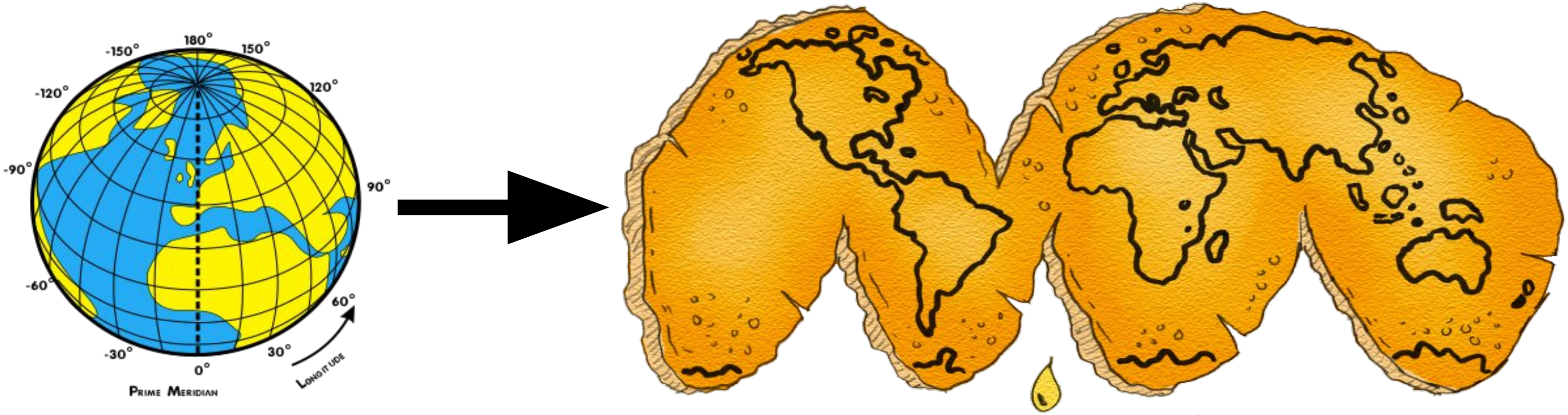


https://commons.wikimedia.org/wiki/File:Latitude_and_Longitude_of_the_Earth.svg

Amherst, MA =
42.3732° N, 72.5199° W
WGS84

Projected CRS's

A **projection** is a mathematical technique to convert and portray features from a spherical surface onto a flat surface.

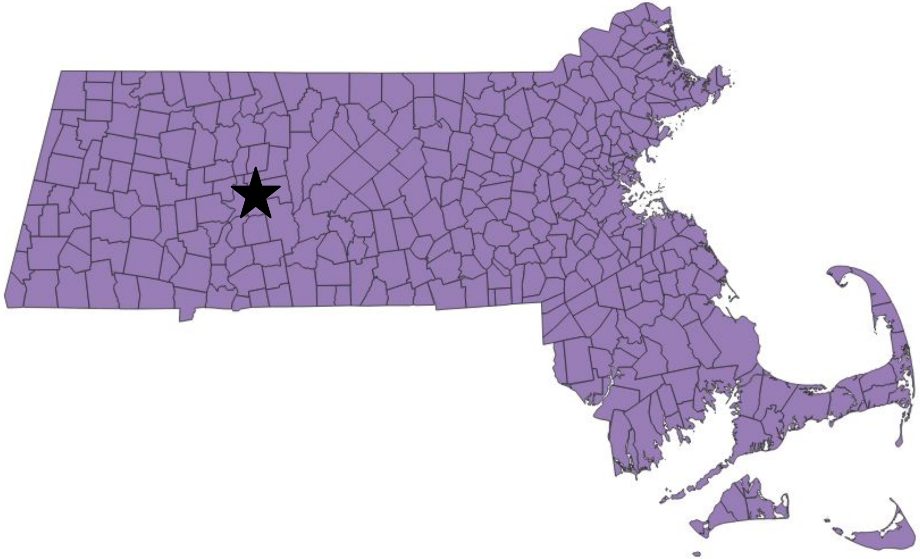


http://dogfoose.com/2011/09/how-to-peel-a-planet/

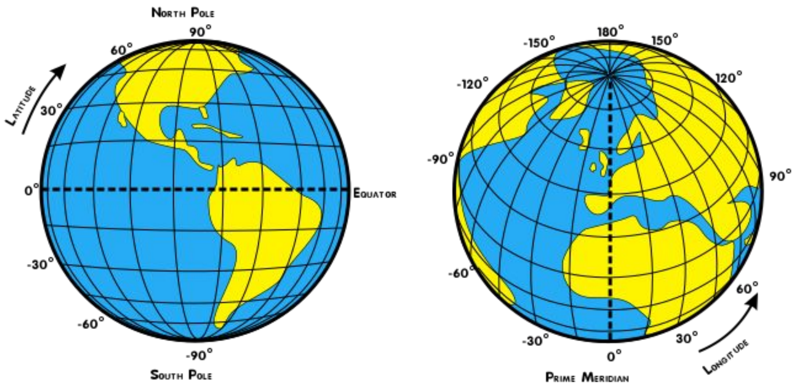
Projected CRS's

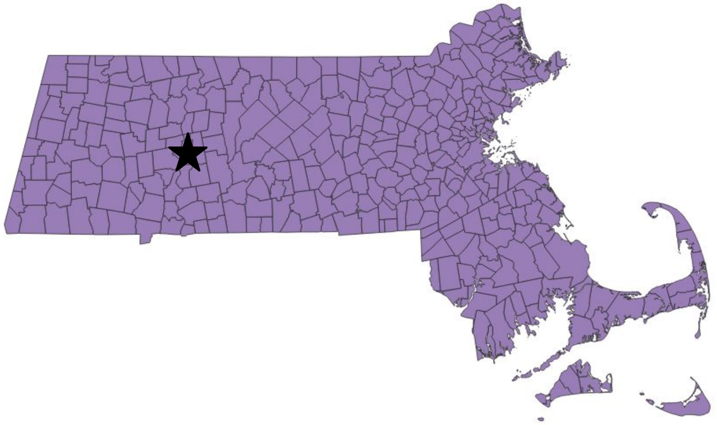
Location is represented as a distance (often in meters) from the CRS's 0 point.

Amherst, MA =
704197 (X), 4694193 (Y)
WGS84 UTM Zone 18N



Coordinates can be converted from a Geographic CRS to a Projected CRS:





Amherst, MA
WGS84
42.3732° N, 72.5199° W

Amherst, MA
WGS84 UTM Zone 18N
704197, 4694193

https://commons.wikimedia.org/wiki/File:Latitude_and_Longitude_of_the_Earth.svg

Projection = distortion

Projections always involve some kind of distortion in **shape**, **distance**, or **area**. Each projection tries to minimize one or all of these types of distortion.

For an interactive experience, check out the Leventhal Map and Education Center's digital exhibition, Bending Lines.

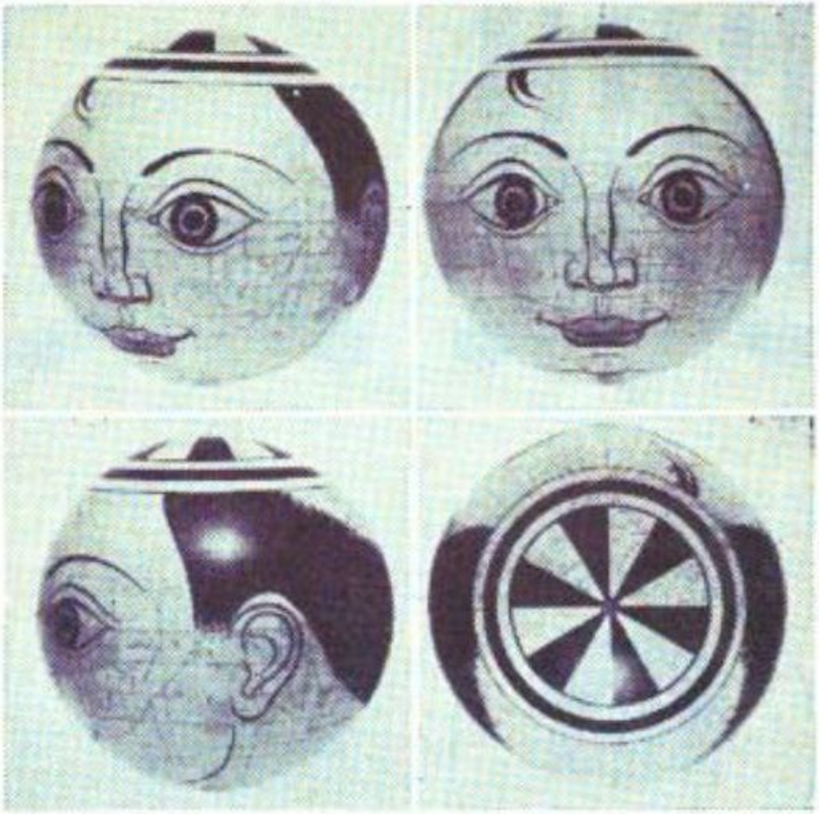


Figure 3. Human head on a geographical globe

Drawn by Boris Artzybasheff, in S. Whittemore Boggs, “Global Relations of the United States” (1954), pp. 905-906.

Source: https://www.leventhalmap.org/digital-exhibitions/bending-lines/how-to-bend/projections/

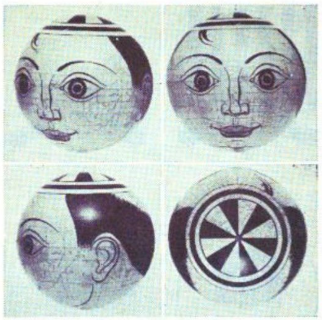


Figure 3. Human head on a geographical globe

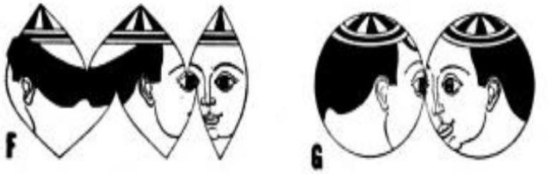


Figure 4. Human head on seven well-known map projections

A. **Mercator projection** (because the North and South Poles are at infinity they cannot be shown on such a map)

B. **Miller cylindrical projection**, a mathematical modification of the Mercator, with all parallels of latitude closer together than on Mercator, and with both geographical poles represented by straight lines as long as the equator

C. **Polar equidistant**, the center of construction being at the North Pole in this instance

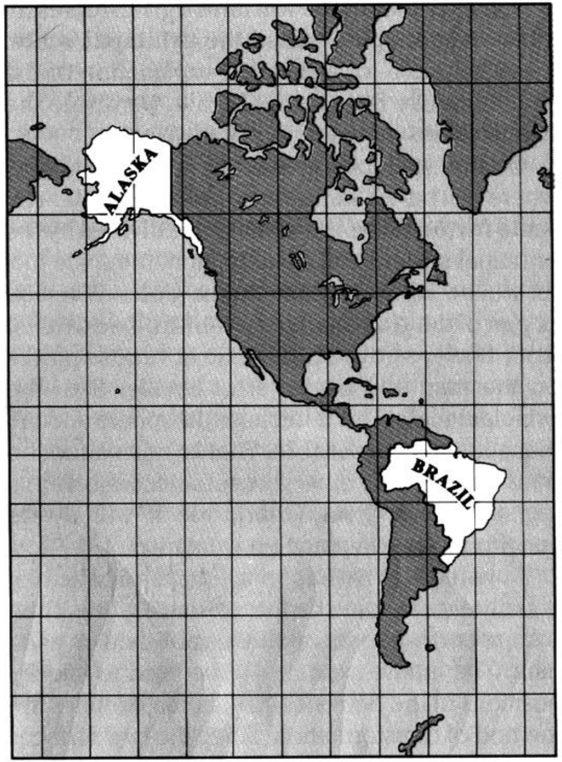
D. **Azimuthal equidistant**, the center of projection being at Washington, D.C.

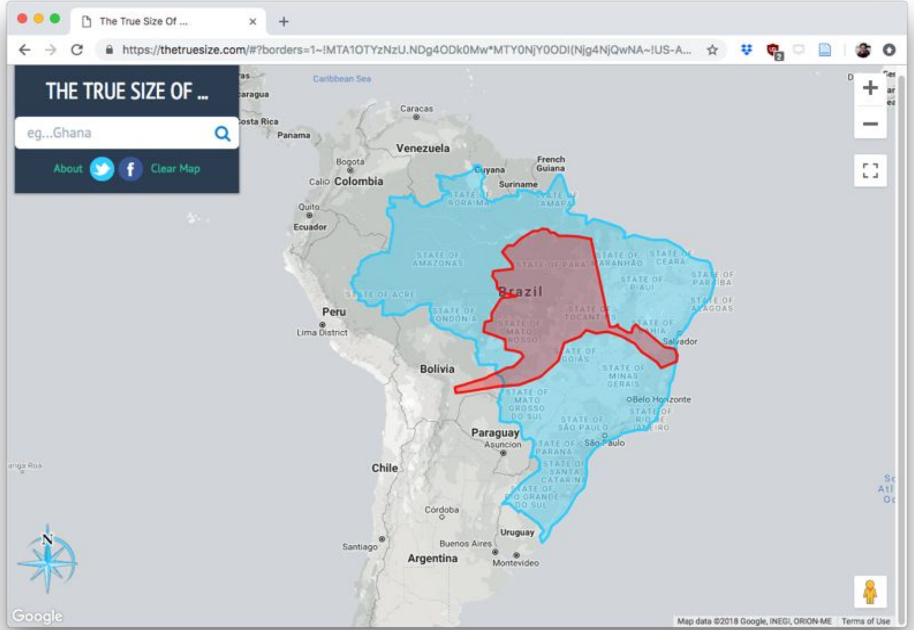
E. **Van der Grinten projection**

F. **Sinusoidal equal-area projection**, as sometimes interrupted to avoid breaking the continents

G. Two **azimuthal equal-area projection** hemispheres

The Mercator projection preserves **shape**, but distorts **area** and **distance**

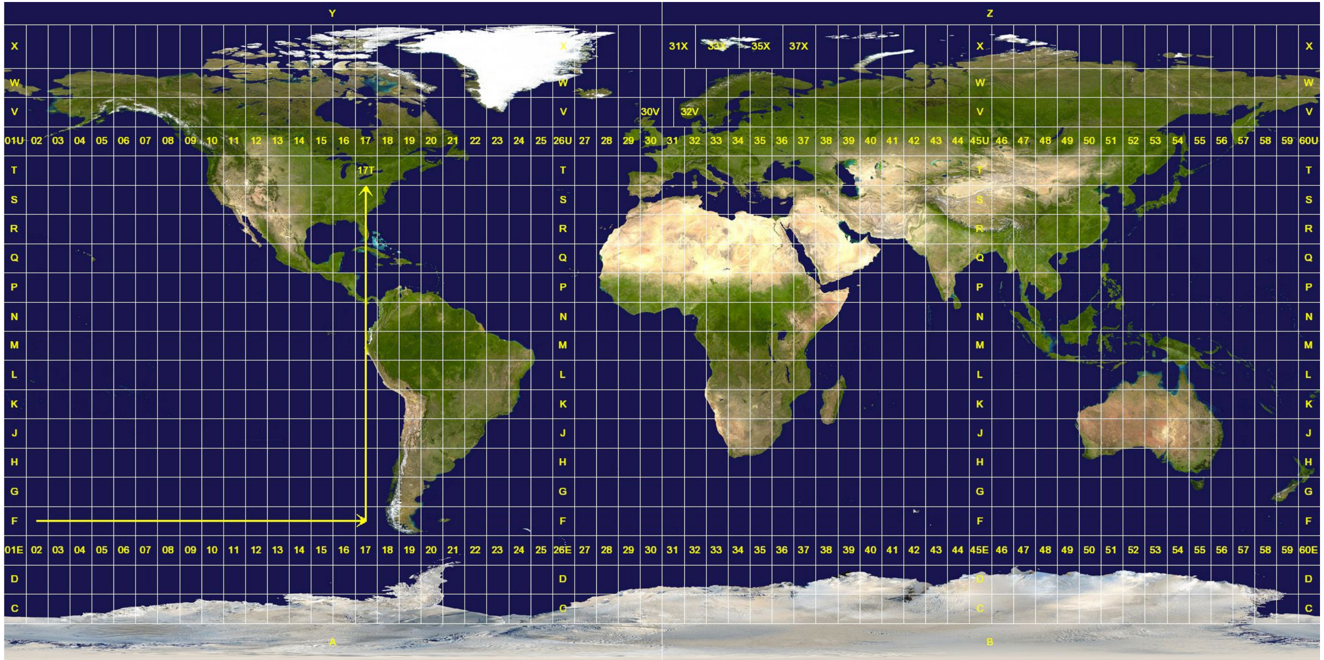




https://www.thetruesize.com/

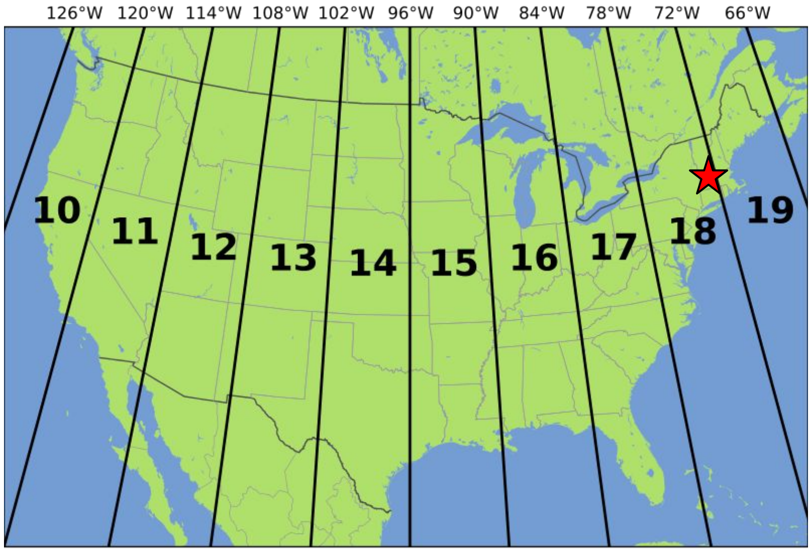
Example projection system: UTM

Universal Transverse Mercator (UTM) divides the surface of the Earth into 60 zones (each 6 degrees wide):

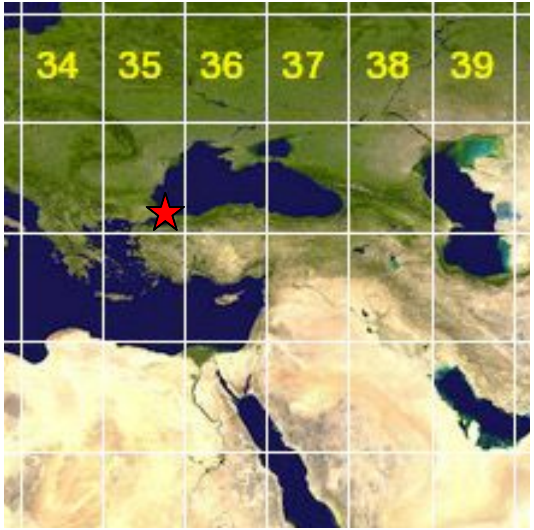


NASA, https://commons.wikimedia.org/wiki/File:Utm-zones.jpg

UTM can be used with different datums:



WGS84 UTM Zone 18

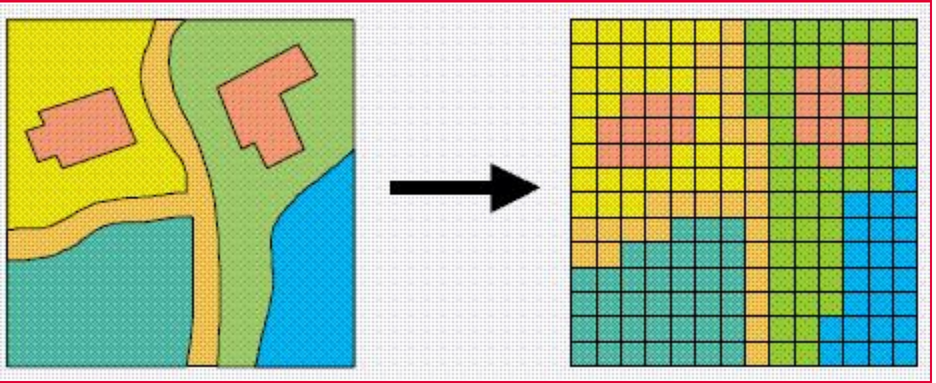


ED50 UTM Zone 35

NASA, https://commons.wikimedia.org/wiki/File:Utm-zones.jpg
https://commons.wikimedia.org/wiki/File:Utm-zones-USA.svg

Summary: Geographic vs. Projected CRS's

	Geographic CRS's	**Projected CRS's**
Length Unit	Degrees	Meters, Feet, etc.
Scope	World	Specific countries / regions
Application	General mapping	Survey, topographic mapping, detailed mapping



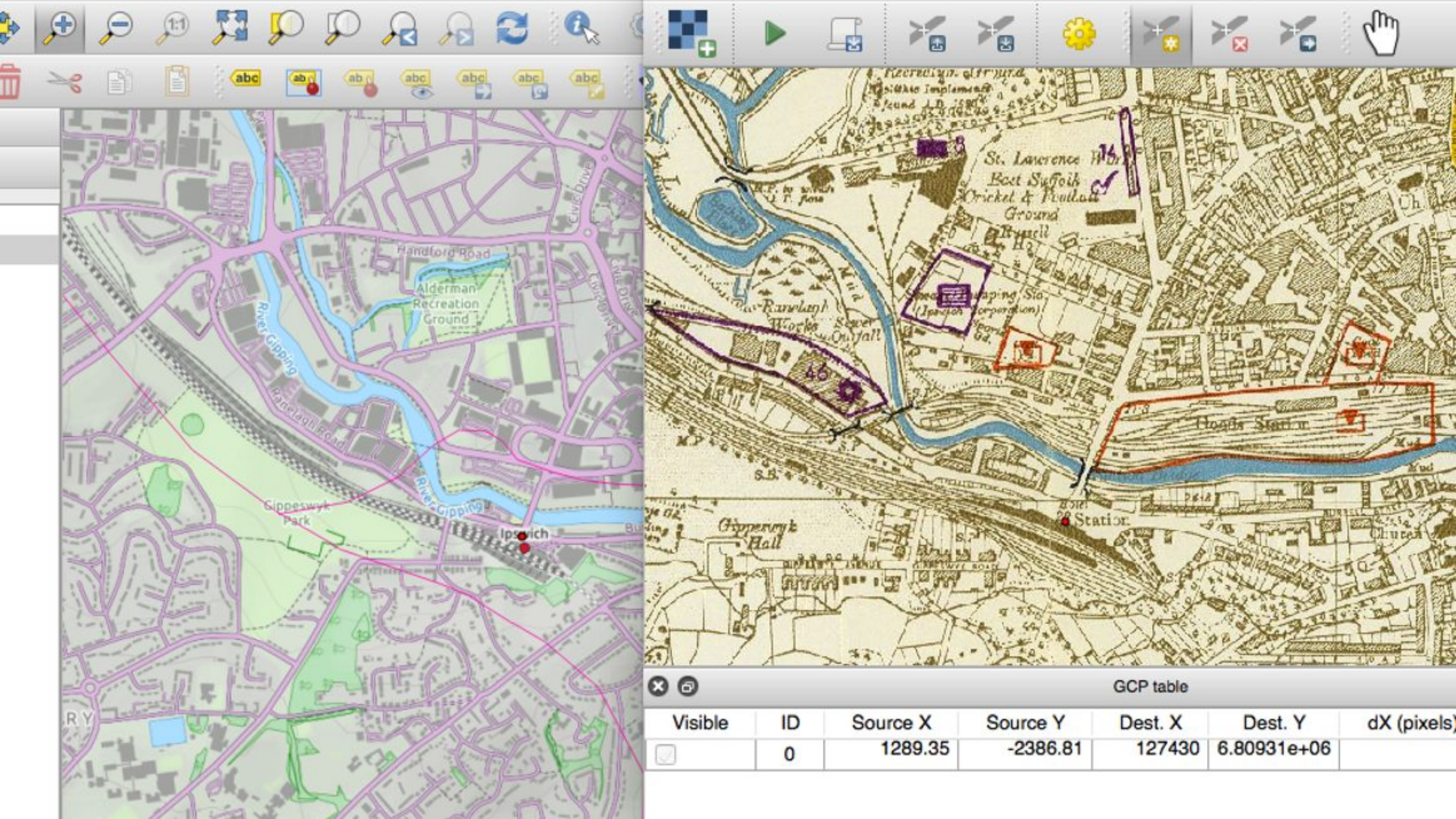
GIS DATA TYPES: RASTER

Raster Data Layers:

- World represented as a static image
- Each pixel has a colour
- Scanned historic maps, satellite pictures etc.

USING RASTER DATA - GEOREFERENCING

- Raster images must be given a *spatial reference* to work within the GIS environment
- Process known as ‘georeferencing’
- (Different to ‘geocoding’ tabular data!)
- Simplest form = defining corner coordinates
- Works for accurately surveyed series maps with known metadata (e.g. national mapping series with regular grid; satellite data)
- Accuracy of map (or parallax of aerial image etc.) mean that warping (‘georectification’) required to make everything match



Handford Road
Alderman Recreation Ground
River Gipping
Ranelagh Road
Civic Drive
Gippeswyk Park
Ipswich
St. Lawrence Works
East Suffolk Cricket & Football Ground
Station
Gippeswyk Hall
GCP table
Visible | ID | Source X | Source Y | Dest. X | Dest. Y | dX (pixels)
0 | 1289.35 | -2386.81 | 127430 | 6.80931e+06

RASTERS – LIDAR

- RADAR, but with a Laser
- Elevation data represented by pixel value
  - Values can be used in analysis
- Two types:
- Digital Surface Model
  - Complete surface – including trees, buildings
- Digital Terrain Model
  - Surface features filtered out (but includes earthworks

RASTERS – LIDAR HILLSHADE

- Representing elevation by pixel value = difficult to see
- Hillshade algorithms apply visibility analysis to create more familiar view
- But no longer directly represent elevation as a value that can be used in analysis

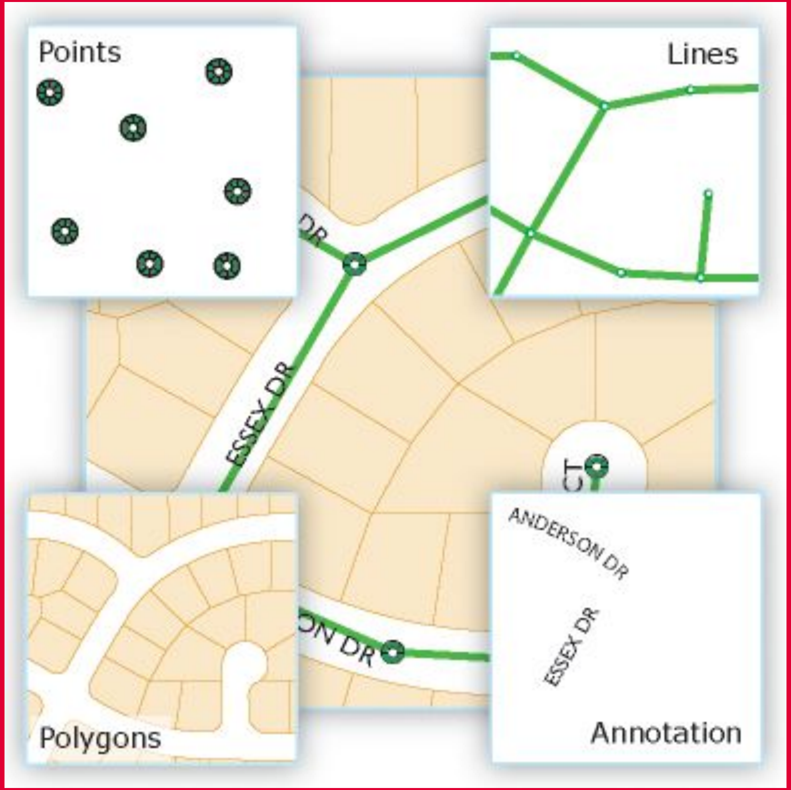
RASTERS - LIDAR

- Layering effects (transparency, blending modes – this is ‘multiply’) can combine rasters
- Georeferenced historic map over hillshade

GIS DATA TYPES: VECTOR

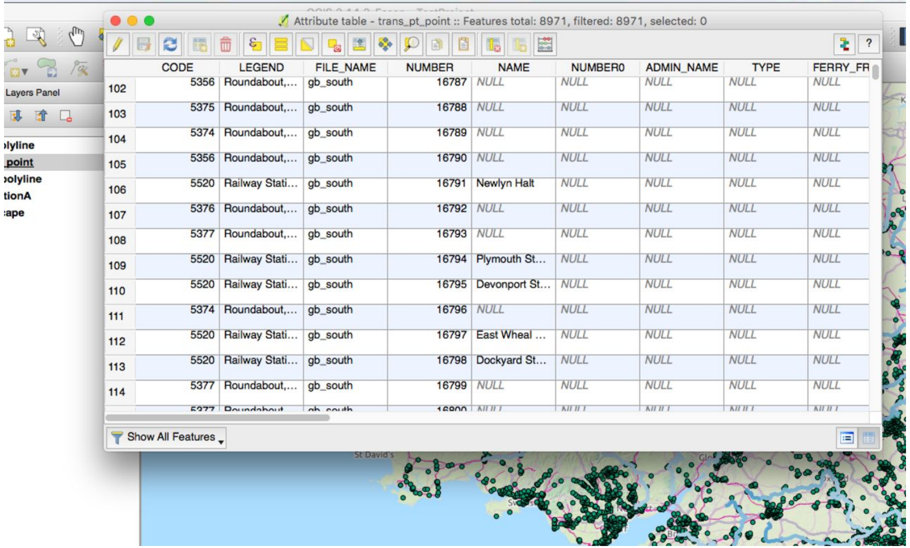
Vector Data

- World represented by a series of coordinates
- Points
- Lines
- Polygons
- Everything can have properties
- Displayed dynamically – appearance can react to properties, zoom level, etc.



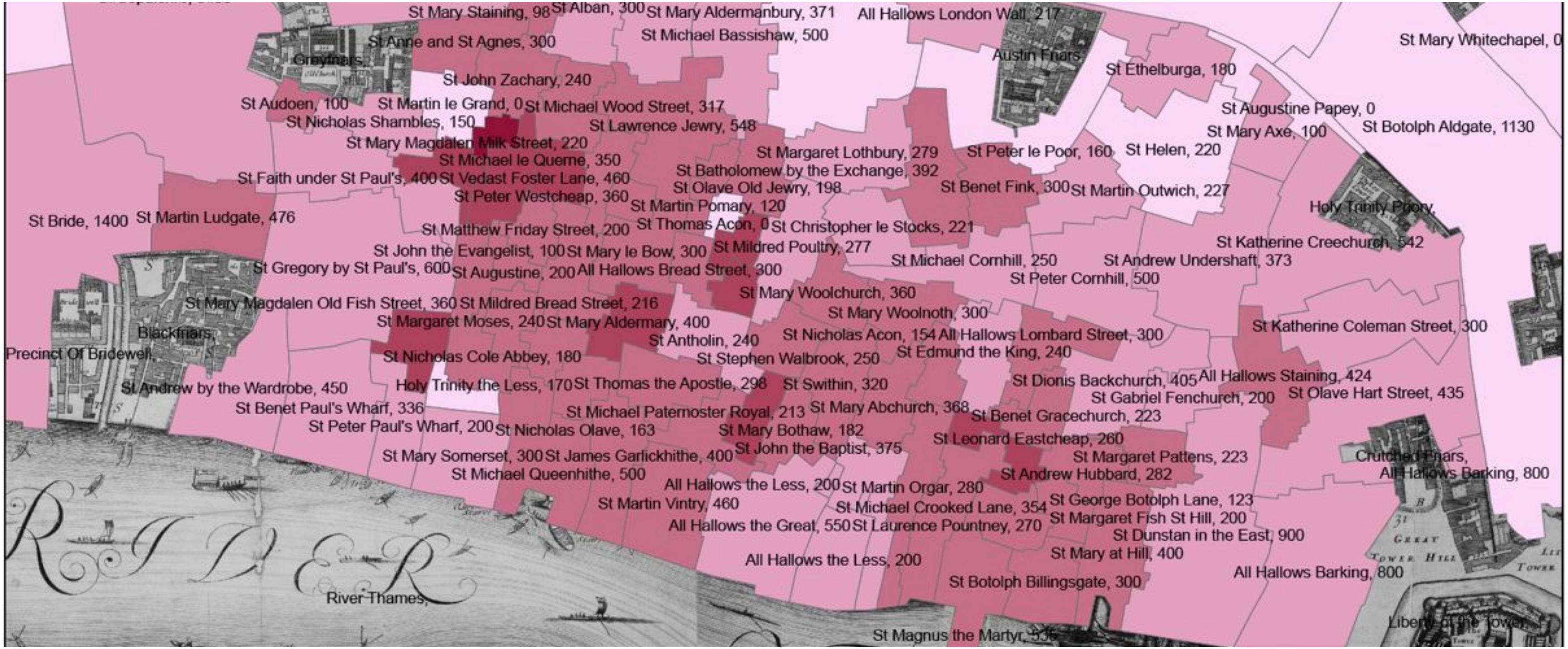
WORKING WITH VECTOR DATA

- When you add existing data, it probably won't display exactly the way you want it
- Vector data usually contains all kinds of other information that is useful
- Map you see is just one way of displaying all of the tabular data in a shapefile!
- You can manipulate the way it is displayed to generate the map you want
- Or combine the pre-existing information with other layers

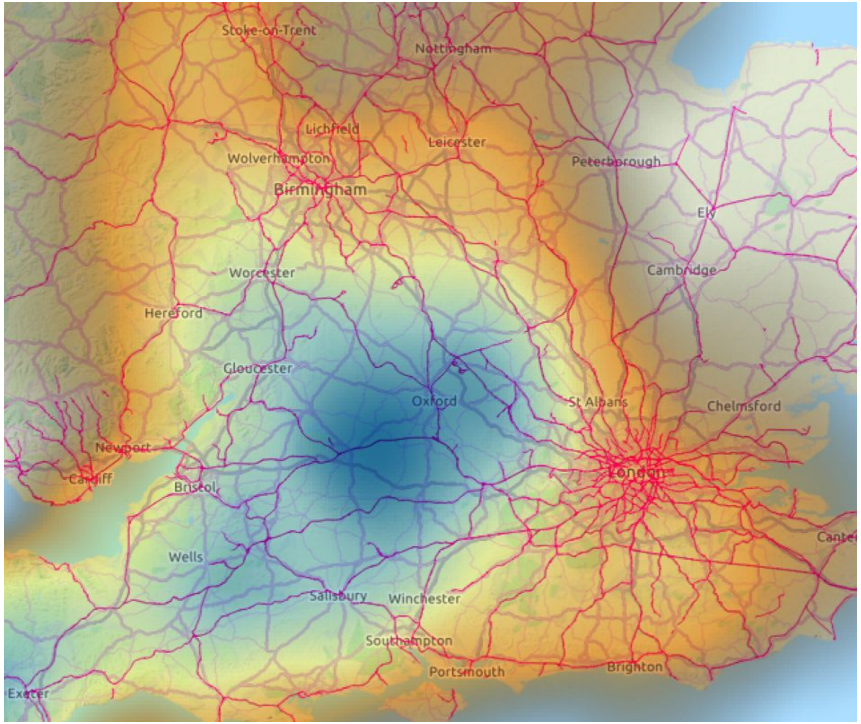


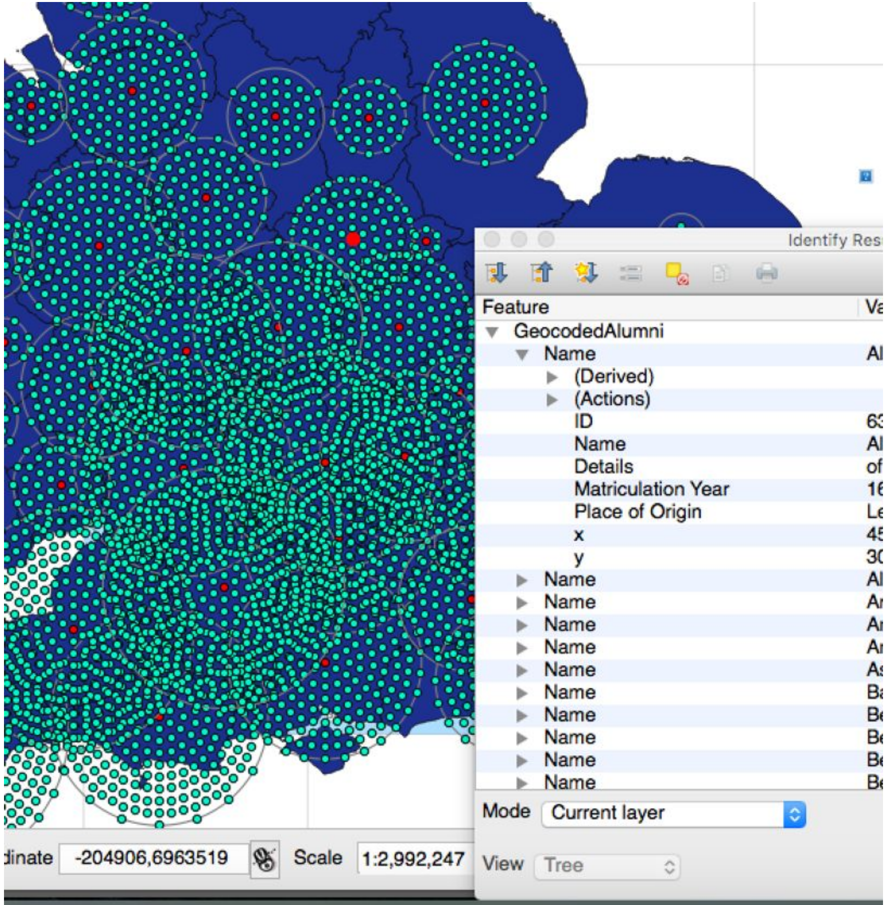
UNIVERSITY OF LONDON

VECTOR SYMBOLOGY



VECTOR SYMBOLOGY - POINTS





Demo

Vector-Based Spatial Queries

Spatial queries

Li and colleagues (2022) selected all archaeological sites within certain elevation ranges to predict risk based on sea-level change

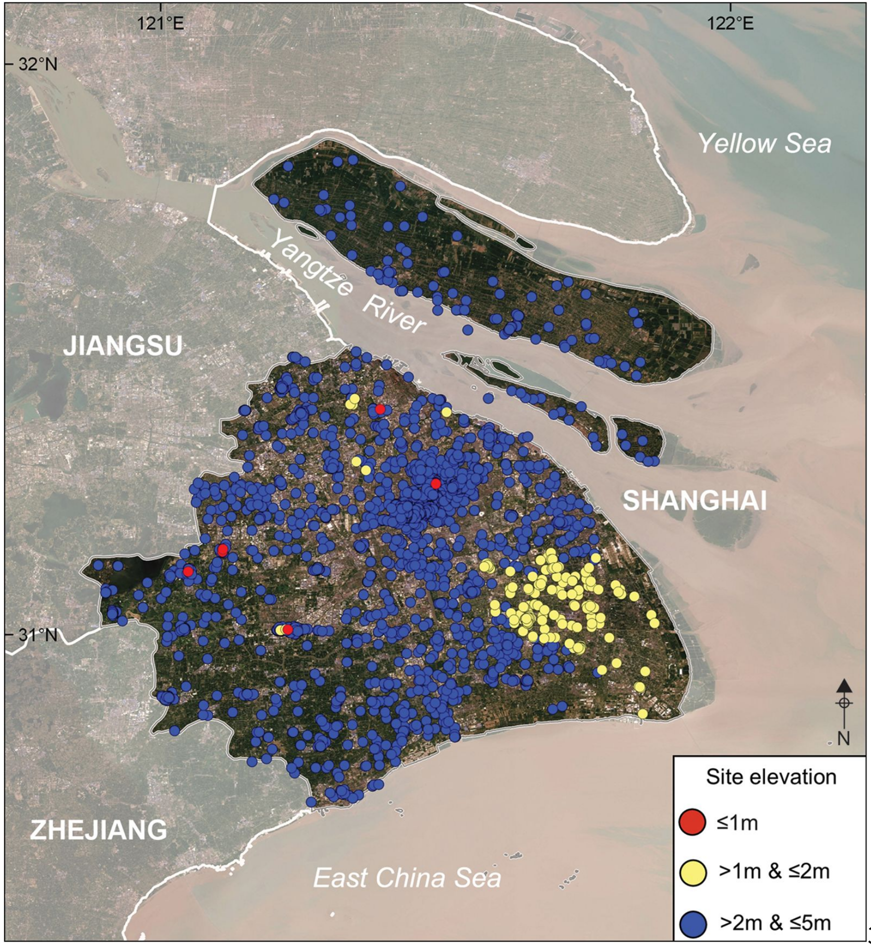
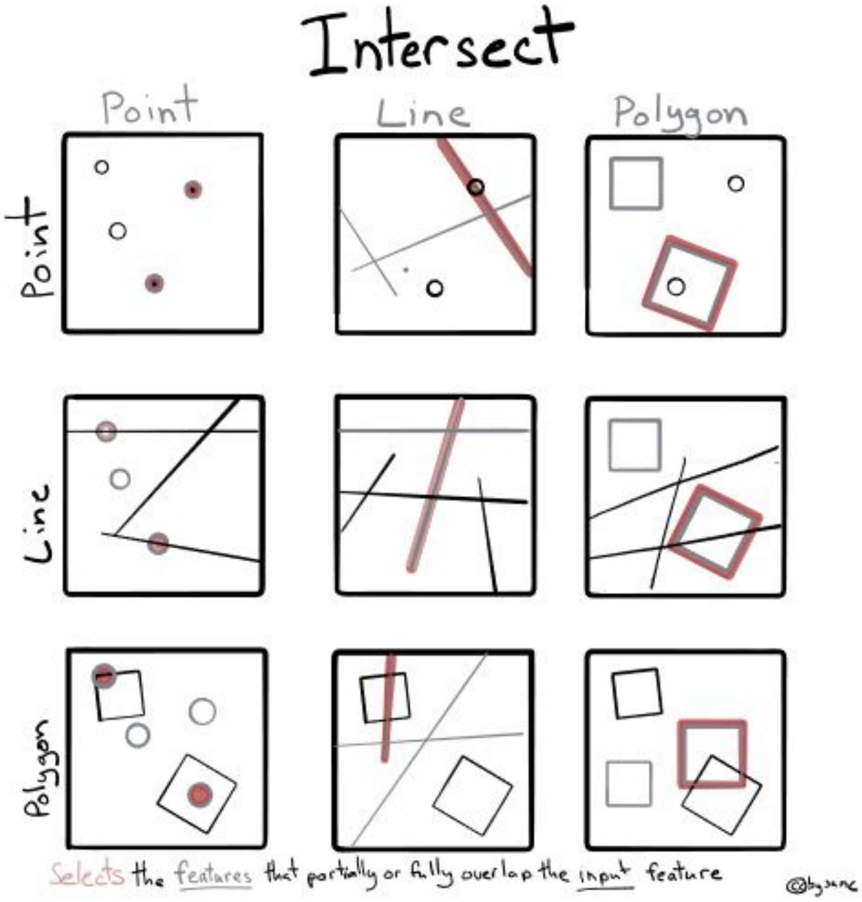
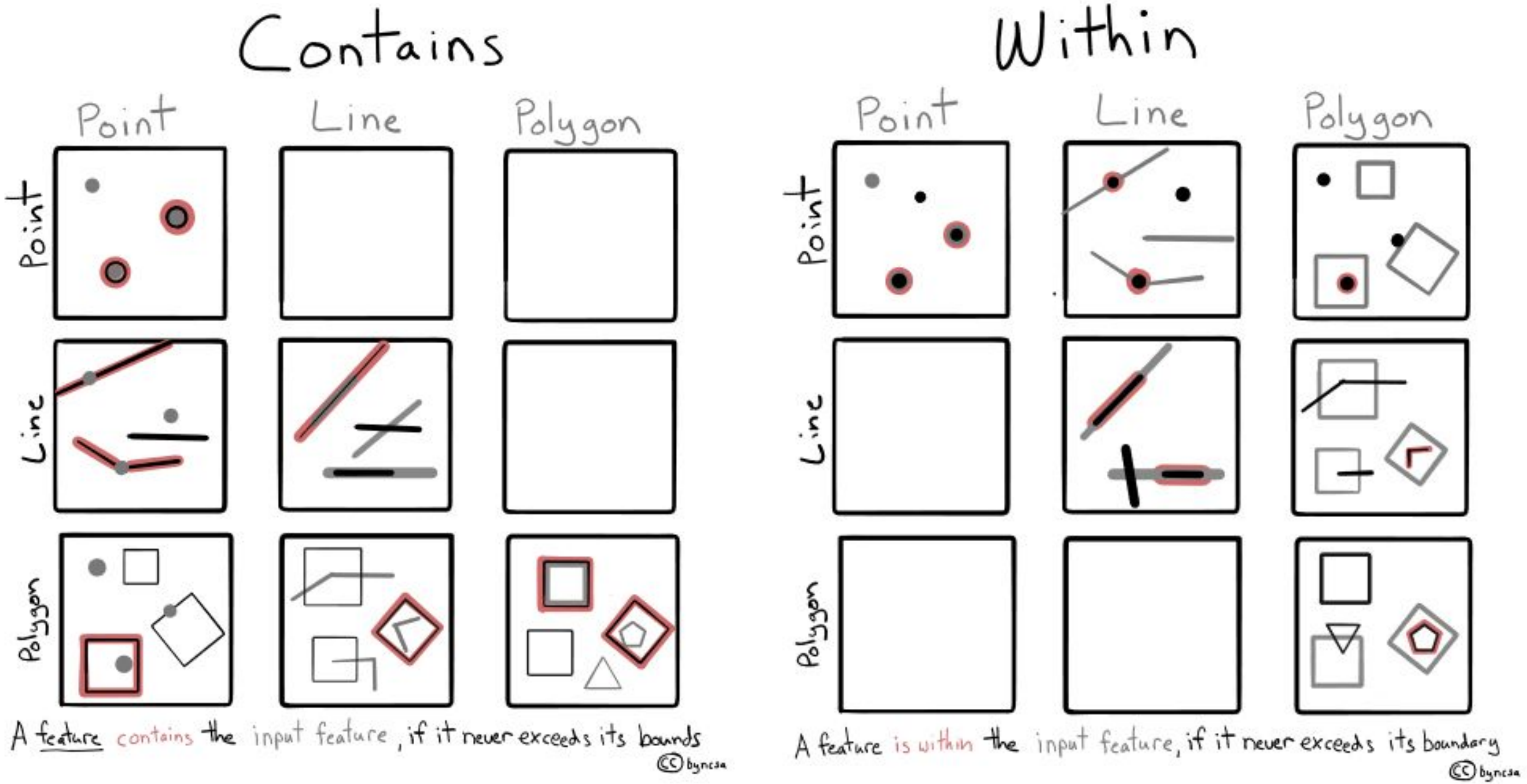


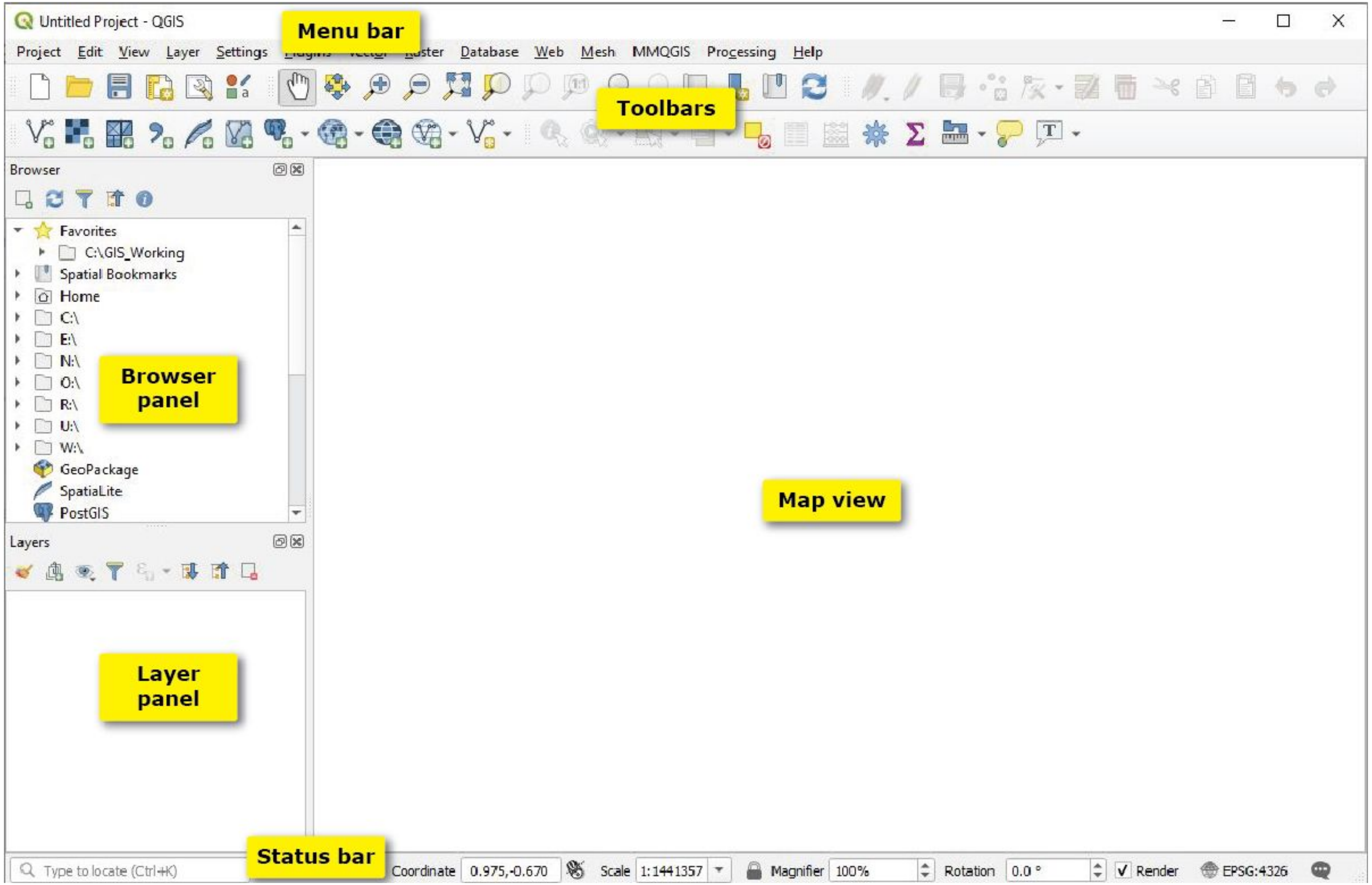
Figure 6 in Li, Y., X. Jia, Z. Liu, L. Zhao, P. Sheng, and M.J. Storozum. 2022. “The Potential Impact of Rising Sea Levels on China's Coastal Cultural Heritage: A GIS Risk Assessment." Antiquity 96(386): 406-421. Available: https://doi.org/10.15184/aqy.2022.1 https://pygis.io/docs/e_spatial_joins.html



https://pygis.io/docs/e_spatial_joins.html



The QGIS Interface



Georeferencing a Historical Map in QGIS

- Where to find maps
- Key terms to know
- Transformation methods
- Resampling methods

Where to find digital historical maps

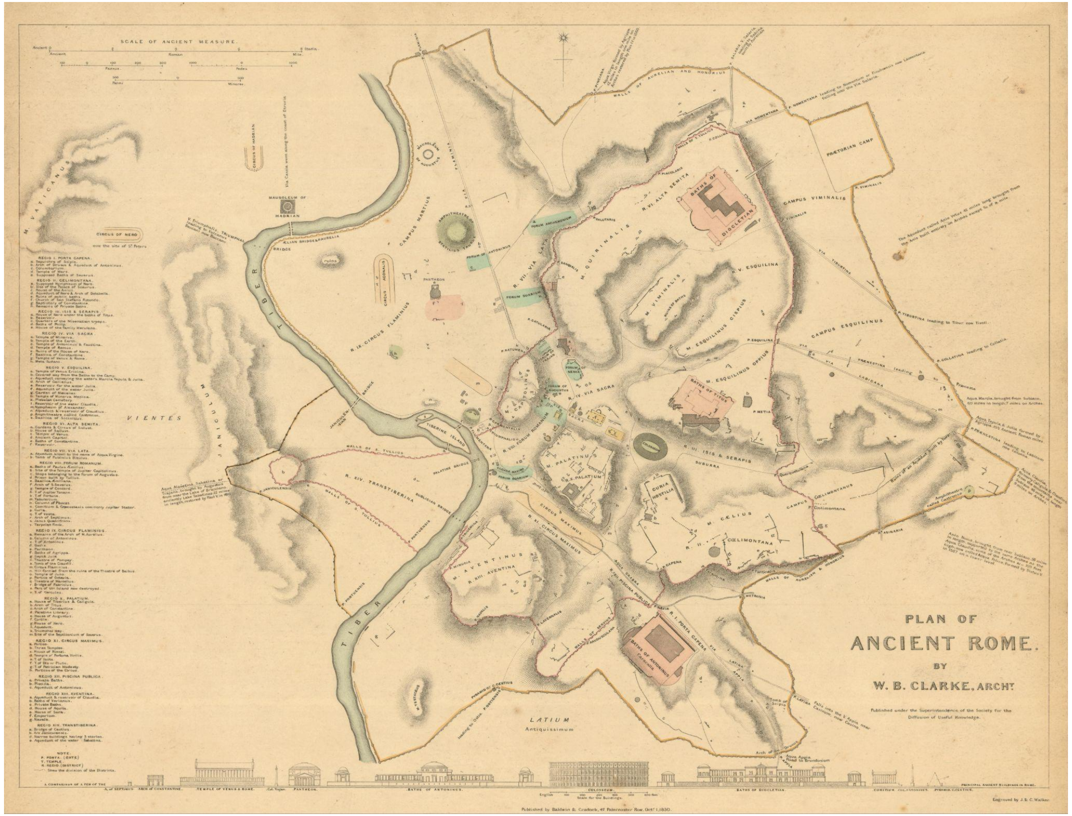
- David Rumsey Map Collection
- Leventhal Map & Education Center Digital Collections
- Old Maps Online (requires free sign-up)
- Geoportals, like Stanford's EarthWorks
- Government and municipal archives
- Library map collections

Some of these maps are already georeferenced! If so, you can download them as GeoTIFFs and add them to QGIS, or you can stream them over the web with a WMS link.

Key terms to know

Source layer

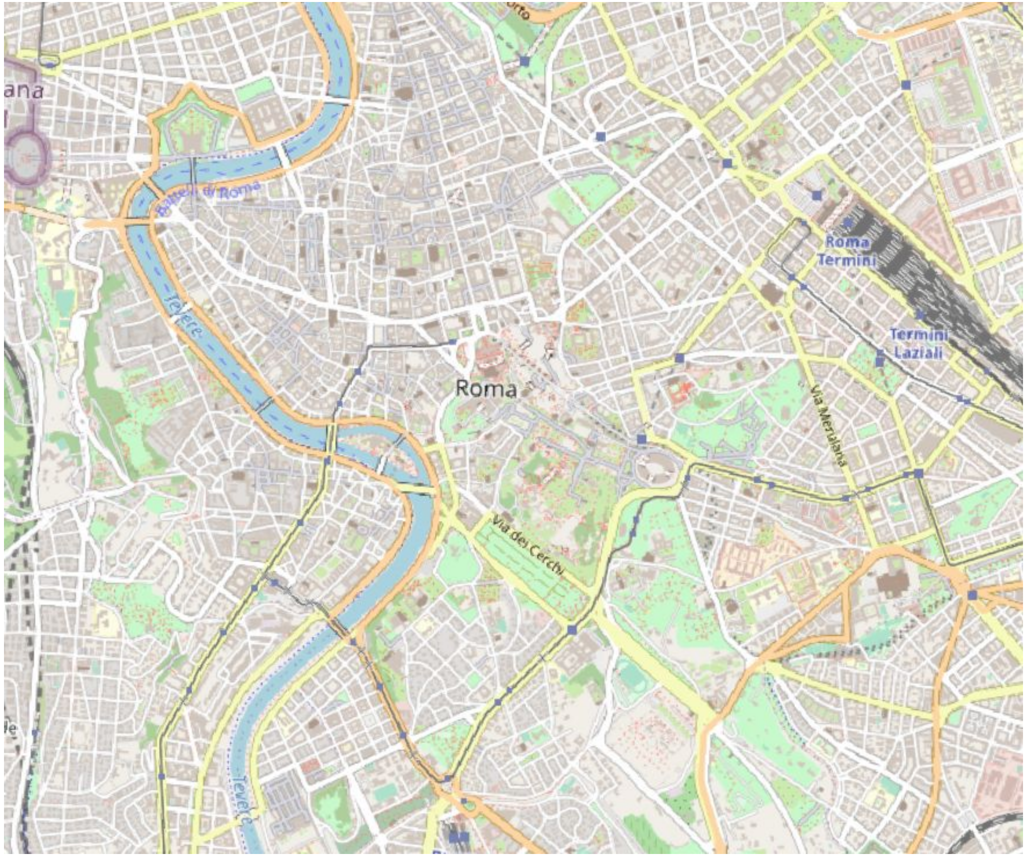
The map or aerial photograph that you will georeference.



Clarke, W. B.. Plan of ancient Rome. London : Published under the superintendence of the Society for the Diffusion of Useful Knowledge : Baldwin & Cradock, 1830. Retrieved from https://earthworks.stanford.edu/catalog/princeton-1c18dj31v

Target layer

An existing spatial dataset that you will use to georeference the source layer.

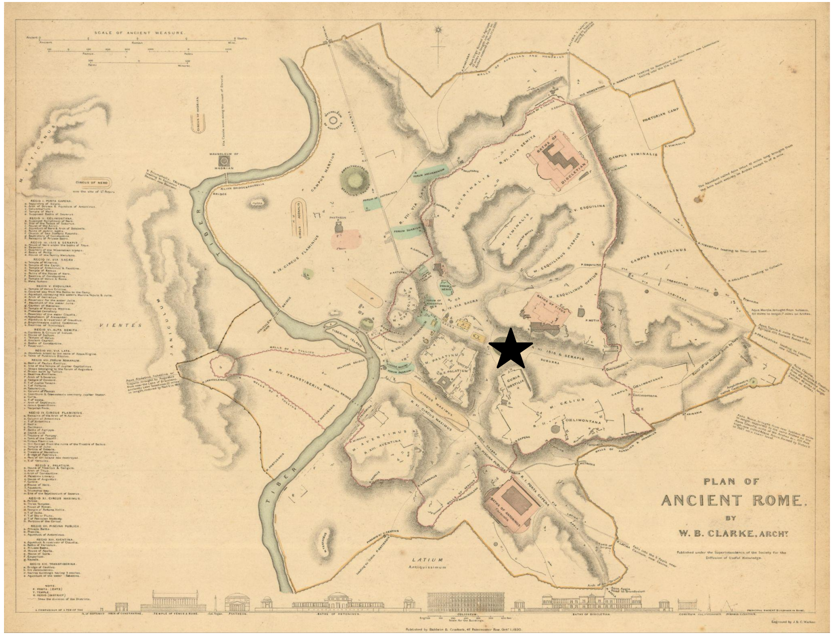


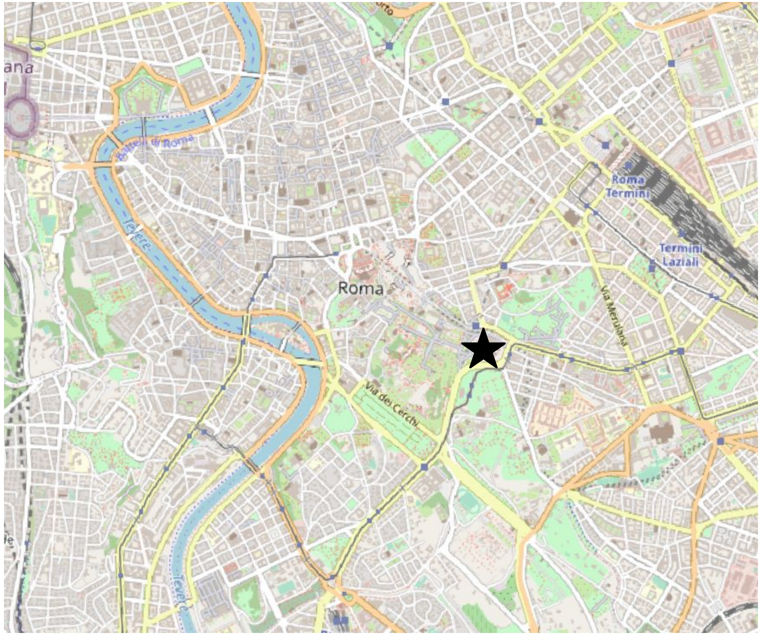
What makes a **good target layer**:

- It should cover the entire extent of the source layer. The image you want to georeference should fit entirely within this target layer.
- It can be a vector dataset (like the outline or boundary of a state) or another raster (like a satellite image).
- The level of detail should be roughly the same as the source layer.

Ground control points (GCPs)

Common points that connect the two layers. You will look for and create GCPs during the georeferencing process.





What makes a **good GCP**:

- Points that stay the same over time, like the corners of buildings, intersections between two roads, or railroad lines.

What to avoid:

- Points that are more likely to move, like the borders of water bodies, individual trees or bushes, or the edges of roads.

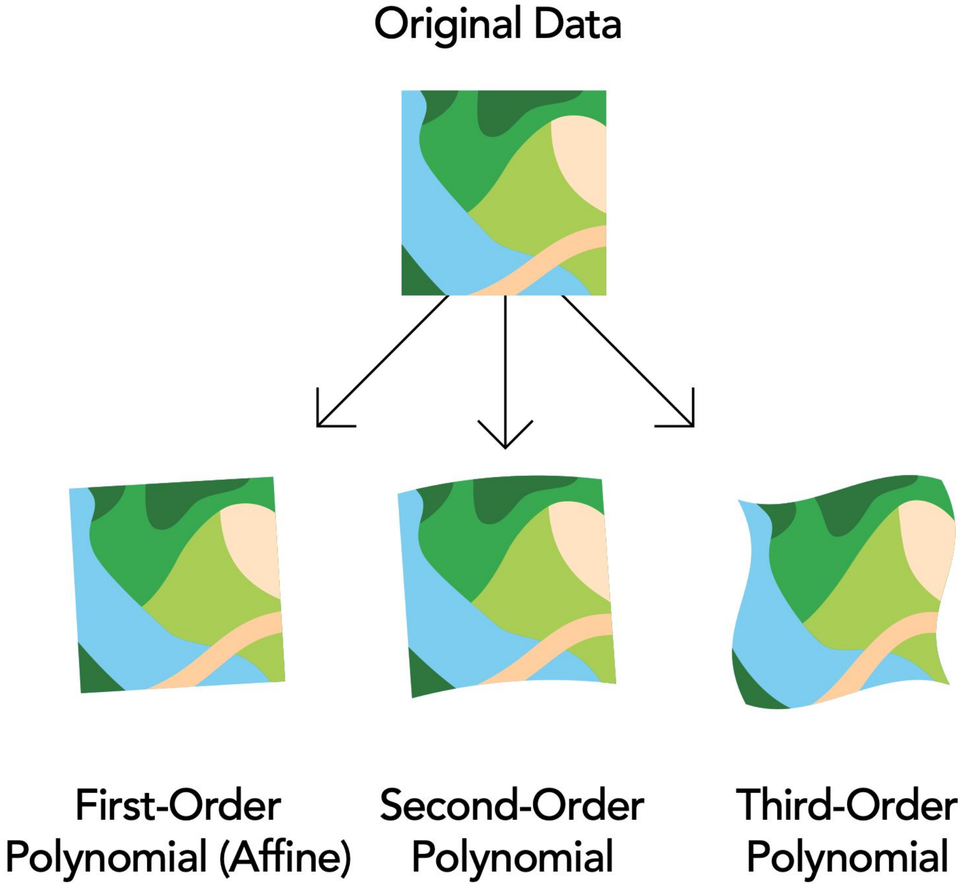
QGIS Exercise: Georeferencing a map – Part 1

1. Download the “Plan of Ancient Rome” by W.B. Clarke: https://earthworks.stanford.edu/catalog/princeton-1c18dj31v
2. Open QGIS, load the OpenStreetMap basemap, and set the project CRS to **EPSG:32633 - WGS84 / UTM Zone 33N**.
3. In the top menu, click **Raster > Georeferencer** to launch the georeferencing tool.
4. Load the historical map into the Georeferencer by clicking the **Open Raster…** button at the top of the tool and navigating to the map you downloaded.
5. Zoom in to a part of the map where you can identify a common point with the target dataset, such as the Colosseum. Click the center of the Colosseum. In the “Enter Map Coordinates” window that appears, select **From Map Canvas**.
6. Click in the same location on the target dataset. Click OK in the window that pops up to add the link to the GCP table.

Transforming the map

Transformation is the next step in the georeferencing workflow after GCPs are added. The computer uses the GCPs to calculate coordinate locations for every single pixel in the source image.

Every 2D representation of a 3D landscape has some amount of distortion. Hand-drawn maps may not be 100% accurate or to scale. Transformation allows the computer to stretch parts of the map to align with the “real” boundaries of the mapped area.



1st and 2nd order polynomials are a pretty good choice for most historical maps

https://www.esri.com/about/newsroom/wp-content/uploads/2018/07/Understanding-Raster-Georeferencing.pdf

How many GCPs do I need?

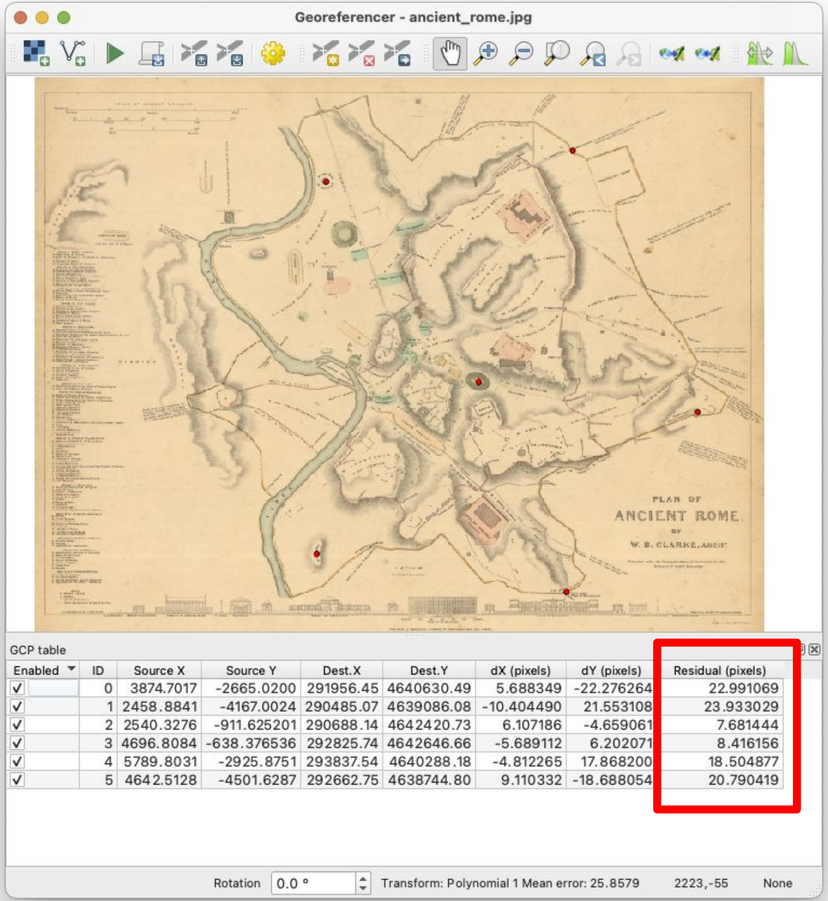
Transformation type	Minimum # of GCPs	What it does
Helmert	2	Shift, scale, rotate
Polynomial 1	**3**	**Shift, scale, rotate**
Polynomial 2	**6**	**Shift, scale, rotate, warp**
Polynomial 3	10	Shift, scale, rotate, warp more
Thin Plate Spline	Lots	Rubber-sheeting; GCP accuracy is guaranteed, but other pixels may have more distorted
Projective	4	Warps lines so they remain straight; good for oblique images

Residual error

As you add GCPs to the georeferencing session, the software constantly tries to readjust the image so that it has the best possible match with target layer.

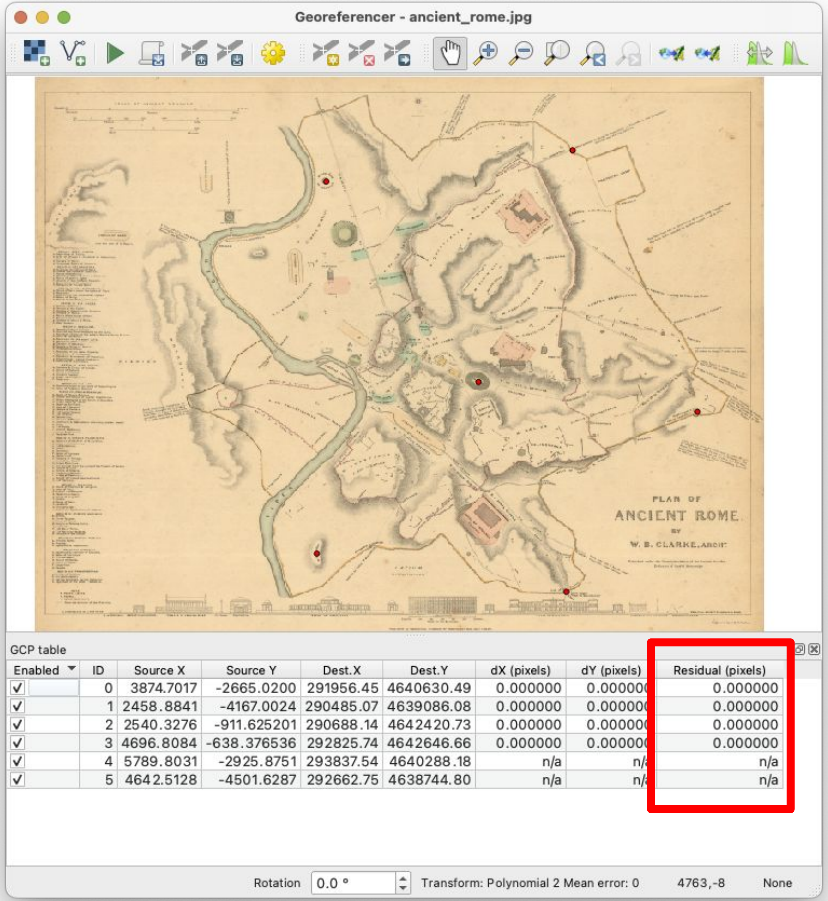
Even perfectly located GCPs will have some **residual error**: the distance (in pixels) from where you placed the GCP to where the software wants to put it.

Polynomial 1



Enabled	ID	Source X	Source Y	Dest.X	Dest.Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	0	3874.7017	-2665.0200	291956.45	4640630.49	5.688349	-22.276264	22.991069
✓	1	2458.8841	-4167.0024	290485.07	4639086.08	-10.404490	21.553108	23.933029
✓	2	2540.3276	-911.625201	290688.14	4642420.73	6.107186	-4.659061	7.681444
✓	3	4696.8084	-638.376536	292825.74	4642646.66	-5.689112	6.202071	8.416156
✓	4	5789.8031	-2925.8751	293837.54	4640288.18	-4.812265	17.868200	18.504877
✓	5	4642.5128	-4501.6287	292662.75	4638744.80	9.110332	-18.688054	20.790419

Polynomial 2



Enabled	ID	Source X	Source Y	Dest.X	Dest.Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	0	3874.7017	-2665.0200	291956.45	4640630.49	0.000000	0.00000	0.000000
✓	1	2458.8841	-4167.0024	290485.07	4639086.08	0.000000	0.00000	0.000000
✓	2	2540.3276	-911.625201	290688.14	4642420.73	0.000000	0.00000	0.000000
✓	3	4696.8084	-638.376536	292825.74	4642646.66	0.000000	0.00000	0.000000
✓	4	5789.8031	-2925.8751	293837.54	4640288.18	n/a	n/	n/a
✓	5	4642.5128	-4501.6287	292662.75	4638744.80	n/a	n/	n/a

QGIS Exercise: Georeferencing a map – Part 2

7. Continue adding GCPs across the map. Make sure they are well distributed, ideally in each of the four corners and throughout the center of the map. Collect **at least 6 GCPs**.

8. Click the **Transformation Settings…** button in the menu bar and set the transformation type (Polynomial 2) and resampling method (Nearest Neighbor). Give the output file the .tif extension to save it as a GeoTIFF. Optional settings:

    a. Compression = LZW (makes the file smaller)
    b. Leave “Use 0 for transparency when needed” unchecked
    c. Check “Save GCP points” to save your GCPs to a text file.

9. Click the **Start Georeferencing** button in the menu bar to run the tool. When finished, the newly georeferenced map will be added to your QGIS session and you can close the georeferencer window.

PLAN OF
ANCIENT ROME
BY
W. B. CLARKE, ARCHT
Cimitero
Monumentale
di Campo
Verano
Garbatella
Parco della
Caffarella