# Using GIS in Behavioural Science and Health

*How you can use GIS in your research*


## Dr Nick Bearman

*nick@nickbearman.com*

*Freelance GIS Trainer & Consultant | UCL Geography*

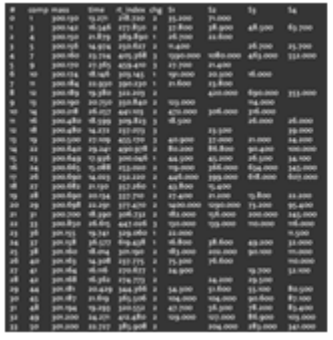

May 18, 2023

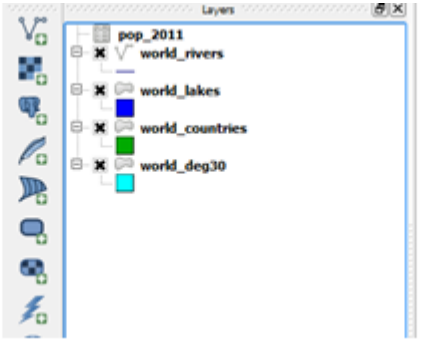
# Using GIS in Behavioural Science and Health

- Welcome!
- Introduce GIS as a tool in your research
- Example of using GIS
- Things to remember
- Strengths and weaknesses
- Where do we find data?
- Spatial Analysis & New Developments
- Software & Common Issues
- Questions / Case Studies

# What is GIS?

- Geographical Information Systems
- Turning (spatial) data into information





- Using this information to answer research questions

→ *"Where are groups of patients more likley to be suffering from obesity?"*

# GIS

- Nearly all human activities & decisions involve geography
  - → the "**where?**"
- Working with geographic information is **different** to working with a **CSV** file
- This is why we need specialized **GIS** software to:
  - → **organize** and **store**
  - → **access** and **retrieve**
  - → **present** and **manage** spatial data
- But ultimately to apply to the solution of our **problem**

# GIS: Systems and Science

- Geographical Information **Systems**

  → The methods, process and technology we use

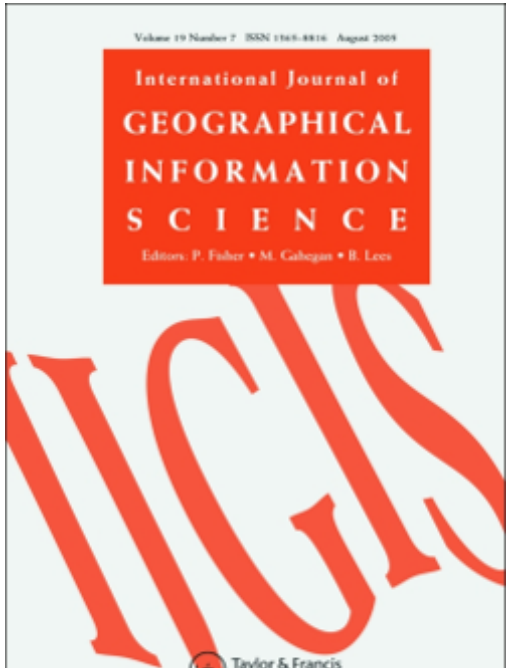
- Geographical Information **Science**

  → The science behind the technology

  → Including development of new technology, methods and processes

- Geographic*al* / Geographic

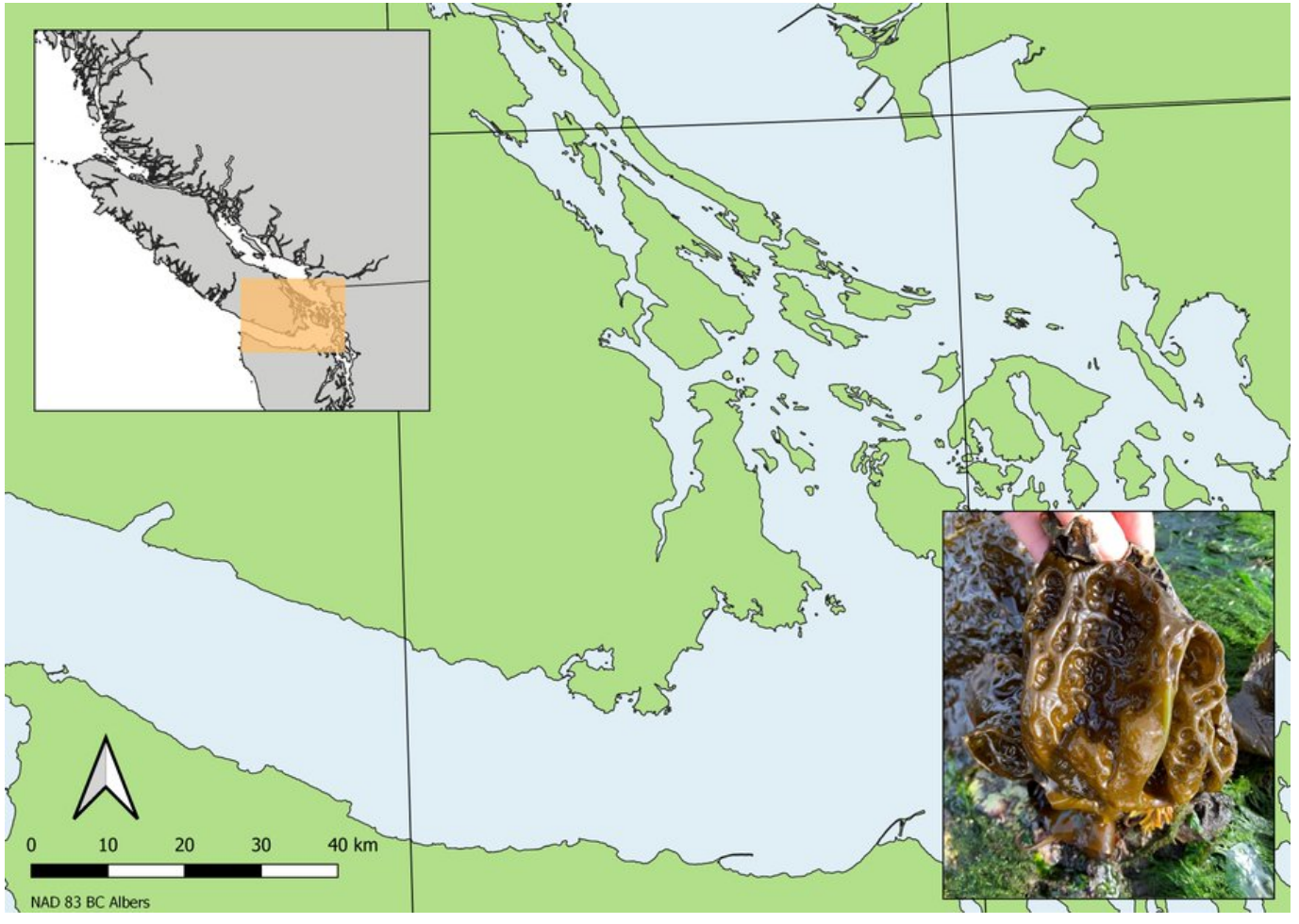
- Geographic Data Science



# Examples of using GIS

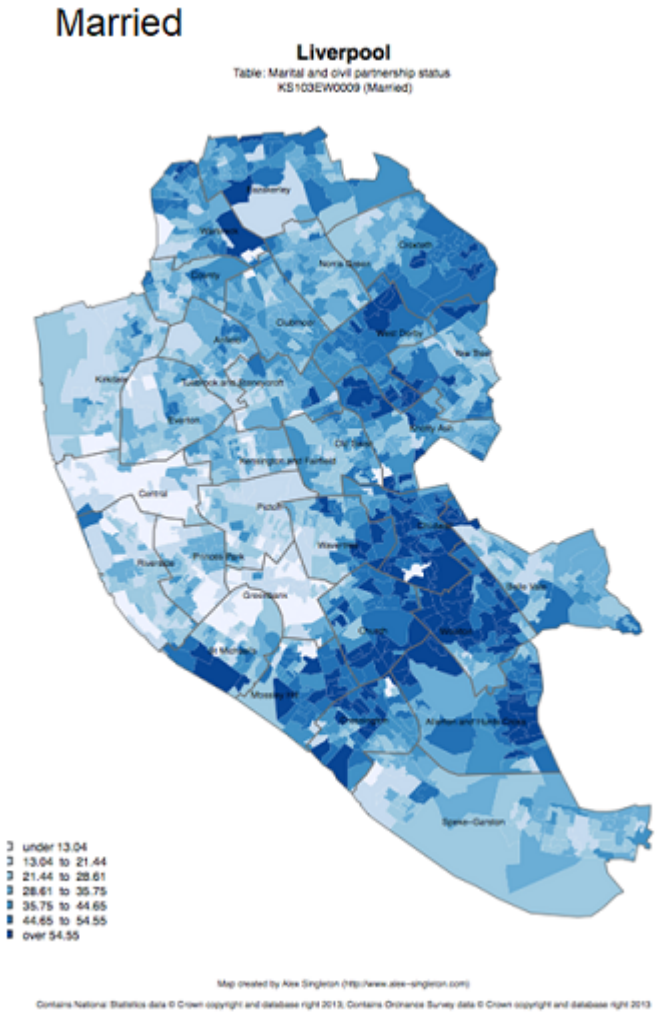
- making maps: your study area
- showing data: choropleth maps, rates of disease
- spatial statistics: is space important? clustering & regression
- spatial analysis: where meets x criteria?

# Making Maps:



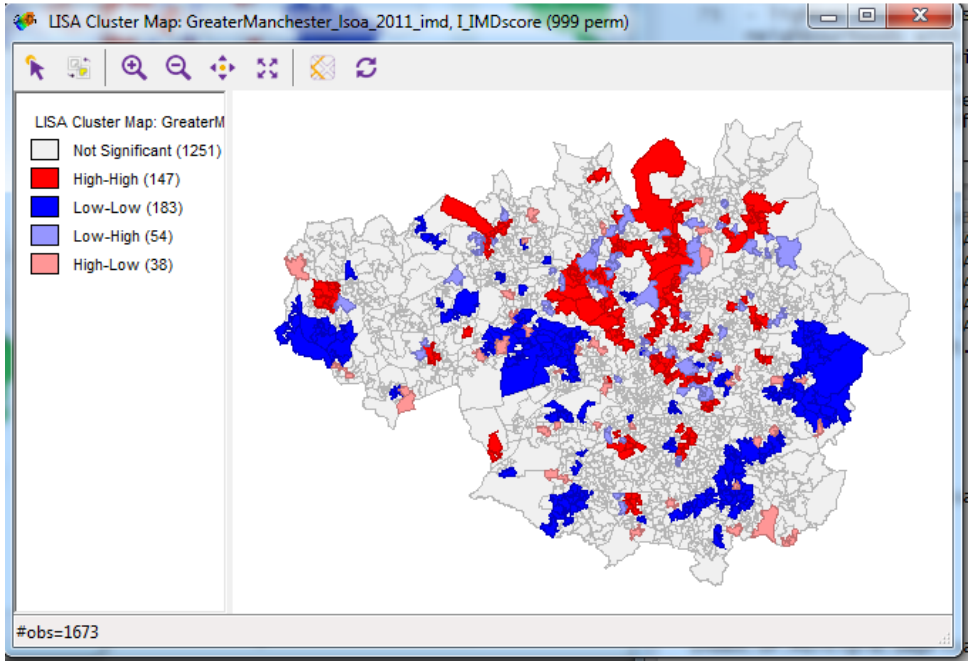

- Your study area
- *Getting excited for my first summer of fieldwork with a study site map I created during today's GIS workshop at #PEEC2022*
- Rebecca Hansen, Twitter

# Showing Data:




- Choropleth maps
- Rates of disease
- % Married, Census data, 2011

# Spatial Statistics:



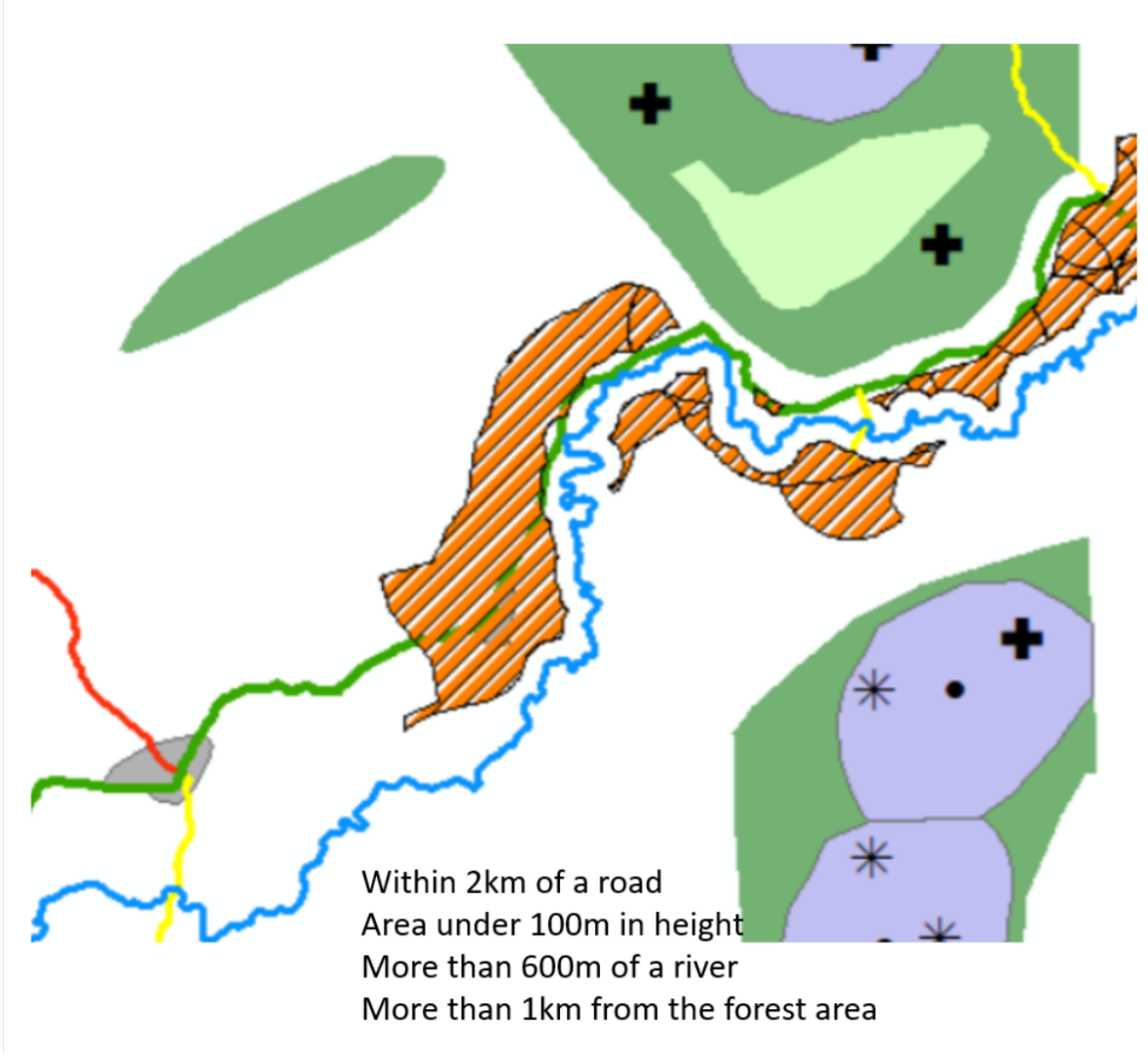

- Is space important? (Spatial autocorrelation / Moran's I)
- Clustering? (Local Moran's I)
- Regression? (What is driving these relationships?)

  → GWR - geographically weigthed regression

  → Full spatial regression

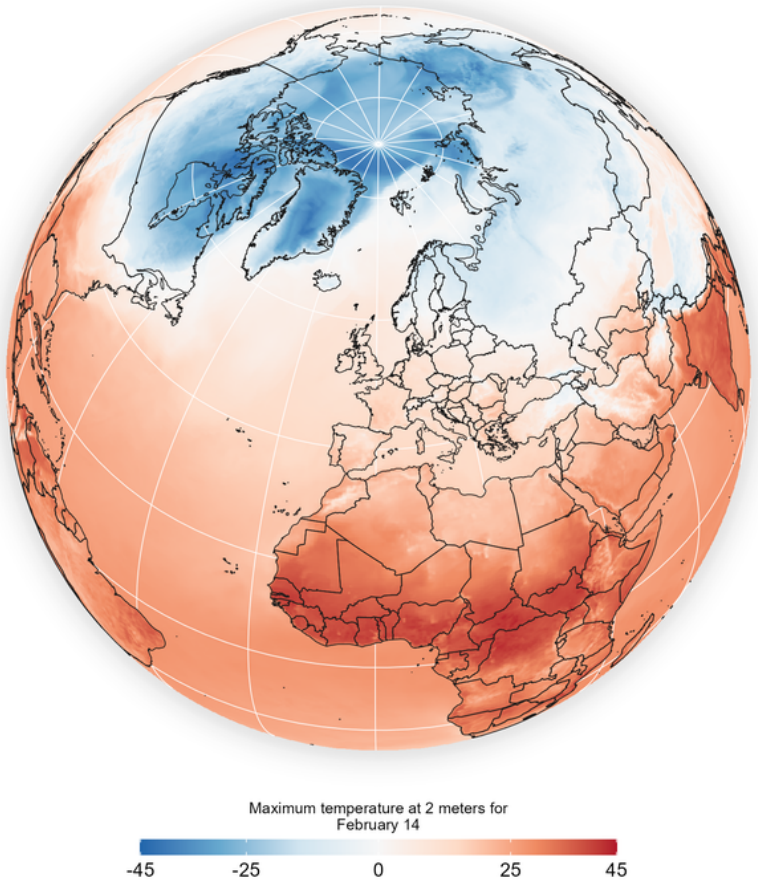
*Map of Greater Manchester, UK, showing clustering areas of high deprivation (red) and low deprivation (blue) (measured using IMD).*

# Spatial Analysis:




- Where meets x criteria?
- Which site meets criteria for a new hospital?
- How can public transport help address deprivation?
- Where should we route a water pipe?

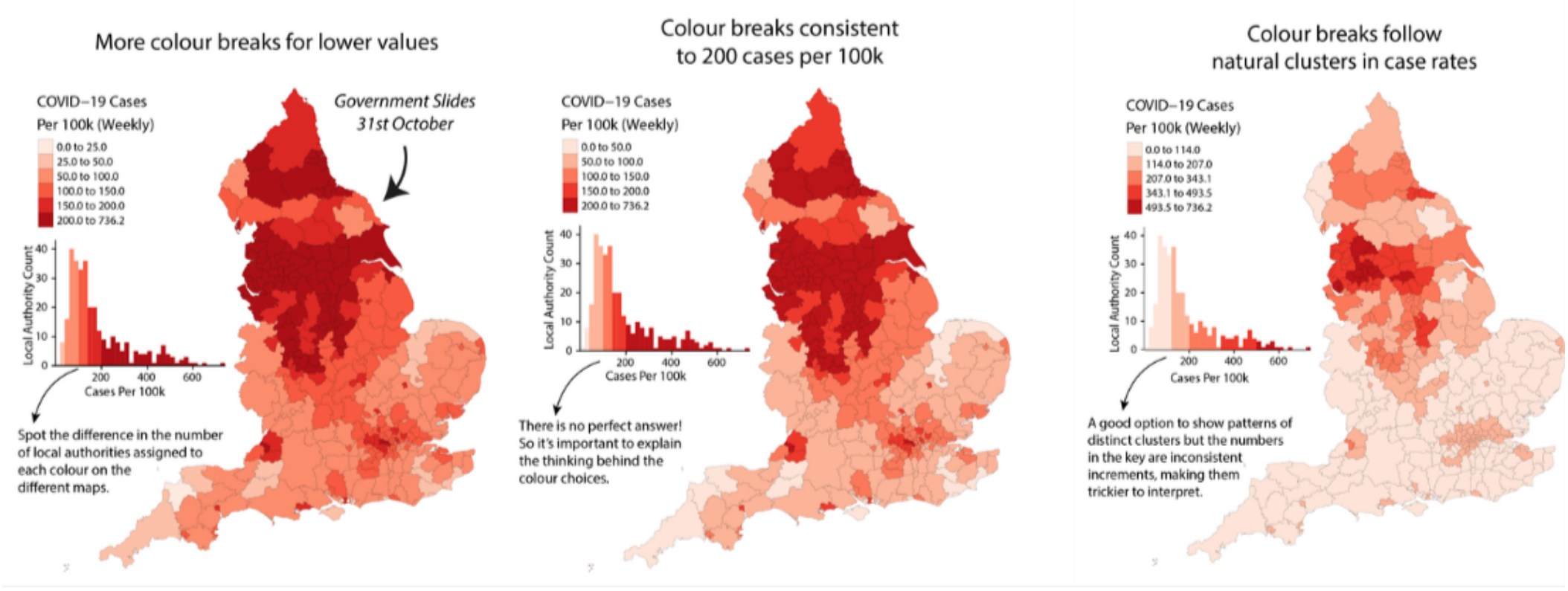
# Weather




Tomorrows Weather - https://dominicroye.github.io/en/2023/tomorrows-weather/?utm_source=puntofisso&utm_medium=email

# Things to remember!

- Like *"correlation <> causation!"*
- Be critical of maps - they are not what they seem

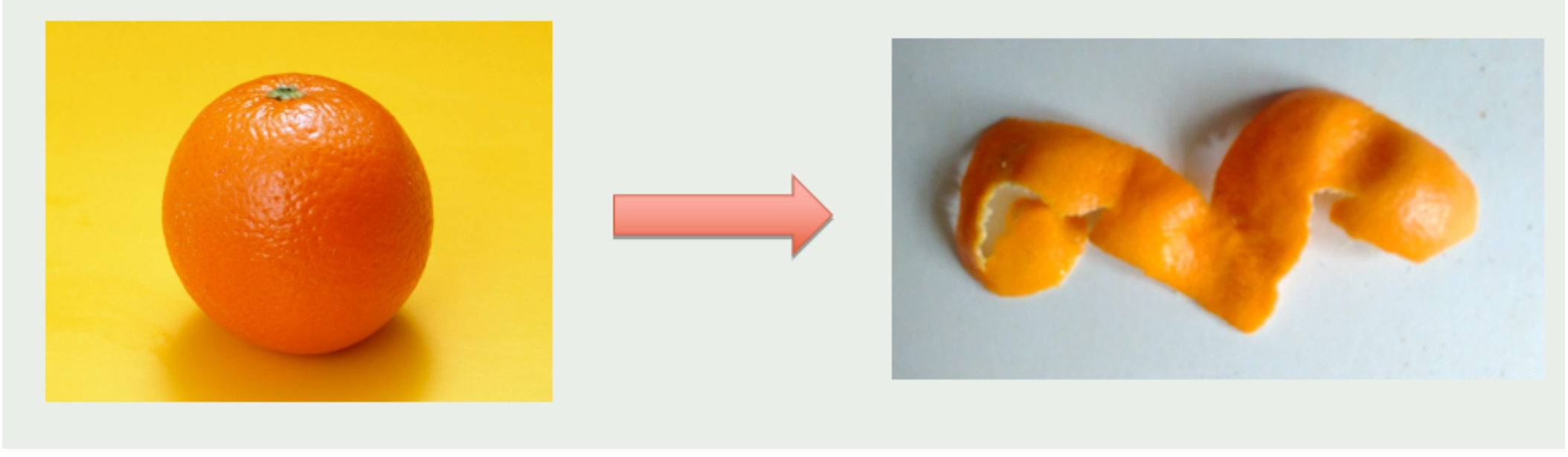


# Things to remember!

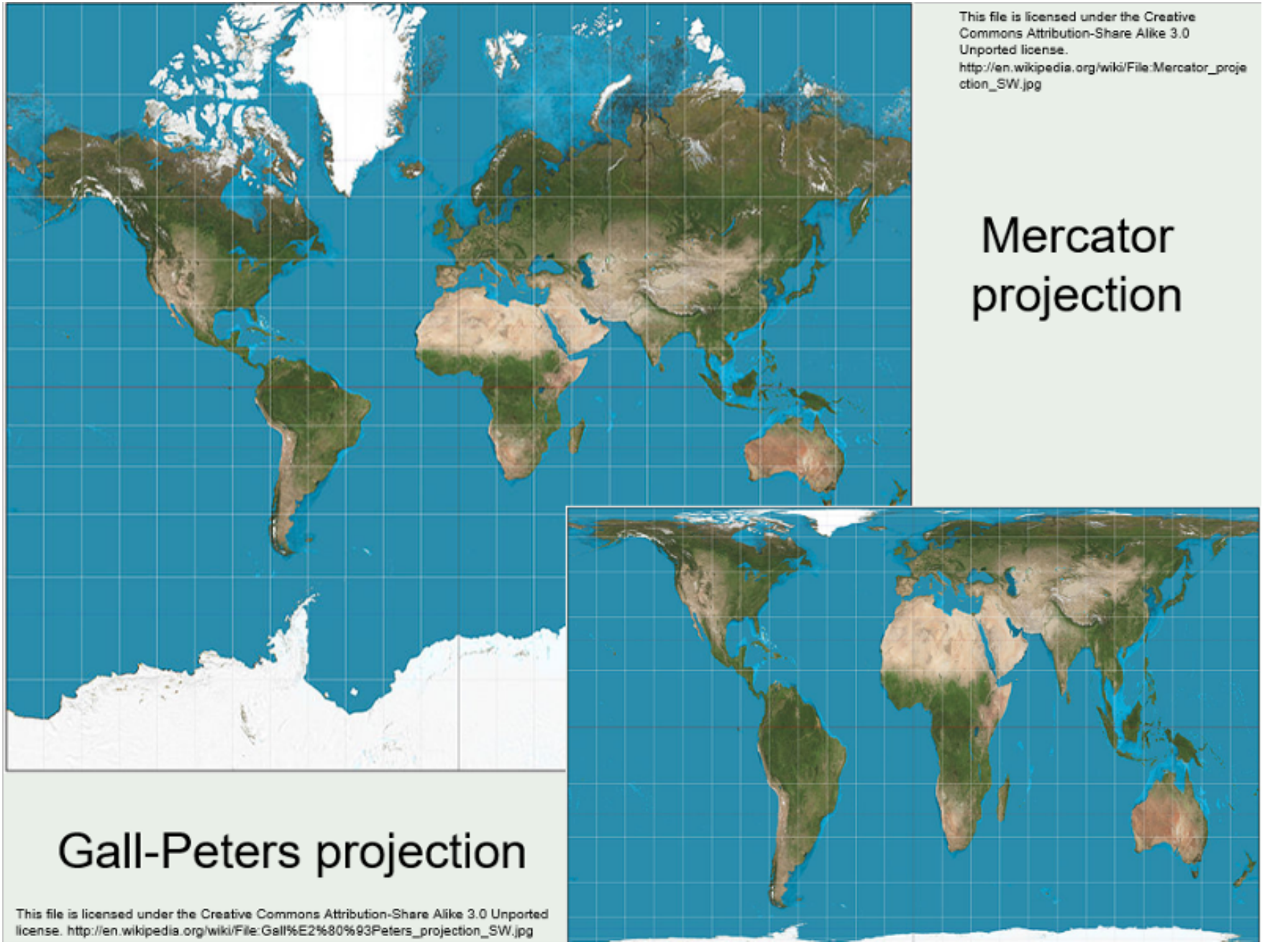
- Projections and Coordinate Systems
- How we show data
- Many many different types

# Projections

- Projection - going from a sphere to a flat surface



http://www.hdwallpapersos.com/orange-fruit-hd-wallpapers.html

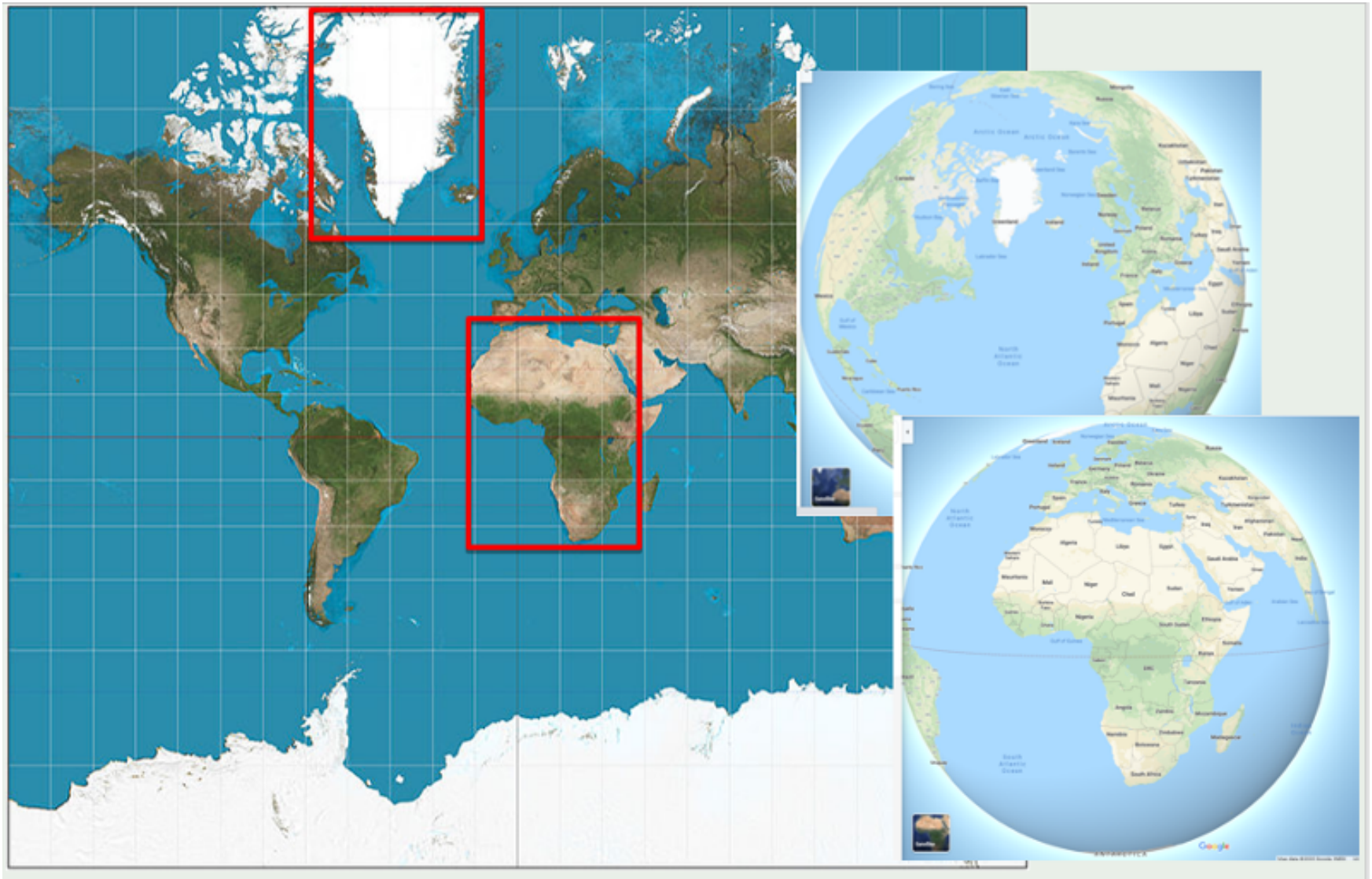


This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.
http://en.wikipedia.org/wiki/File:Mercator_projection_SW.jpg

Mercator projection

Gall-Peters projection

This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. http://en.wikipedia.org/wiki/File:Gall%E2%80%93Peters_projection_SW.jpg



Africa: 30 million sq km

Greenland: 2 million sq km

# Mercator

- *The son of a poor shoemaker near Antwerp, Belgium*
- *The Father of Modern Cartography*
- *By Mercator's Projection the navigators of the succeeding centuries sailed on their voyages of discovery*
- 1569

Sefton Park

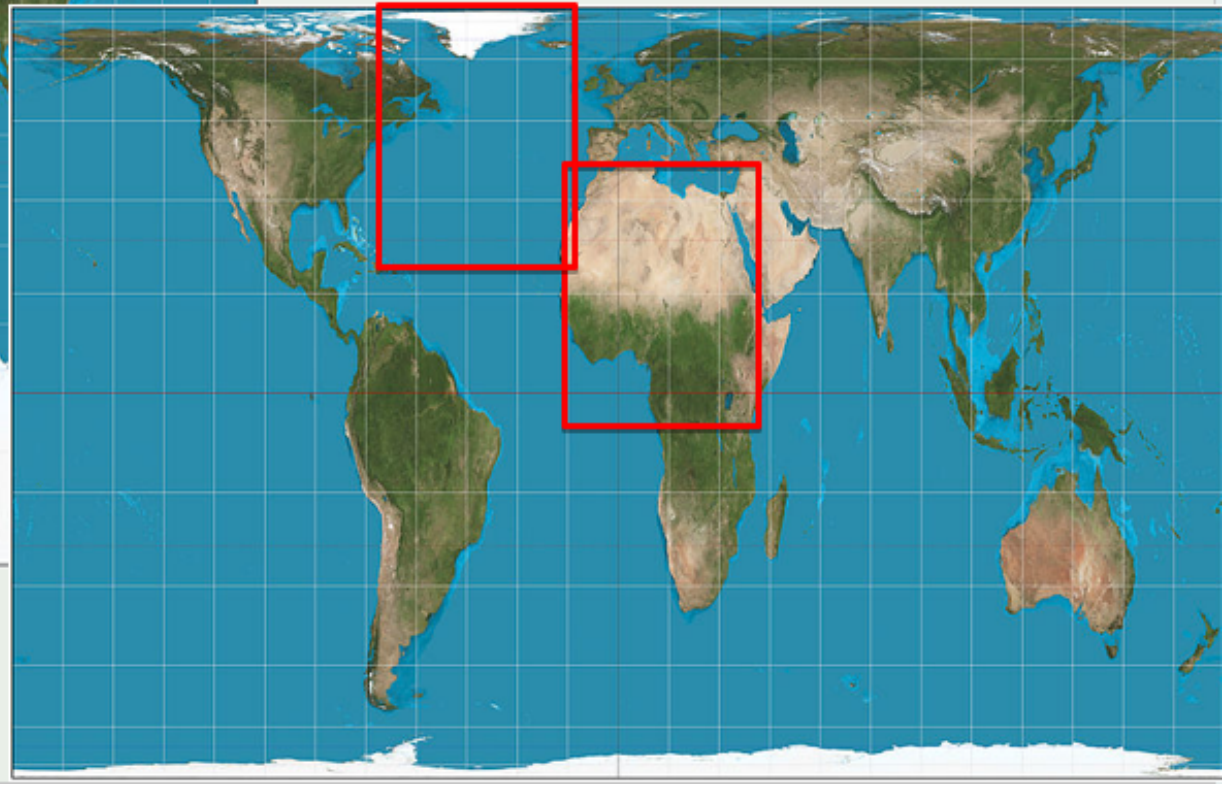
Gall-Peters projection

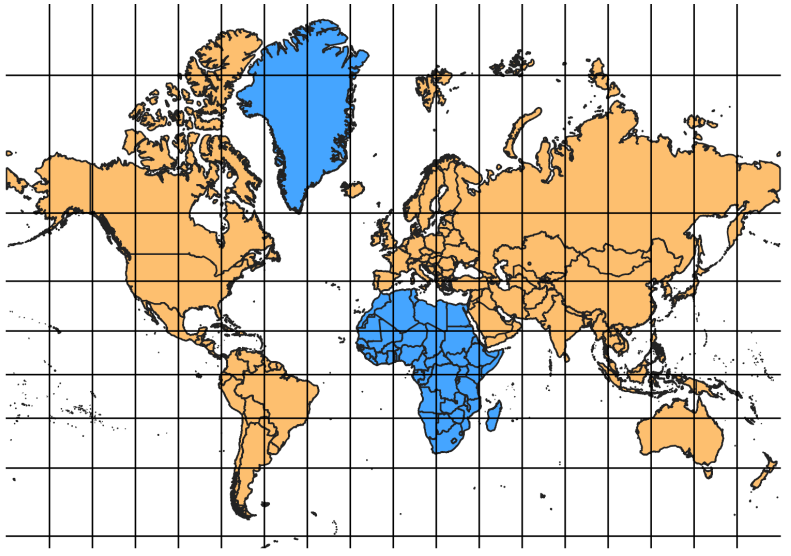
# Mercator projection

Reprojection

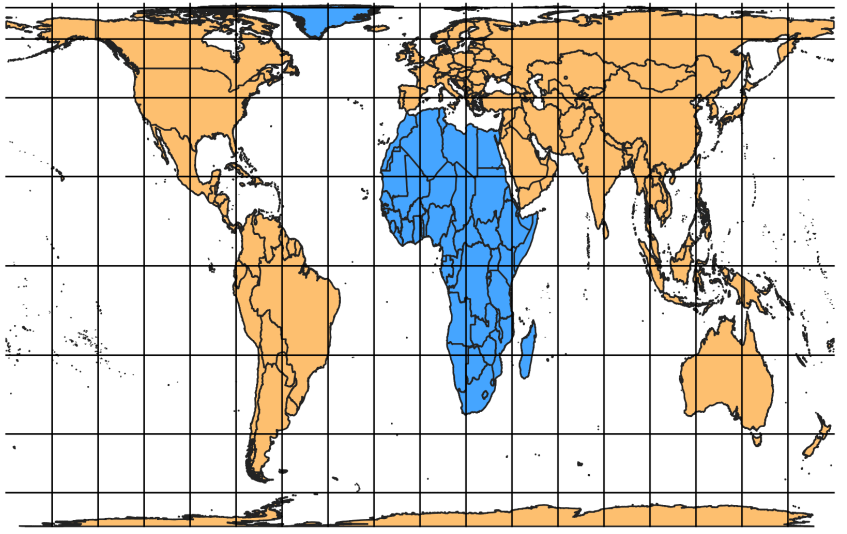
Can't maintain area and angles

Mercator: **angles**
Peters: **area**

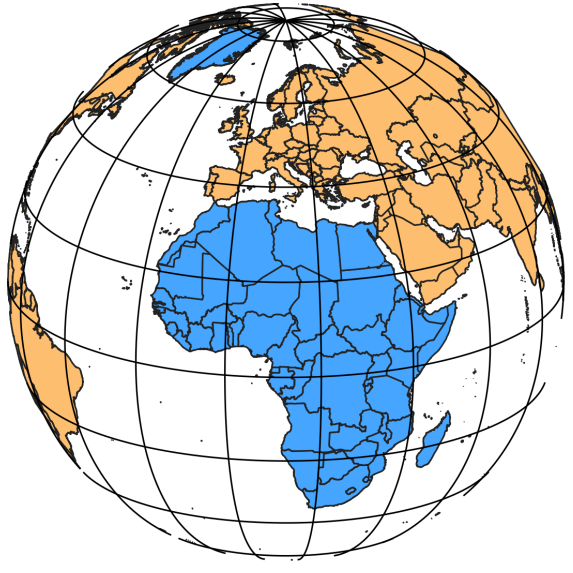




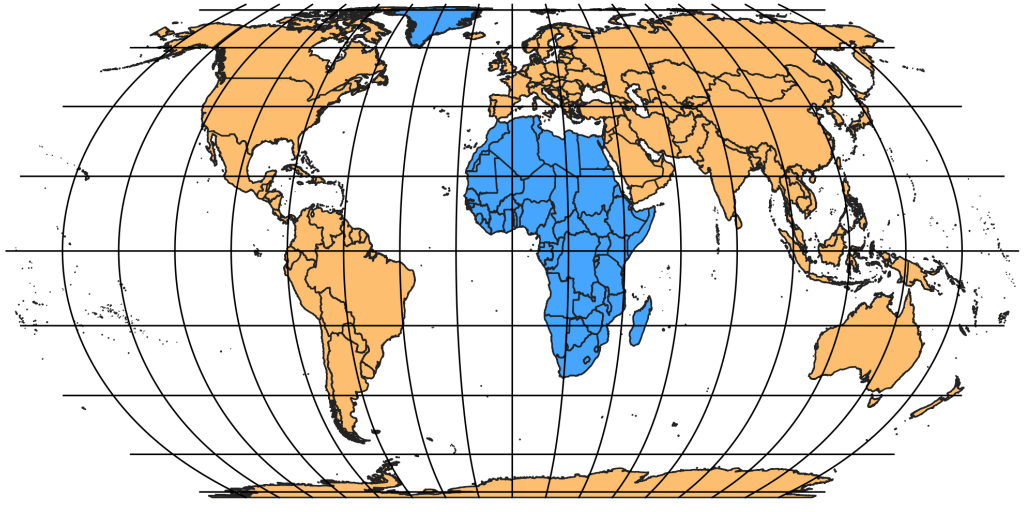
a) Mercator EPSG: 3395



b) Peters EPSG: SR-ORG:22



c) Adjusted version of The World From Space ESRI: 102038

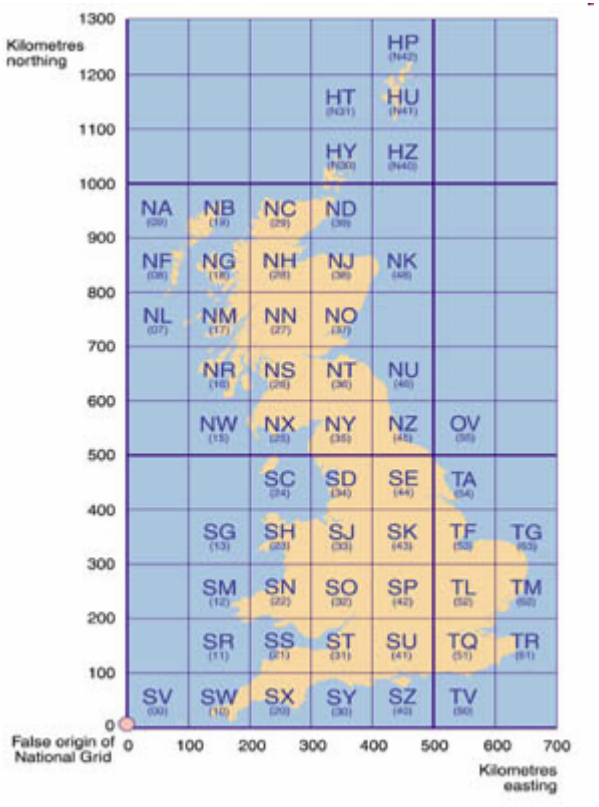


d) Equal Earth Greenwich EPSG: 8857

# Coordinate Systems

How we specify location

# Coordinate Systems




- British National Grid
- *Easting: 619301*
- *Northing: 307416*

# Strengths and Weaknesses

- You need **data** to be able to do anything
- Availability of data is key
- As is quality of data, who collected it, bias, etc.
-
- Can be quite quantitative - not everything can be quantified!
- Some qualitative GIS, but very much the junior partner
- Participatory GIS an interesting area

# Data

- where does it come from?
- **spatial** & **non-spatial** data
- **spatial** - the geographic areas
- **attribute data** - the spreadsheet

# Joining

*Polygons*

| Area ID |
|---|
| 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 14 |
| 7 |
| 8 |
| 11 |

*Data*

# Joining

*Polygons*

| Area ID |
|---|
| 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 14 |
| 7 |
| 8 |
| 11 |

*Data*

| Area ID | Deprivation |
|---|---|
| 1 | High |
| 2 | High |
| 3 | High |
| 4 | Average |
| 5 | Average |
| 6 | Average |
| 7 | Low |
| 8 | Low |
| 9 | Low |
| 10 | High |
| 11 | Low |
| 12 | High |
| 13 | High |
| 14 | Average |
| 15 | Average |

*Lookup*

# Joining

*Polygons*

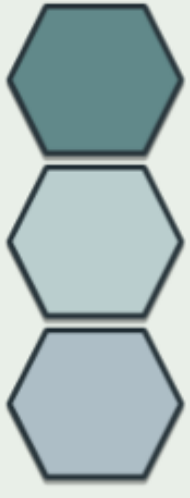
| Area ID |
|---|
| 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 14 |
| 7 |
| 8 |
| 11 |

*Data*

| Area ID | Deprivation |
|---|---|
| 1 | High |
| 2 | High |
| 3 | High |
| 4 | Average |
| 5 | Average |
| 6 | Average |
| 7 | Low |
| 8 | Low |
| 9 | Low |
| 10 | High |
| 11 | Low |
| 12 | High |
| 13 | High |
| 14 | Average |
| 15 | Average |

*Lookup*

| Area ID | Deprivation |
|---|---|
| 1 | High |
| 2 | High |
| 3 | High |
| 4 | Average |
| 5 | Average |
| 14 | Average |
| 7 | Low |
| 8 | Low |
| 11 | Low |



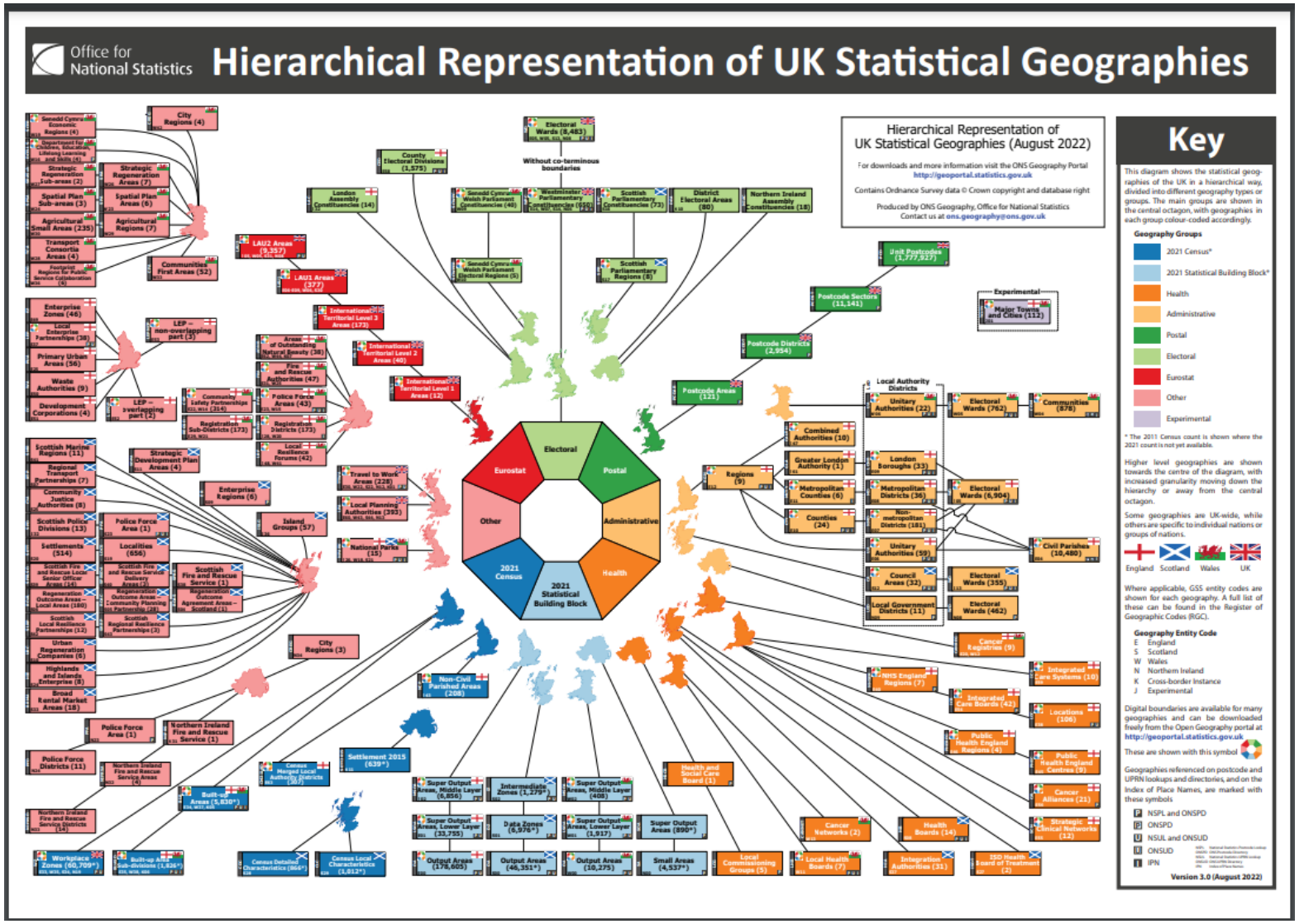
High Deprivation

Average Deprivation

Low Deprivation

# Spatial Units

- Many many different spatial units
- Countries, counties
- Local authorities, GOR
- MSOA, LSOA, OA

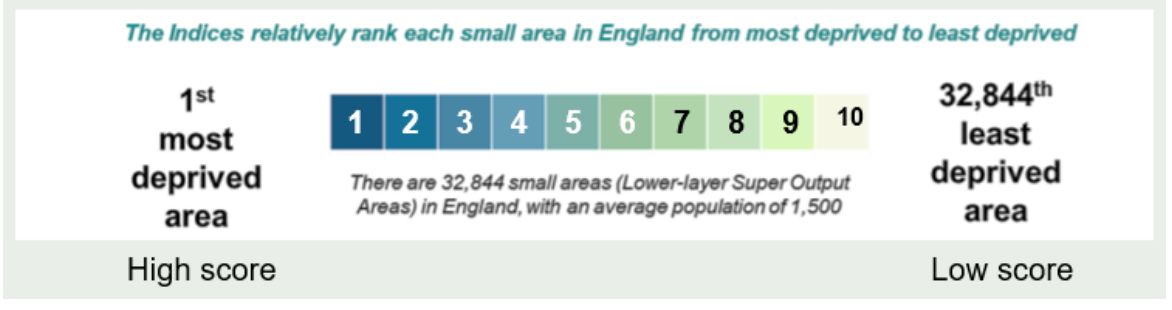



Hierarchical Representation of UK

Statistical Geographies (August

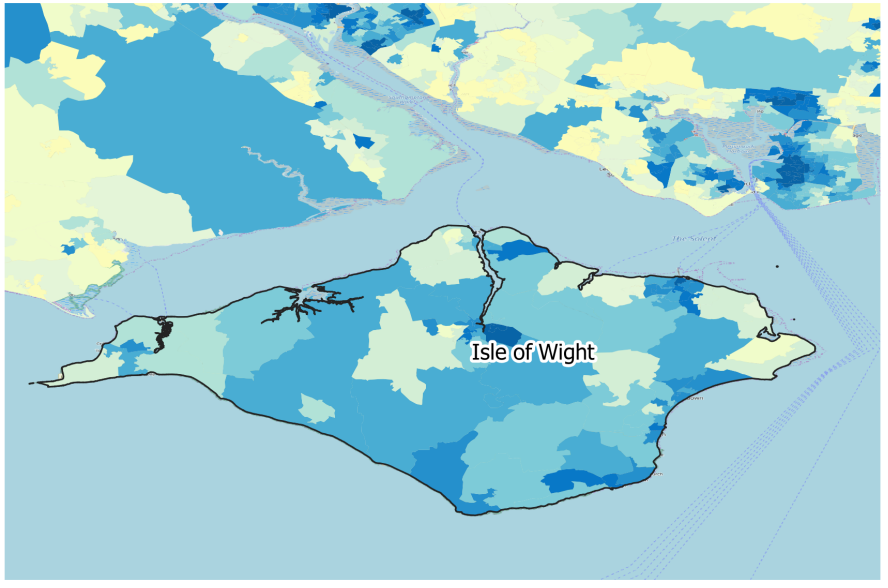
# IMD - Index of Multiple Deprivation

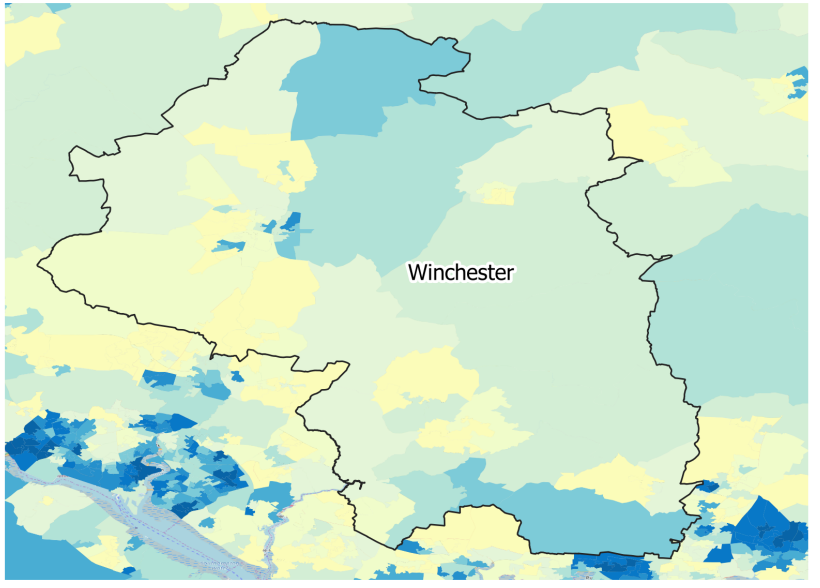
- A measure of deprivation
- 2019 (England)
- 2019 (Wales)
- 2020 (Scotland)
- Rank, score & decile

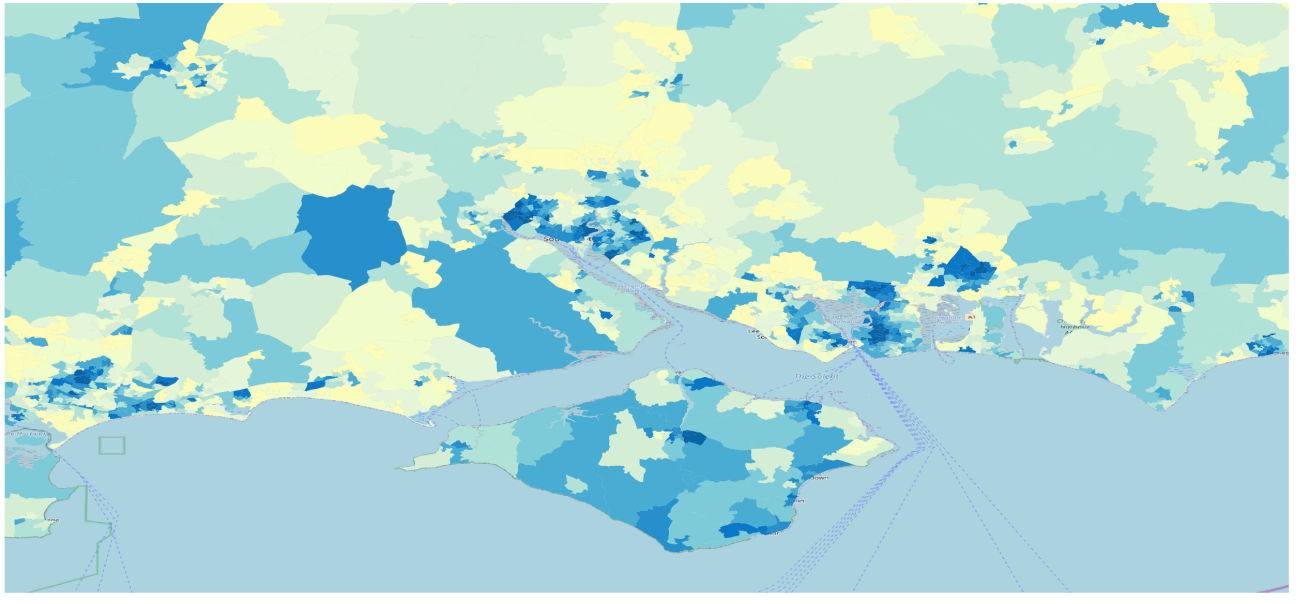



England: 7 domains

- Income
- Employment
- Health
- Education
- Access to Services & Housing
- Crime
- Living Environment



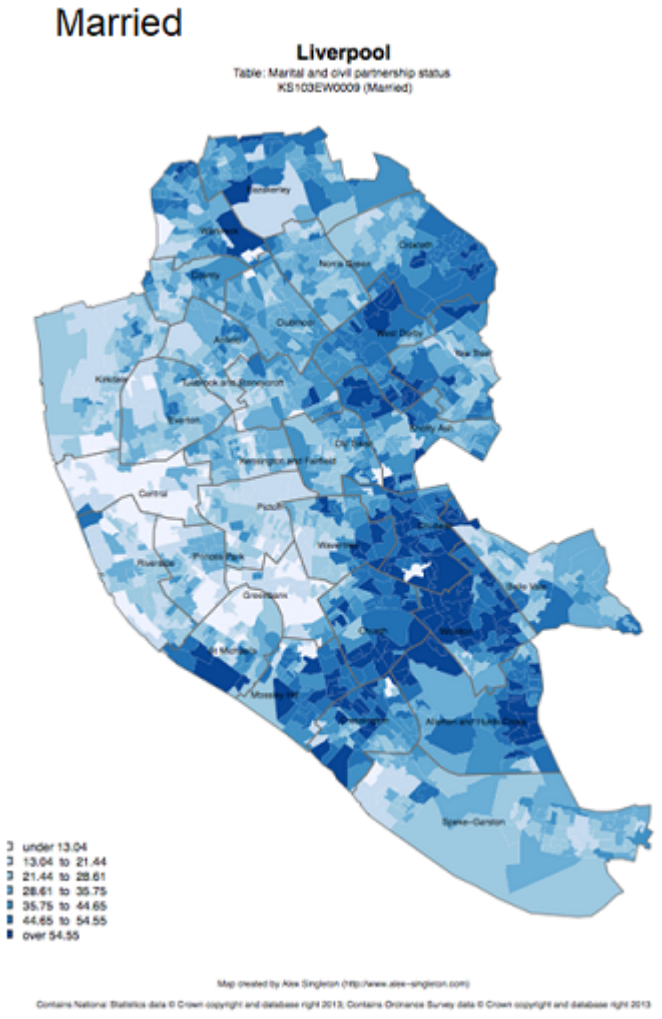



IMD Deprivation Decile
1 More deprived
2
3
4
5
6
7
8
9
10 Less deprived
Local Authority boundaries

0 5 10 15 20 25 km

MHCLD, 2019. Contains National Statistics data
© Crown copyright and database right 2020.

# Census Data




- Population
- Age groups
- Gender
- Marital status
- Ethnicity
- Health
- ....

# GIS Software

- ArcGIS Pro most popular commerical product
- QGIS most popular open source
- R / RStudio open source, command line based
- Reproducible analysis becoming more important
- *All can be useful*

# New developments

- AI
  - → Some automated interpreting of satellite data
  - → Houses / buildings / water / fields
- LLM / ChatGPT / Bard
  - → Clear learning / assessment challenges
  - → Also clear learning potential
  - → Potentially really useful for writing code

# Common issues in GIS

- **Coordinate systems**

  → "9 out of 10 problems when using GIS results from using the wrong coordinate system"

- Finding (the right) **data**

  → "You can easily spend 50% of the time of your project looking for and bringing in data"

- **Postcodes** as geocoding

  → "WC1E 6BT" is a great way of asking people where they live, a residential postcode unit is about 10 - 15 houses

  → But (e.g. in Wales) this can vary from 0.6 ha (a football field) to 5404 ha (20 sq miles, ~ Ealing)

# Moving forward

- Books
  - → GIS: Research Methods | *First chapter free*
  - → Cartography: An Introduction
  - → Discover QGIS 3.x
  - → Using R as a GIS | preview chapter
- Training courses | June *Intro to R* | July *Essex Summer School*
- Data | Free GIS Data | GeoBoundaries
- Looking at other people's maps | #30DayMapChallenge

# Thank you :-)

- Questions
- Case studies

  → *“Nick will pick one or two case studies to talk through the process of how we might go about mapping or doing spatial analysis on those projects”*