**Working with GIS**

GIS – **G**eographic **I**nformation **S**ystems – is a massive field; we're just scratching the surface here.

But it's possible to actually create real maps that no one has made before pretty easily by drawing together elements from various map **layers** online. This does not involve programming, but *does* require understanding some high-level concepts that can be tricky to grasp at first.

## The most import concepts for working with GIS

Here are some of the most important features

**Layers**

A GIS map is built up out of **layers**. In the online version, you can turn these on and off; in the final representation, the way your maps appear will depend on the layers that are visible.

Layers are an abstraction of teh actual file formats that you might use. In practice, they can be populated from existing data.

**Shapefiles** the general descriptive term used for vector data; originally an internal standard for ARCgis, but now the general interchange format for geographic data.

**Working with layers in QGIS**

Right click on a layer to examine it. The two most important options are:

- Attribute table, which lets you see and sometimes edit the data about each item in a vector layer, and
- Properties, which lets you change the appearance of a layer: map cities by population density, for instance.

**Raster and Vector**

**Raster data**

Raster data is image map data: it represents points as colors generally.

Raster files usually have an image suffix (often `.tif` or `.tiff`).

If you georectify a historical map, it will

They are the equivalent of a text file that has been scanned, but that hasn't had OCR performed on it.

But sometimes raster data is actual raw data on its own. Imagine, for example, that you take a series of estimates of the altitude for every point in a country. The most natural way to represent this is as a single value for each point on the grid. You can think of this as an image, but it also *directly* encodes the raw data.

**Vector data.**

Vector files, on the other hand, are the geospatial equivalent of text that has been OCR'ed. They represent space as a series of lines that can be manipulated. We are looking in class at street data, for example, where each street is represented as a single line. (In a raster dataset, a street might be several grey bands across.

Vector data may not be as pretty as raster data in many cases, but it makes it possible to represent *things*, not just space abstractly. A street can have a name; a building can have

**Types of Vector layers.**

There are three basic types of vector layers:

- Points are things that take up only a single point in the data. In a crime database, an individual mugging might be represented as a point; in a rough list of cities for mapping on the surface of the globe, each city might be a point.
- Lines (like streets)
- Polygons (like buildings or countries; things that have an exterior.)

**Projections**

The last, and by *far* the most frustrating thing you will encounter, are various different *projections* used to describe the shape of the earth.
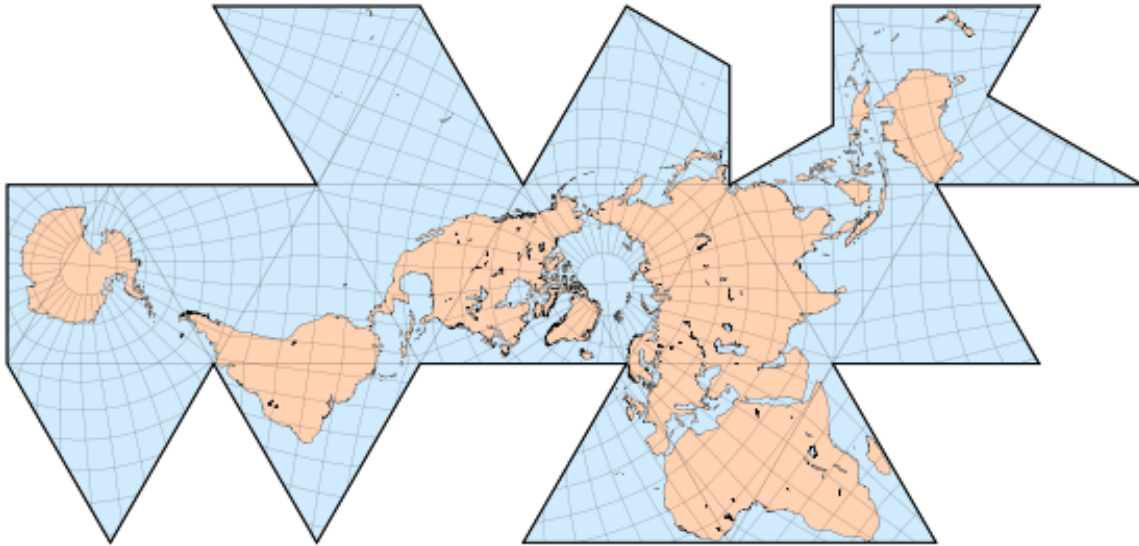


There is no right way to represent a spherical planet on a flat page. So each of the files that you use may have a different base projection built in. These have funny names assigned by world geographical bodies, like "WGS84." Internally in QGIS (and in a variety of programs) a projection is represented as a "Coordinate Reference System," or CRS.[1]

---

[1]Technically, a CRS is not exactly the same thing as a projection; a projection just tells how to represent the sphere on a plane, while a CRS includes a lot of other information, like units of measurement (some systems use meters from a fixed point; others use latitude-or-longitude points; others use miles; and so forth).

**Fixing broken CRS coordinates in QGIS.**

If you encounter a problem where things aren't appearing, odds are that it has to do with broken CRS interpretation.

Here's what do when this happens.

0. Bang your head against a wall for 90 seconds. Scream "Gosh darn it! Why is this so frustrating" into a pillow.
1. Confirm you have a CRS problem. The easiest way is to right-click (control-click on a Mac) on one of your layers and select "zoom to layer extent." Confirm you can see the layer; then waggle your cursor over the map and look at the numbers that show your location on the bottom bar. Then do the same for your other layers. If the numbers look extremely different, you probably have a CRS problem. If not, you might just be misaligned.
2. Get rid of all but one of the layers; control click to delete it from the view. If it includes unsaved work, save it first.
3. Go to `Project -> Project Properties -> CRS` in the menu. (This is also accessible by clicking the coordinate system in the lower right hand corner of the map view). Make sure the box that says "Enable 'on the fly' CRS transformation" is checked.
4. Right click on your remaining layer: select "set project CRS from layer"
5. Reimport the layer that didn't display before. It should now display.

These projections *can* all be combined into a single system.

## GIS Software

### ArcGIS

The industry standard is ArcGIS, made by the company ESRI. It's expensive! That's about the only downside. If you take a GIS class, they often use Arc.

### QGIS

QGIS is free, and not all that bad. It's continually getting better, in fact; and you should be completely able We're using it so you get used to something you'll be able to keep using without a site license.

As you saw, installation is a pain.

### Web mapping platform.

There are a variety of web mapping technologies that are taking different areas of the humanities and journalism by storm. Hypercities builds on some of the Google platform; Stanford/Orbis builds on D3.js, which is hard to use (requires coding in Javascript) but extremely powerful for interactive maps. It's what, most notably, any map you've seen on the New York Times in the last few years uses.

CartoDB is an online platform that allows relatively easy integration of datasets with a detailed base map of the entire world. For interactive online applications, it can be much easier to set up.

## Finding Map Layers

- Large scale
  - The best source of large scale (state/country level or larger) is Natural Earth Data
    * Cities
    * Countries
- Smaller Scale: Examples
  - Massachusetts towns.
  - More Massachusetts data than you can shake a stick at
  - Every building in Massachusetts
  - Railroads in the US in 1860
  - Raster topo maps of the United States

## Georectification

Georectification is the process of warping a map so that it can be used

Some examples of various sorts:

- Northeastern's Boston Maps Project
- The Interactive Nolli Map of Rome
- My map of the Boston subway

## Geolocation

The process of matching places in English (or Latin, or whatever) to latitude-longitude coordinates you can place on a map.

Example: Mitch Fraas and Ben Schmidt, Mapping the State of the Union