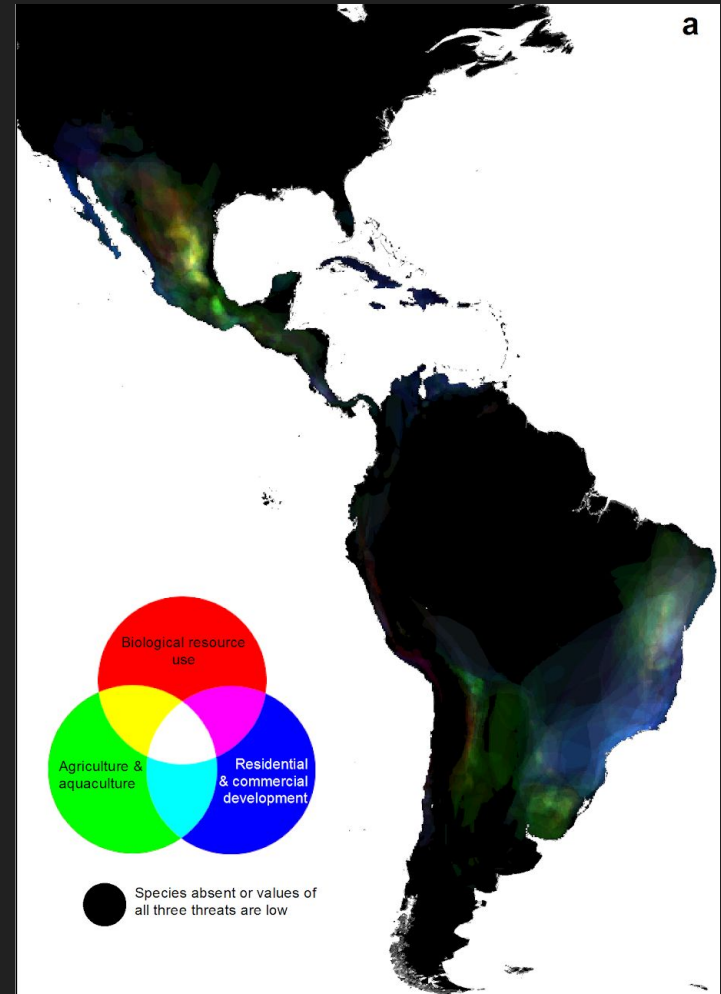


Introduction to Spatial Data & Analysis

James Duffy
Chris Yeomans
David March

James Duffy

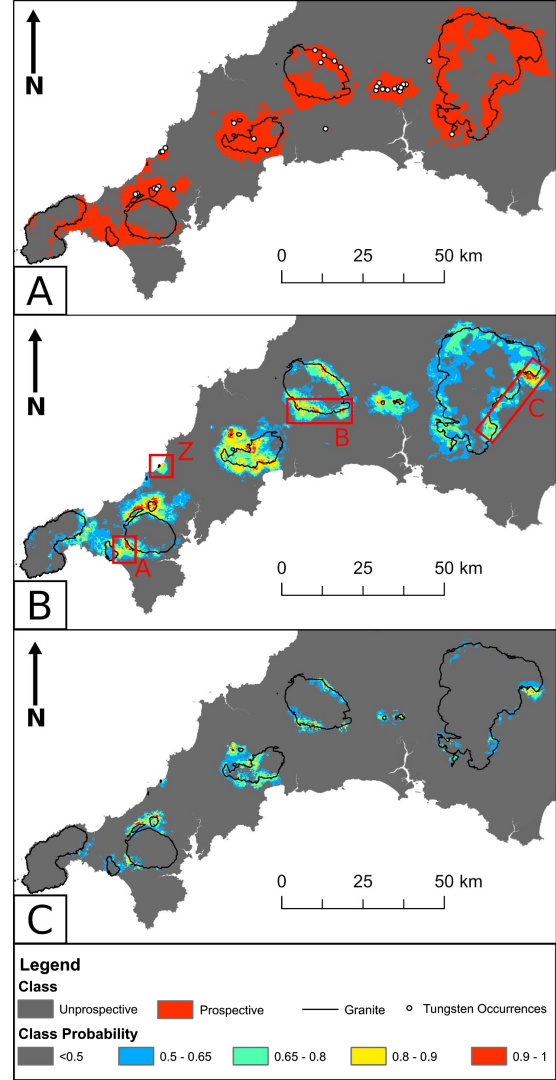
- 9 years of experience using GIS for scientific research
- Broad range of data - global to 'hyper-local'
- Big fan of open-source (but realise ESRI products are 'standard' in some places)
- Currently working on novel remote sensing applications for microclimate modelling



Chris Yeomans

- 6 years experience in GIS, 4 years experience with R
- Focus on spatial data over Cornwall for geoscience applications
- PhD in Geology with core focus on machine learning techniques (in R)
- Now researcher in geothermal energy exploration

Interested in geostatistics, open-source software & programming, dabbling in compositional data analysis (CoDA), failing to learn Python.



David March

- 13 years of experience using GIS and R
- Oceanography, Fisheries, Marine spatial planning, Spatial Data Infrastructures
- Movement Ecology, Statistical modelling, Geostatistics, Machine learning



Course Structure & Rationale

P1

Principles of spatial data

Intro to QGIS - with a focus on map-making

Additional QGIS tasks [if time]

Catch-up on R skills [optional]

P2

Spatial operations in R (vector focus)

P3

Spatial operations in R (raster focus)

P4

Future trends - online tutorials

Bring your own data?

Spatial Data

Thematic Maps

Sampling
Locations

Base Maps

Species Ranges

Surface Maps

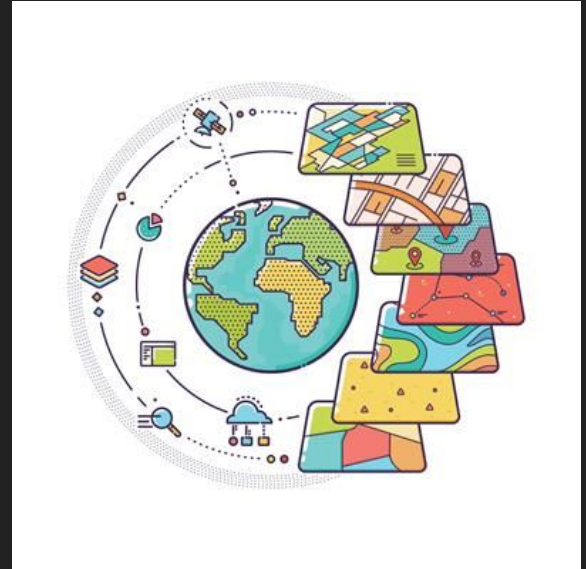
Political
Delineations

Natural
Boundaries

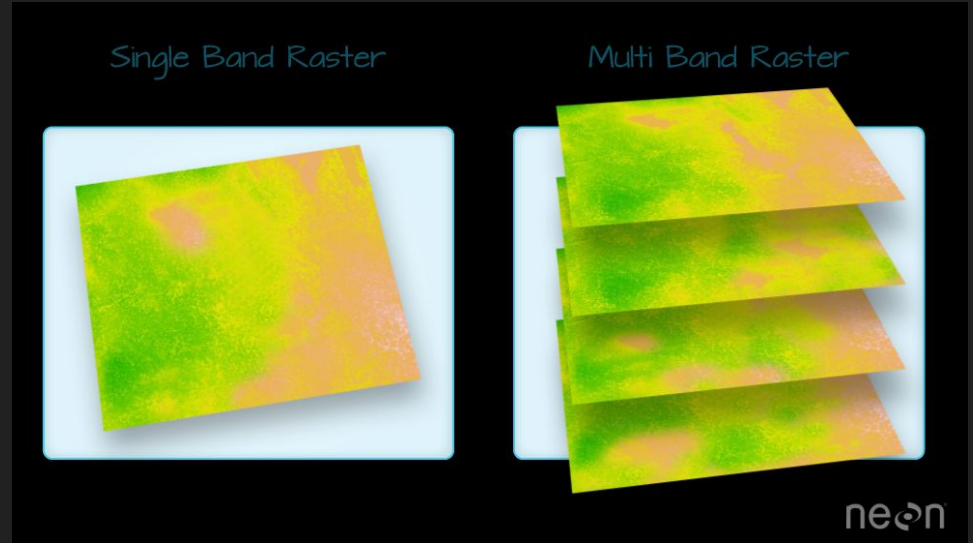
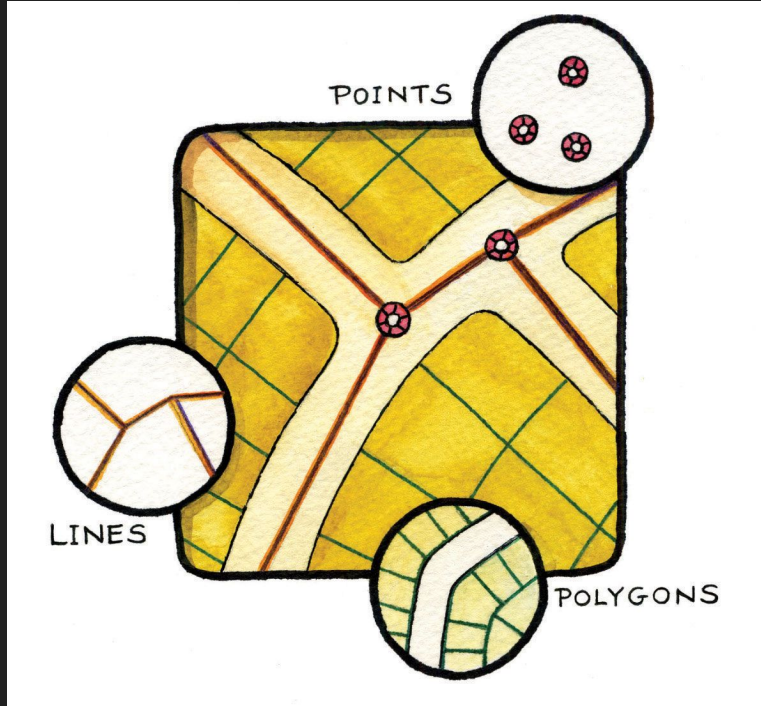
Natural
Boundaries

Animal tracks

Ultimately, data with
location information that
can therefore be **mapped**

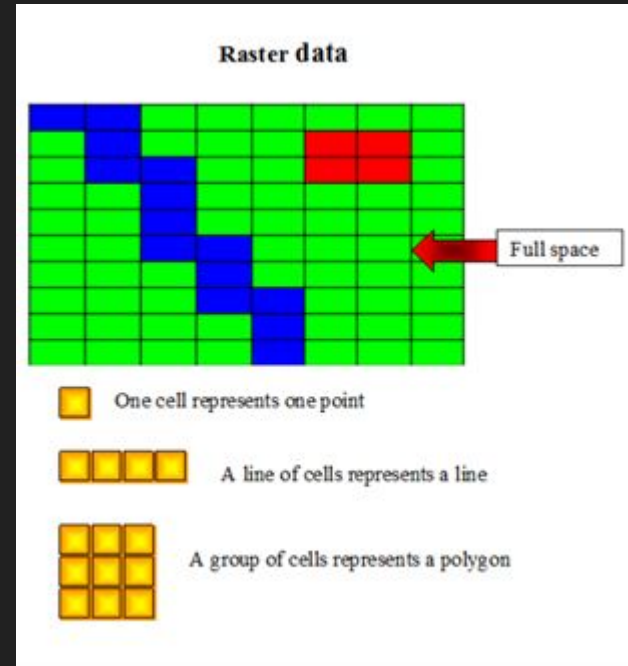
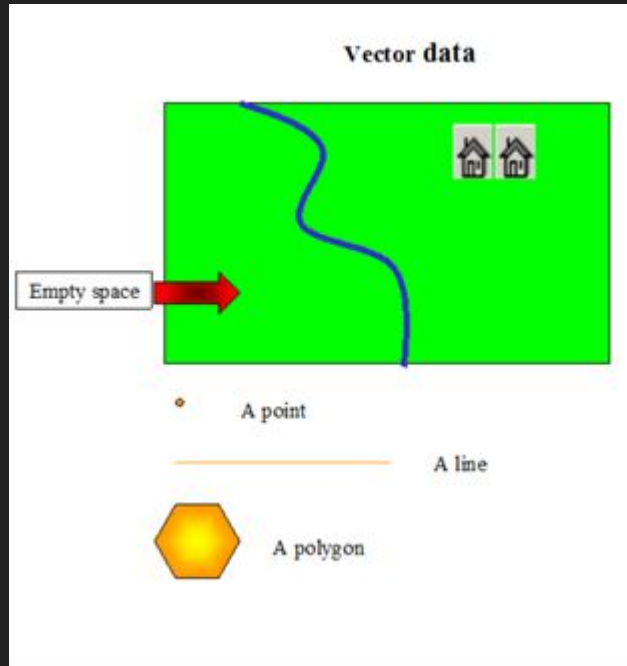


Data Types - Vector and Raster



<https://www.neonscience.org/dc-raster-data-r>

Data Types - Vector and Raster



Vector Data

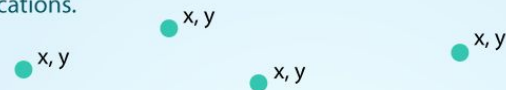
- Resolution-less
- Comprised of 'features'
- Common data formats include:

shapefile / geopackage / CSV /
GeoJSON



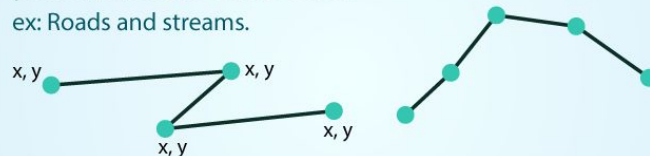
POINTS: Individual x, y locations.

ex: Center point of plot locations, tower locations, sampling locations.



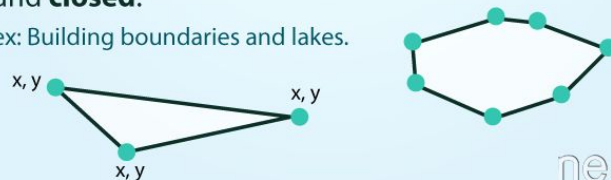
LINES: Composed of many (at least 2) vertices, or points, that are connected.

ex: Roads and streams.



POLYGONS: 3 or more vertices that are connected and **closed**.

ex: Building boundaries and lakes.



neon

Vector Data: The Attribute Table

- A mini-database stored within the vector file
- Each feature is a row and each column an attribute
- Attributes can be any type of data (numeric/character/boolean)
- Most commonly used with SQL to filter/select data based on attributes

ne_10m_admin_0_countries :: Features Total: 255, Filtered: 255, Selected: 0

	featurecla	scalerrank	LABELRANK	SOVEREIGNT	SOV_A3	ADM0_DIF	LEVEL	TYPE	ADMIN	ADM0_A3
1	Admin-0 country	5	5	United Kingdom	GB1	1	2	Dependency	Falkland Islands	FLK
2	Admin-0 country	5	6	Denmark	DN1	1	2	Dependency	Faroe Islands	FRO
3	Admin-0 country	6	5	Australia	AU1	1	2	Dependency	Indian Ocean T...	IOA
4	Admin-0 country	5	5	United Kingdom	GB1	1	2	Dependency	British Indian O...	IOT
5	Admin-0 country	3	6	Singapore	SGP	0	2	Sovereign coun...	Singapore	SGP
6	Admin-0 country	6	5	Australia	AU1	1	2	Dependency	Norfolk Island	NFK
7	Admin-0 country	6	4	New Zealand	NZ1	1	2	Dependency	Cook Islands	COK
8	Admin-0 country	5	4	Tonga	TON	0	2	Sovereign coun...	Tonga	TON
9	Admin-0 country	5	4	France	FR1	1	2	Dependency	Wallis and Futu...	WLF

Raster Data

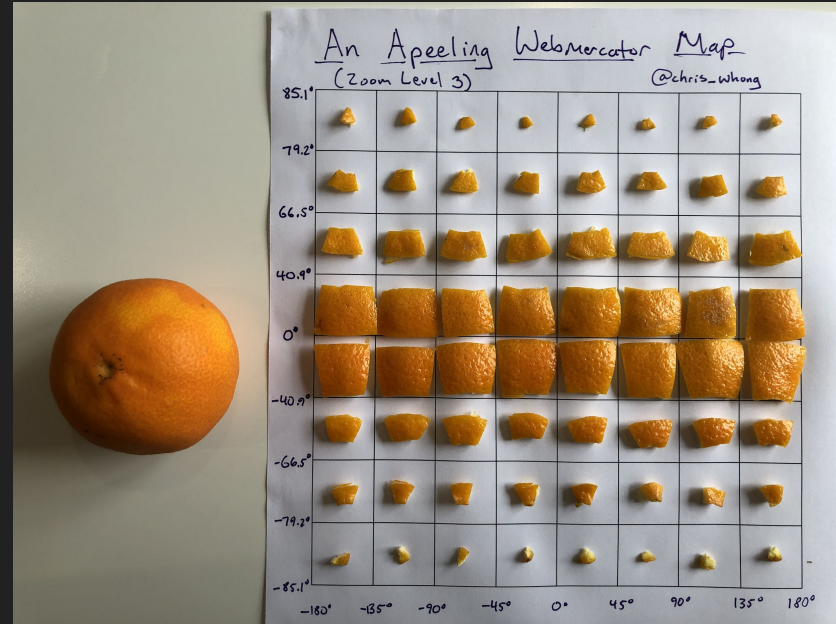
1	1	1	1	1	1	1	1	1	1
1	2	2	2	2	2	2	2	2	1
1	2	3	3	3	3	3	3	2	1
1	2	3	4	4	4	4	3	2	1
1	2	3	4	5	5	4	3	2	1
1	2	3	4	5	5	4	3	2	1
1	2	3	4	4	4	4	3	2	1
1	2	3	3	3	3	3	3	2	1
1	2	2	2	2	2	2	2	2	1
1	1	1	1	1	1	1	1	1	1

- Essentially matrices of numbers with n dimensions.
- Have set resolution/dimensions/extent
- Continuous or discrete
- Typically computationally more efficient to work with
- Common data formats include:

GeoTiff / GRID / ASCII / NETCDF /
JPEG2000



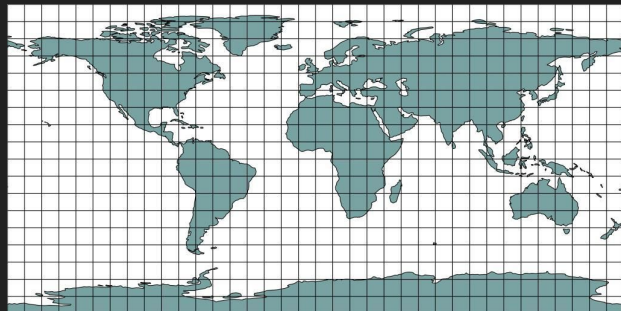
Projections



<https://medium.com/@chris.m.whong/visualizing-the-distortion-of-webmercator-map-s-with-an-orange-peel-cb04460b6415>

Projections

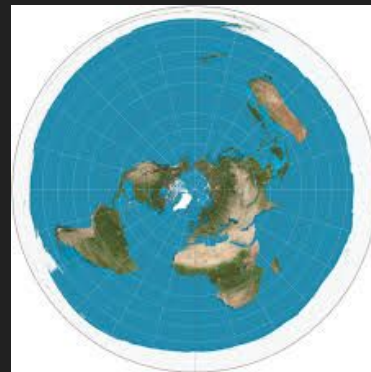
- Geographic vs. Projected projection systems
- A few examples - WGS84 / BNG / Polar
- So many to choose from - need to think carefully about how to marry up your datasets
- Changing projections can distort data, and required resampling. Is this sensible?



Stackoverflow



Wikipedia



Wikipedia

Software Choice - decisions decisions...



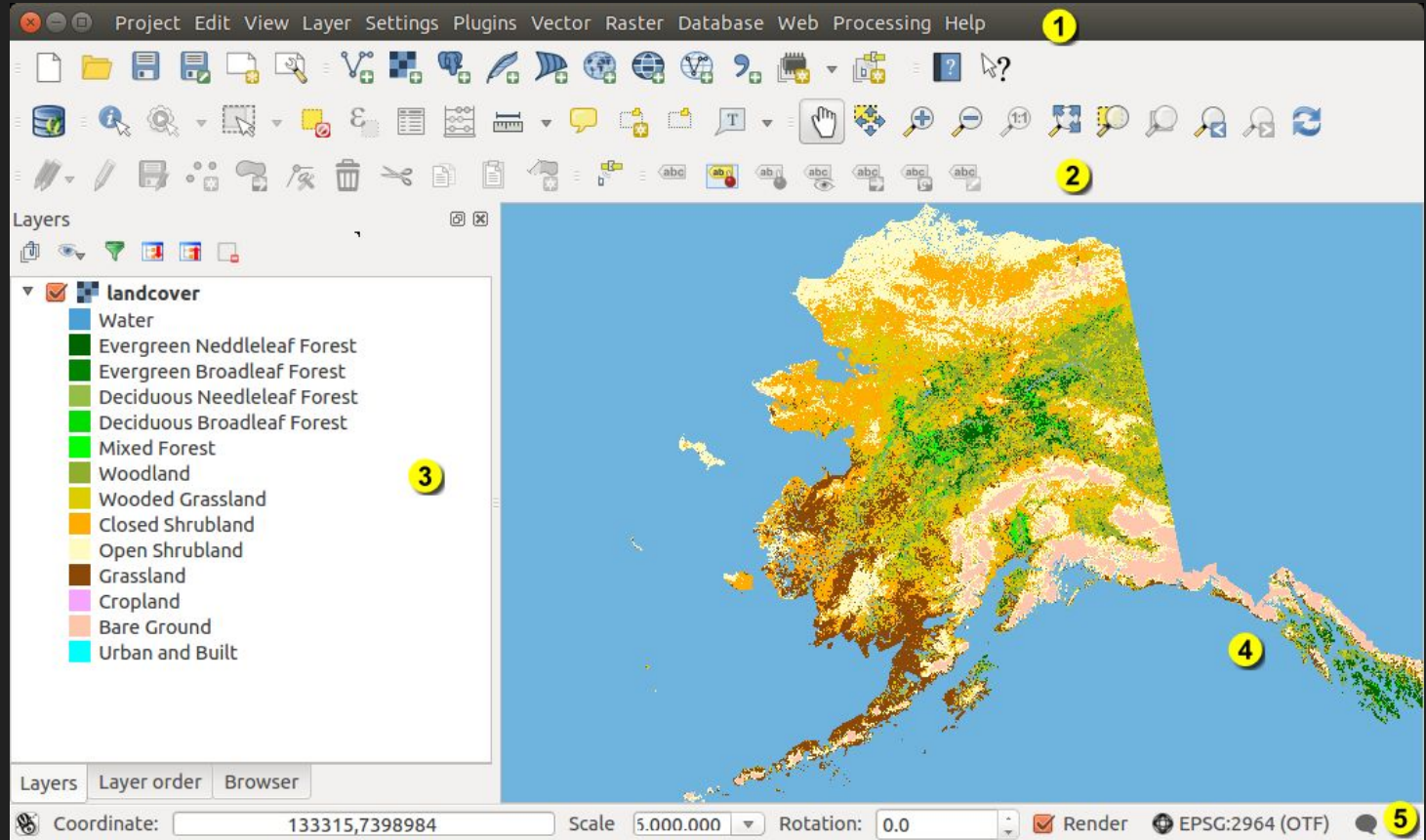
What is QGIS & why use it?

- A free and open source Geographical Information System (GIS)
- Graphical User Interface (GUI) & Programming Interfaces
- Vast array of processing tools + 100s of plugins
- Links to other software packages (e.g. GRASS / SAGA)
- Create high quality maps and figures



The GUI

1. Menu Bar
2. Toolbars
3. Panels
4. Map View
5. Status Bar



Importing your data into QGIS & Tutorial

- Download qgis_data.zip - from github repo
- Vector data - points.gpkg
- Raster data - rast.tif

Now onto a tutorial: https://www.qgistutorials.com/en/docs/3/making_a_map.html

Useful Resources

- <https://www.qgistutorials.com>
- <https://www.globalforestwatch.org/howto/gis/gis-spatial-analysis-in-qgis.html>
- <https://docs.qgis.org/3.4/en/docs/>

Acknowledgements

- Jo Wood - providing content for slides
- JJ Valletta & TJ McKinley - for their time and efforts in setting up the data analytics program