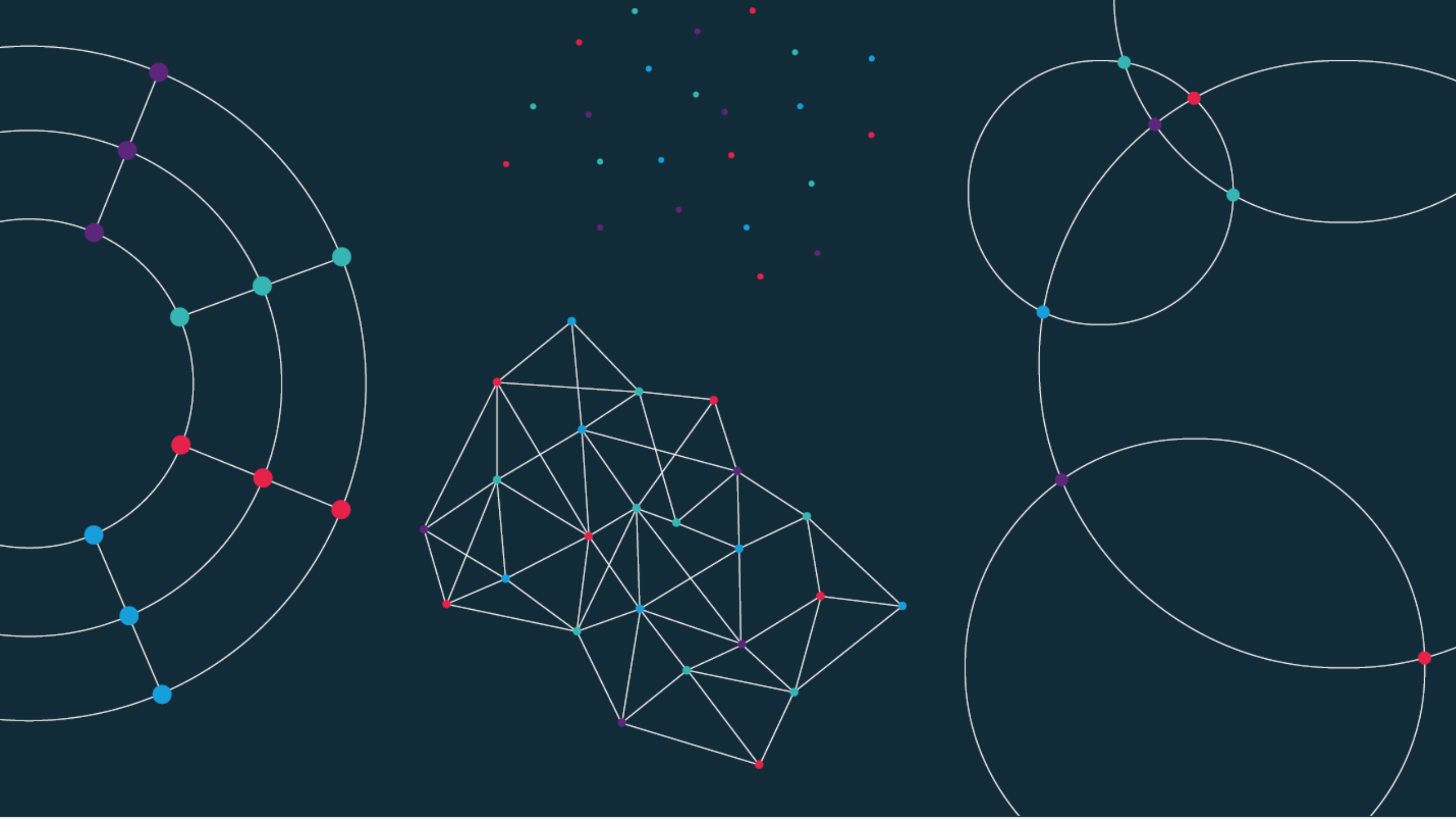


Practical Data Science.

For economists

Week 8.

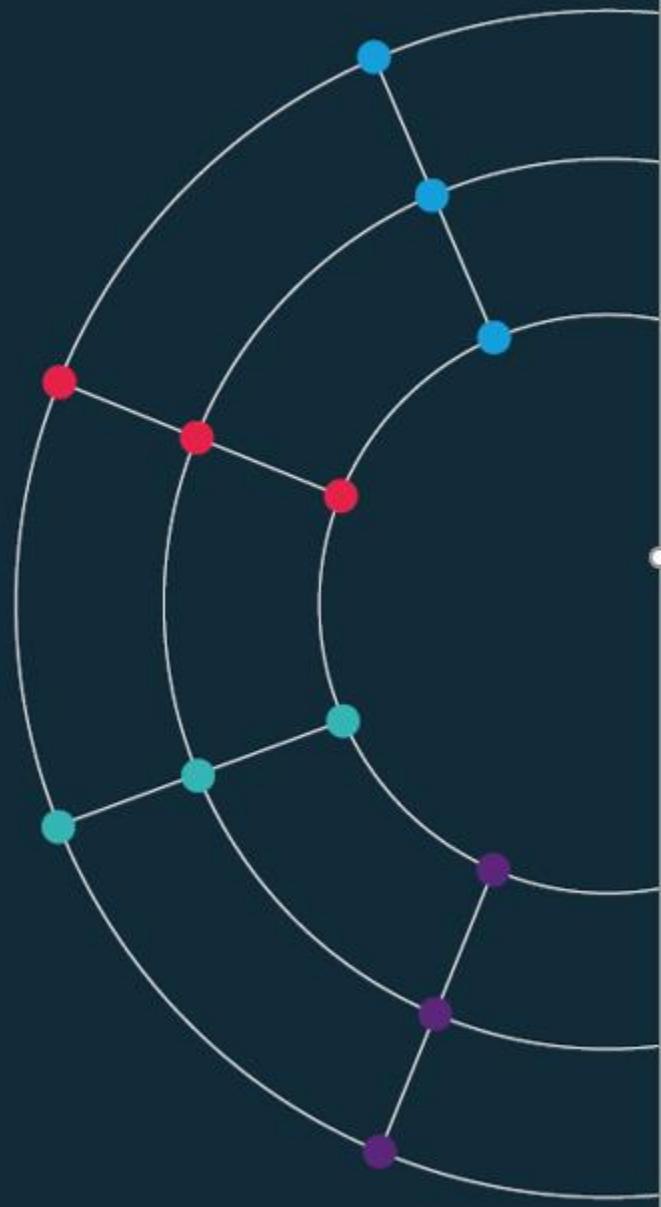
University of Bristol
Autumn 2024

Week 8.

- Motivation. Some maps from the ECO team
- History. Origins of social cartography.
- Mapping today. Modern examples.
- How to make maps. The nuts and bolts of cartography.
- Code-along. Building your first maps.

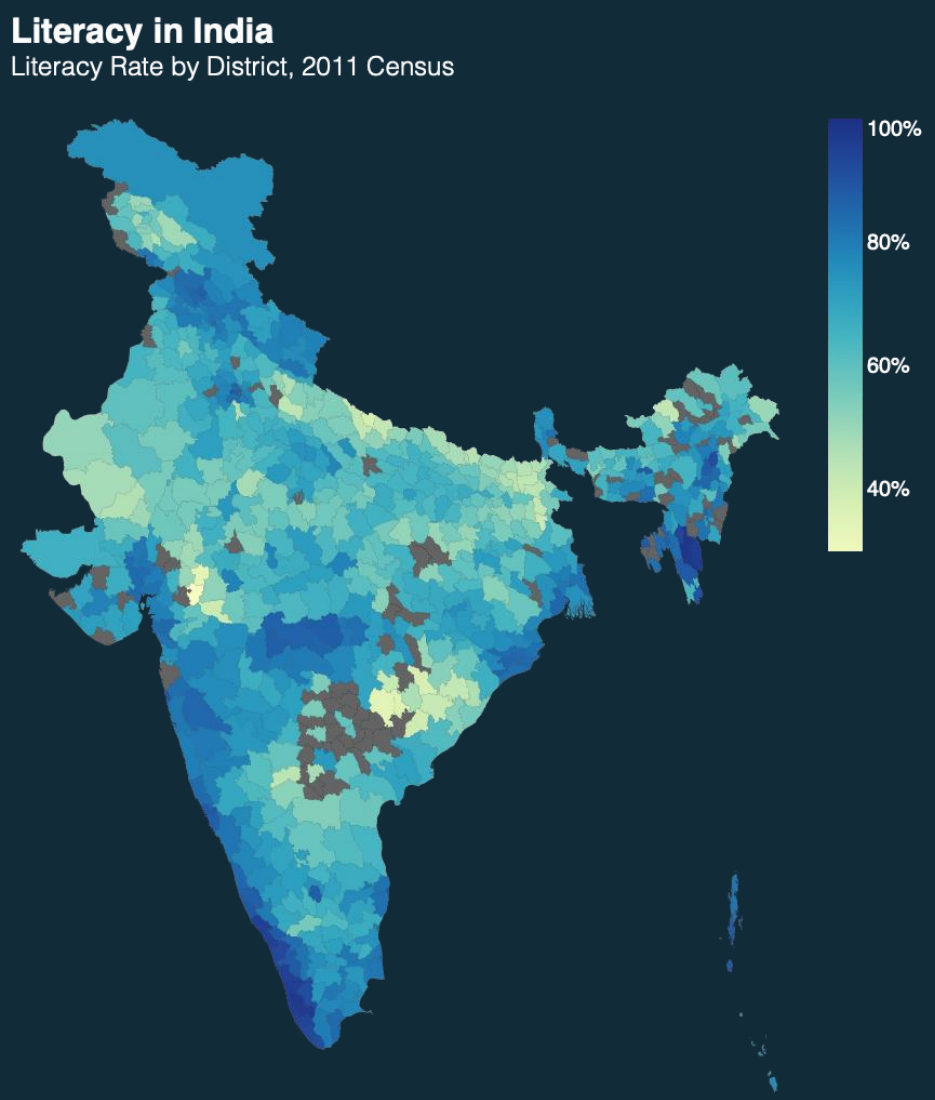
8. Maps.

Data in geographic context



The world’s biggest economy.

India



Every ten years, the Indian state conducts the largest census of any democracy.

The last was completed in 2011 – the 2021 census has been delayed post-Covid.

Two and a half million census workers visited thousands of towns and 600,000 villages to ask 27 questions of 1.2 billion people.

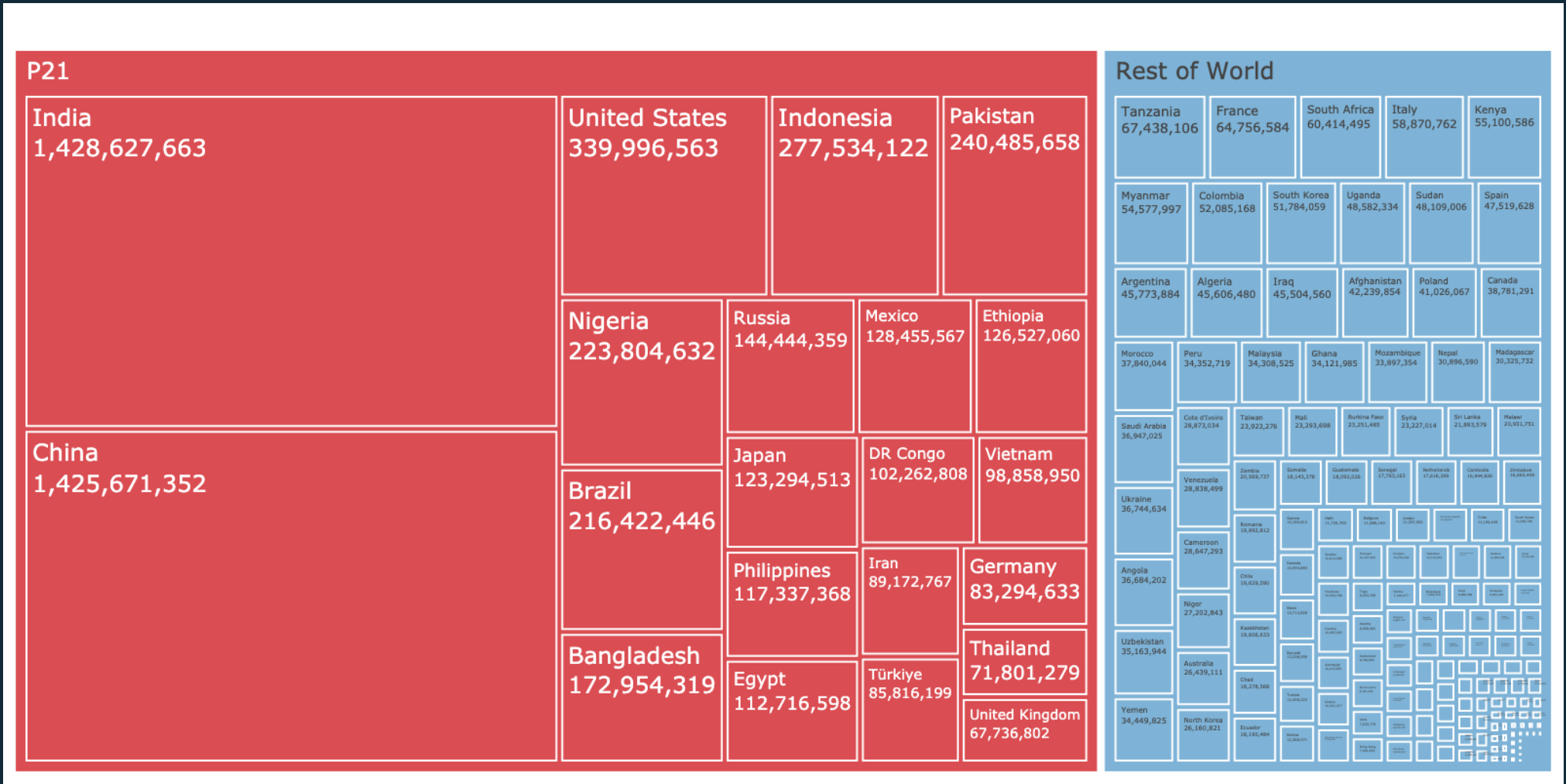
This map shows the literacy rate of each of India’s districts.

20 economies that drive the world.

Plus the UK, number 21…

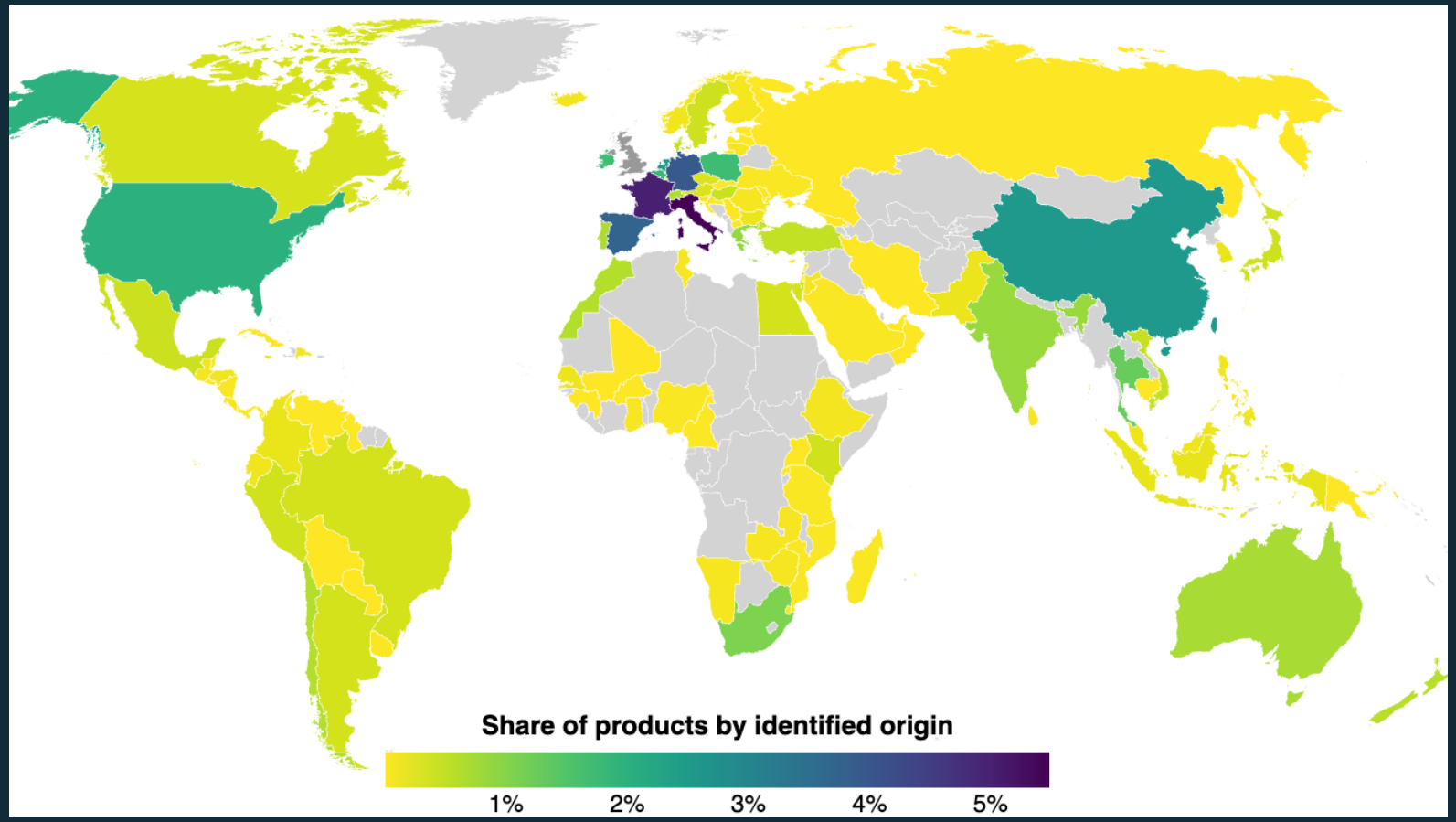
Economies that drive the world.

21 countries make up 71% of the global population



Where does Britain's food come from?

Our research.



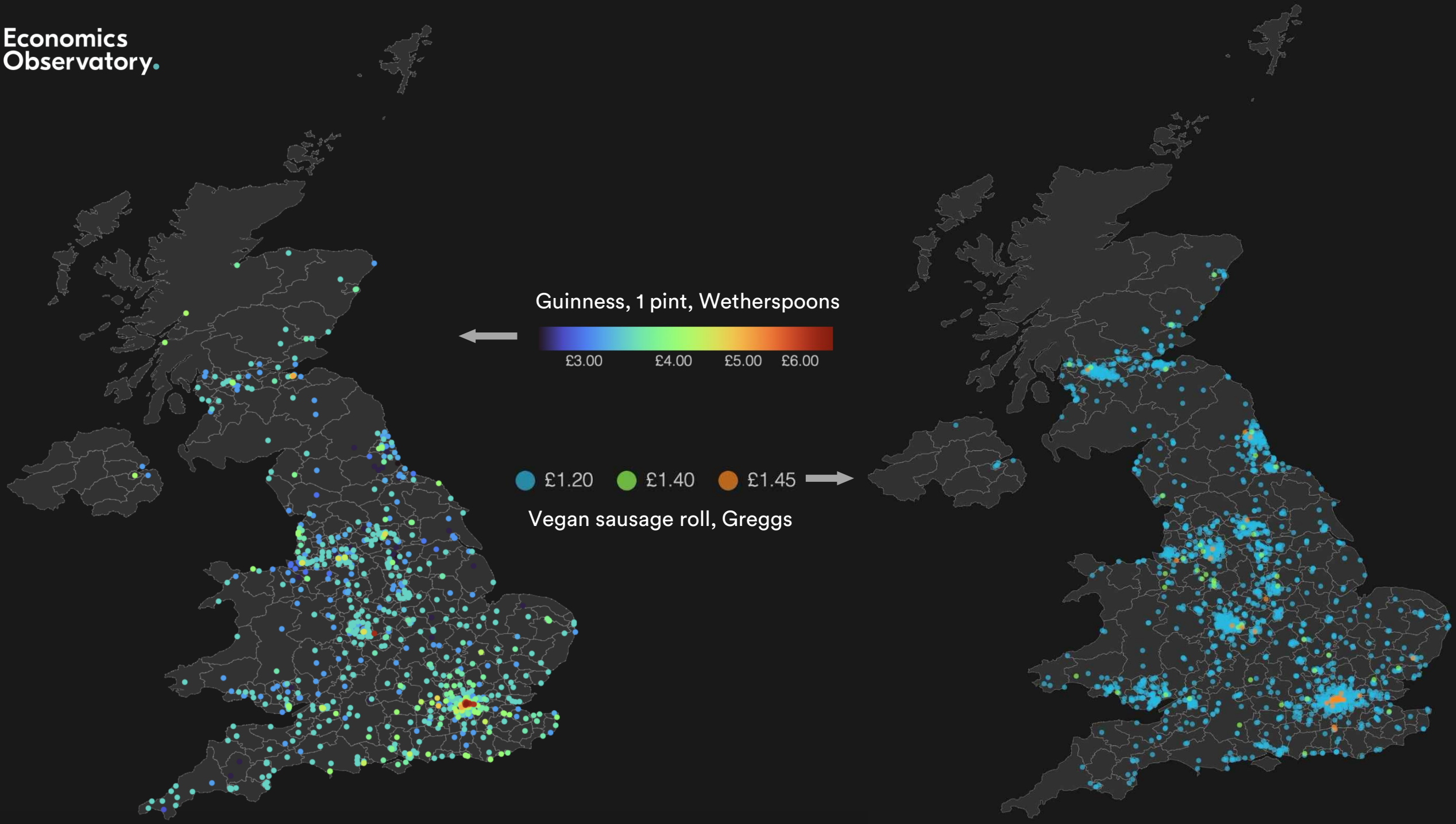
Across 67,000 products from 6 UK supermarkets.

Just under half are of UK origin (49%).

Top five foreign sources of food (in order) are Italy, France, Germany, China and Spain.

Prodcuts sourced from at least 71 countries.

Economics Observatory.



8.1 History.

Why do humans make maps?

Cartography.

A potted history

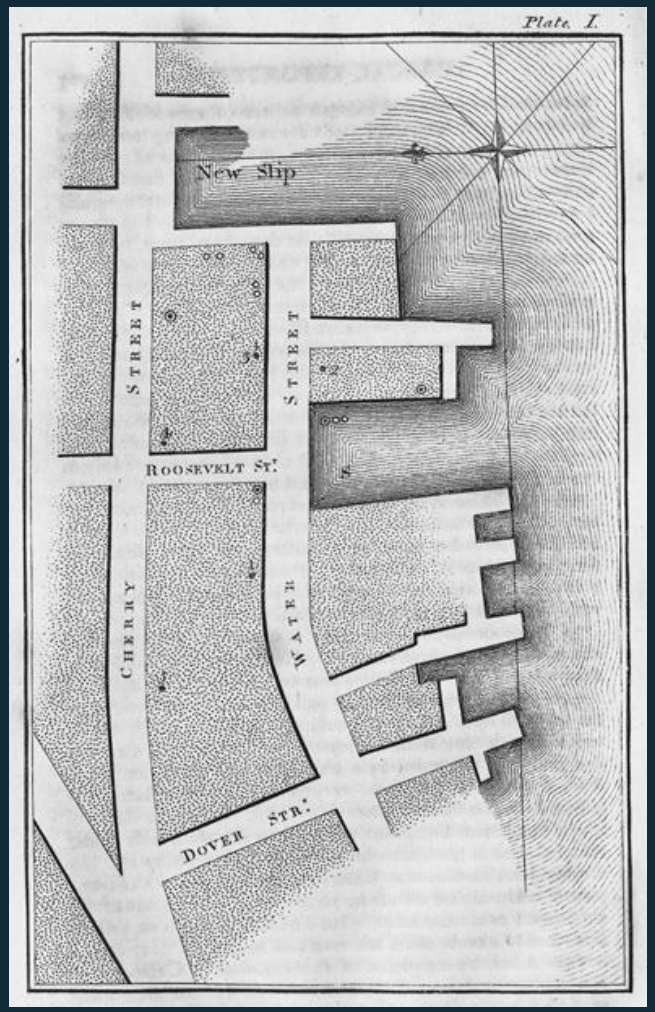
- 600 BC. Babylonian World Map.
- 801. Tang Dynasty. Hainei Huayi Tu (海内華夷圖)
  Map of Chinese and non-Chinese Territories in the World
- 1527. Diogo Ribeiro (Portuguese), the *Pádron Real*.
- 1569. Gerardus Mercator (Flemish). World Map.

In this course we are not interested in maps on their own. Rather we want to plot economic and social data on maps. This is known, historically, as social cartography.

The historical focus has been social concerns: illness, poverty, and deprivation.

Valentine Seaman- Yellow Fever

Dr Valentine Seaman (1796). An Inquiry into the Cause of the Prevalence of the Yellow Fever in New-York



The first example of plotting illness on a map.

1795 yellow fever epidemic.

Cause of the spread was unknown – was it ships from tropics?

[It is spread by mosquitos]

Seaman mapped ships and waste sites, along with cases of the disease.

Cutting edge study (but did not correctly identify mosquitos as the carrier).

Global Health - Cholera

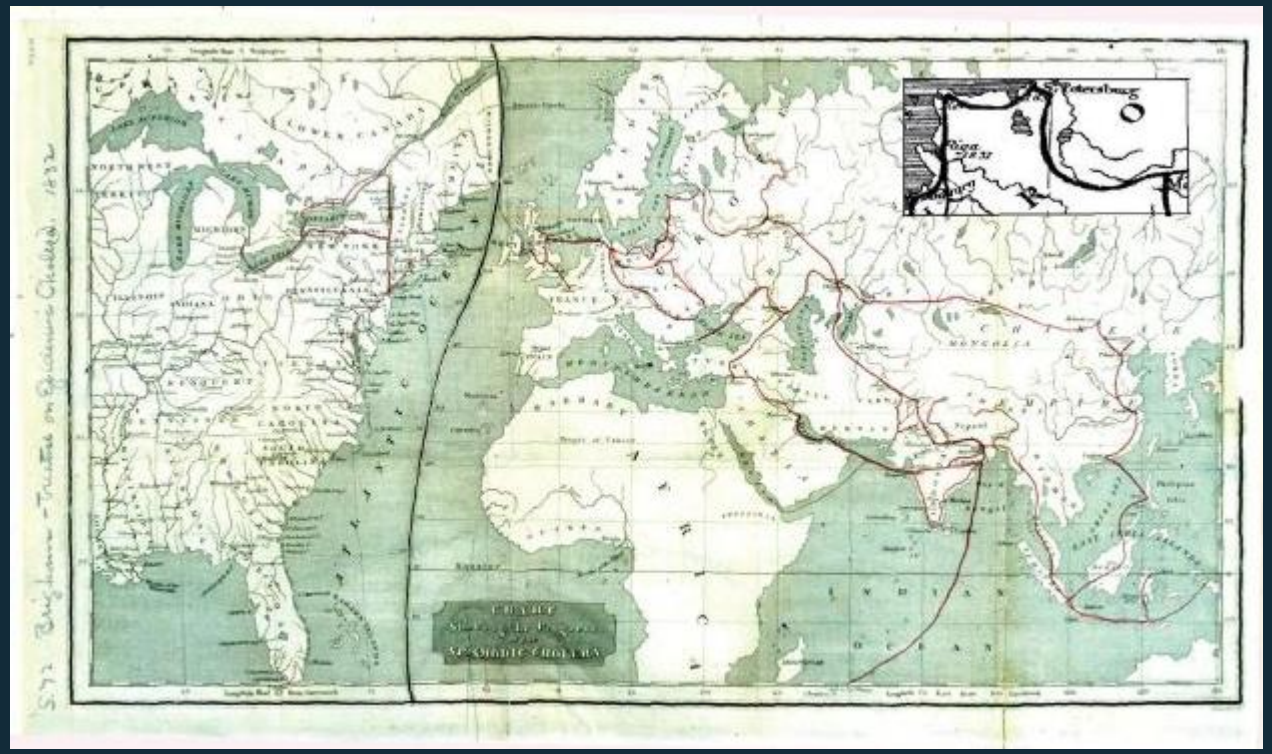
The Lancet (1831).

Maps were used in the first example of global health analytics.

This one plots towns and cities known to have cholera outbreaks.

Amariah Brigham - Cholera

Amariah Brigham (1831). A Treatise on Epidemic Cholera



Theme emerging: early map makers are trying to establish causality

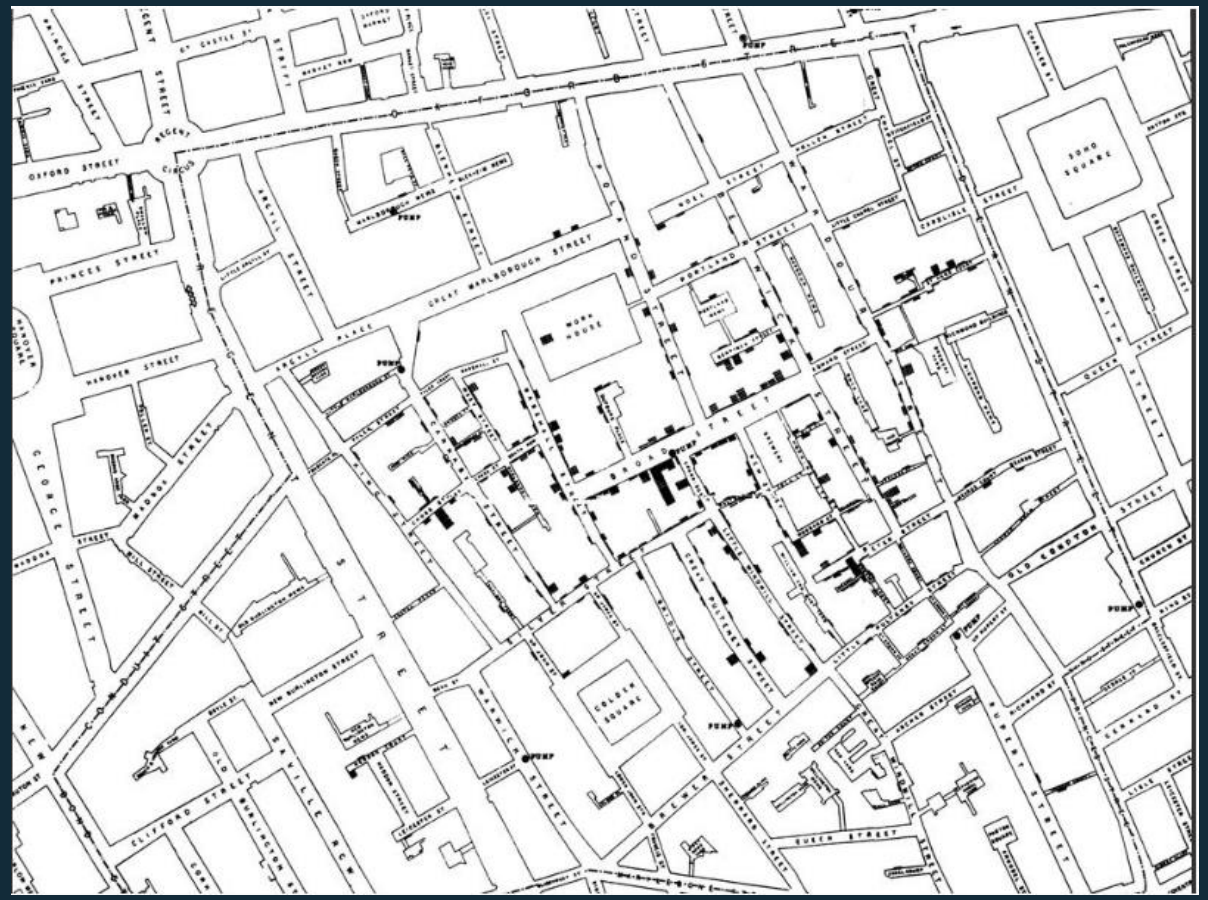
The hypothesis: a link between human travel and disease.

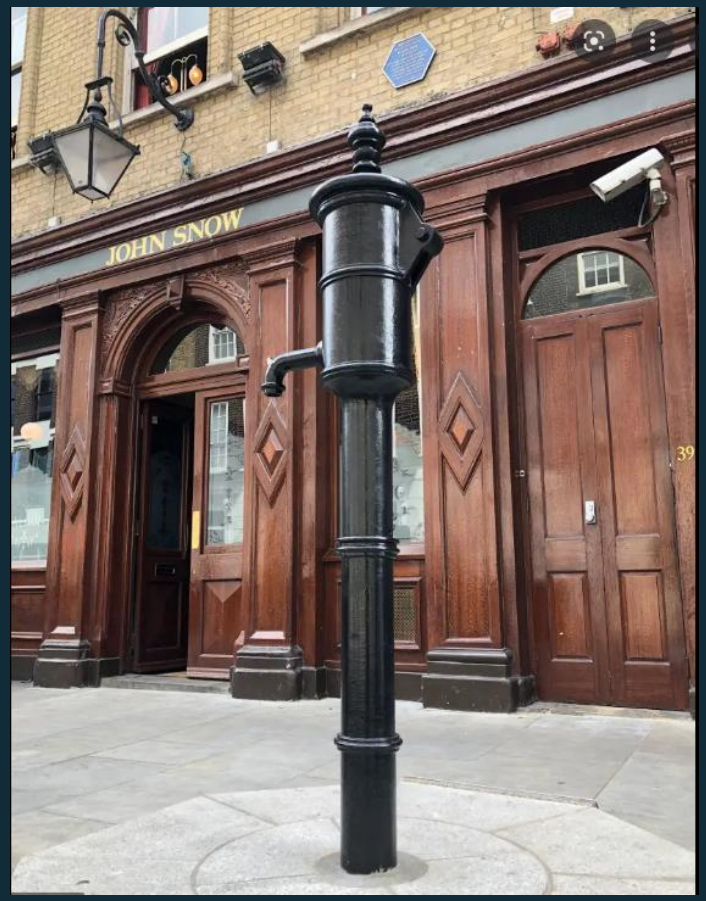
The finding: trade and illness followed the same routes across continents.

[Cholera is a bacterial infection, and transmitted through food or water contaminated with human waste]

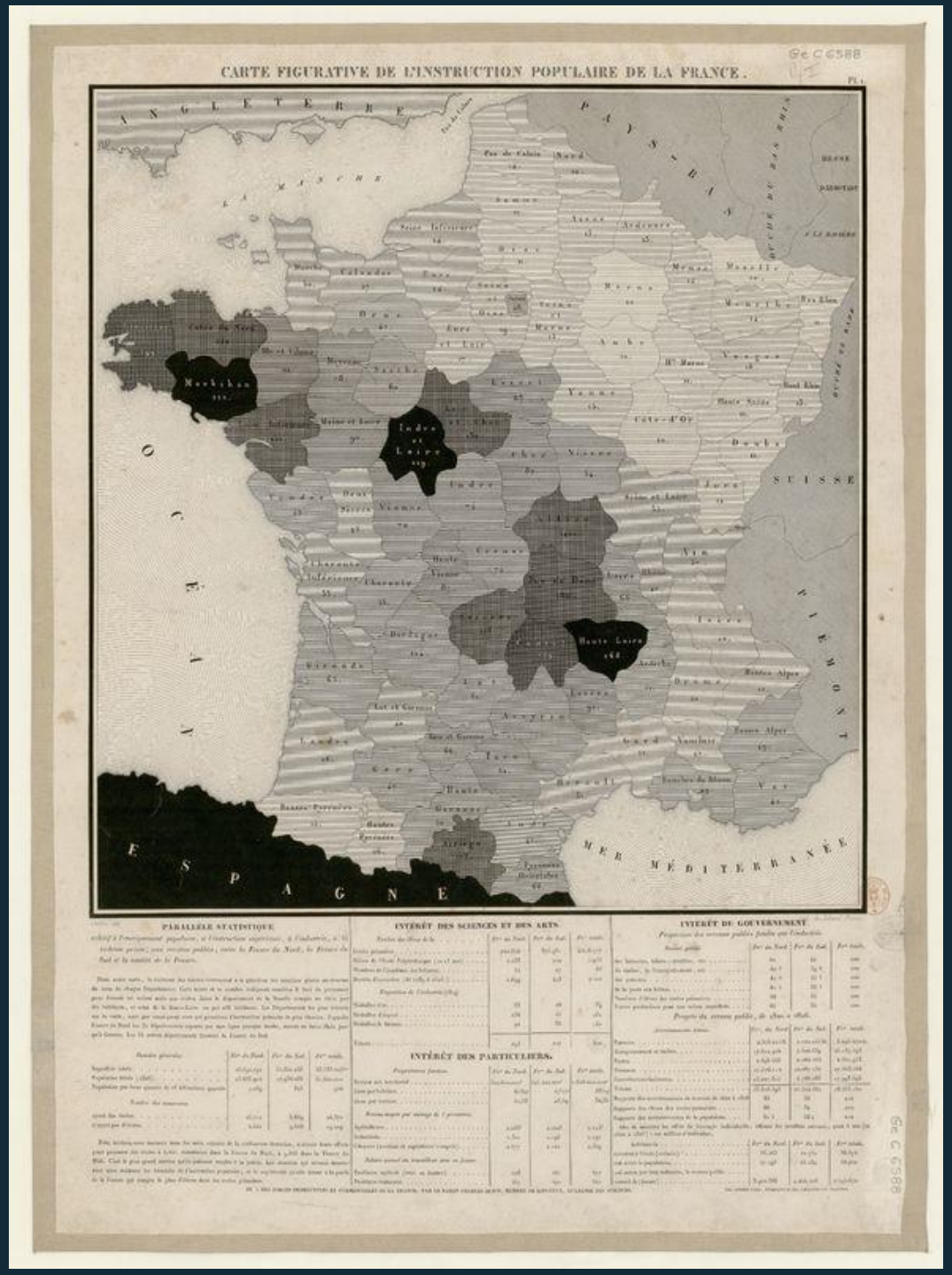
John Snow - Cholera

Dr. John Snow, father of epidemiology, 1854.



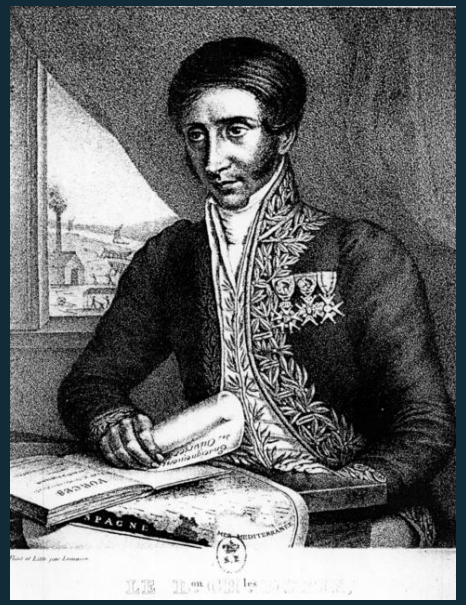


https://www.theguardian.com/news/datablog/interactive/2013/mar/15/cholera-map-john-snow-recreated



Pierre Charles Dupin

The emergence of tinted maps (1827)



Choropleth map:

- A term from 1938 (geographer John Kirtland Wright)
- From Greek: *choros* 'area/region', *plethos* 'multitude'.
- Earliest versions had simpler name: cartes teintées.

To make visible the main difference, I had the idea to give to the various départements shades all the more dark since they sent less pupils to schools. Dupin (1827)

What explains low literacy rates?

Charles Booth – Poverty

Inquiry into the Life and Labour of the People in London

Vital historical survey

Undertaken between 1886 and 1903

Examined working class life C19th

Notes and data have survived

A milestone in development of the methodology of social investigation in the United Kingdom.

https://booth.lse.ac.uk/map/14/-0.1203/51.5212/100/0

Augustus Petermann – Population

Mapping Census data (1849)

Augustus Petermann was a prominent German geographer known for his pioneering work in thematic cartography.

He worked in the UK producing population maps.

This map from the 1841 census shows the population distribution of the British Isles.

Modern incarnation, the latest ONS mapping tool:

ONS Census Maps

Gilbert (1958) “Pioneer Maps” - https://www.jstor.org/stable/pdf/1790244.pdf

Hans Haacke

Manhattan. Maps and photos used in a work of art and social commentary

Shapolsky et al. Manhattan Real Estate Holdings, a Real-Time Social System, as of May 1, 1971.

https://whitney.org/collection/works/29487

London – the Greenwich meridian

Where time starts

By the late 19th century, sailors navigated around the world relying on maps centred on their country of origin.

There became the need for a consistent global reference point.

Simultaneously, railways and telegrams linked cities in minutes, while each city kept its own time zone — Bristol, for example, eleven minutes behind London.

The Greenwich Meridian standardized navigation and unified global time and positioning.

Modern version – the IERS reference meridian
https://www.iers.org/

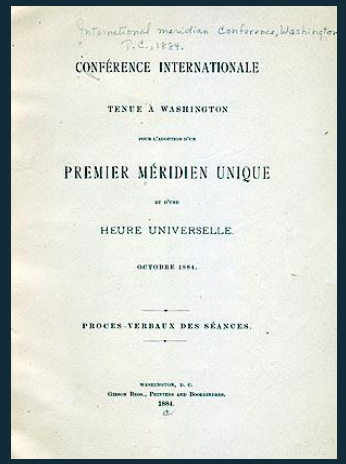
Rival standards

The Thames and the Seine

Paris Meridian

A long-standing rival, running through the Paris Observatory.

It was defined in 1667, and was essential for the triangulation mapping of France.



International Meridian Conference

The Greenwich meridian became a global standard in 1884, when it was adopted at the conference in Washington DC.

Refining the standard

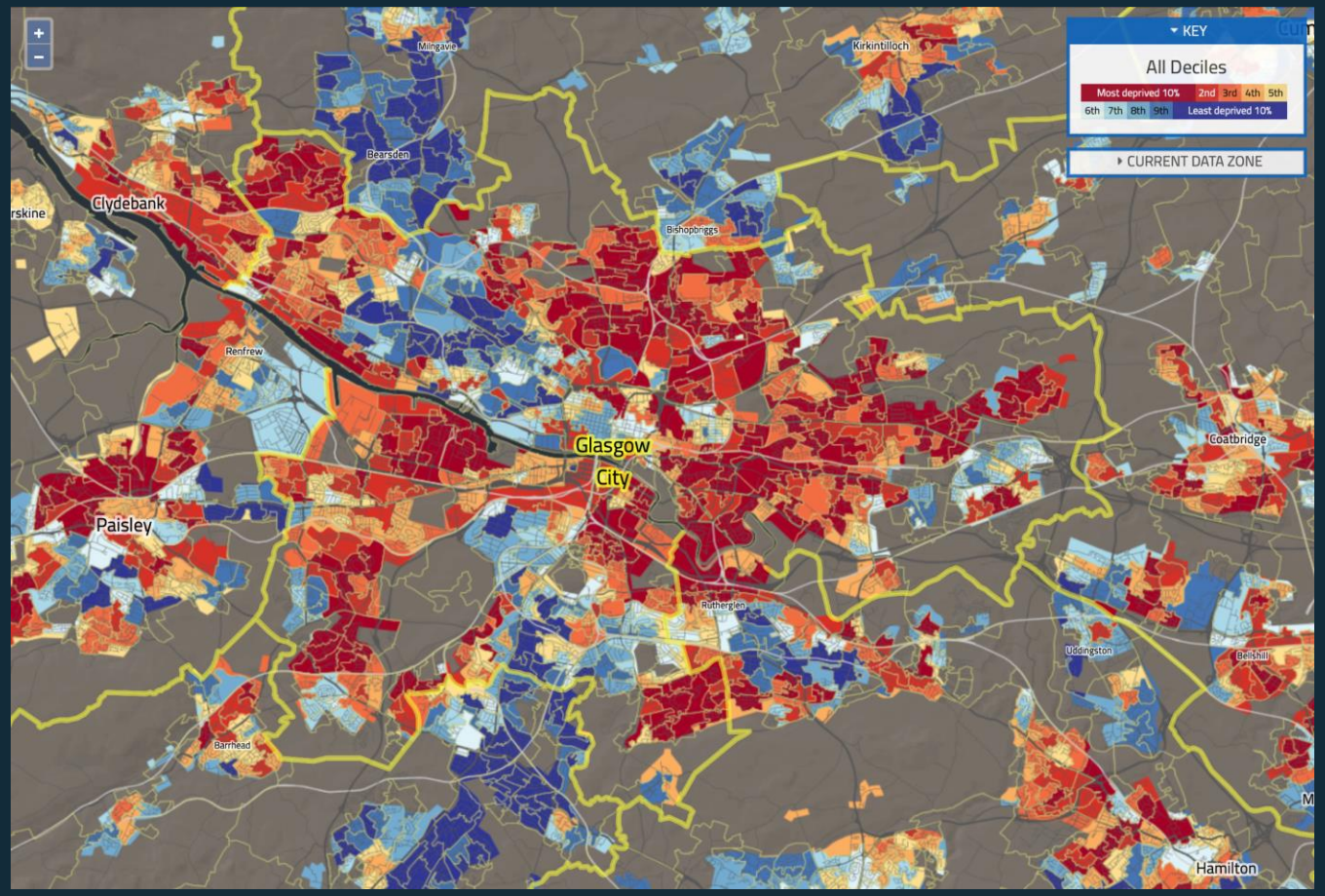
The IERS Reference Meridian, which is a refinement and formalization of the Greenwich Meridian, passes 102 meters east of the Greenwich Meridian. It passes through the Earth's centre of mass rather than aligning with the gravitational direction at Greenwich.

8.2 Modern mapping.

Uses of maps in economic and policy settings

Glasgow – Deprivation

Population Health in Glasgow



Scottish Index of Multiple Deprivation

Glasgow, Scotland's largest city by population, is one of the UK's most deprived areas.

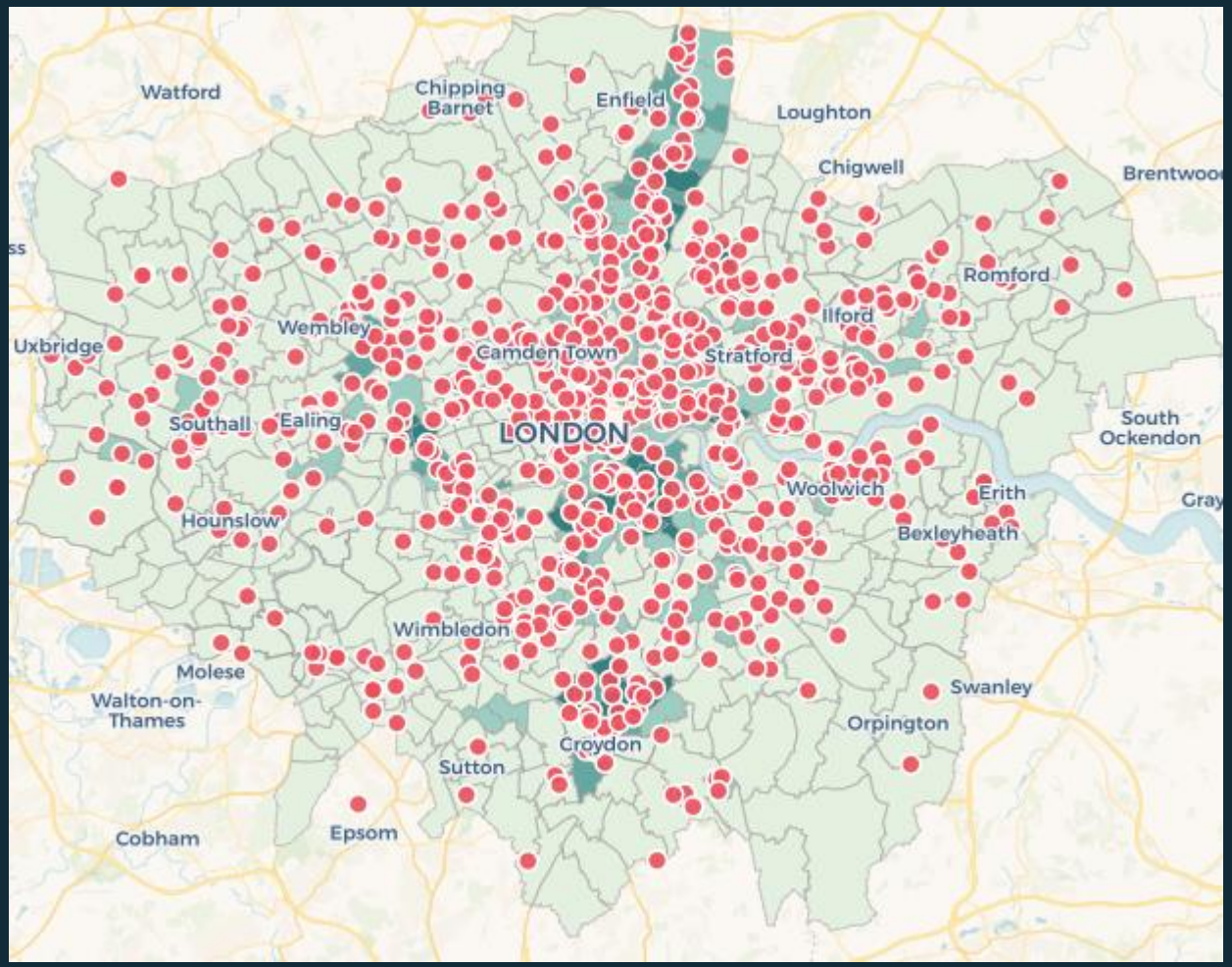
Demographers and epidemiologists have long debated a 'Glasgow effect

Its life expectancy is among the lowest in Europe and lower than expected, even after accounting for poverty.

This map shows large neighbourhood variations in deprivation.

London – Crime

Tom Kirchmaier's work with the Police



https://sites.google.com/site/tomkirchmaier/links-data

Manhattan – Shadows

The rise of super tall building and the problem of predicting shadows

https://www.mas.org/interactive_features/accidental-skyline-shadows/

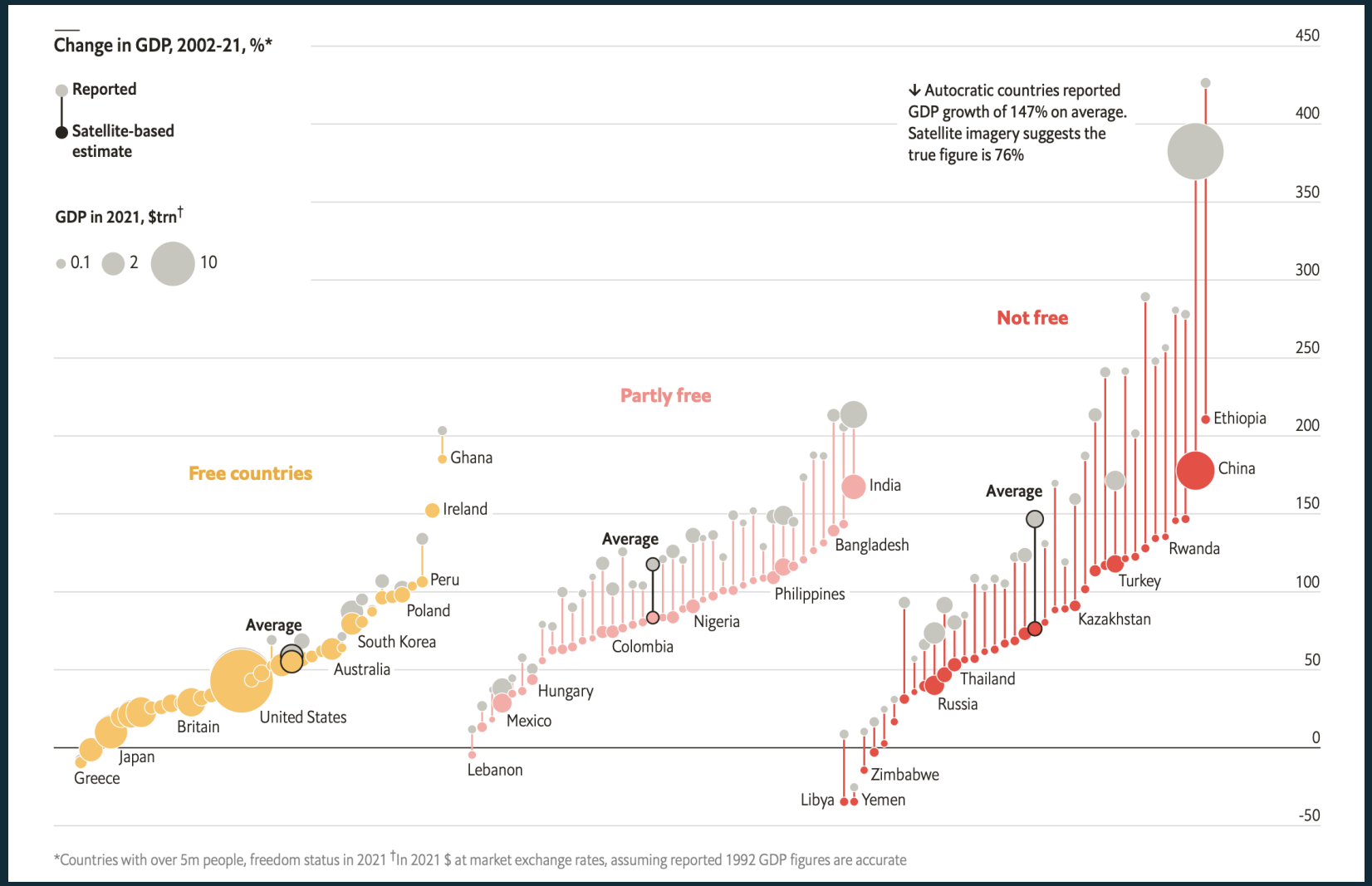
NASA – The Earth at Night

Nasa, and others, produce detailed night-light satellite imagery:

https://earthobservatory.nasa.gov/features/NightLights/page3.php

Nightlight

Economic (mis)measurement



A study of lights at night suggests dictators lie about economic growth, The Economist

Nightlight data can be used as a proxy for economic activity.

This is useful where data is missing, or can't be trusted.

Martinez (2022) uses nightlight data to show that autocracies systemically overreport their economic performance.

Modern maps.

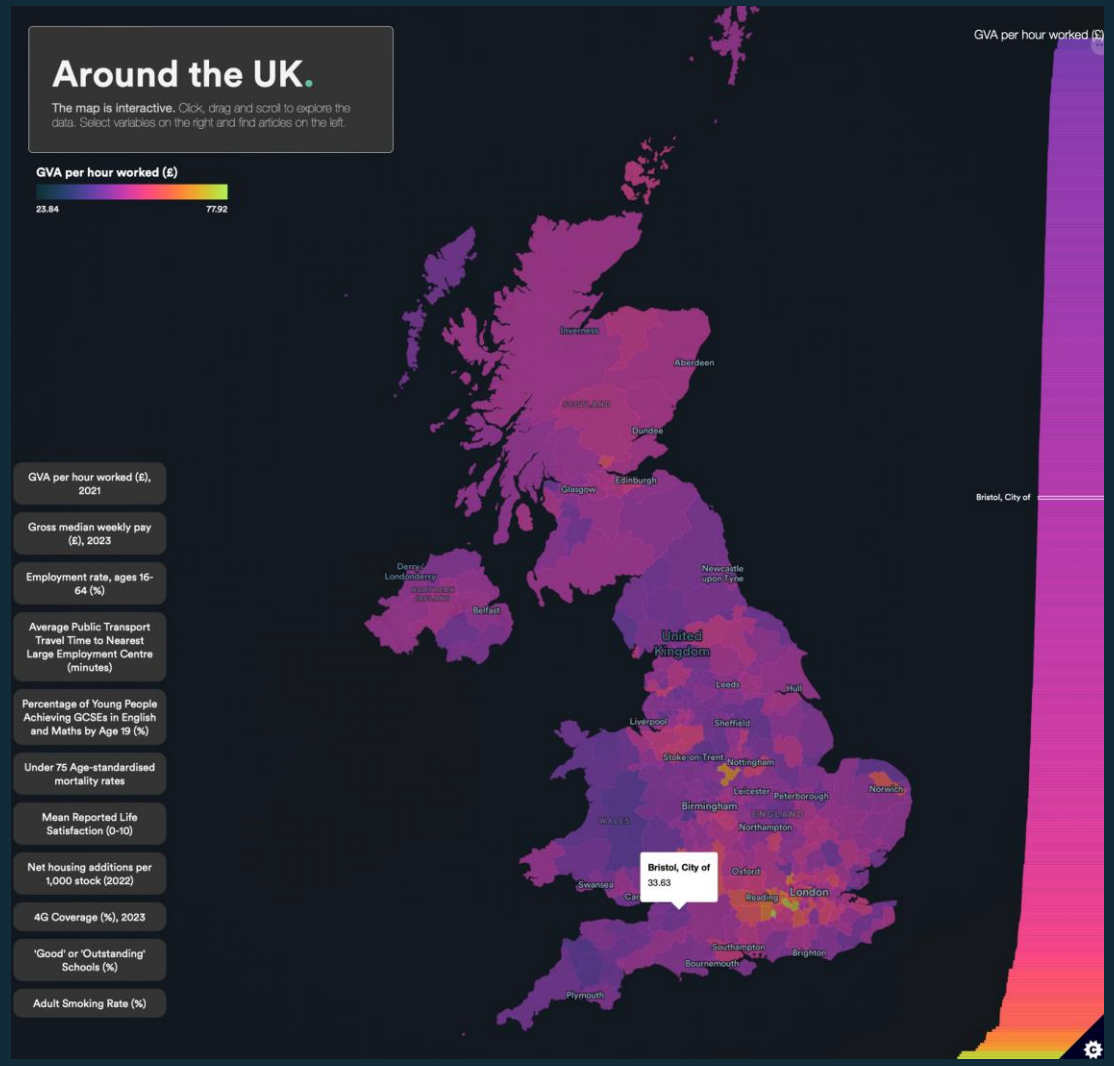
More interesting examples

Justice Map – Crime in the US
justicemap.org

Deforestation
globalforestwatch.org/map/

ONS Census Maps
ons.gov.uk/census/maps/

ECO Around the UK
economicsobservatory.com/data-hub/sandbox/map

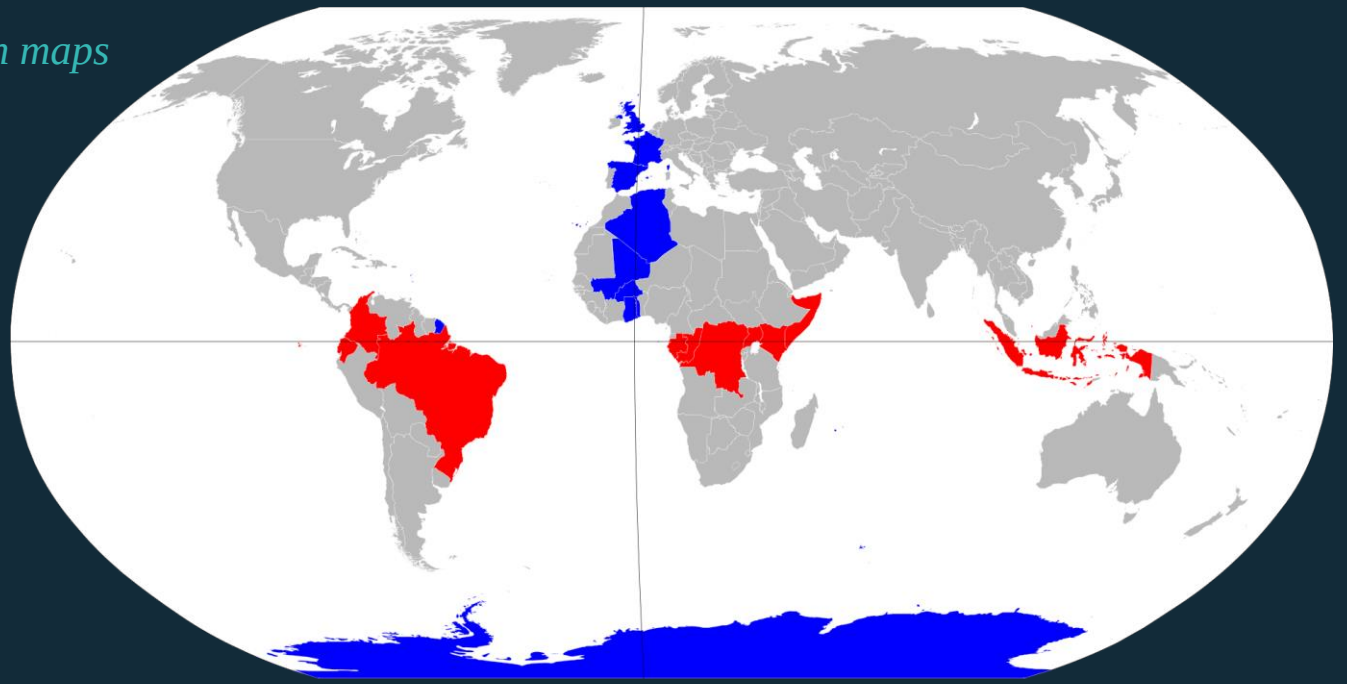


8.3 Map making.

The theory and tools you need to make maps

Key terms.

Some of the key terms that you will see when building your own maps



Longitude

- Appear as vertical lines. Measured from Prime Meridian in degrees.
- IERS Reference Meridian (close to Greenwich Meridian), used in Global Positioning System (GPS).
- 0° at the Prime Meridian, from +180° Eastward to – 180° Westward.

Latitude

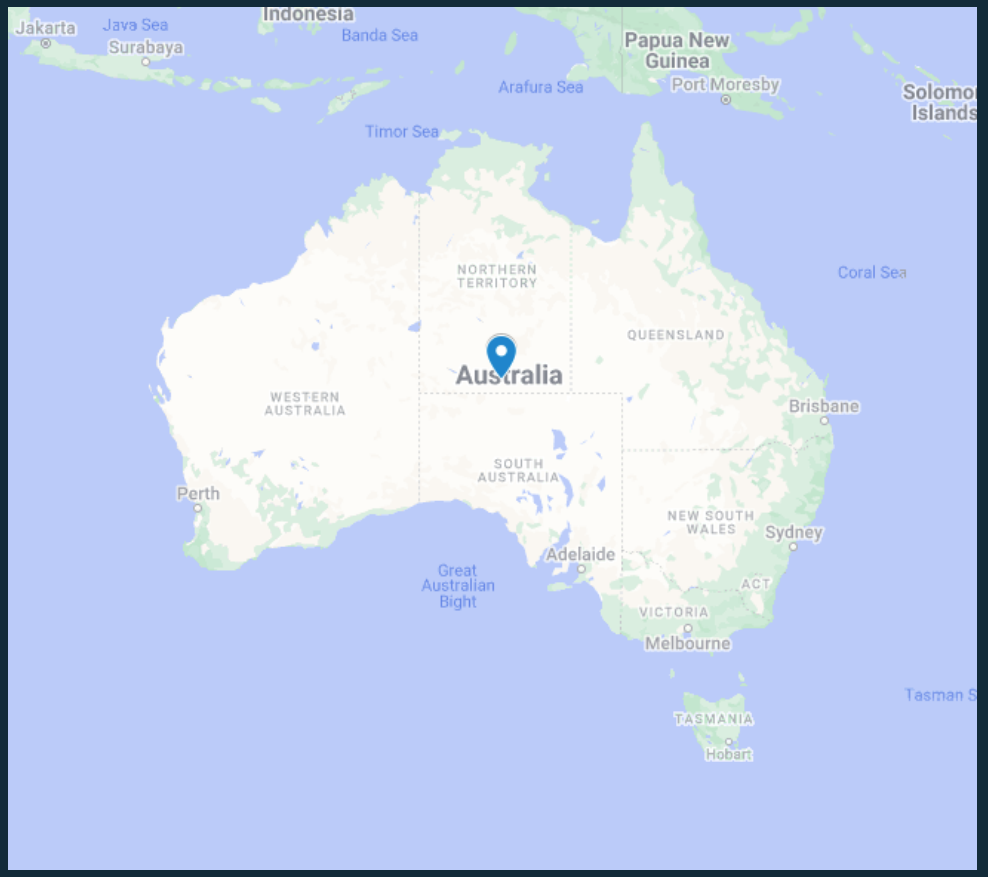
- Appear as horizontal lines. Measured as degrees.
- -90° at North Pole, 0° at the Equator, +90° at the South Pole.

Our world map.

On Google maps

Which is bigger?

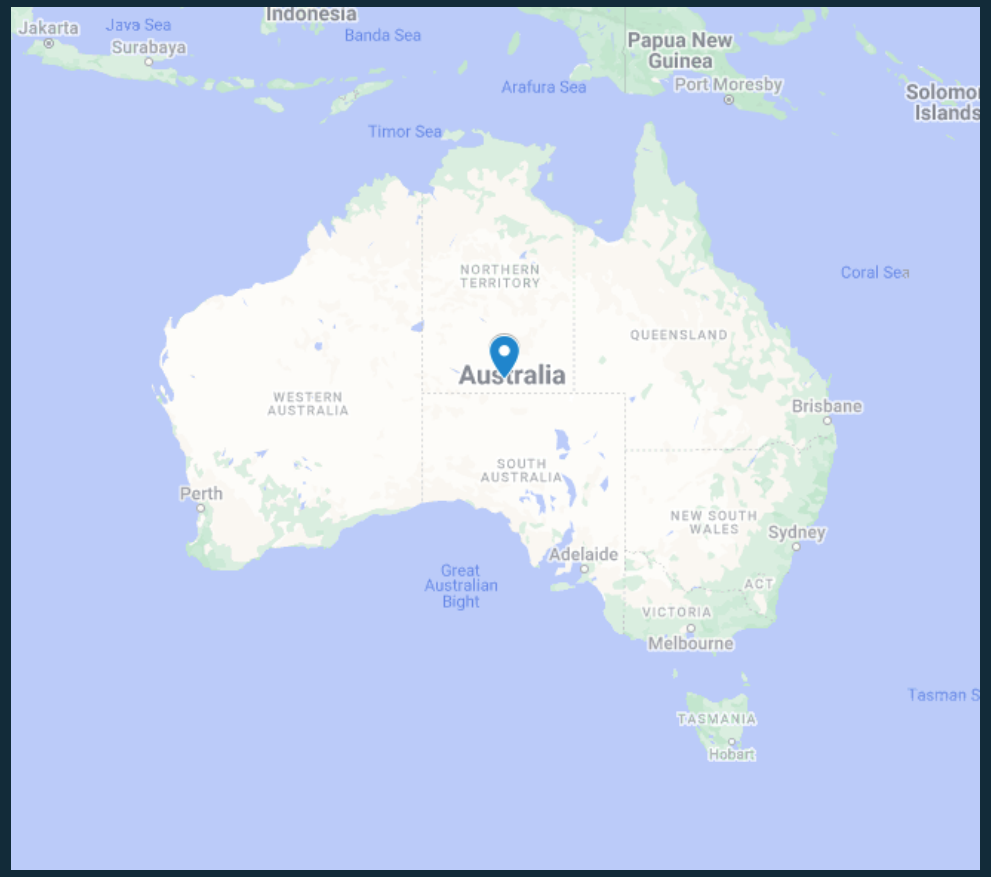
Greenland vs Australia



Which is bigger?

Greenland vs Australia

2.166 million km²



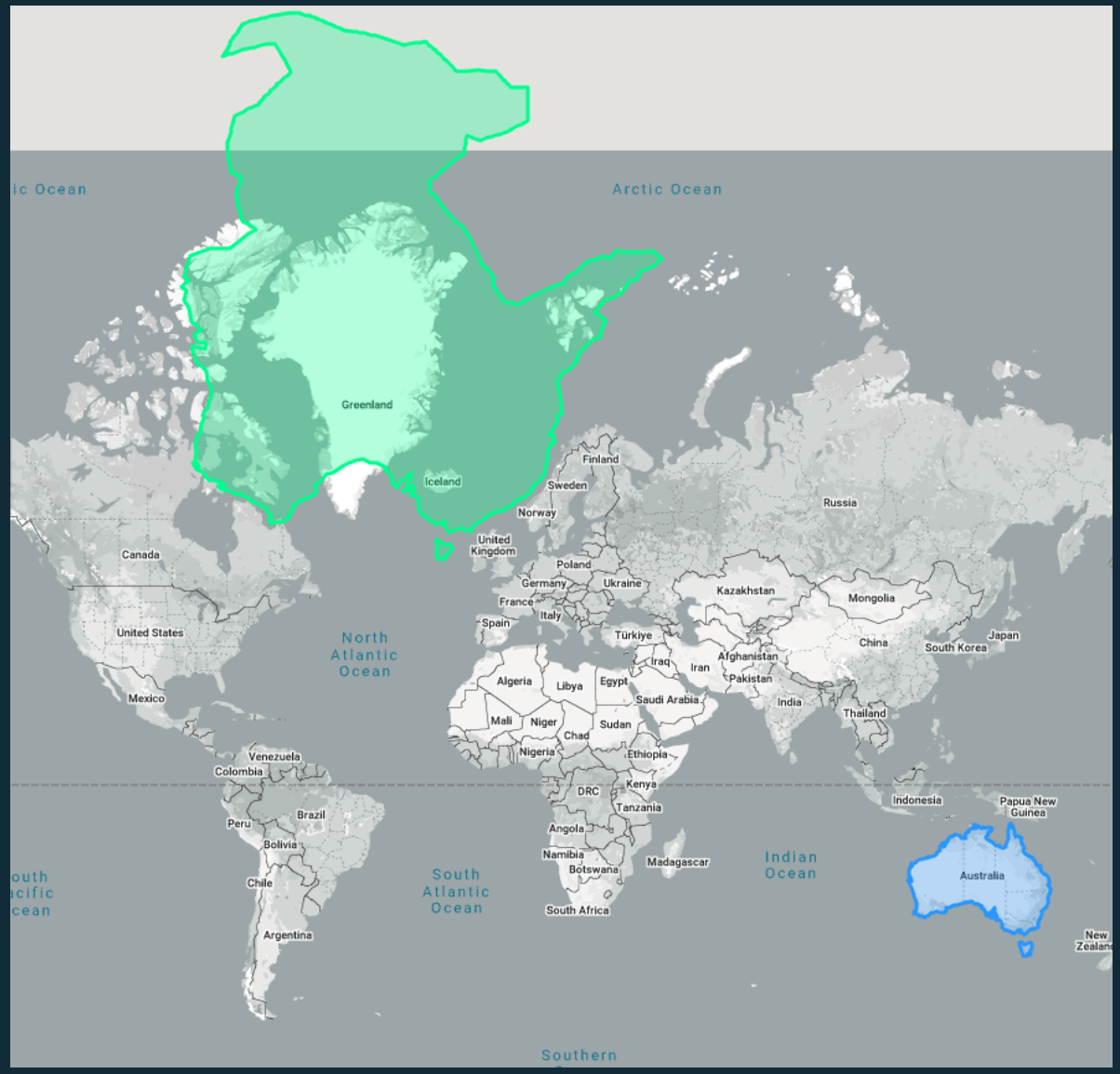
7.741 million km²

Which is bigger?

Greenland vs Australia

The maps we're used to are **distorted**.

Try this tool: thetruesize.com



True size of Africa.

Bigger than you thought.

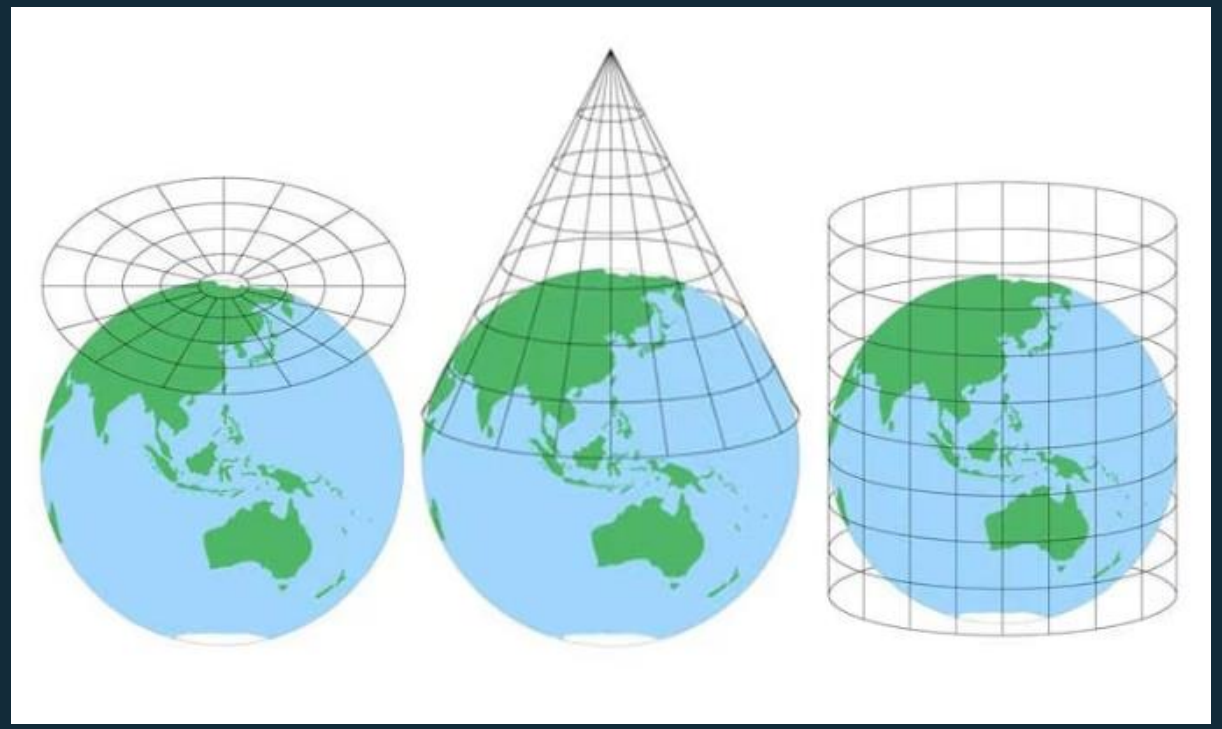
Here is USA, China, India, Mexico, Türkiye, Ukraine, France, Spain, Sweden, Norway, Japan, Germany, Finland, Poland, Italy, Romania, UK, Greece, Portugal, Austria, Ireland, Croatia, Netherlands, Switzerland, Belgium

Combined, these 25 countries are roughly the same size as Africa.

Africa: **30.4 million km²**

Projections.

Turning a sphere into a flat surface



Map projections represent spherical data on a flat surface.

In 1827, mathematician Carl Friedrich Gauss proved that all map projections must introduce distortion. It is impossible to flatten a curved surface while preserving distance, angle, and area simultaneously.

Every map projection is, therefore, a trade-off. Common approaches are:

- Azimuthal Equidistant
- Conic
- Cylindrical (e.g. Mercator)

https://blog.mapchart.net/misc/quick-guide-to-map-projections/

Projections.

Common choices

Mercator

Azimuthal Equidistant

Projections.

Common choices

Conic (equal area)

Equal Earth

Projections.

Picking a projection

Equal Earth

Each approach trades off distance, angle, and area.

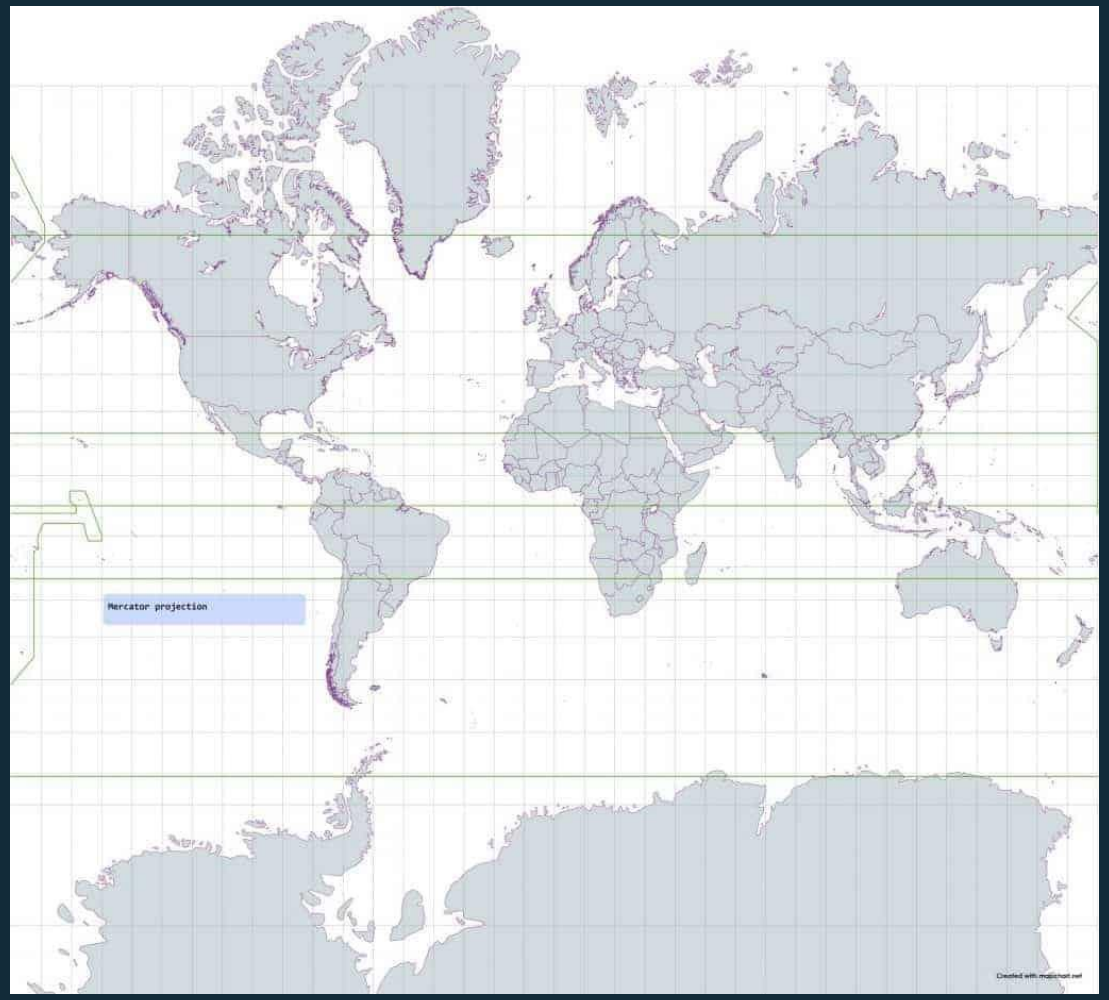
Projection	Distance	Angle	Area
Conic (Equal Area)	Good	Poor	Good
Equal Earth	Moderate	Moderate	Good
Mercator	Poor	Good	Poor
Azimuthal Equidistant	Good	Moderate	Poor

The Equal Earth projection, designed to balance shape and area with moderate accuracy in distance, is a decent middle ground.

The right projection, though, is context-dependent. For example, if visualising navigation paths, then the Mercator projection works well. But if equal area representation is needed, then Conic (Equal Area) or Equal Earth are better suited

Projections.

All projections distort…



Projections.

Mercator projection, with true shape overlaid.

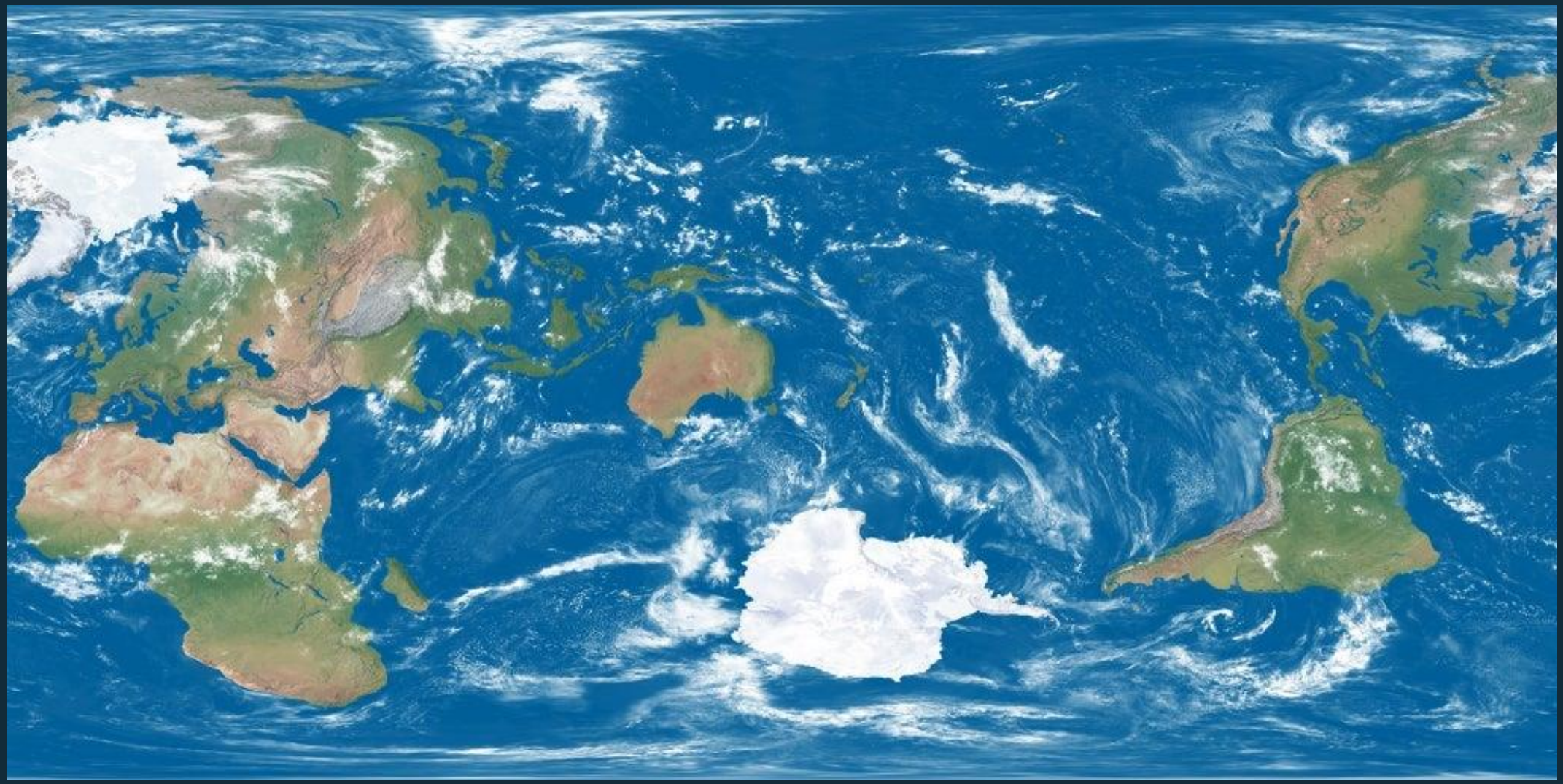
Consider carefully how representation impacts perception: the Mercator projection significantly distorts the Southern Hemisphere.

For example, Africa is 14 times the size of Greenland but appears similarly sized with the Mercator projection.

https://twitter.com/neilrkaye

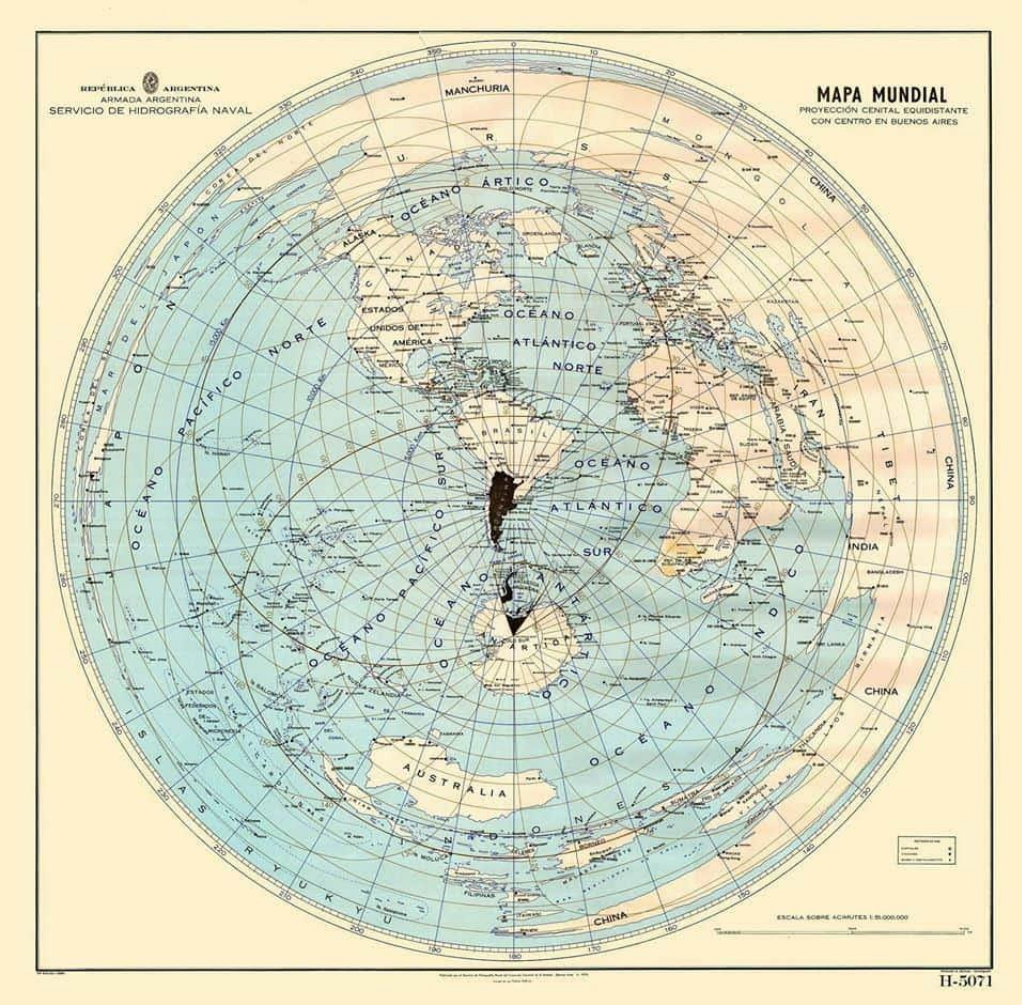
Projections.

Earth centred on NZ



Projections.

World centred on Buenos Aires



Conclusions:

- All projections distort
- The larger the area, the greater the distortion at the edges in many projections.
- Think about the analytical implications of your projection. How does it drive your story?

Data 1.

GeoJSON and TopoJSON. The data you need to draw your map

Based on the JSON format. {“key”: “value”} pairs

GeoJSON data in the form of arrays, [X, Y]. Sets of arrays can be used for:

- Points
- Lines
- Polygons

TopoJSON data in the form of *arcs.* These are sequences of points.

- Topo from topology, not from topography. [Not about mountains].
- Arcs are only defined once, so storage much smaller.
- Files more complicated, and not intuitive to read with human eye.
- Not so good with user-collected points
- Very good when regions line up next to one another, as in political/state maps.

Base maps

Data sources

The base TopoJSON files are open source:

- US-Atlas. https://github.com/topojson/us-atlas
- UK files. https://martinjc.github.io/UK-GeoJSON/


```
{
    "type": "Topology",
    "objects": {
        "wards": {
            "type": "GeometryCollection",
            "crs": {
                "type": "name",
                "properties": {
                    "name": "urn:ogc:def:crs:OGC:1.3:CRS84"
                }
            },
            "geometries": [
                {
                    "type": "MultiPolygon",
                    "properties": {
                        "WD13CD": "W05000041",
                        "WD13CDO": "00NCQA",
                        "WD13NM": "Aberdaron",
                        "WD13NMW": "Aberdaron"
                    },
                    "id": "W05000041",
                    "arcs": [
                        [
                            [
                                0,
                                1,
                                2
                            ]
                        ],
                        [
                            [
                                3
                            ]
                        ],
                        [
                            [
                                4
                            ]
                        ],
                        [
                            [
                                5
                            ]
                        ],
                        [
                            [
                                6
```

Base maps

First simple map

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "data": {
    "url": "https://raw.githubusercontent.com/RDeconomist/RDeconomist.github.io/main/charts/maps/geoWalWards.json",
        "format": {"type": "topojson",
        "feature": "wards"}},
  "projection": {"type": "mercator"},
  "mark": {"type": "geoshape"}
}
```

Regions

How local do you want to be?

United Kingdom | International Terretorial Levels

Maps of the UK - Richard Davies and Josh Hellings.

These UK maps demonstrate differing levels of detail.

United Kingdom
ITL1 regions, Mercator Projection

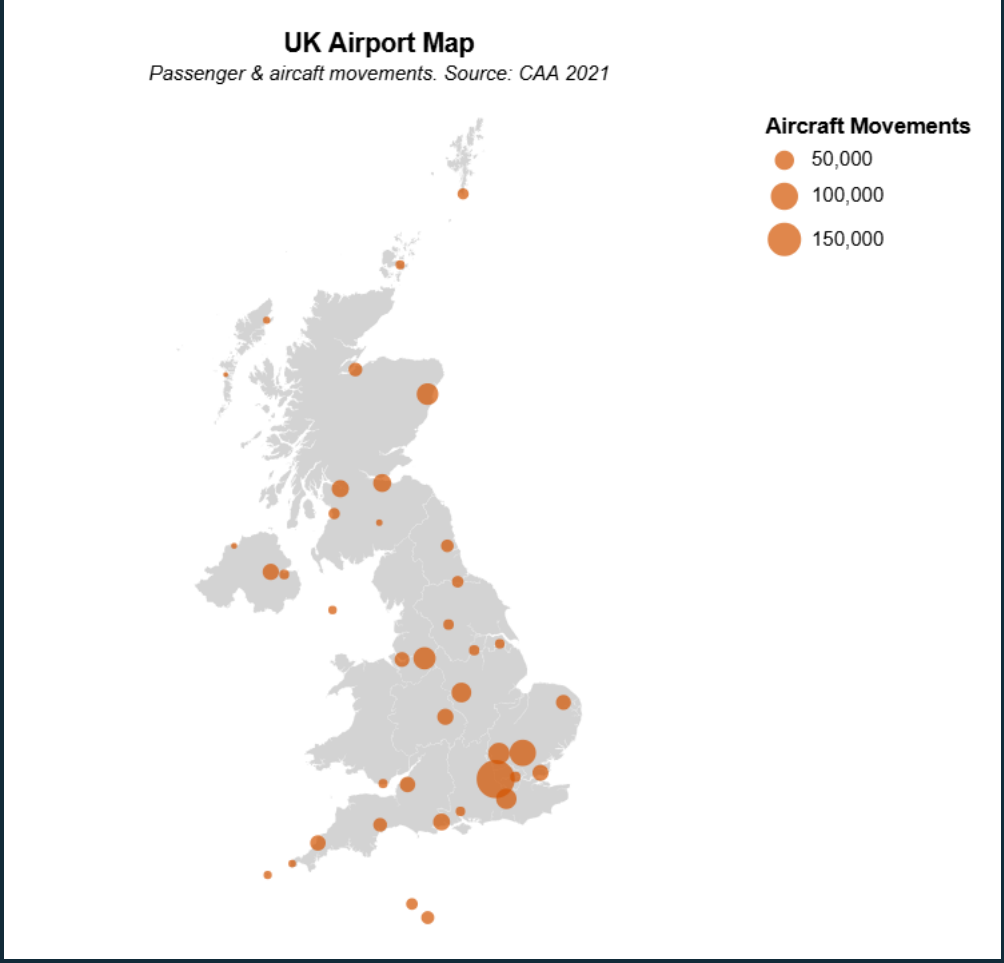
United Kingdom
ITL2 regions, Mercator Projection

United Kingdom
ITL3 regions, Mercator Projection

These maps show the NUTS-1, -2 and -3 breakdown for the UK. Also the ITL1, 2 and 3.

Maps + coordinates

Plotting specific points on a map



If our data is **point** rather than **area** based, we can just plot a points map.

Two data sources:

- UK map (Geo or TopoJson)
- CSV with data + coordinates

Airport	latitude	longitude	Passengers	Aircraft Movements
London Heathrow	51.470022	-0.454296	19393886	195336
London Stansted	51.886018	0.238866	7148200	93316
London Gatwick	51.153662	-0.182063	6261814	55817
Manchester	53.355437	-2.277298	6085103	66310
London Luton	51.876265	-0.371747	4674800	61703
Edinburgh	55.950785	-3.361453	3024960	43674
Birmingham	52.452382	-1.743507	2482430	35411
Belfast International	54.655670	-6.216864	2328276	35411

Maps + coordinates

Plotting specific points on a map

Our chart will have two layers:

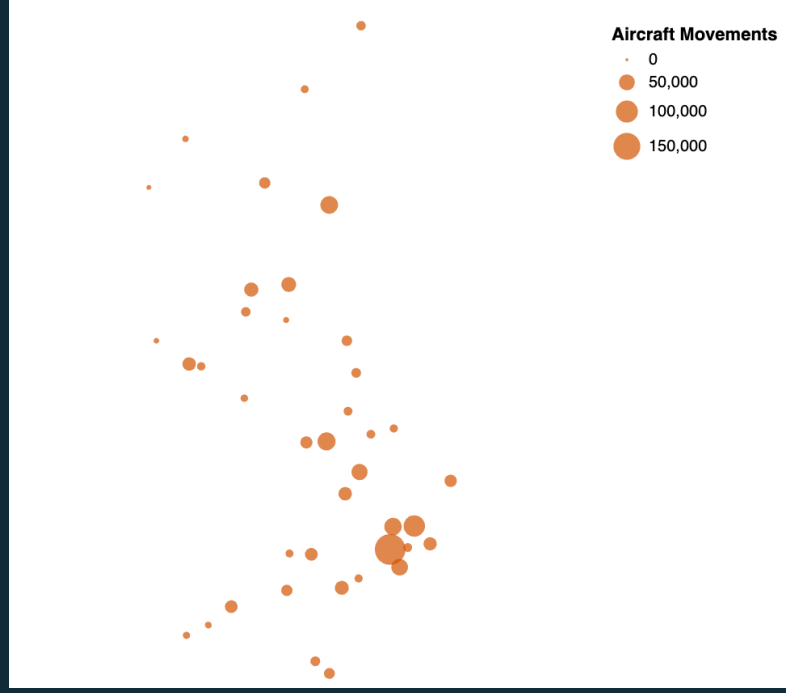
- “geoshape”: Base map outlining UK and regions
- “points”: location of each observation

In our **geoshape** layer, we set our set our projection type.
We don’t need any encodings to show the map.

```
{
    "data": {"url": "ITL1_2021_20m.geojson", "format": {"property": "features"}},
    "projection": {"type": "mercator"},
    "mark": {
        "type": "geoshape",
        "fill": "lightgray",
        "stroke": "white",
        "strokeWidth":0.1
    },
}
```

Maps + coordinates

Plotting specific points on a map



Our chart will have two layers:

- “geoshape”: Base map outlining UK and regions
- “points”: location of each observation

In our **points** layer, we **encode** our latitude and longitude to these visual channels.

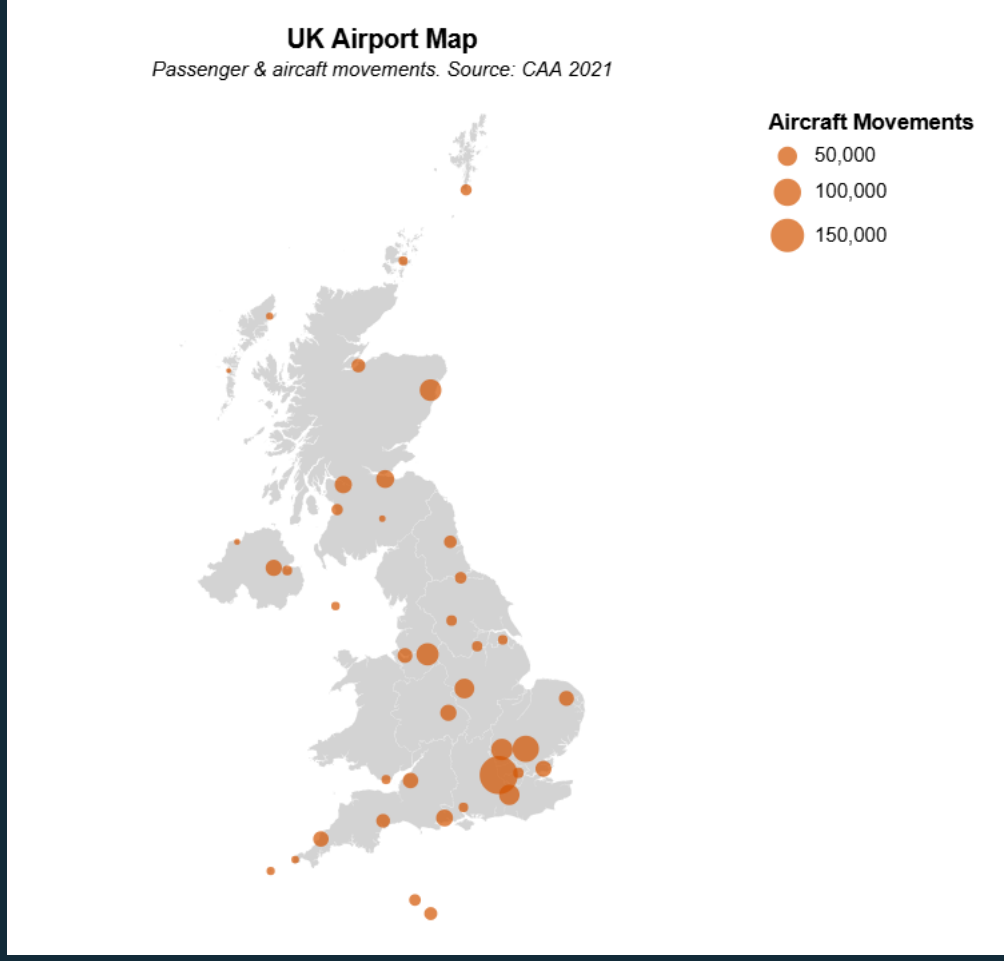
```
{
    "data": {"url": "UK_airports2.csv"},
    "mark": {”type": "circle”},
    "encoding": {
        "longitude": {"field": "longitude", "type": "quantitative"},
        "latitude": {"field": "latitude", "type": "quantitative"},
        "size": {"field": "Aircraft Movements", "type": "quantitative"},
    }
}
```

Maps + coordinates

Plotting specific points on a map

Using a layer, we combine the map and the points into a single chart.

```
"layer": [
        {
                MAP CHART
        }
        {
                POINTS CHART
        }
]
```



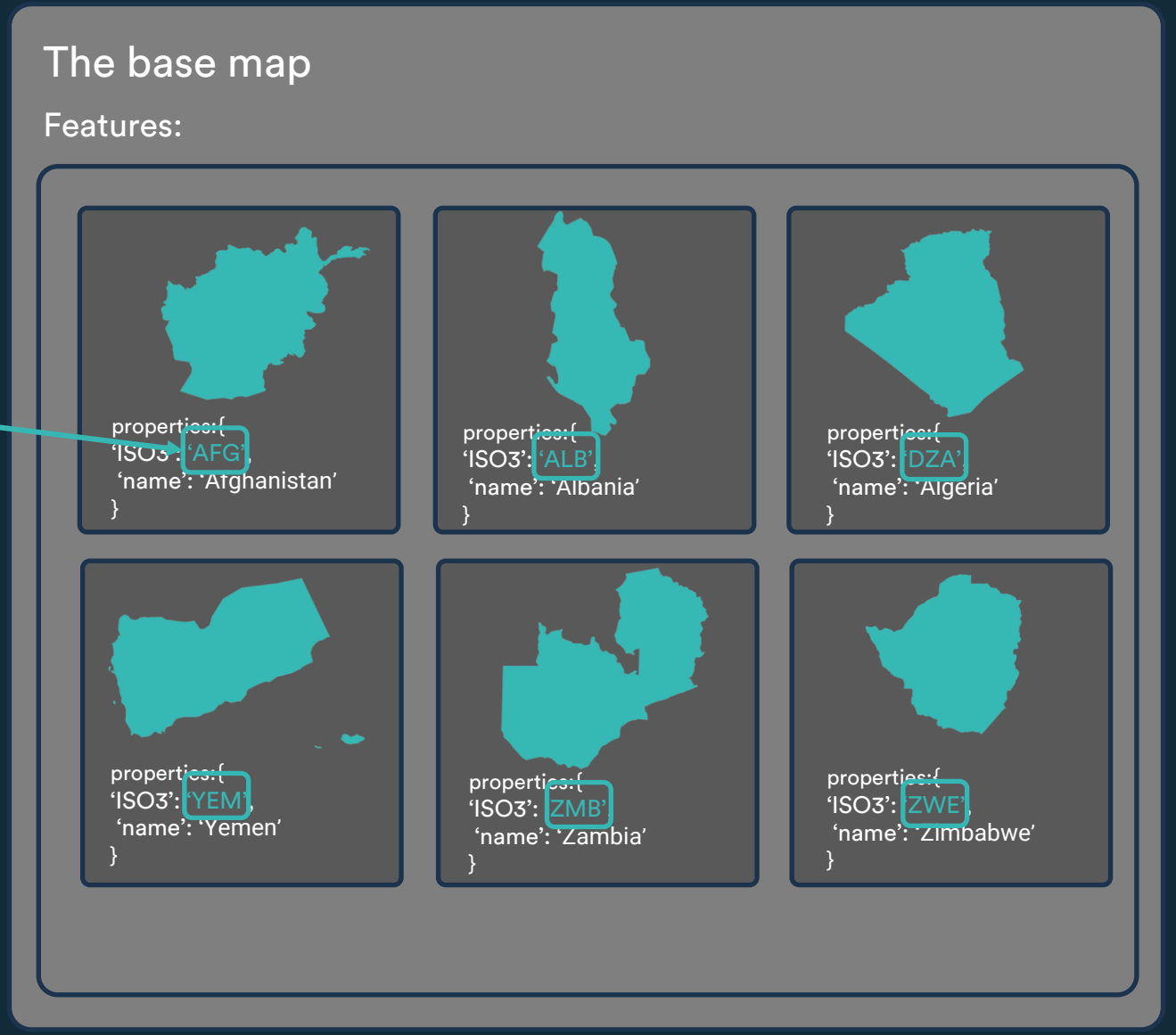
Maps + data

Making a choropleth

We form choropleth maps by linking numeric entries with geographic data.

Country	ISO3	value
Afghanistan	AFG	1665
Albania	ALB	15709
Algeria	DZA	4216
...	...	...
Yemen, Rep.	YEM	594
Zambia	ZMB	3555
Zimbabwe	ZWE	2323

To do this, we use a common key to match the two datasets.



Maps + data

Making a choropleth

We form choropleth maps by linking numeric entries with geographic data.

Country	ISO3	value
Afghanistan	AFG	1665
Albania	ALB	15709
Algeria	DZA	4216
...	...	...
Yemen, Rep.	YEM	594
Zambia	ZMB	3555
Zimbabwe	ZWE	2323

In Vega-lite, we do this with a lookup transform:

```
"transform": [
      {
            "lookup": "properties.ISO3",
            "from": {
                  "key": "ISO3",
                  "fields": ["country", "value"],
                  "data": {
                  "url": ”[...]/data.csv"
            }
            }
      }
]
```

Maps + data

Making a choropleth

We form choropleth maps by linking numeric entries with geographic data.

Country	ISO3	value
Afghanistan	AFG	1665
Albania	ALB	15709
Algeria	DZA	4216
...	...	...
Yemen, Rep.	YEM	594
Zambia	ZMB	3555
Zimbabwe	ZWE	2323

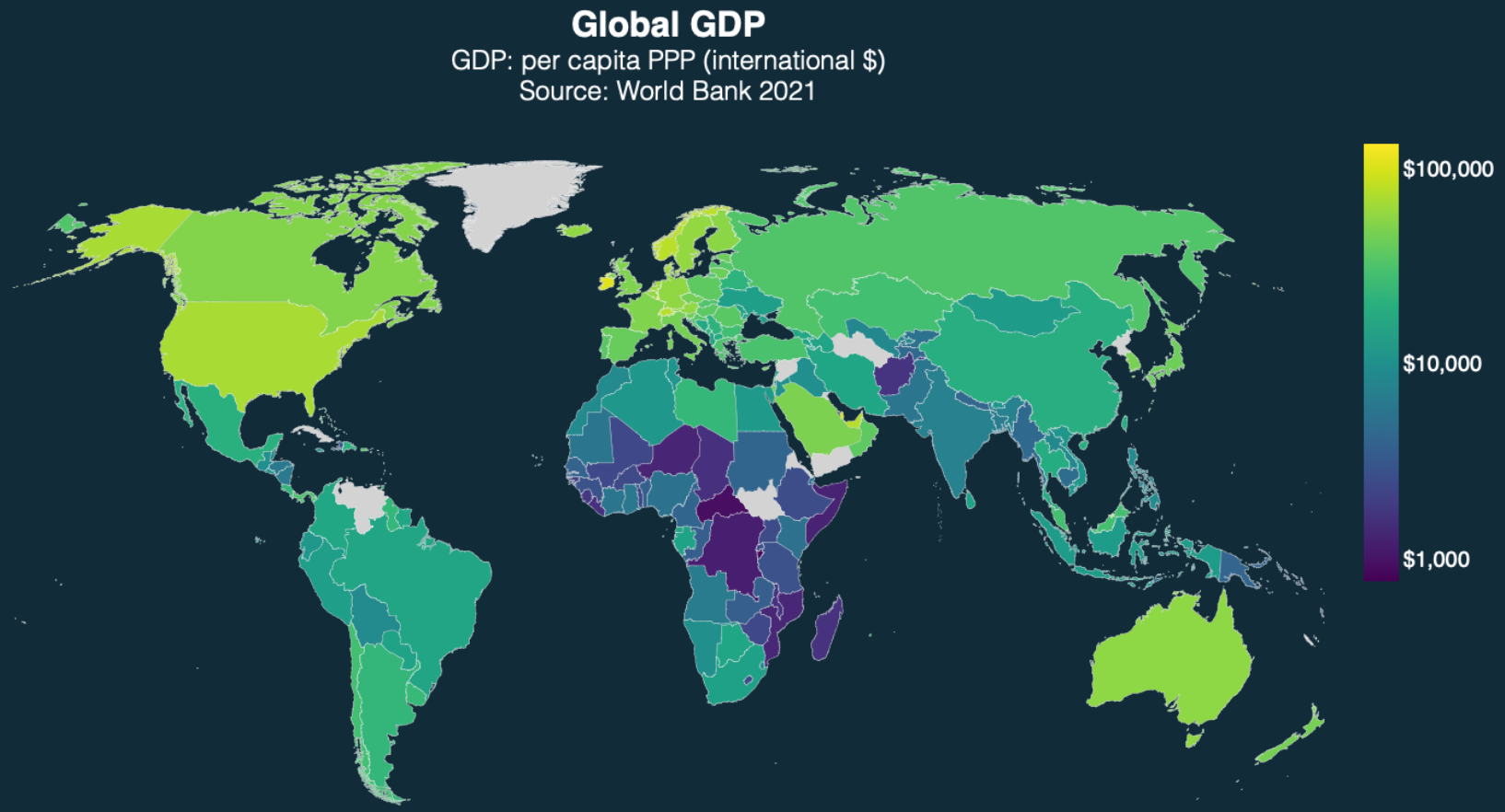
In Vega-lite, we do this with a lookup transform:

```
"transform": [
      {
            "lookup": "properties.ISO3",
            "from": {
                  "key": "ISO3",
                  "fields": ["country", "value"],
                  "data": {
                  "url": ”[...]/data.csv"
            }
            }
      }
]
```

Maps + data

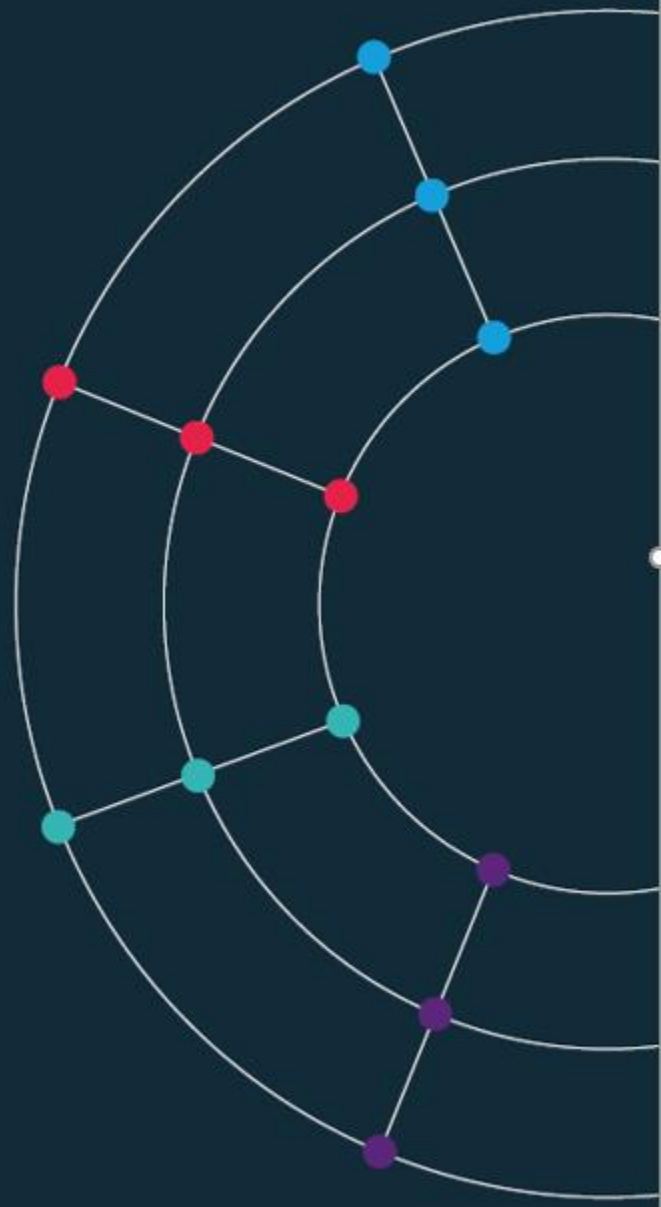
Making a choropleth

By matching every country, we create a choropleth:



Practical session.

https://rapidcharts.io/maps



Break.

Take a look at the maps on the rapidcharts site

https://www.rdeconomist.github.io/maps

Copy the following two files to your GitHub:

File: github.com/jhellingsdata/RADataHub/blob/main/misc/Guinness_Map/UK_LADs.topojson

Raw: github.com/jhellingsdata/RADataHub/raw/refs/heads/main/misc/Guinness_Map/UK_LADs.topojson

File: github.com/jhellingsdata/RADataHub/blob/main/misc/Guinness_Map/pint_data.csv

Raw: github.com/jhellingsdata/RADataHub/raw/refs/heads/main/misc/Guinness_Map/pint_data.csv

England

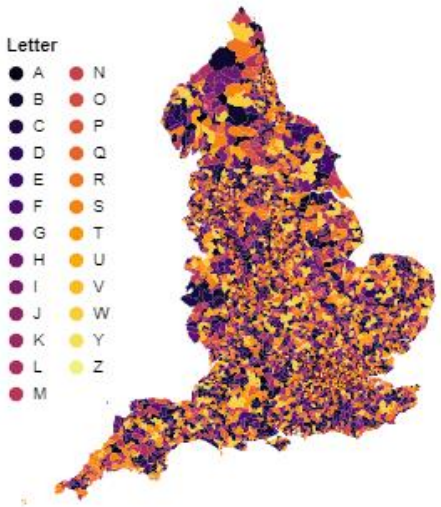
Start to build more interactive maps using tooltips and year selectors.

England map

Electoral Wards, Mercator Projection

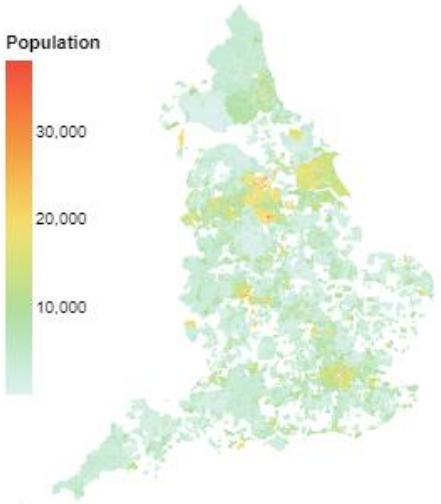
England - place names

Local authorities by starting letter of place name



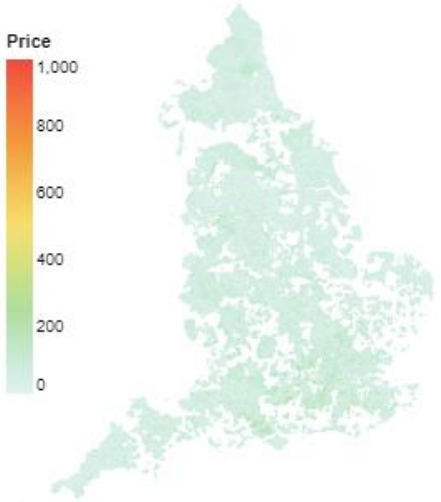
England - population

Population by electoral ward. Source: ONS



England: house prices

Median price, by electoral ward, Source: ONS



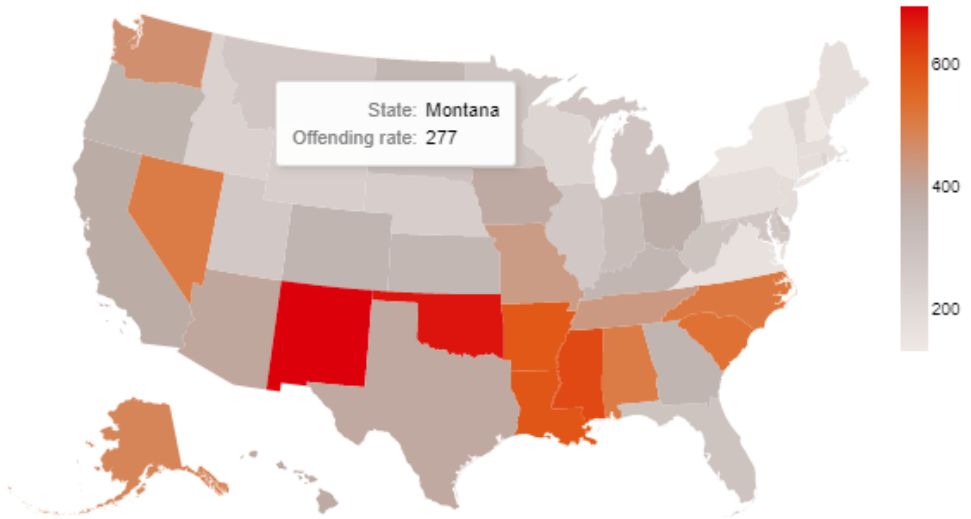
Select HP_1995

USA - base map

Albers USA projection

Crime in the USA

Offences per 100k population. Source FBI, 2019



Select Burglary