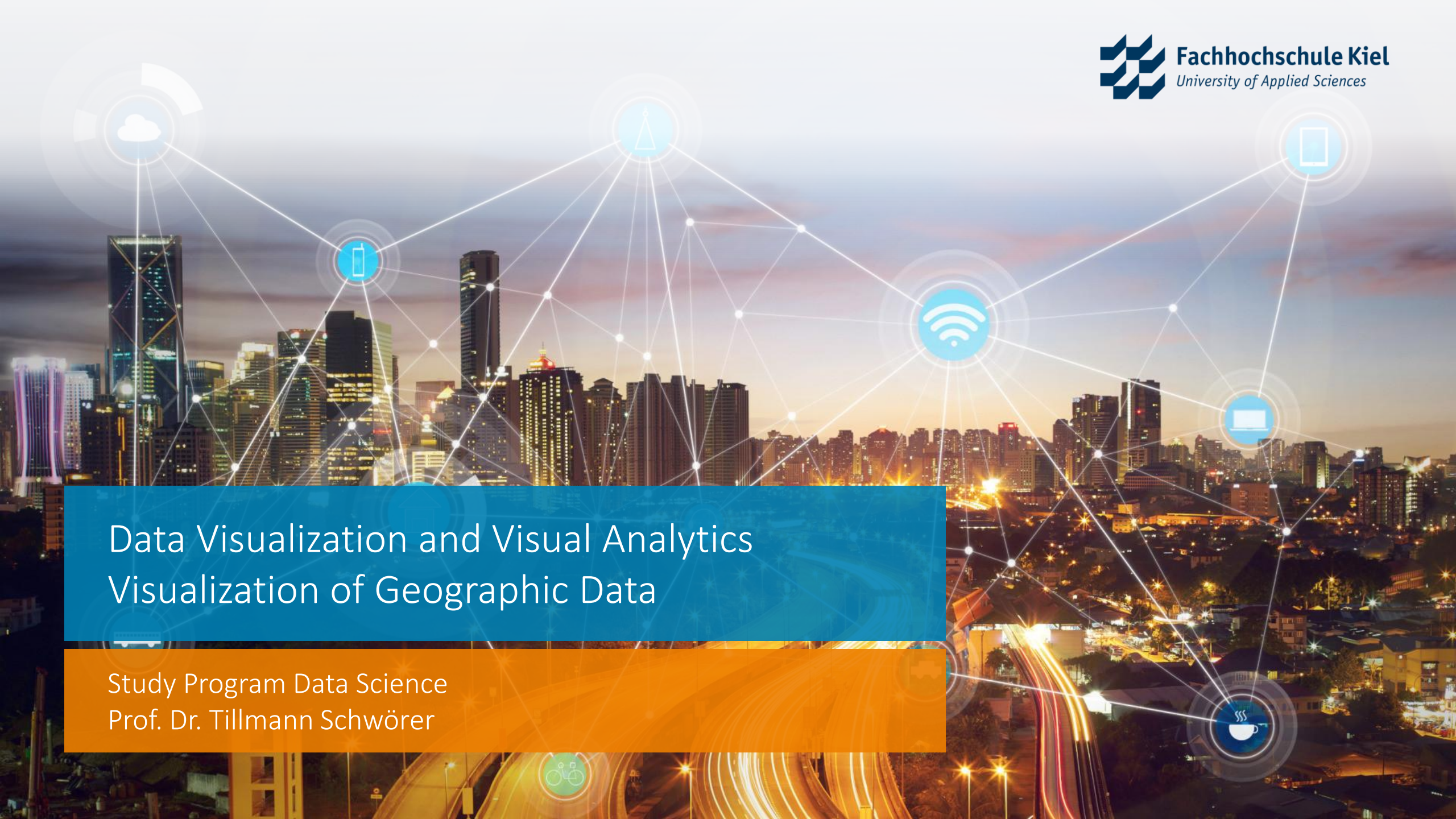


Fachhochschule Kiel
University of Applied Sciences

Data Visualization and Visual Analytics
Visualization of Geographic Data

Study Program Data Science
Prof. Dr. Tillmann Schwörer

Lecture Roadmap

Data Domains

Comparing Categories | Relationships | Geospatial | Time |
Part-to-whole | Distributions | Uncertainty | …

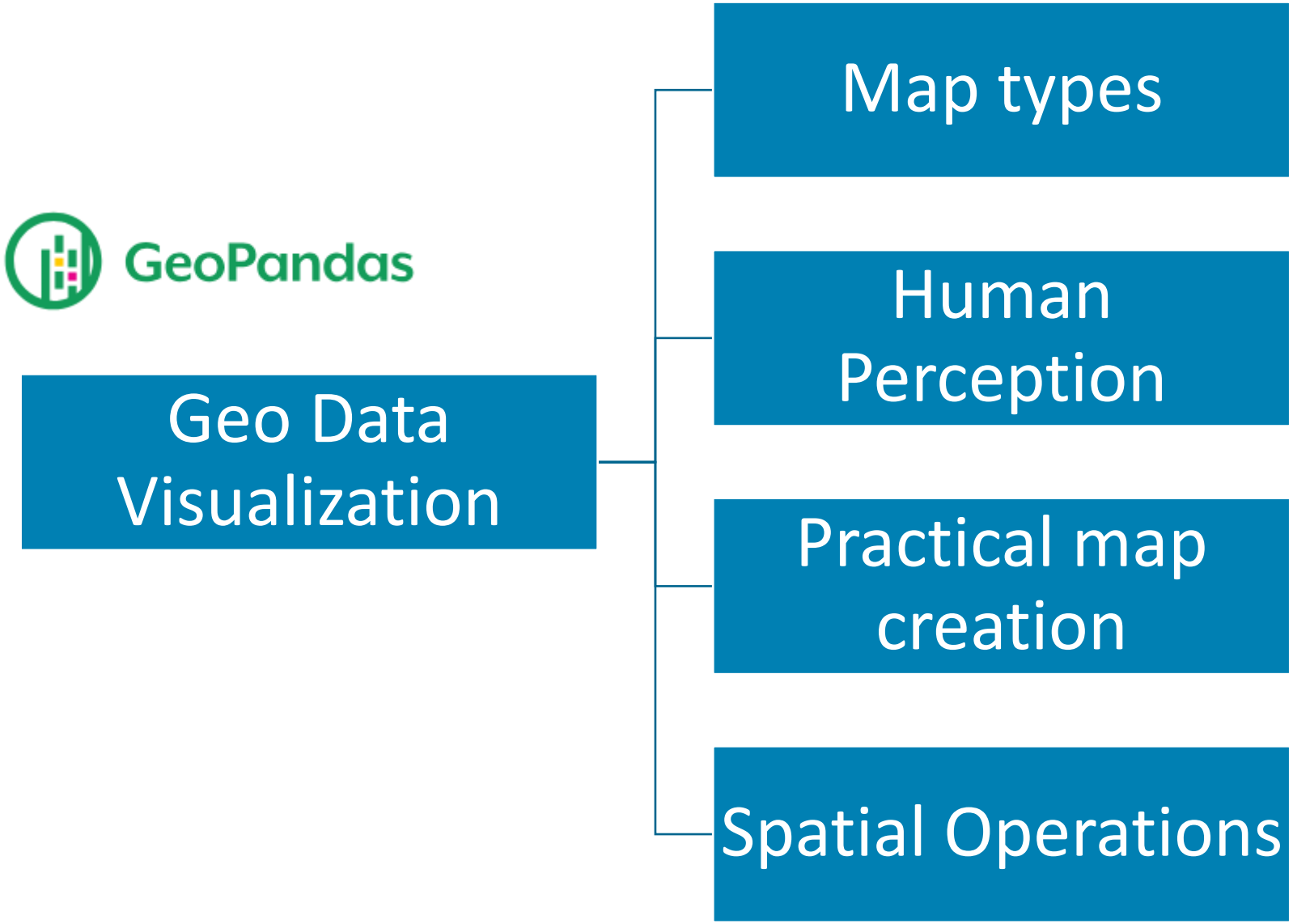
Storytelling

Perception + Visualization Design

Python + Tools

Interactive Visualization

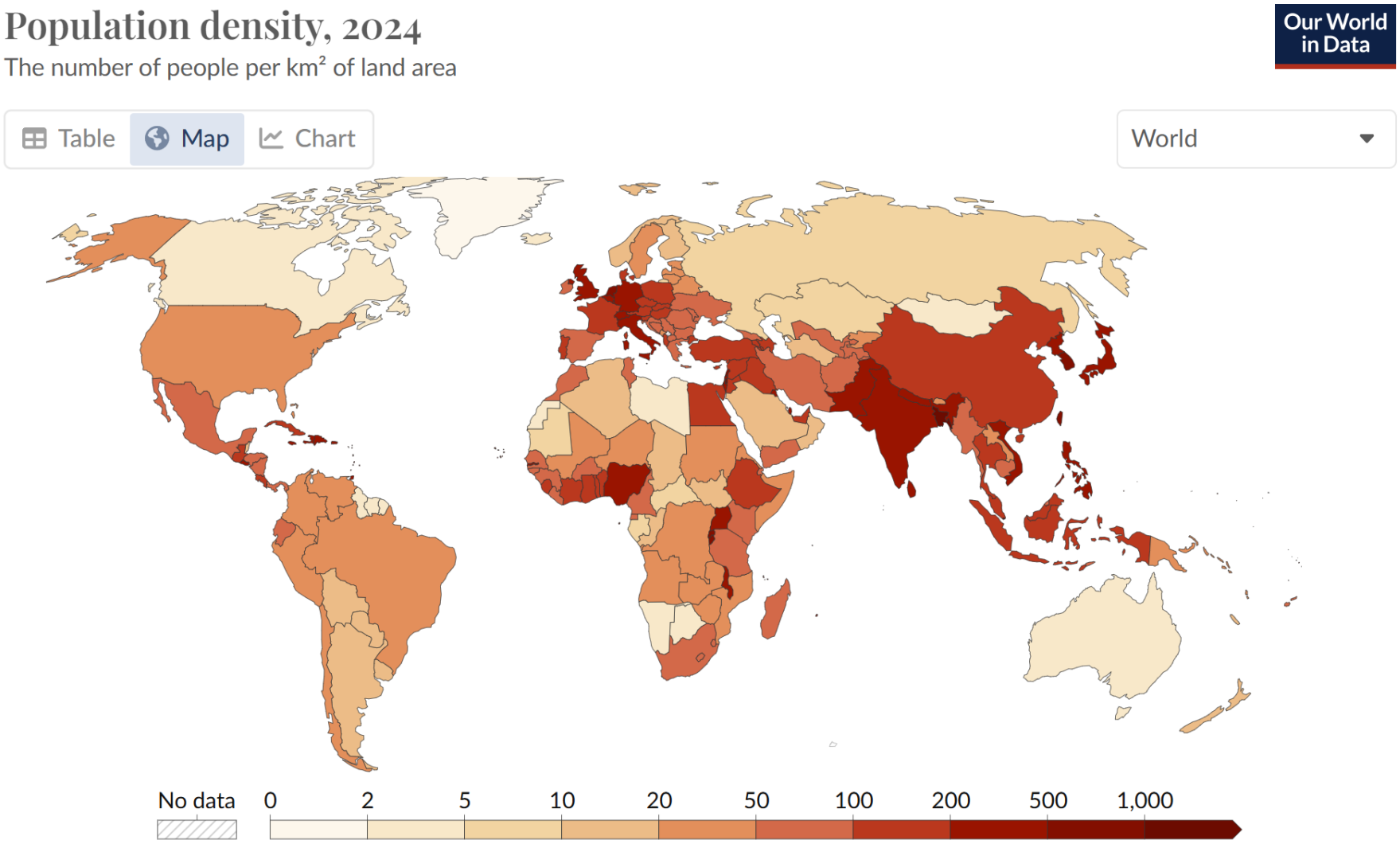
Roadmap


GeoPandas
Geo Data Visualization
Map types
Human Perception
Practical map creation
Spatial Operations

Map types

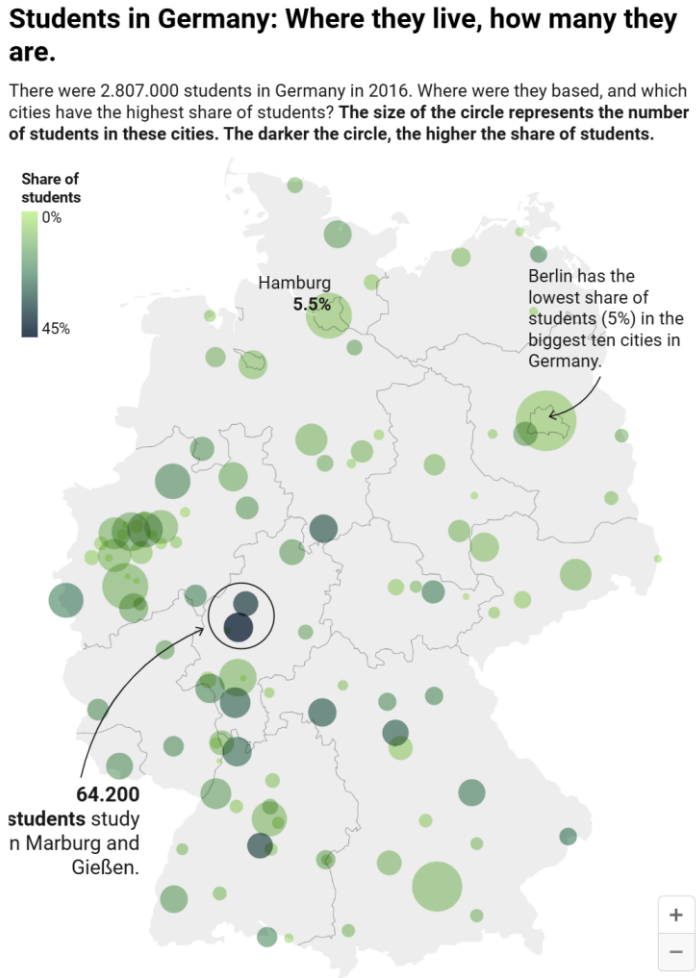
Choropleth map

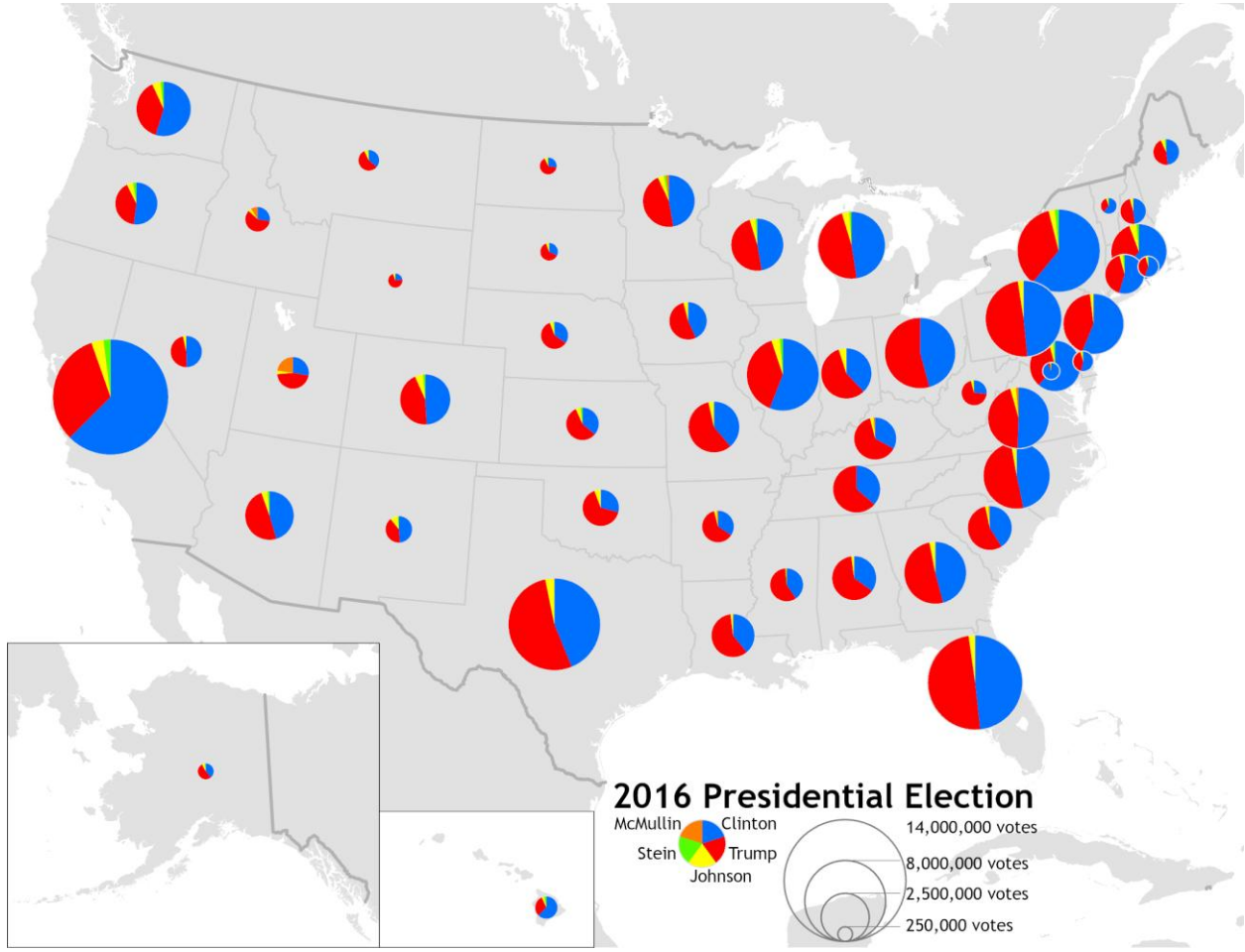
Use colored areas to represent data values



Symbol map

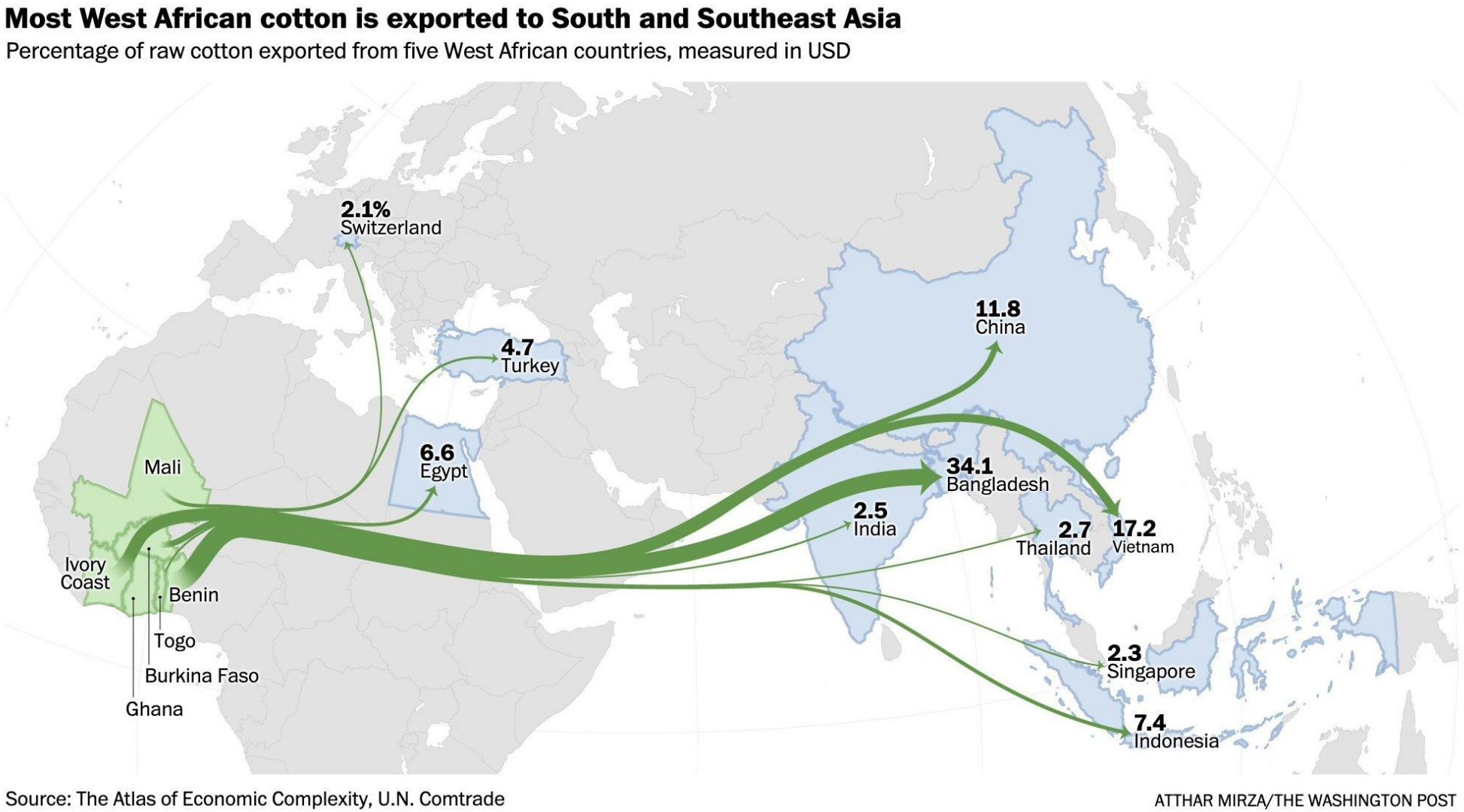
Use symbol size, shape or color to represent data values



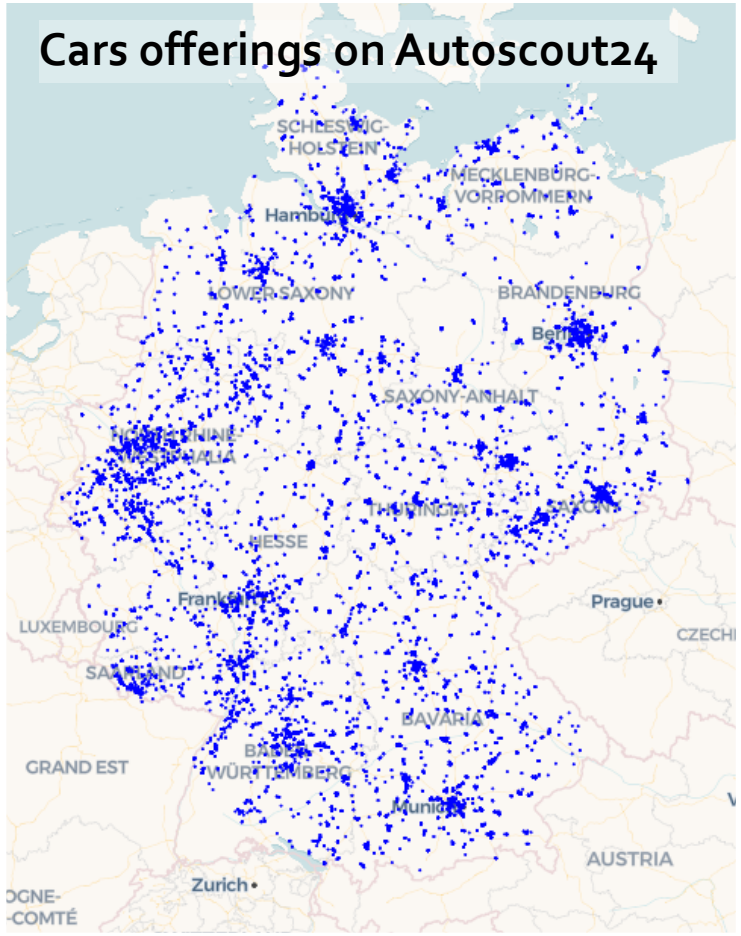


Flow map

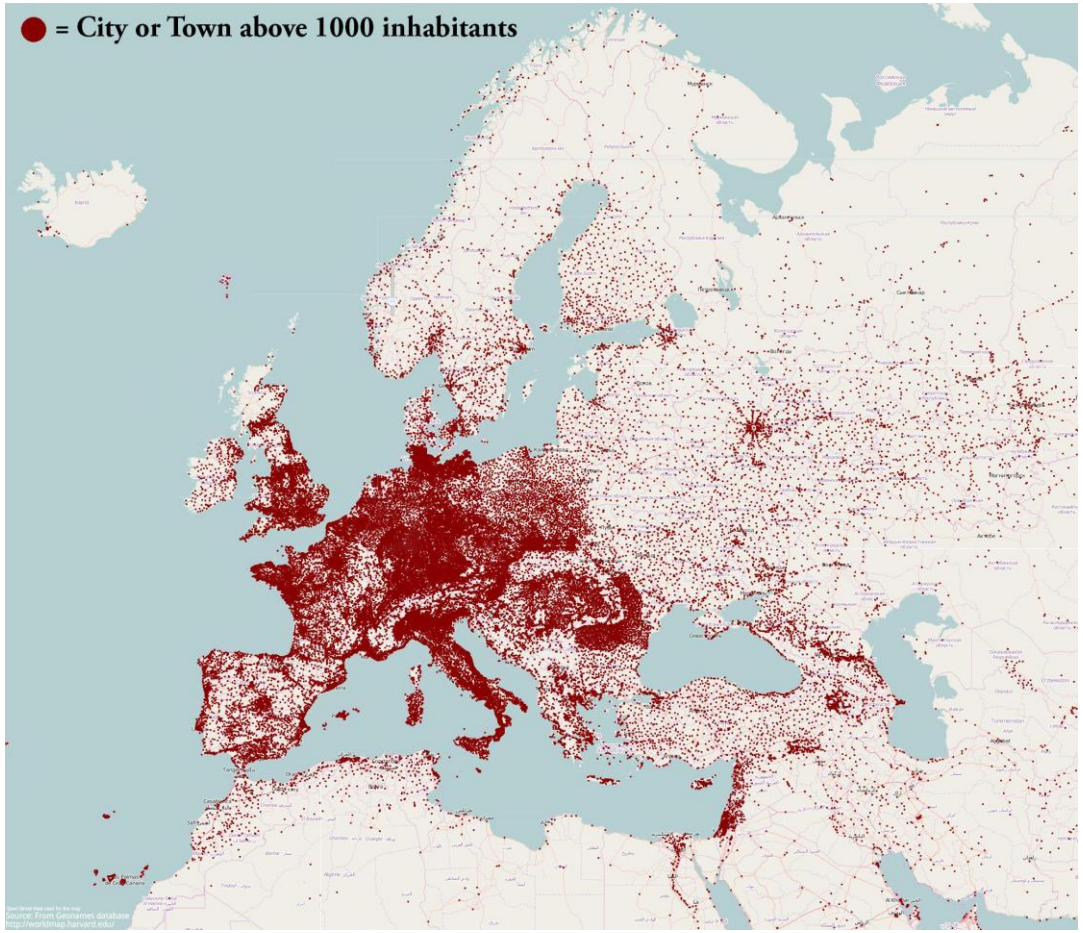
Use line width to represent flow magnitude



Dot density map



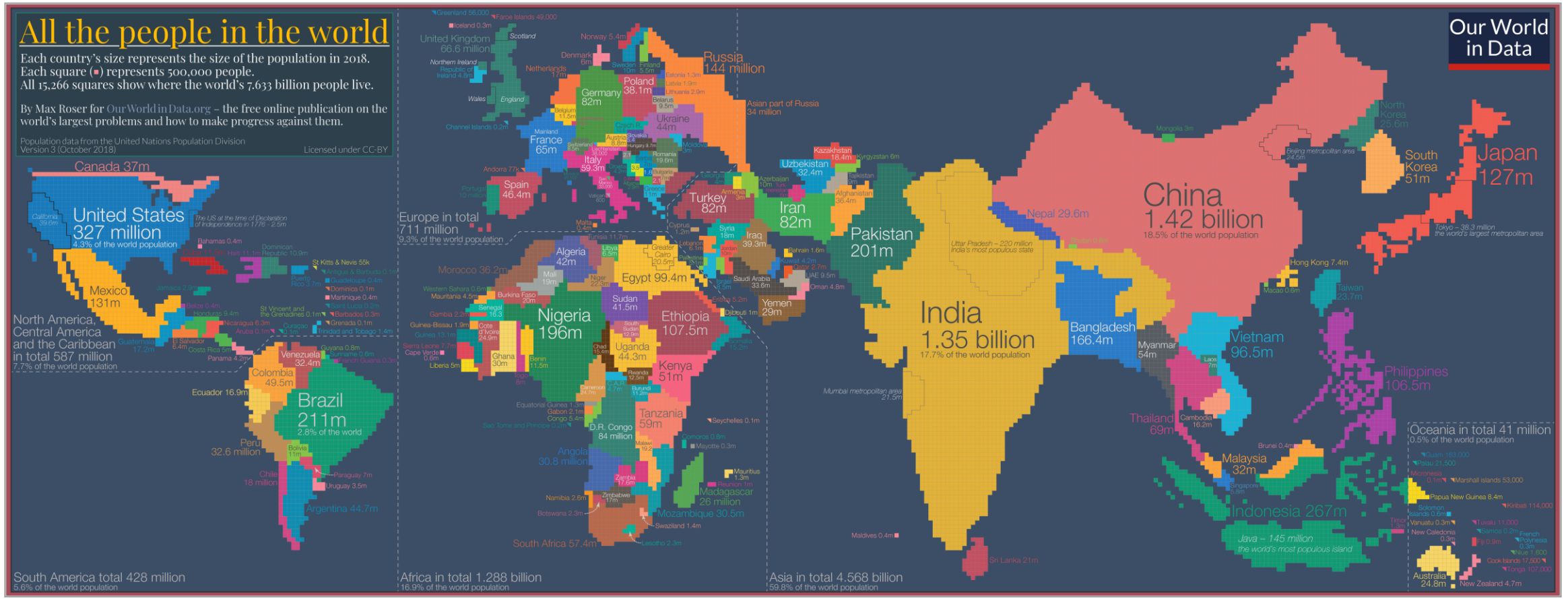
1 dot = 1 car



1 dot = 1000+ people

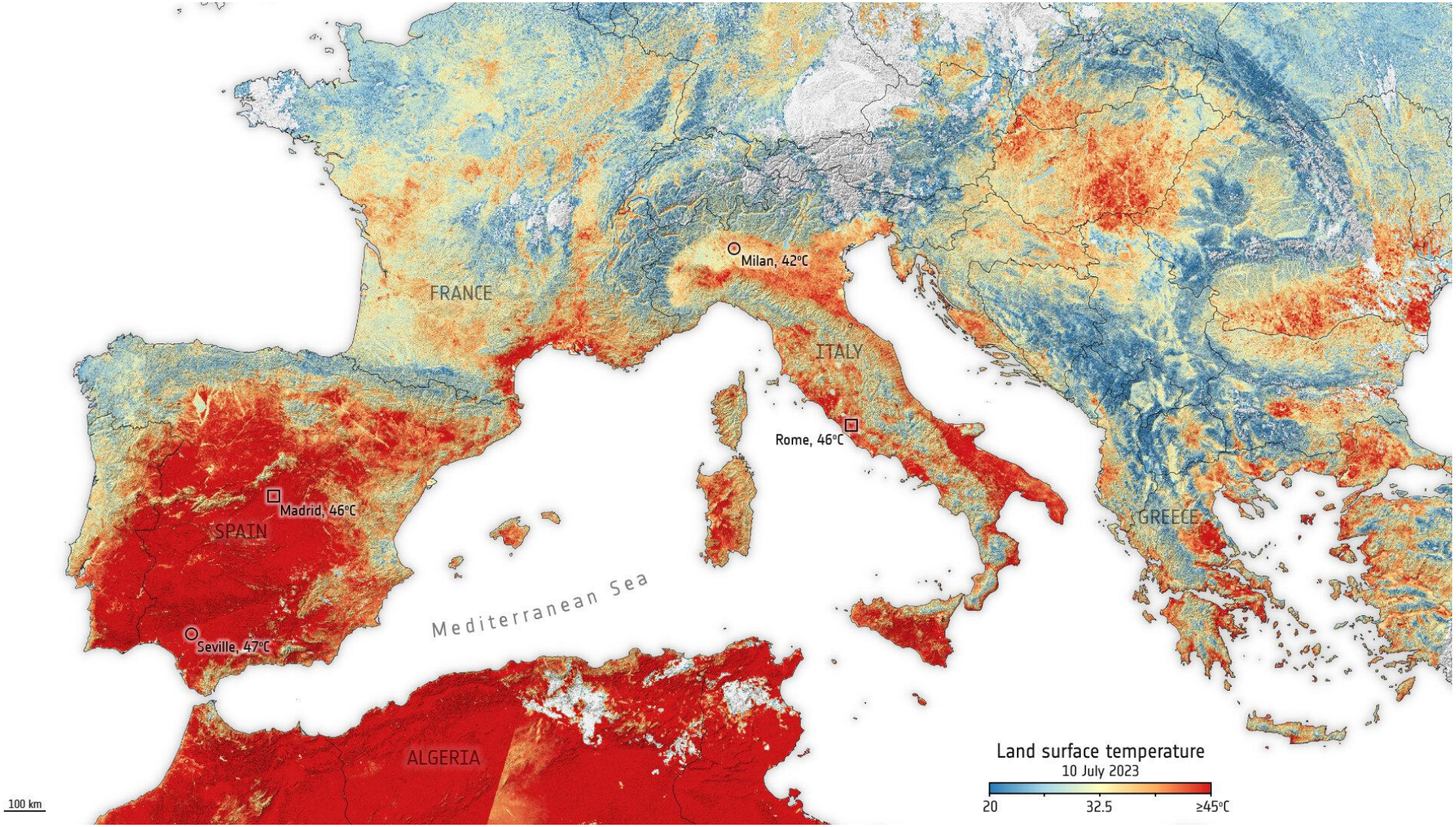
Cartogram

Use size of geographic area to represent data values



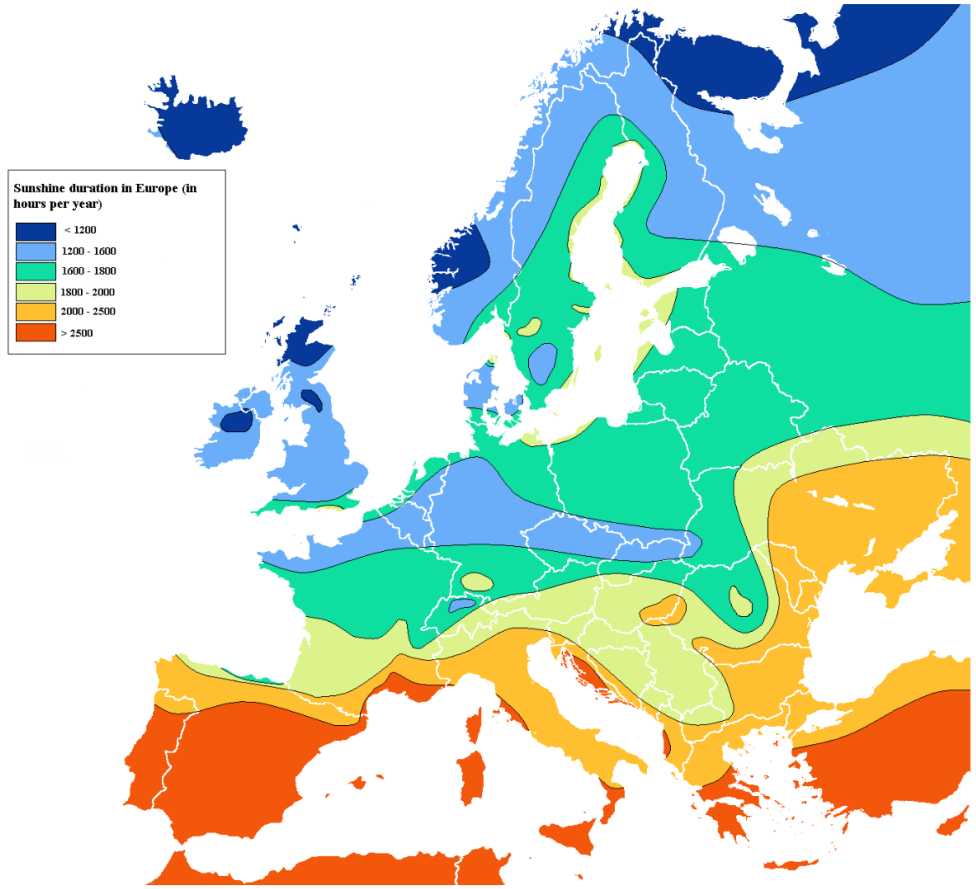
Raster map

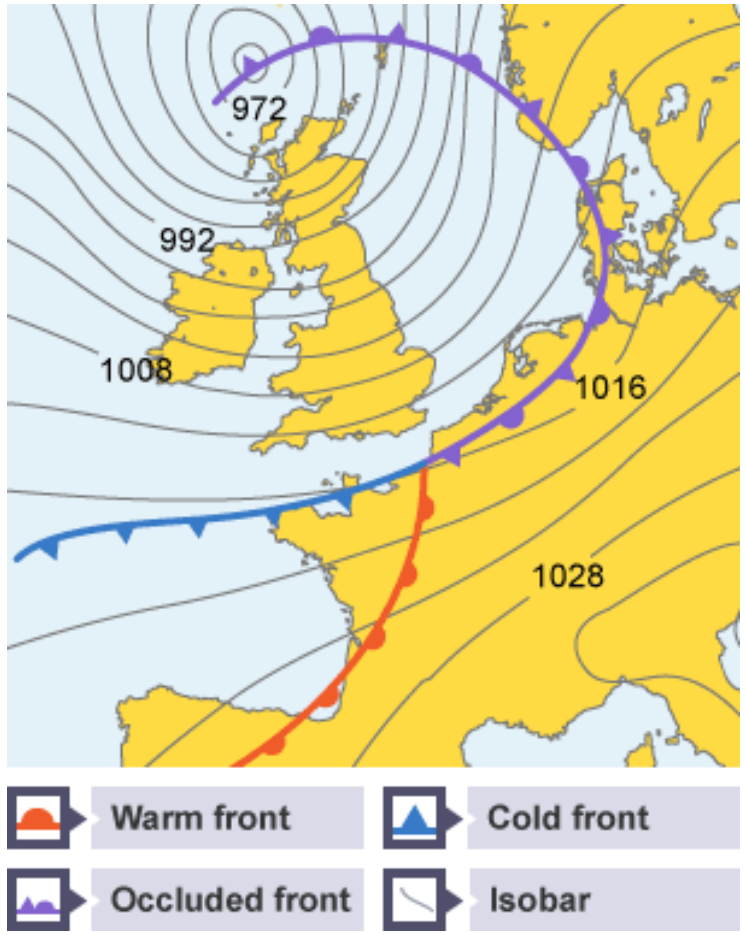
Use color of single pixels to represent data values for continuous phenomena



Isoline Map

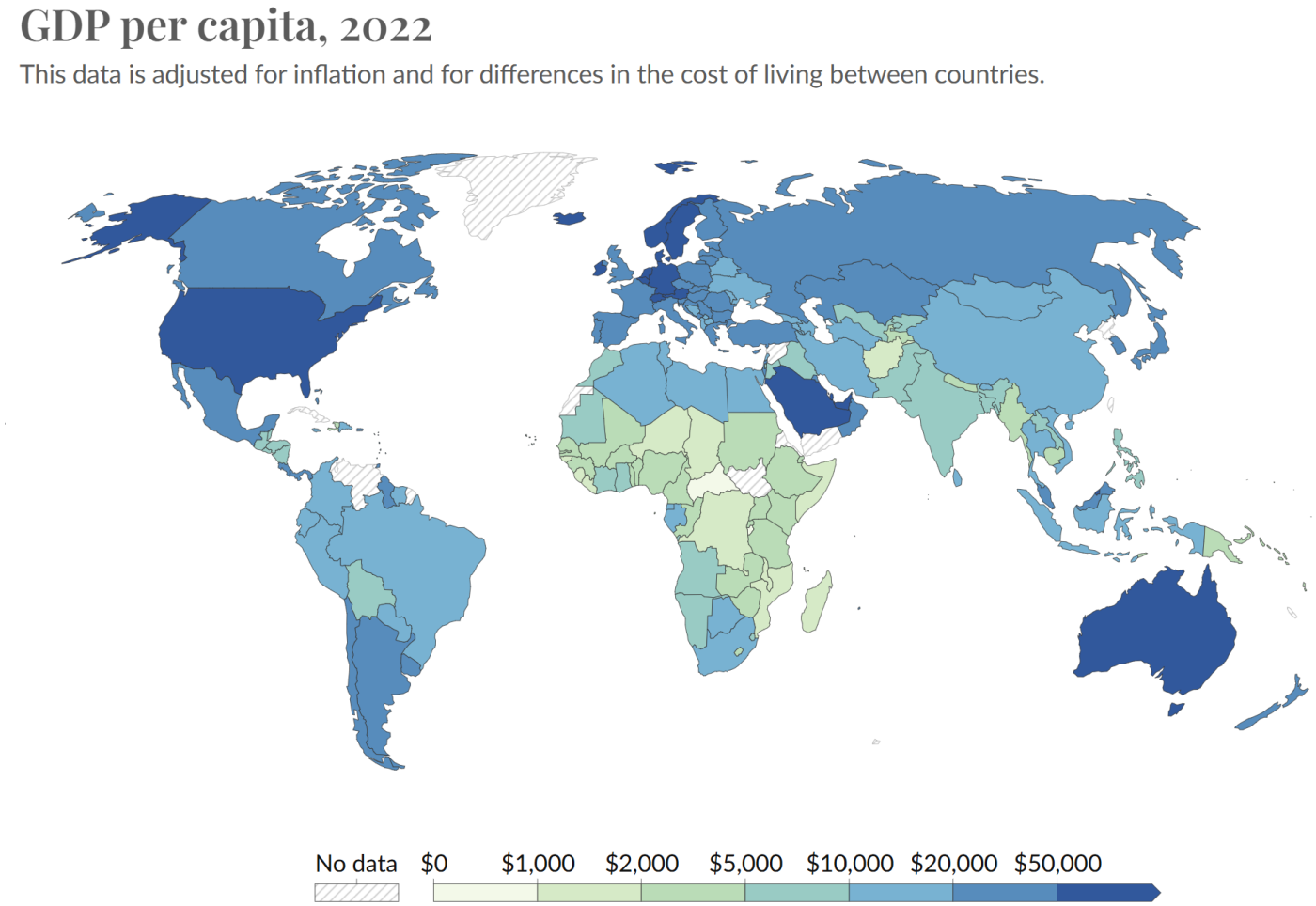
Use Lines to represent areas with equal values of a continuous phenomenon





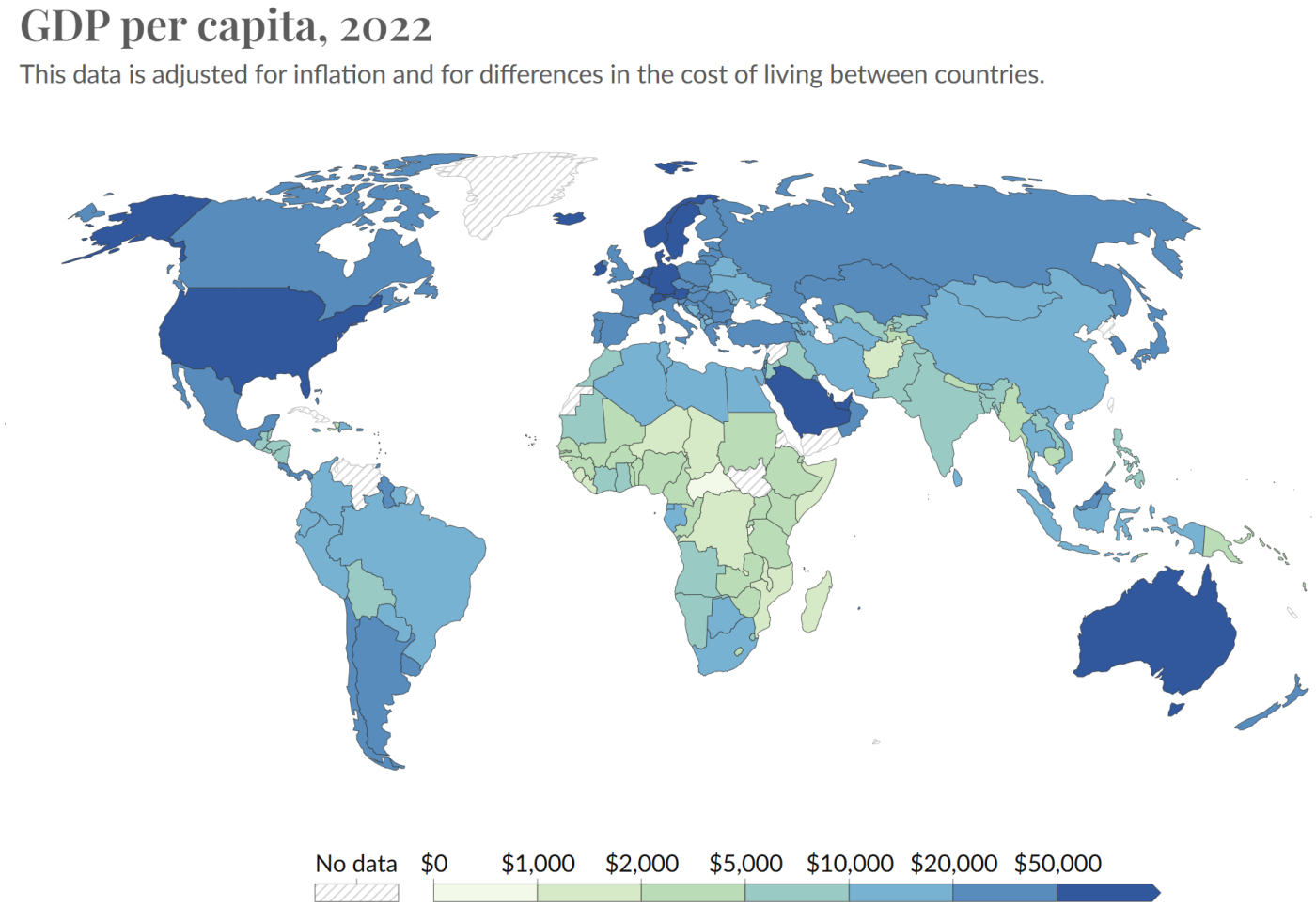
Human perception

Advantages of choropleth maps



- Complete and compact representation of many data points
- Familiarity
- Useful if strong geographic patterns are present

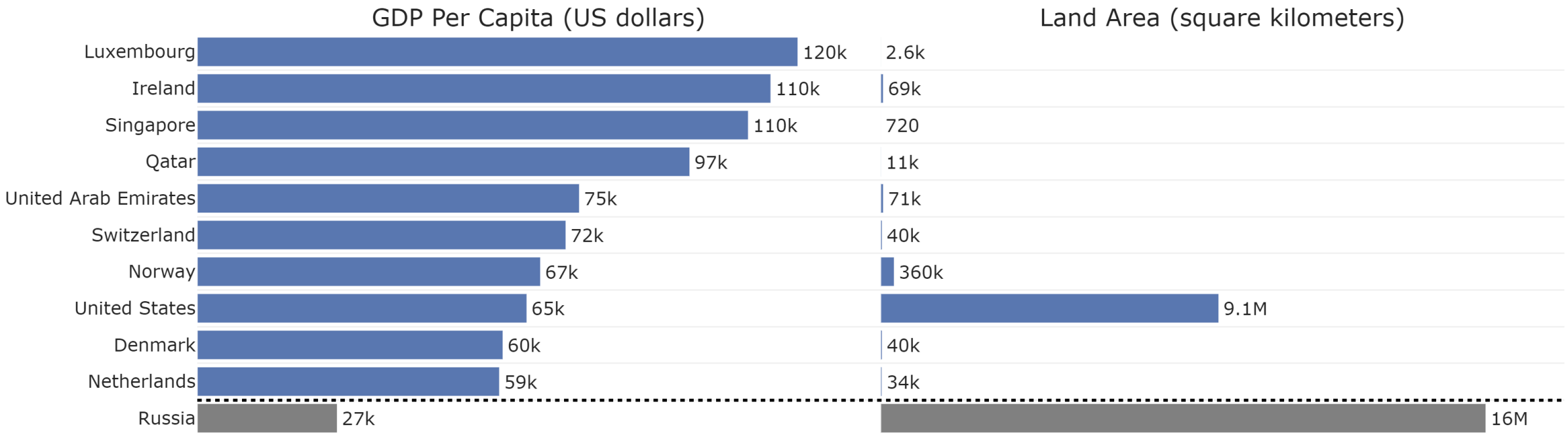
Disadvantages?



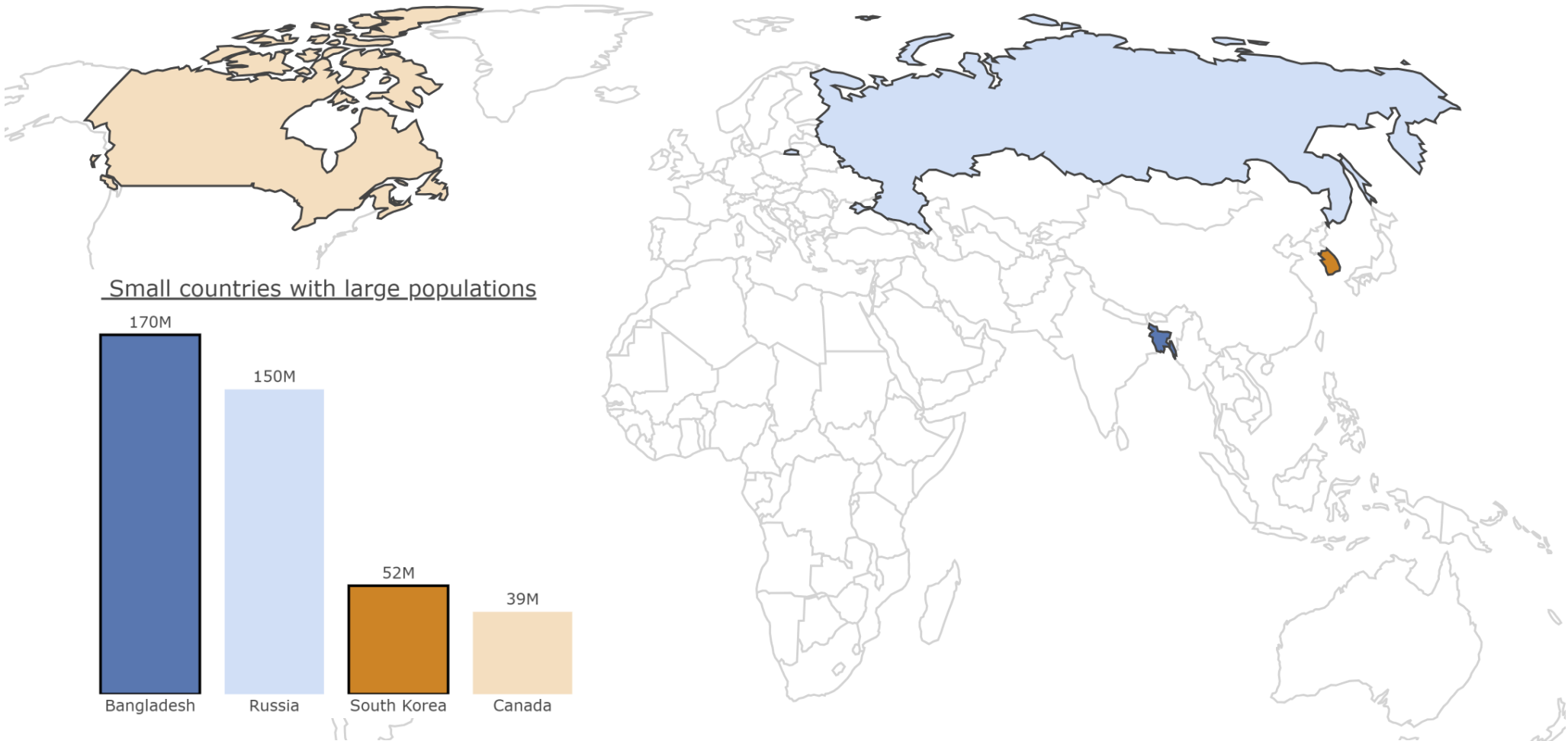
What is the GDP per capita of Singapur, Luxembourg, and Denmark?

Bias towards large countries

- Only large countries are perceived well
- Most countries are small →they are hardly perceived on the map



Area is a questionable proxy for importance

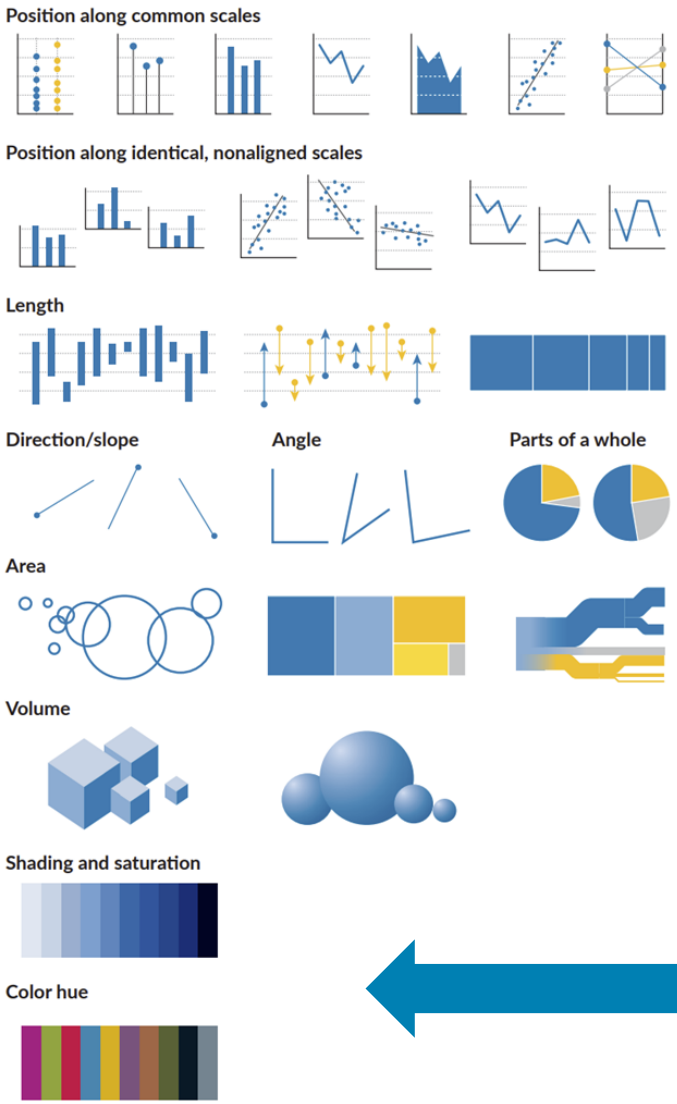


Bangladesh has less than 1% of the area of **Russia**, but a larger population

South Korea has about 1% of the area of **Canada**, but a larger population

Inaccurate color perception

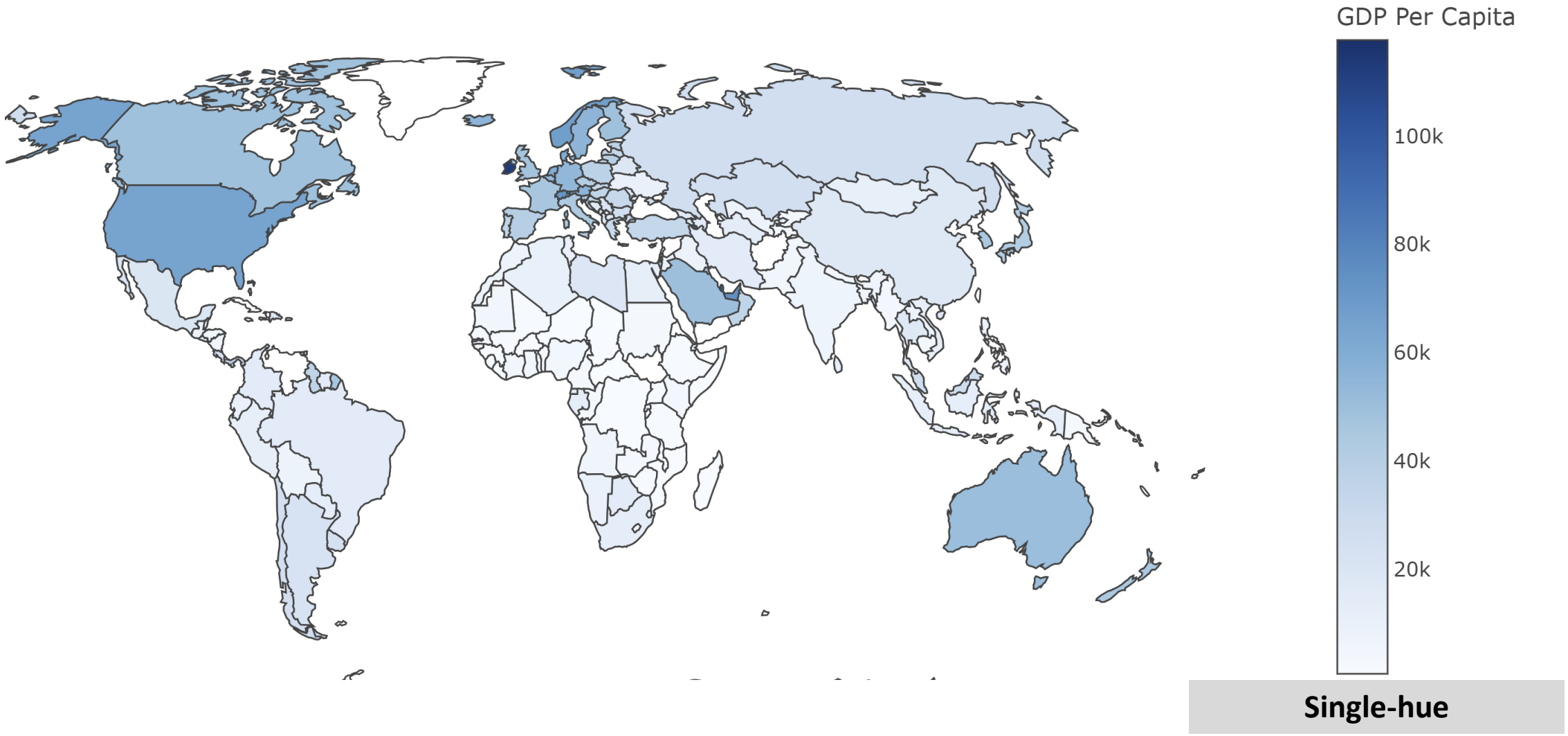
More accurate



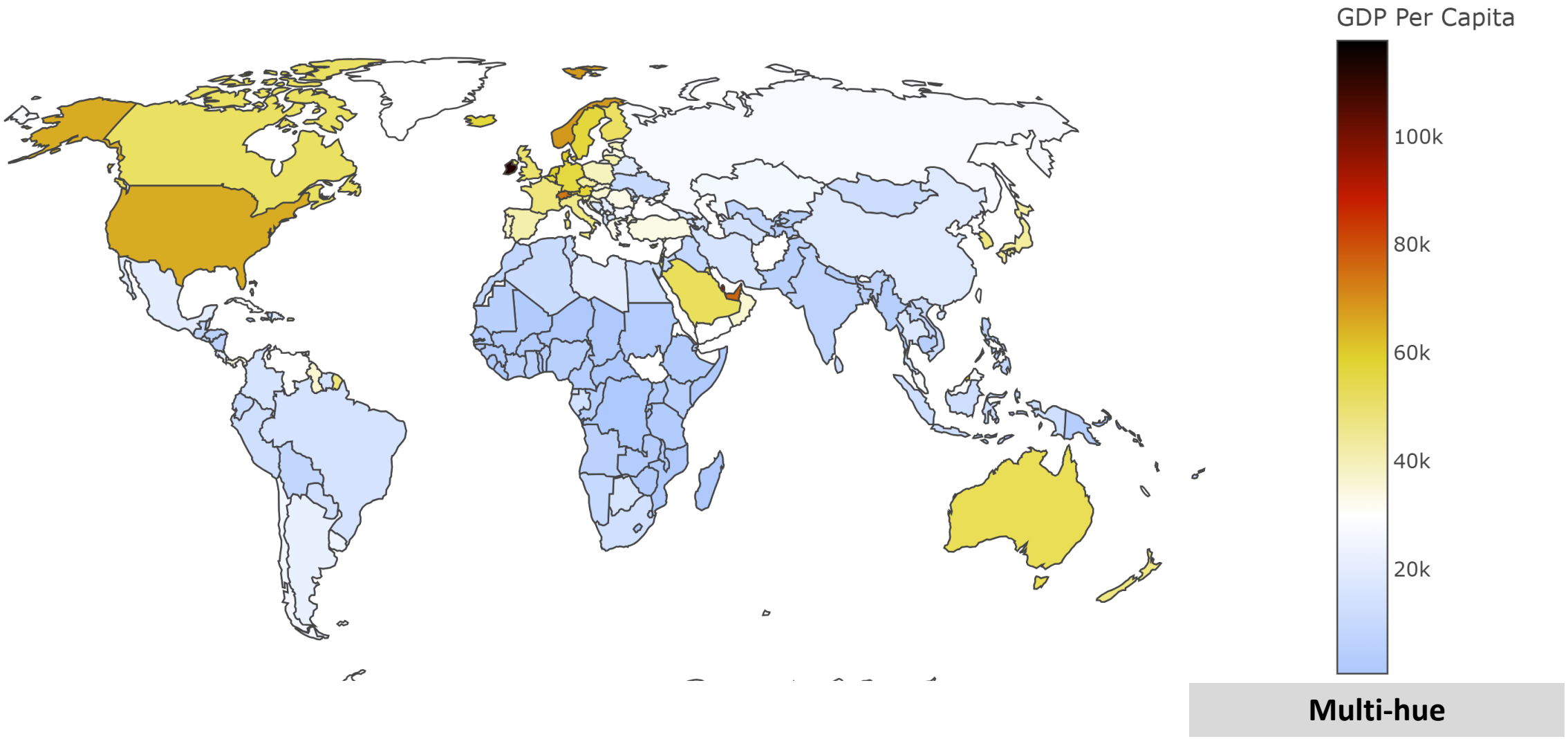
Less accurate

Choropleth maps encode data using colors, but humans are bad in decoding colors

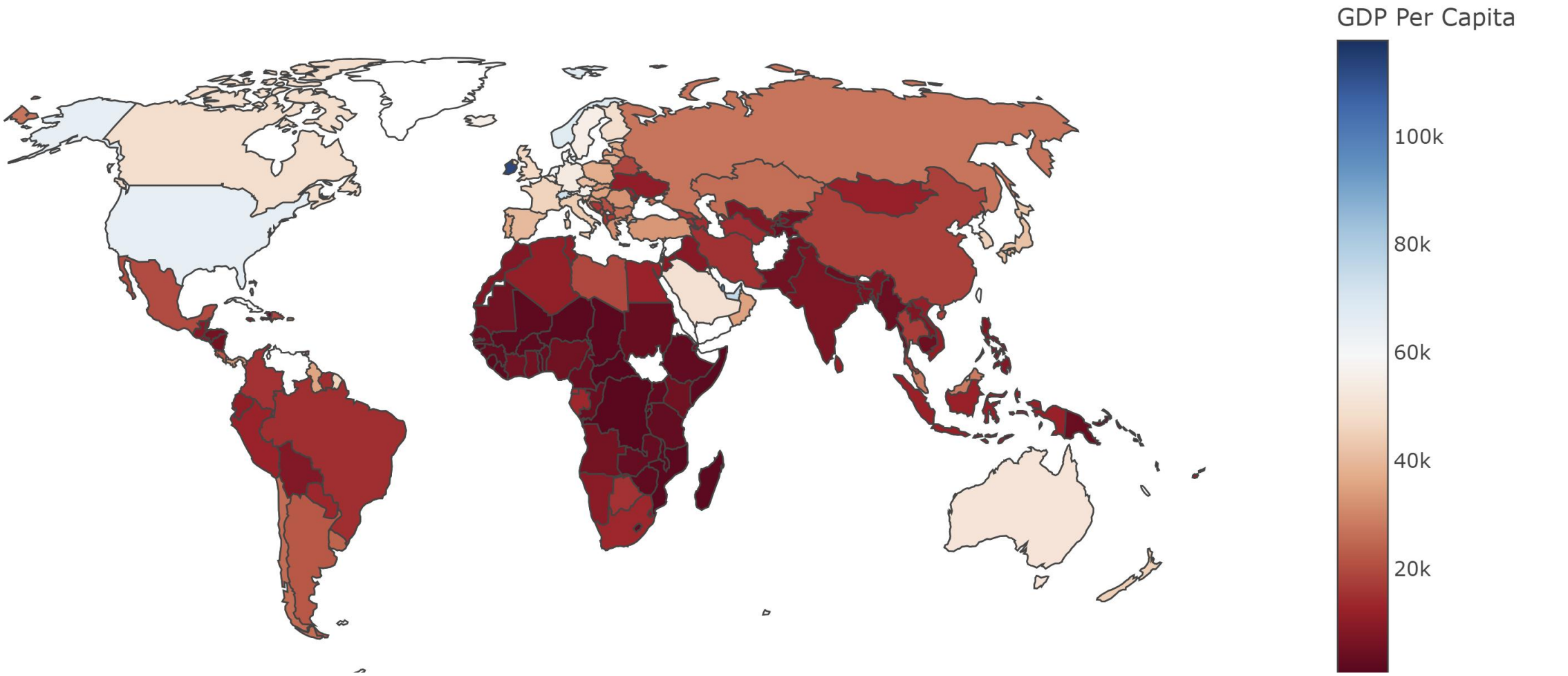
Color scales affect interpretation



Color scales affect interpretation

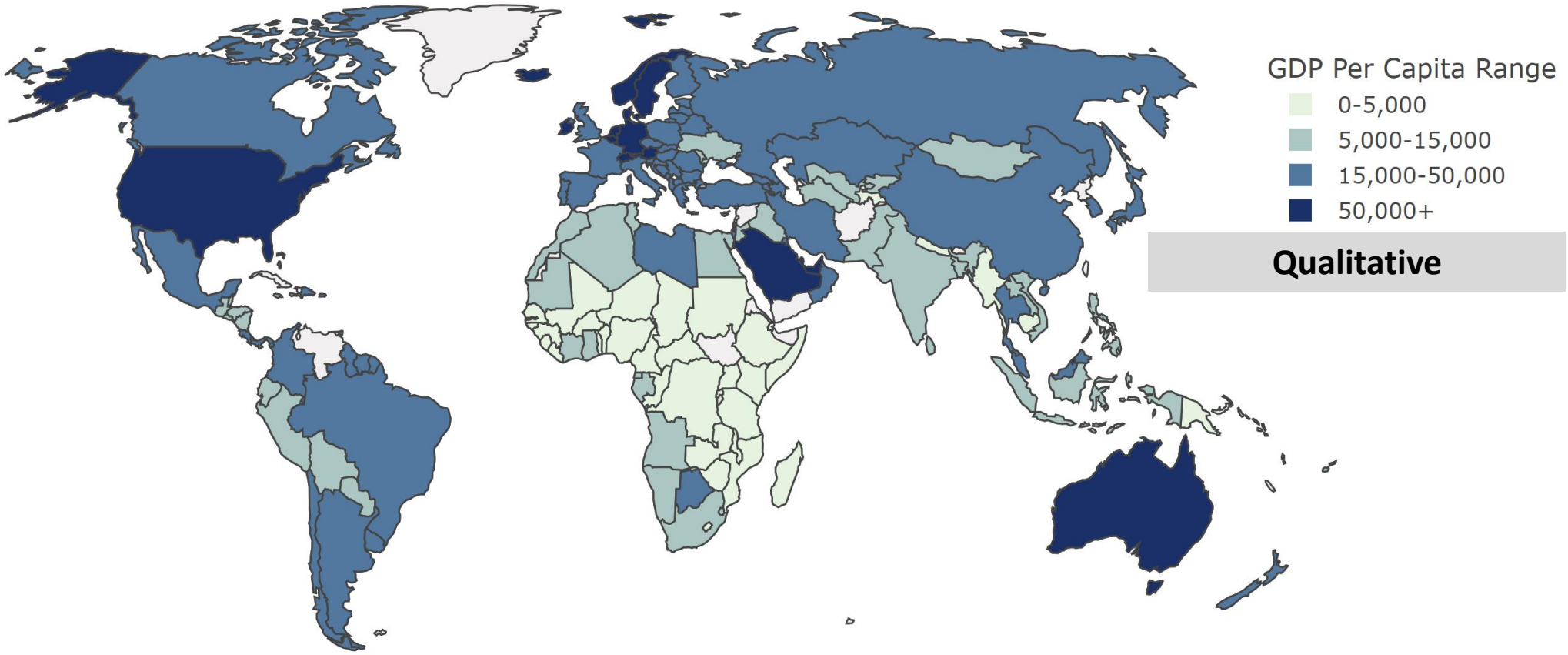


Color scales affect interpretation



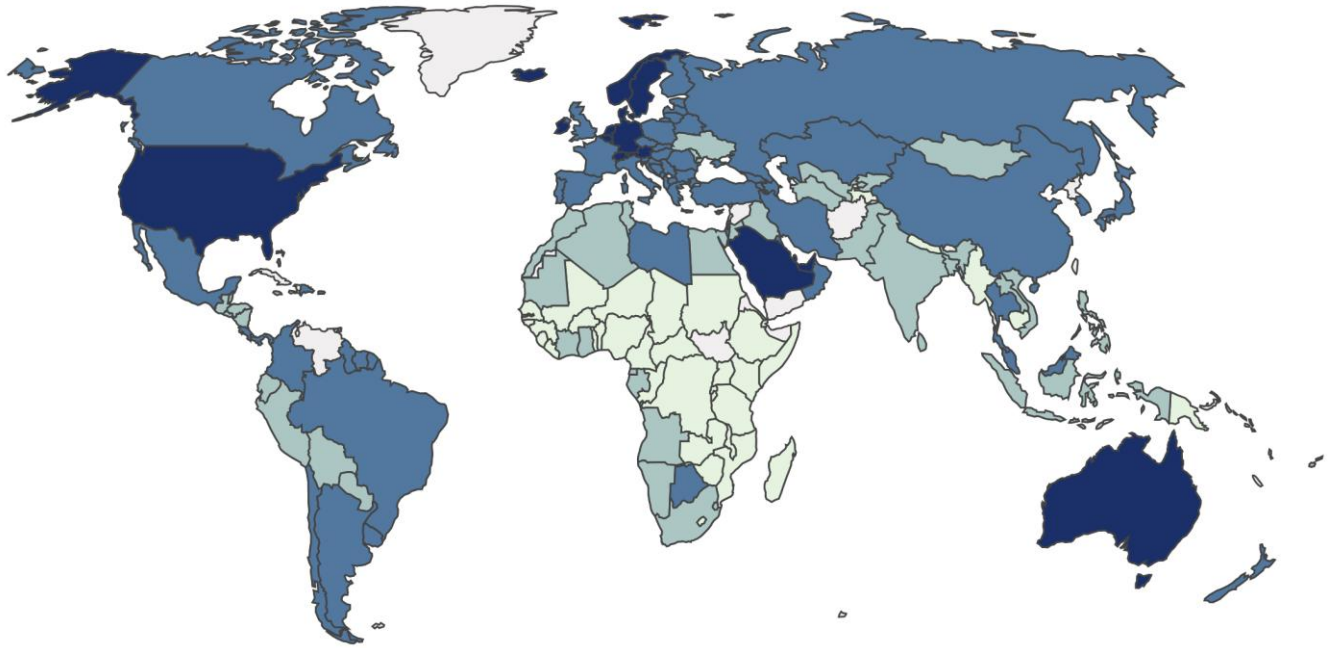
Diverging

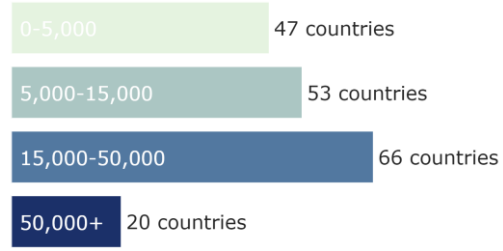
Color scales affect interpretation

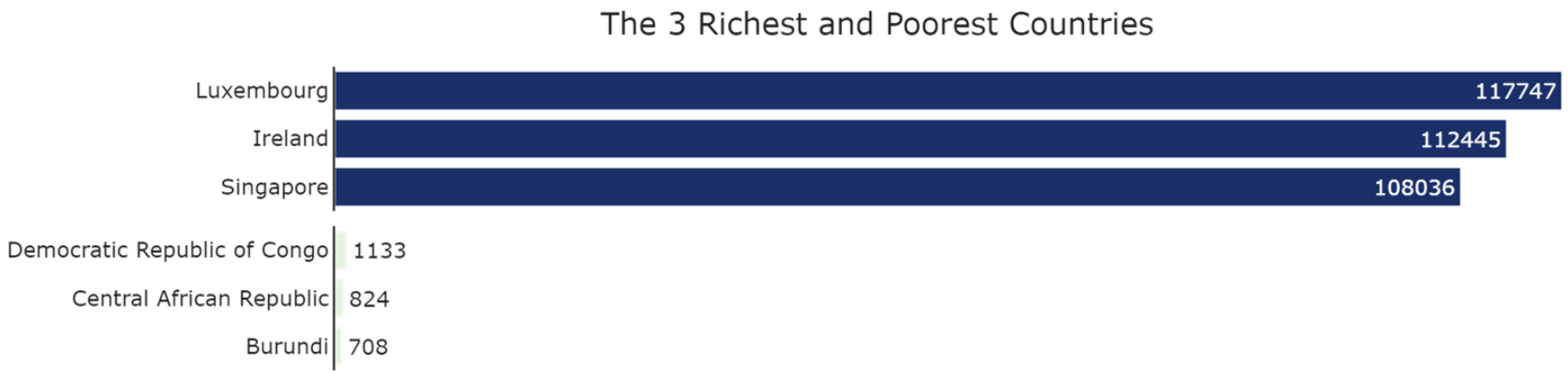


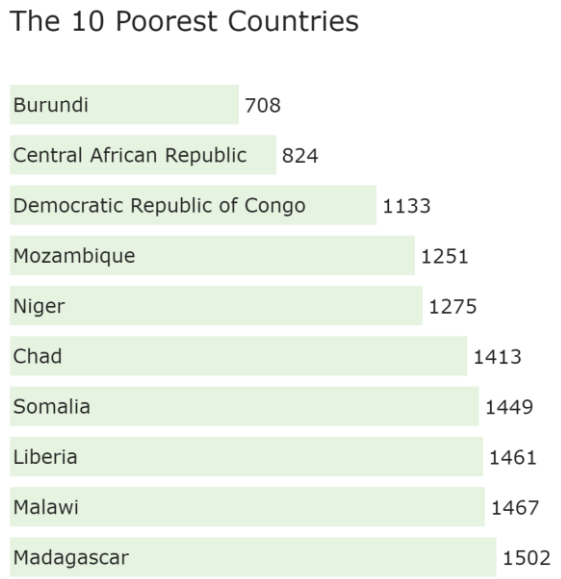
Add complementary visualizations

Maps can show high-level geographic patterns – for details we need additional plots

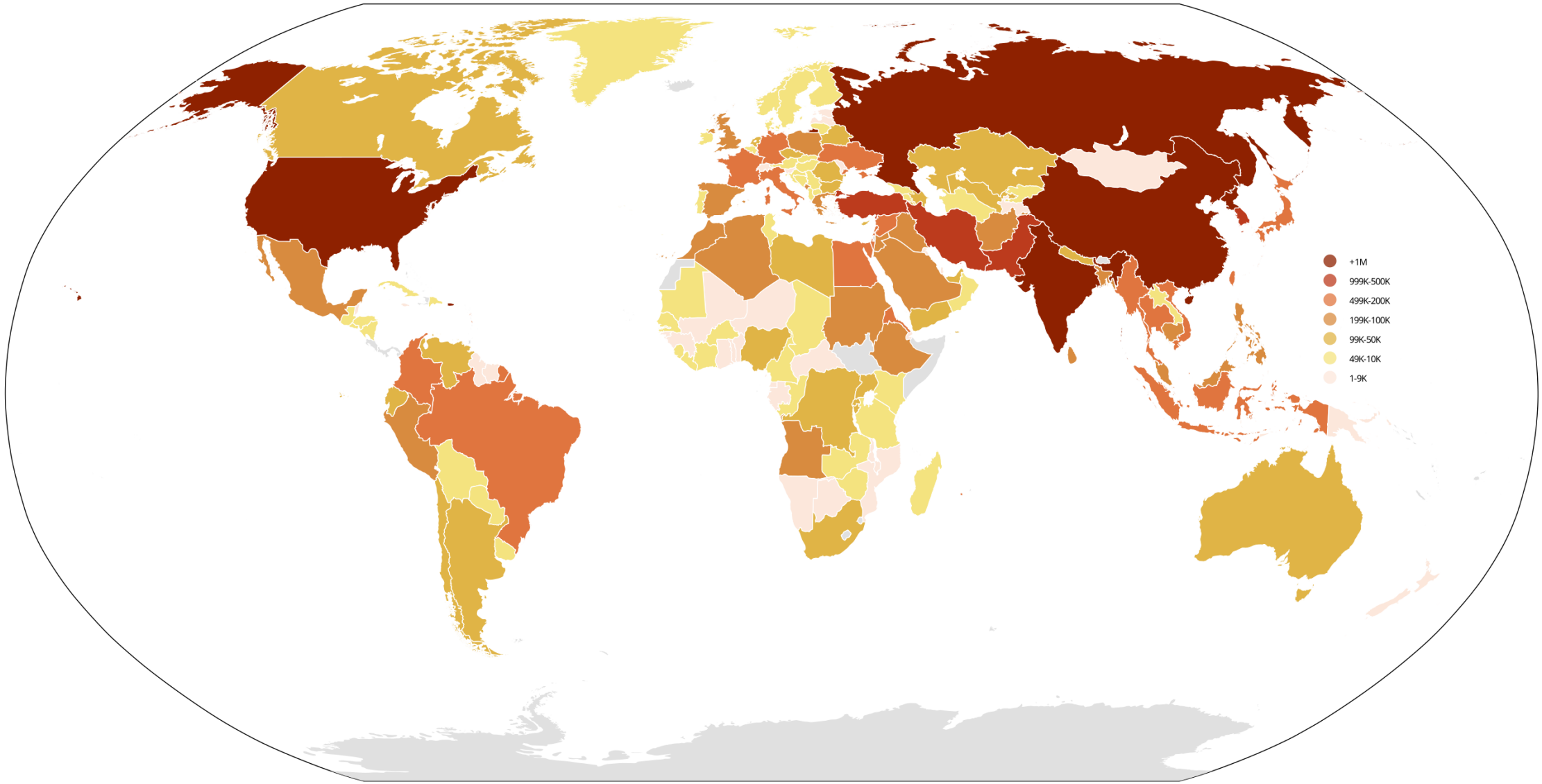




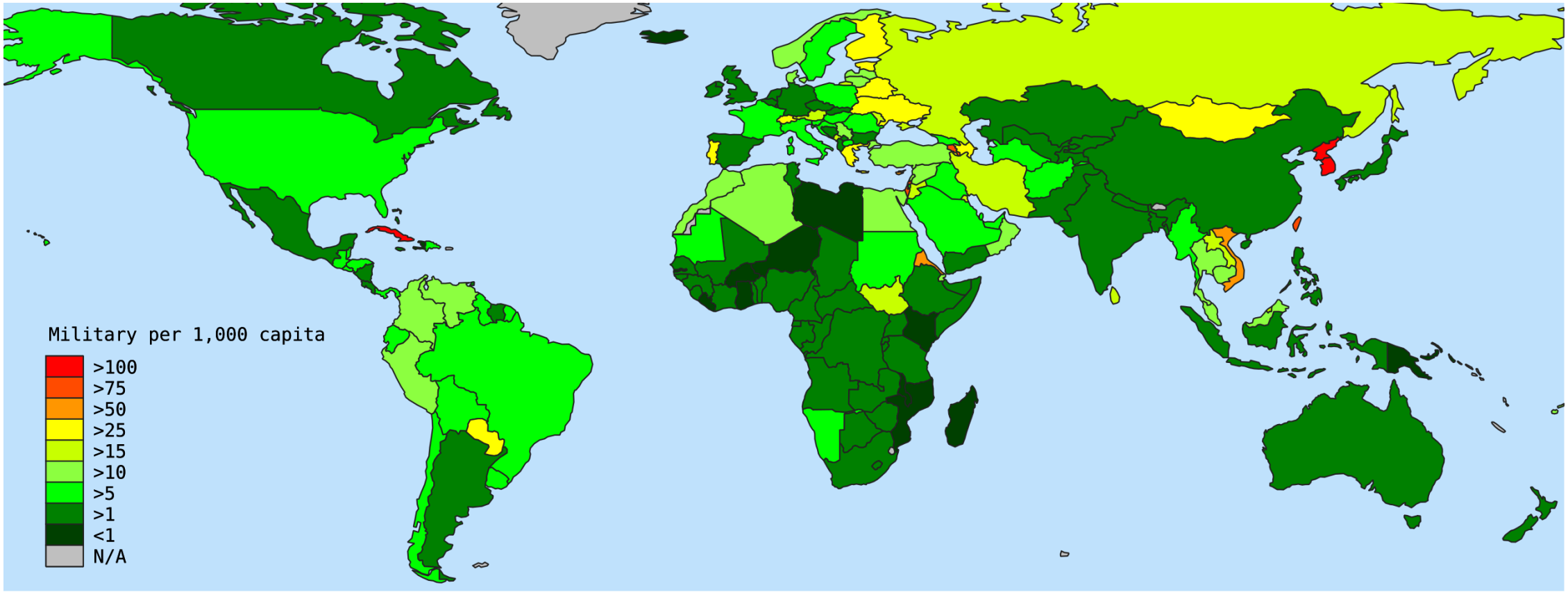




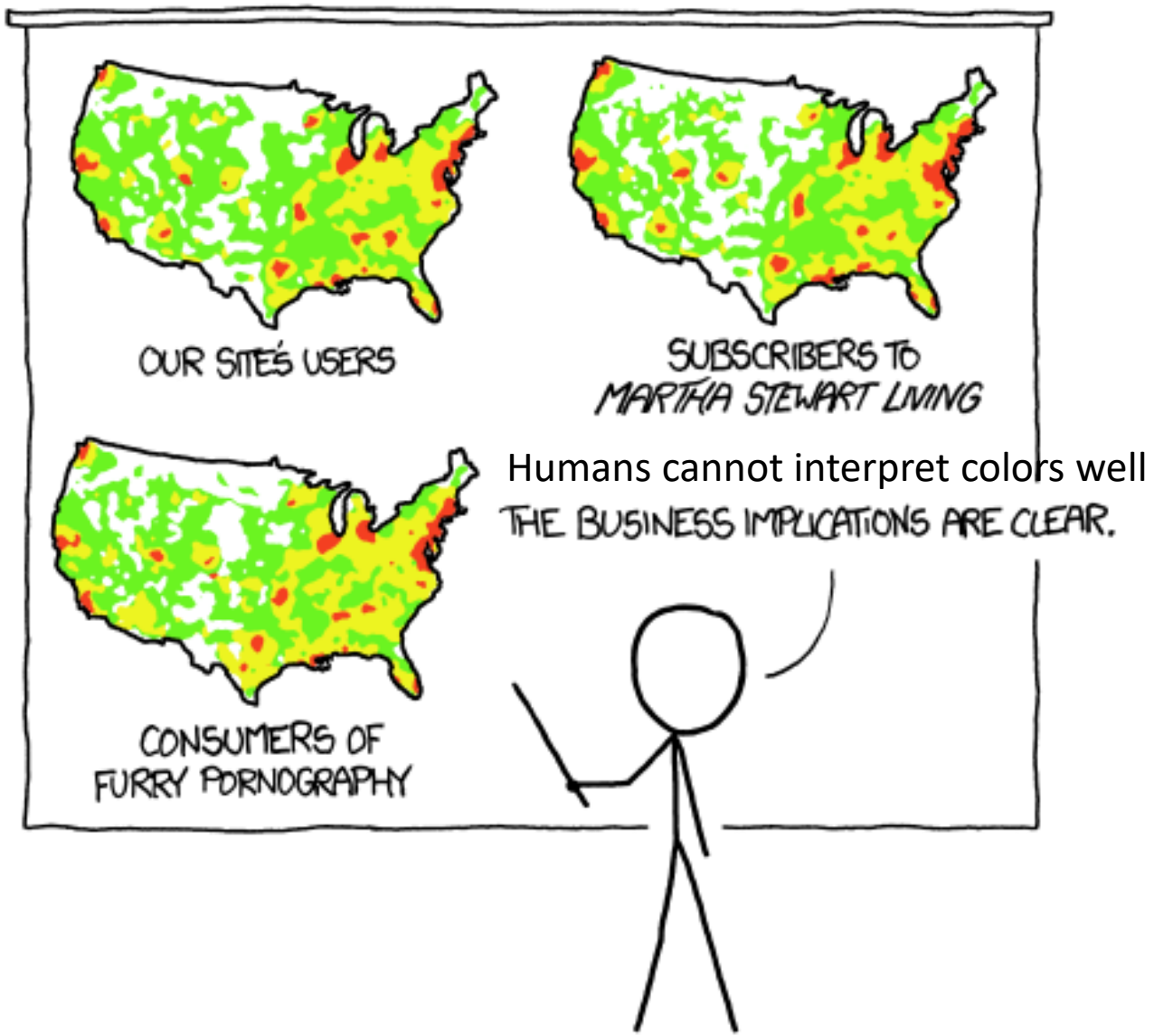
Number of soldiers



Number of soldiers per 1000 inhabitants



Normalize your data!



Often data correlates with population size or land mass

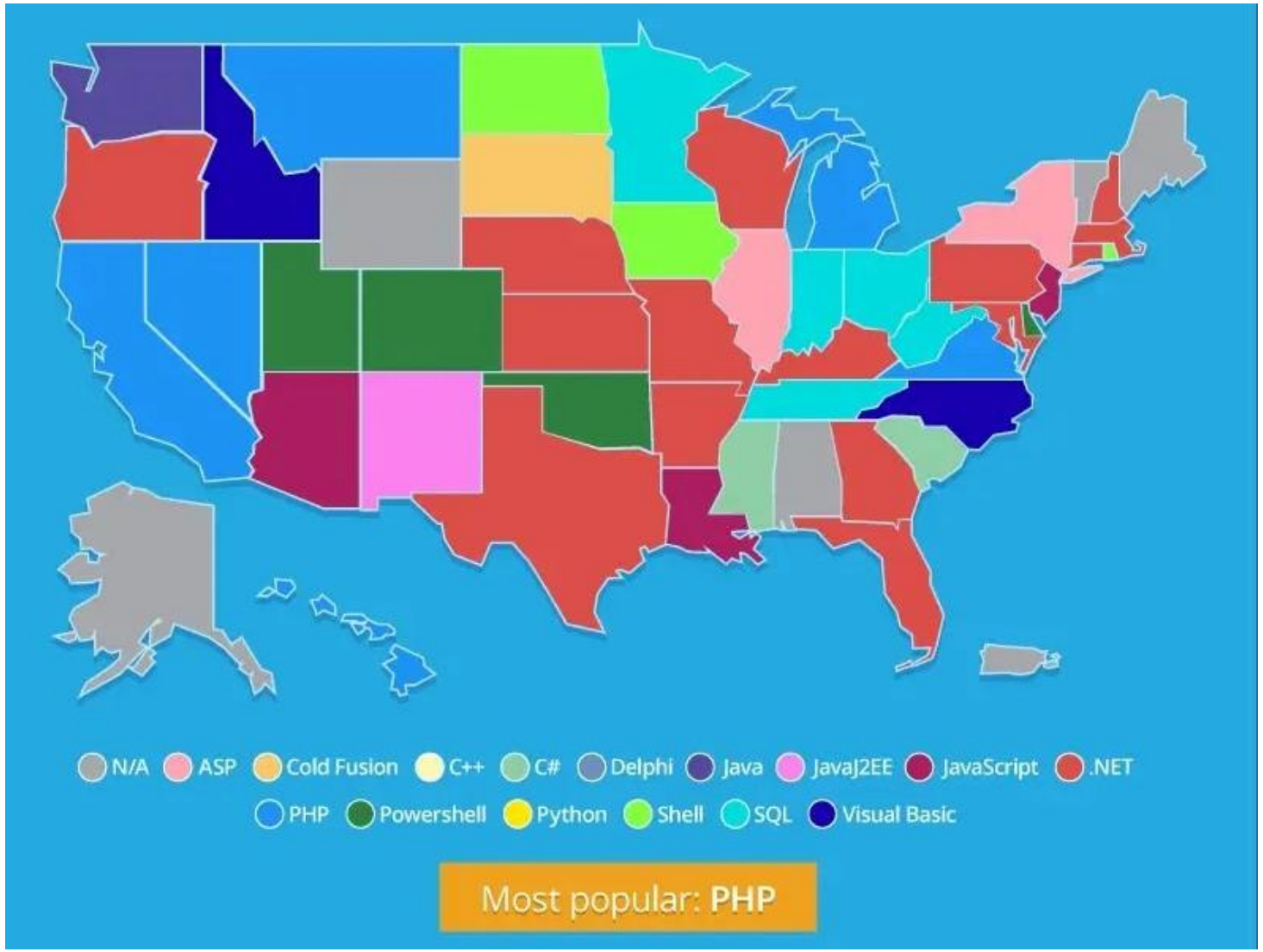
→ **To make data comparable, normalize adequately!**

- per capita
- per land area

Otherwise all maps essentially just show population/country size

Sources: https://xkcd.com/1138/

When to avoid maps

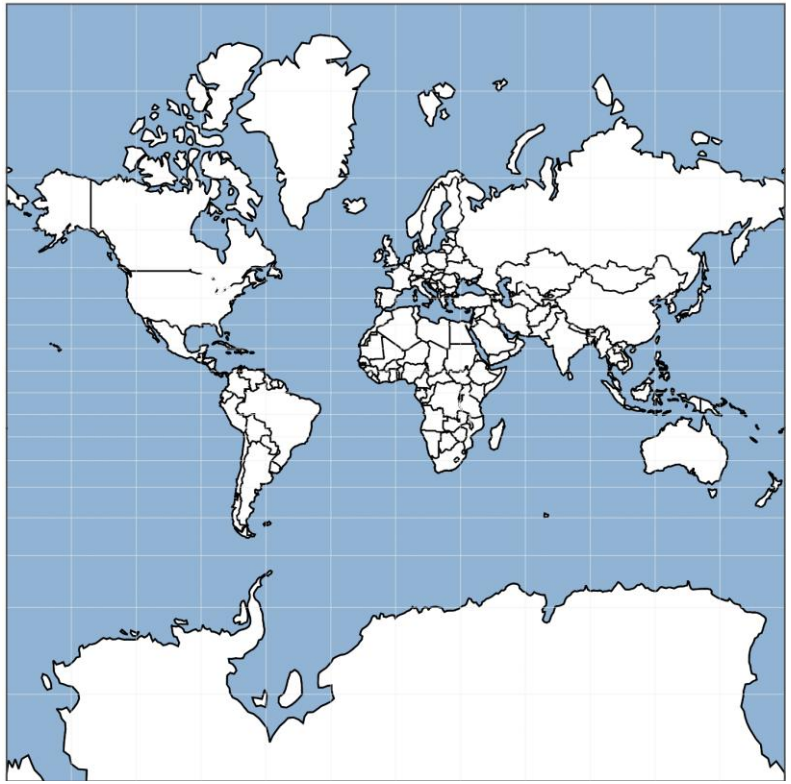


If you present data in form of a map, the audience will expect geographic patterns.

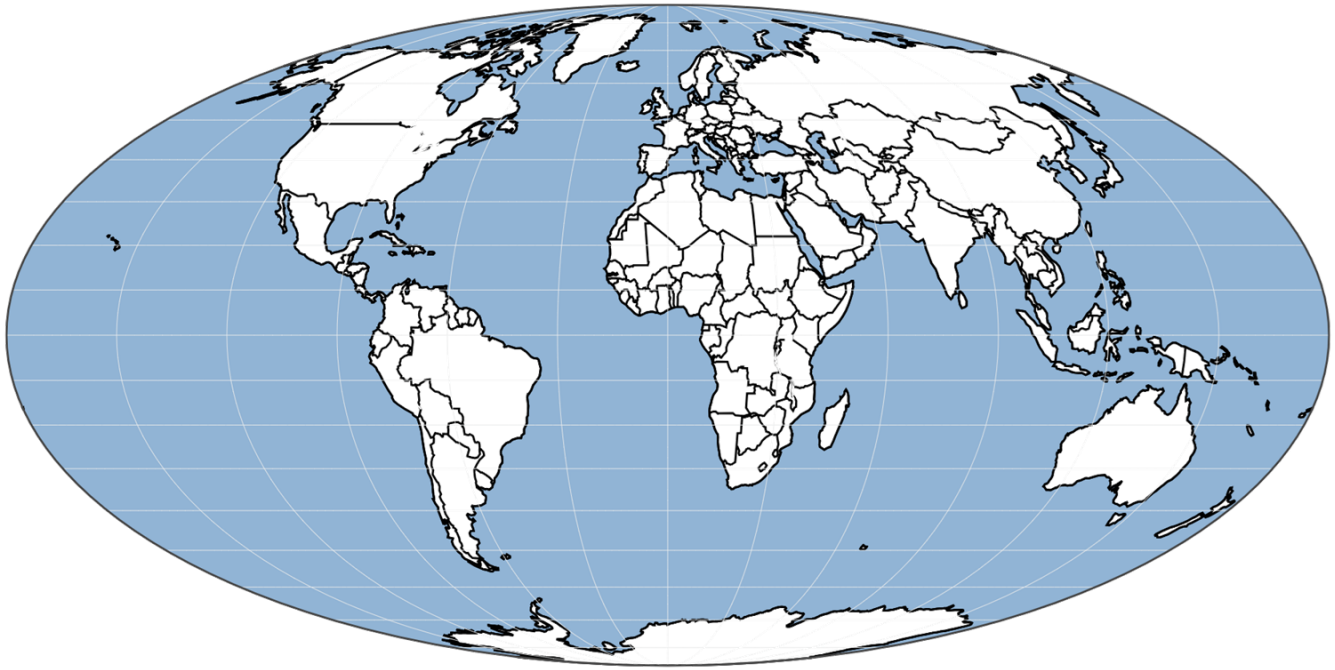
→ In the **absence of clear geographic patterns,** the audience will be confused

→ Consider a differerent plot type!

Map projections distort perception

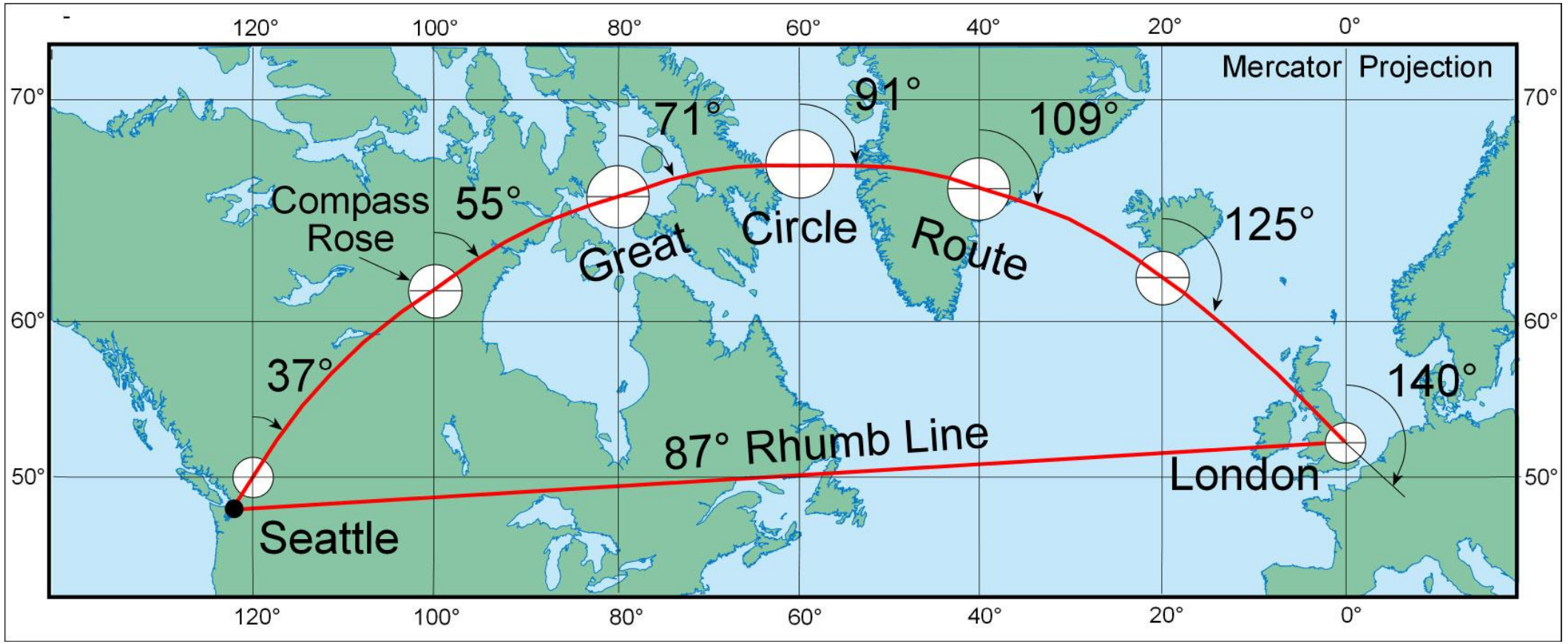


Mercator



Mollweide

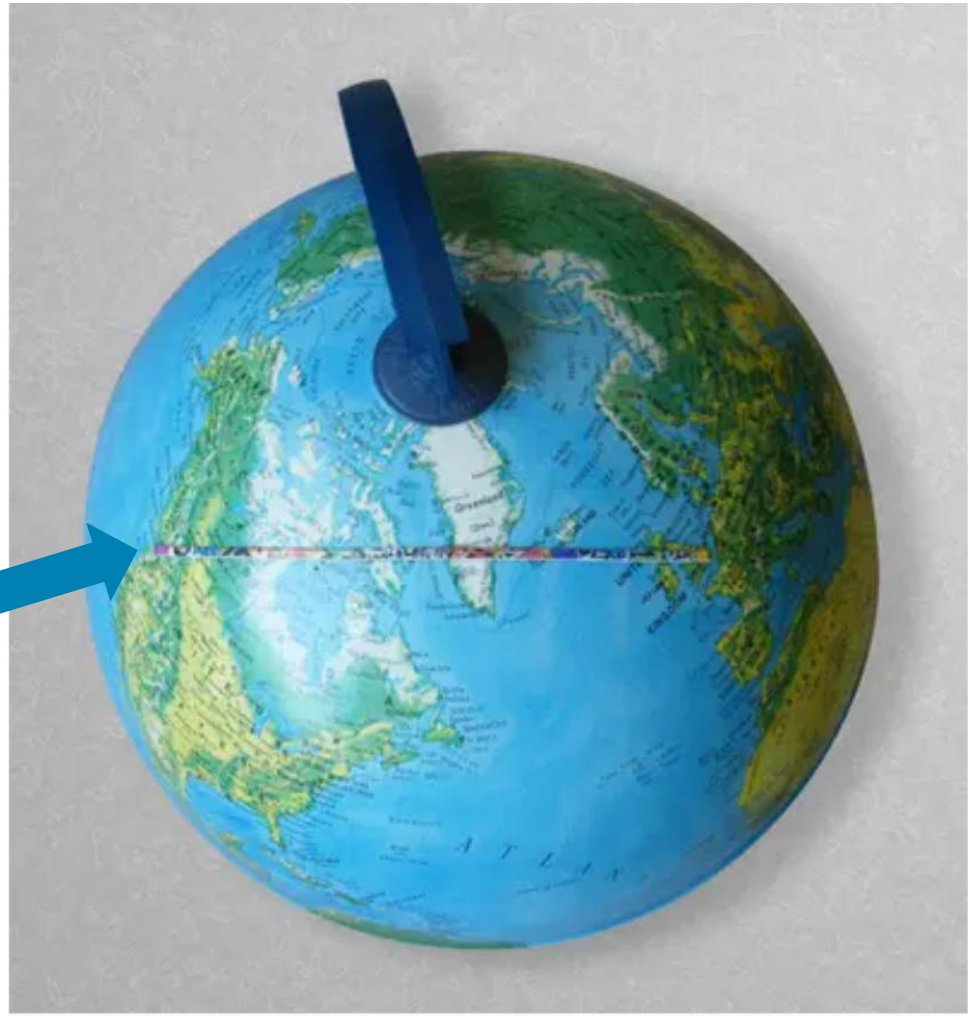
Why the Mercator projection was historically important



Sailors could represent their course on the Mercator map as a straight line (rhumb line), and then simply sail one constant compass bearing (here: 87°) along this line

Mercator distorts distances

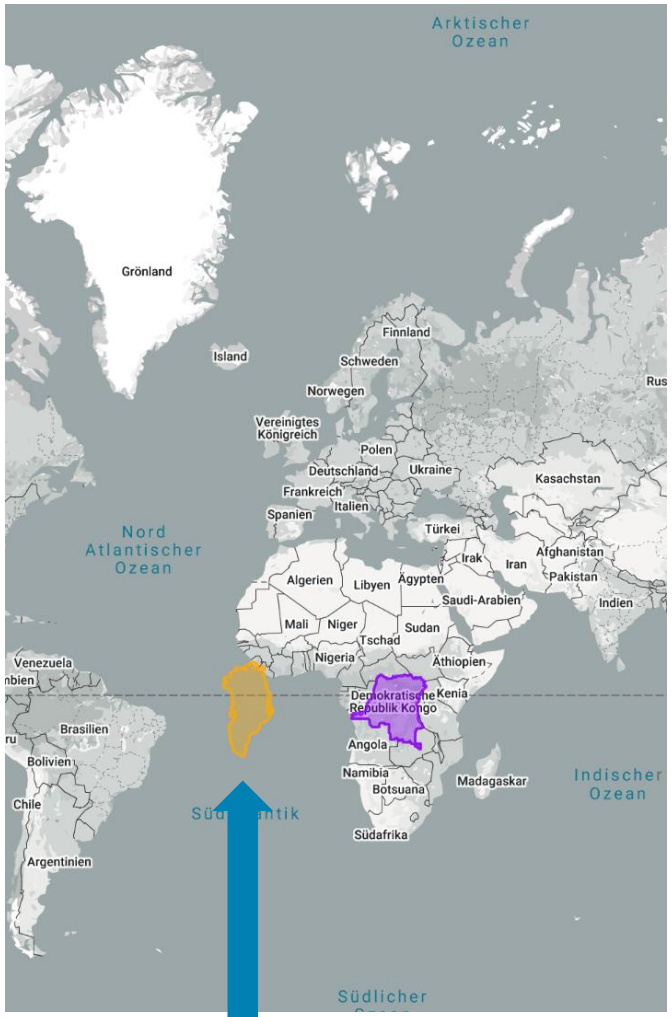
- The shortest path from London to Seattle (rhumb line) on the Mercator map does not represent the shortes route in real life (great circle route)
- Do travel the shortest route seaman would need to adjust the compass bearing constantly



A ‘great circle’ route between any two places is the shortest distance, i.e. it is ‘as the crow flies’.

Mercator distorts areas

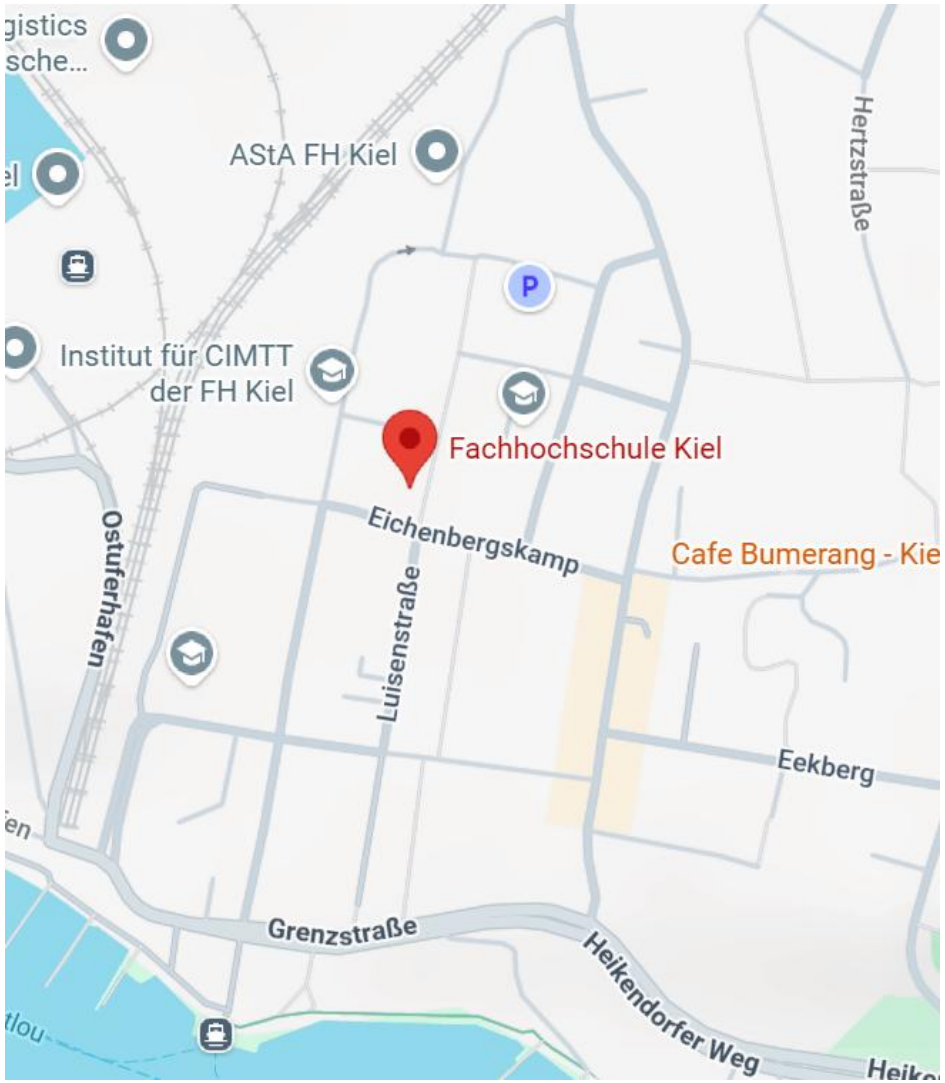
Mercator‘s Greenland



True size of Greenland

See: www.thetruesize.com

Web maps use Web Mercator Projection



- **Google Maps**, **Open Street Maps** and **Tableau** use a close variant of the classic Mercator projection: Web Mercator (EPSG:3857)
- Preserves local shapes and angles → street intersections look the same as in real life
- Ideal for computers: the Earth can be divided into square tiles that can be loaded, shifted, and zoomed very efficiently

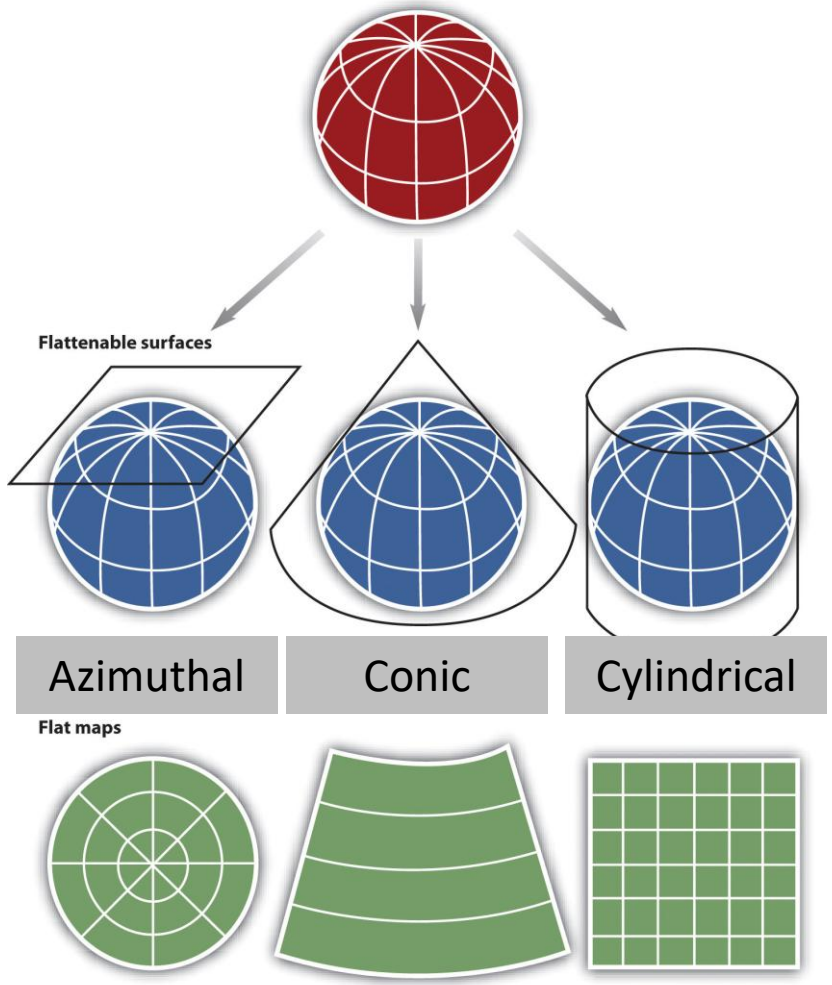
(see also: How zoomable maps work)

Distortions are inevitable

We want to represent the 3-dimensional earth on 2 dimensions

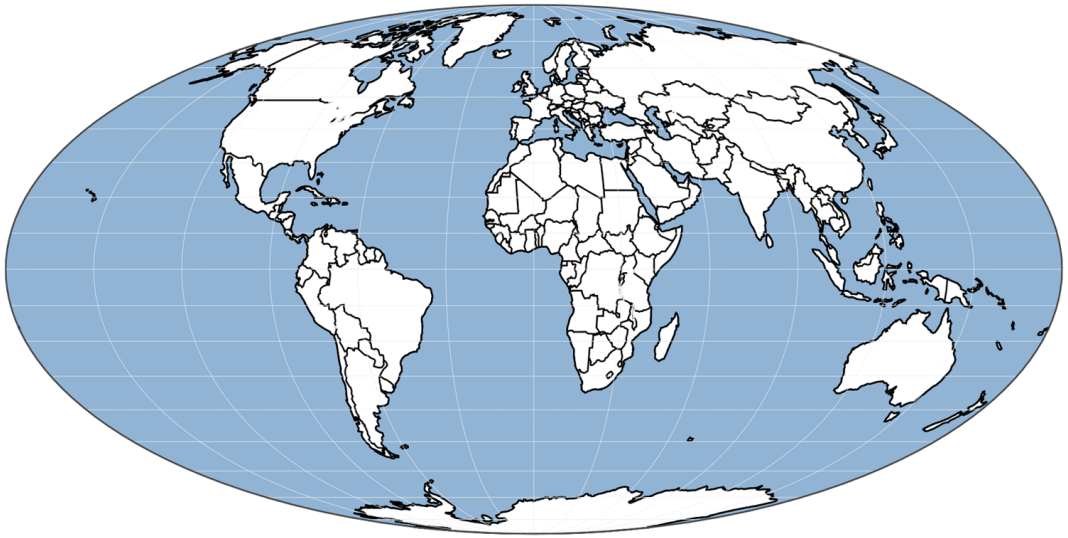
No matter how we try to do this, it will lead to distortions in at least one of:

- area
- distance
- local shapes

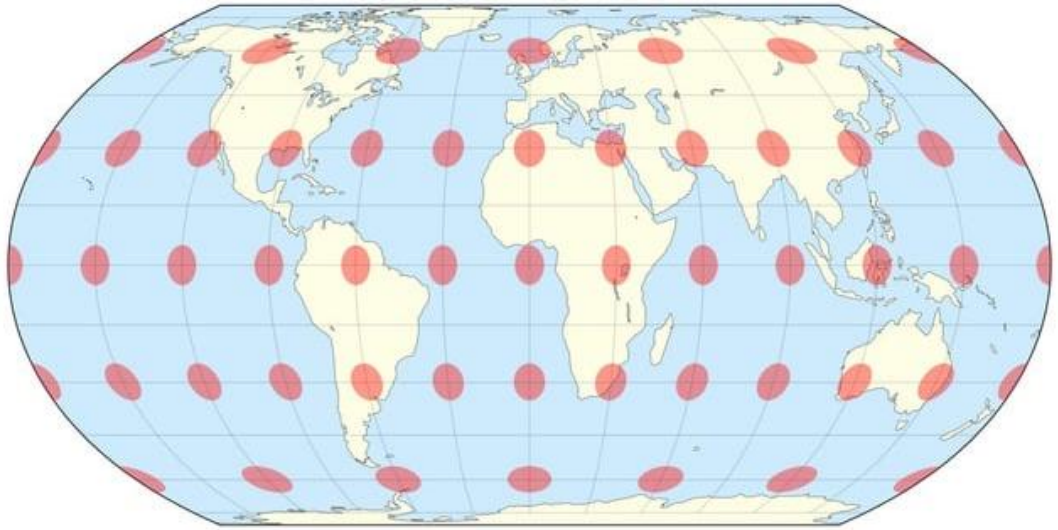


→ Video: Why all maps are false

Recommended projections for the entire world



Mollweide
EPSG:54009



Equal Earth
EPSG:8857

Projections should represent areas faithfully (→ equal area projections)

Recommended projection for Europe

LAEA Europe
EPSG:3035

LAEA Europe is used by Eurostat, and other official agencies

- Preserves areas perfectly
- Minimal distortions of shape (towards outer regions)

Recommended projection for Germany

EPSG:25832

Given these perceptual problems of map visualizations – what can we take away?

Take away messages

Map or no map?

- Use maps if the data shows interesting geographic patterns
- Don‘t use maps for absolute values (that correlate with population or size) → normalize your data
- Maps can show high-level patterns, but for details consider complementary visualizations

Projection

- There are tradeoffs, so we have to make compromises
- Choose projection where area proportions are represented faithfully
- For a map that zooms into a small region, local shapes should be correct and look familiar

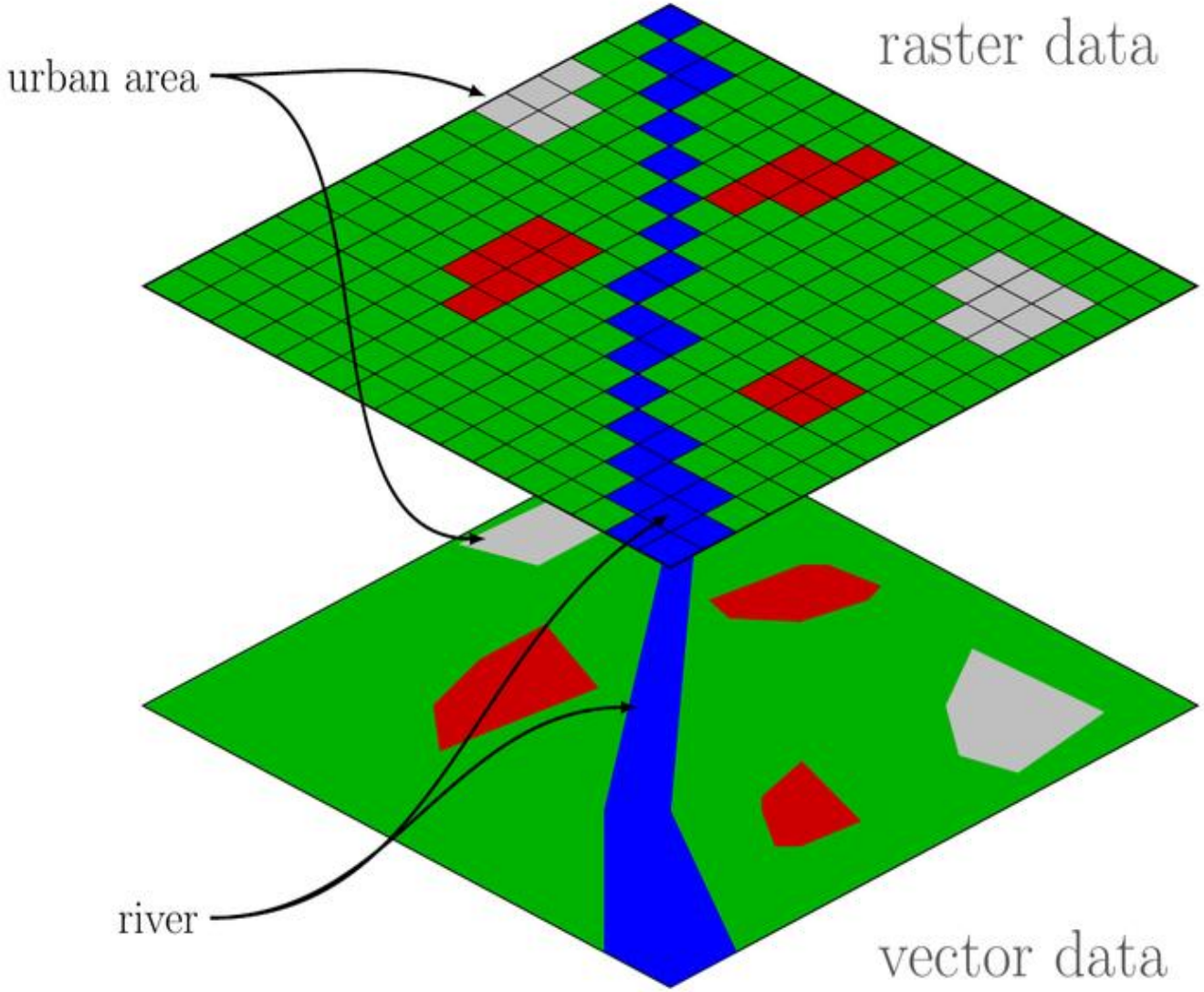
Take away messages

Colors

- Choose a color scale that makes differences visible
- Be aware of the perceptual biases due to color scaling → use complementary visualizations to give viewers also an accurate impression of data values
- **Useful tool**: https://colorbrewer2.org/

Practical map creation

Geographic data models

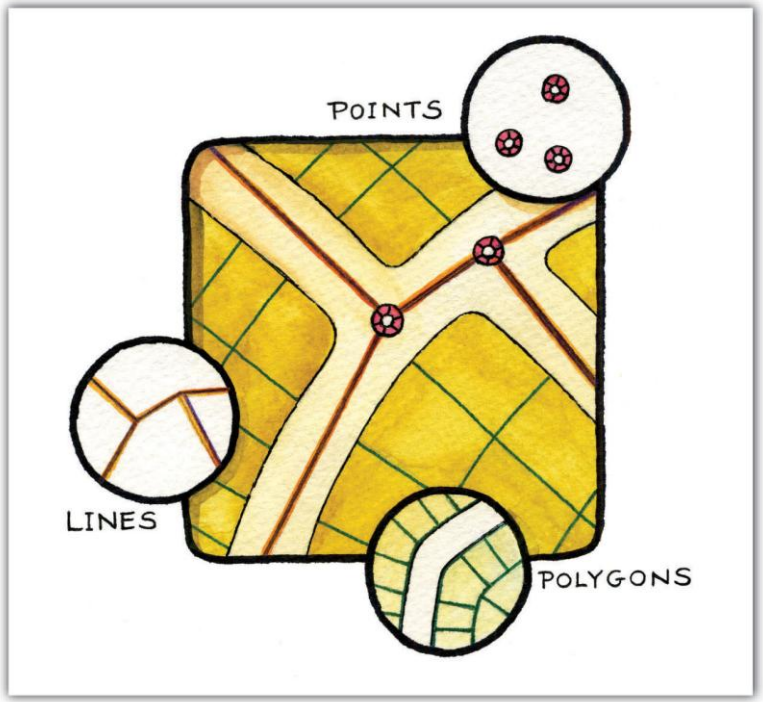


Our focus

Vector data

- Vector: discrete geometric entity described by their coordinates
- Typical use cases: thematic maps

Features	Geometry	Attributes
Bus stop	Point	
Amazonas	Line	Length, Temperature
Kiel	Polygon	Population
UK	Multipolygon	Area, GDP

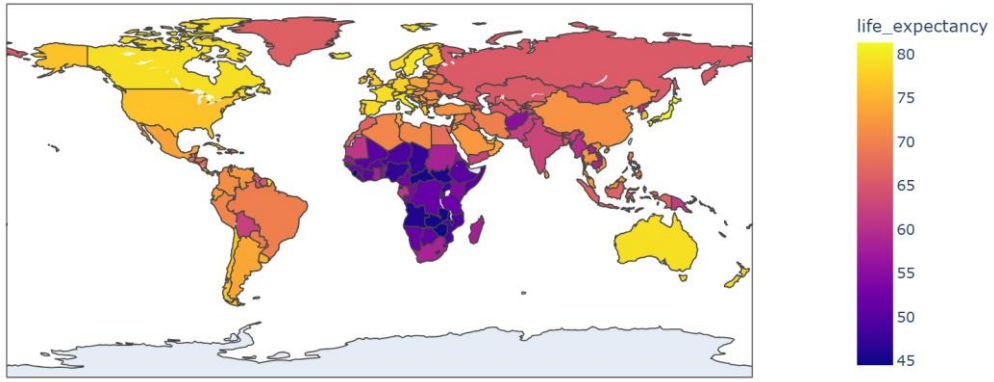


Simple choropleth maps with Plotly

iso3	country	life_expectancy
FRA	France	79.056098
CHN	China	71.881000
DEU	Germany	77.926829
USA	United States	76.636585
ITA	Italy	79.778049

- **Plotly** has built-in vector data for countries
- Your data must have ISO3 country codes

```
px.choropleth(df, locations="iso3", color="life_expectancy")
```



Other situations?

1. **For all other use cases, we need to find and integrate suitable geographic vector data ourselves!**
2. **Plotly is an option – but GeoPandas + Matlotlib is often more powerful**

Typical workflow in other cases

1. Find suitable vector data for your entities

DataFrame (Pandas)

iso3	country	life_expectancy
FRA	France	79.056098
CHN	China	71.881000
DEU	Germany	77.926829
USA	United States	76.636585
ITA	Italy	79.778049

2. Represent as GeoDataFrame (GeoPandas)

iso3	geometry
CHN	MULTIPOLYGON (((78.91769 33.38626, 78.91595 33...
FRA	MULTIPOLYGON (((-54.11153 2.11427, -54.13491 2...
DEU	MULTIPOLYGON (((13.81572 48.76643, 13.78586 48...
ITA	MULTIPOLYGON (((7.02208 45.92526, 7.06694 45.8...
USA	MULTIPOLYGON (((-122.75302 48.99251, -122.6532...

3. Merge with your DataFrame

4. Visualize

Vector data for world

- Naturalearthdata: cultural and physical vector data of the entire world

10 50 110	**Coastline** – ocean coastline, including major islands. Coastline is matched to land and water polygons.
10 50 110	**Land** – Land polygons including major islands
10 50 110	**Ocean** – Ocean polygon split into contiguous pieces.
10	**Minor Islands** – additional small ocean islands ranked to two levels of relative importance.
10	**Reefs** – major coral reefs from WDB2.
10 50 110	**Physical region features** – polygon and point labels of major physical features.
10 50 110	**Rivers and Lake Centerlines** – ranked by relative importance. Includes name and line width attributes. Don't want minor lakes? Turn on their centerlines to avoid unseemly data gaps.

Vector data for Europe

Eurostat Vector Data

Level	Name	Example
NUTS 0	Country	Germany
NUTS 1	State	Schleswig-Holstein
NUTS 2	Region	-
NUTS 3	District	Kiel

Datasets available for download

NUTS year *

NUTS 2024

File format *

SHP

Geometry type *

Polygons (RG)

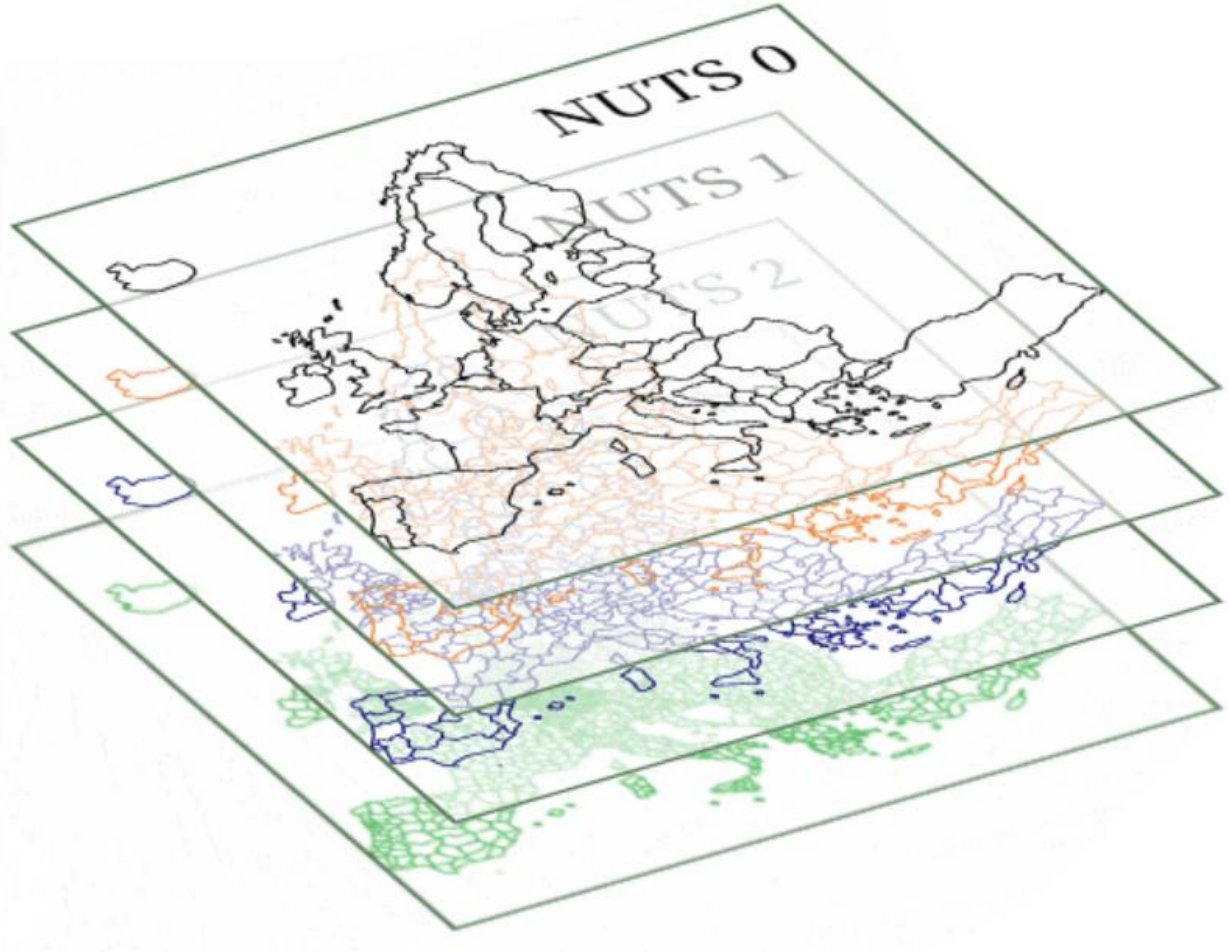
Scale *

20M

Coordinate reference system *

EPSG: 3035

Download



Vector data for Schleswig-Holstein

- Schleswig-Holstein vector data
- Shapefiles on various topics at state, region, or city level

Vectors from OpenStreetMap

- **Tailored download via Overpass API**
  - Python Package: osmnx
  - https://wiki.openstreetmap.org/wiki/Overpass_API
  - https://overpass-turbo.eu/
- **Bulk download**
  - Python package: pyrosm
  - Continents, countries, regions, cities: https://download.geofabrik.de/

Vector data file formats

Shapefile:

Collection of multiple files

- .shp: shape
- .shx: shape index for fast search
- .dbf: attributes
- .prj: projection
- …

Geojson and Topojson

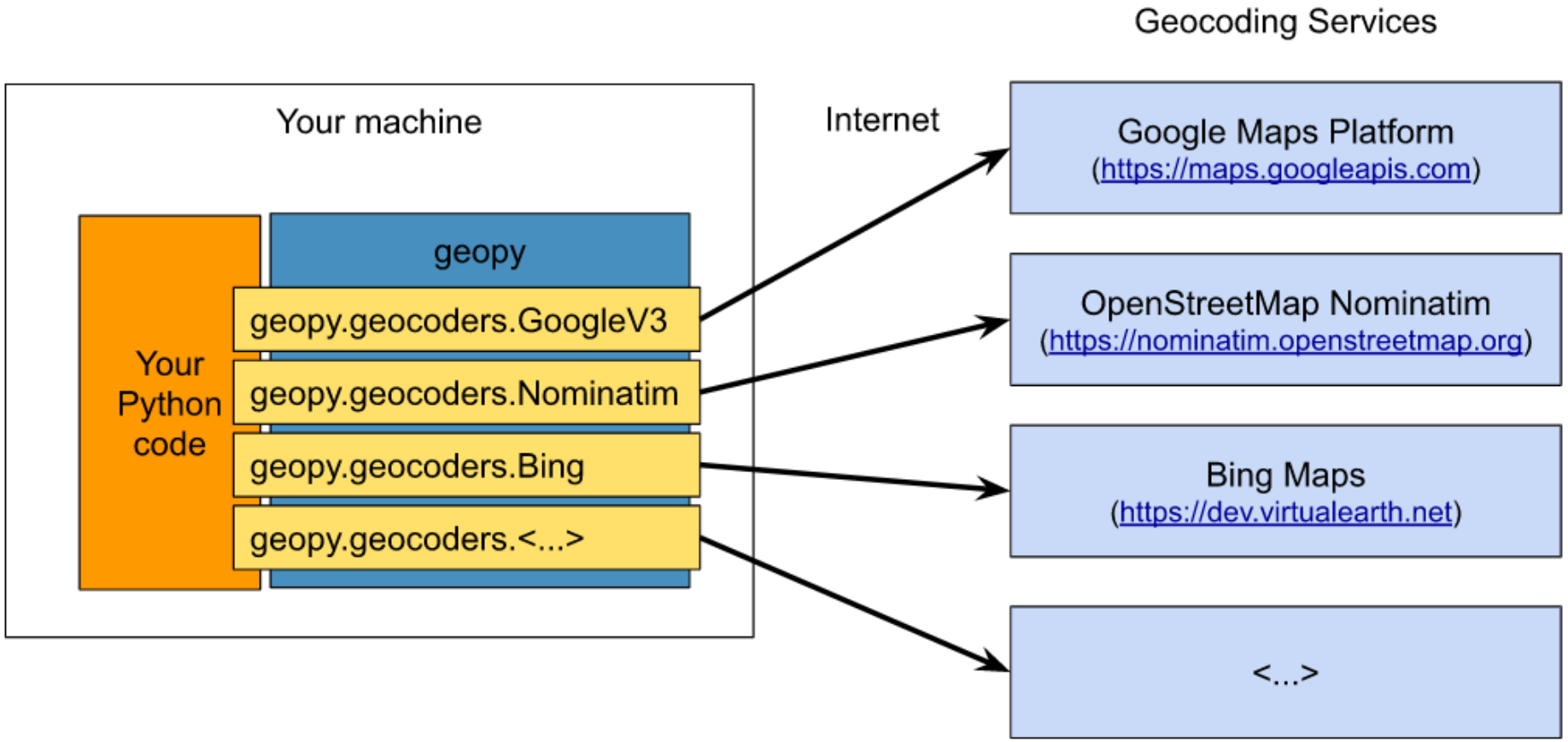
```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [125.6, 10.1]
  },
  "properties": {
    "name": "Dinagat Islands"
  }
}
```

Spatial Operations

Data Source: Geocoding

Geocoding: Place name or address → coordinates

Reverse geocoding: coordinates: → address or place name

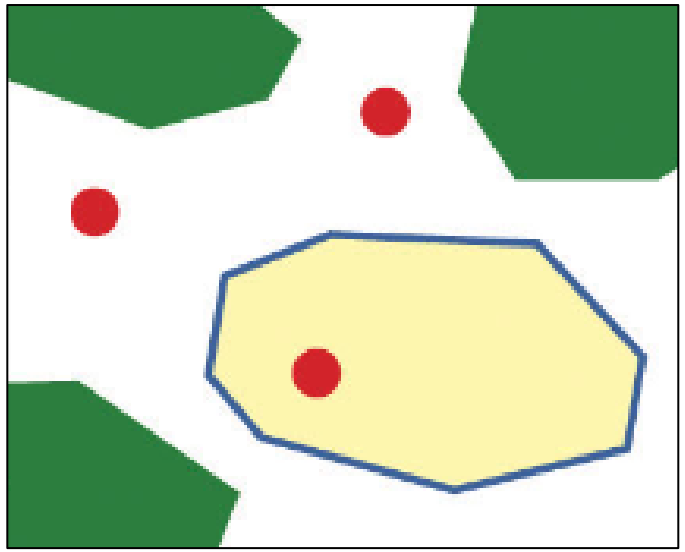


Source: https://geopy.readthedocs.io/en/stable/#module-geopy.geocoders

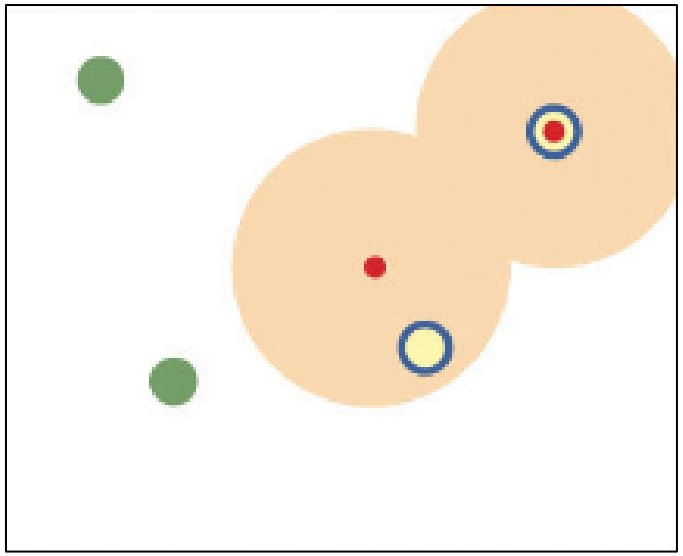
More spatial operations

- Calculate:
  - Distance
  - Length
  - Area
  - Shortes route
  - Location of Centroid
  - …

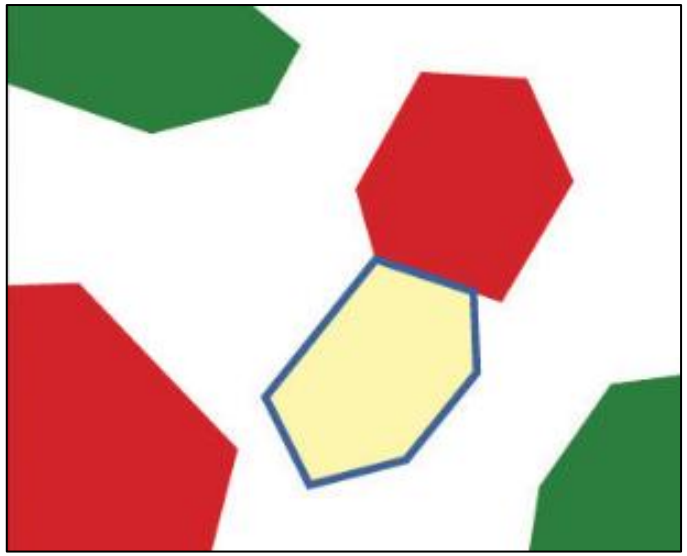
Topological relations



Contain
Which of the given bus stops are located within Kiel?



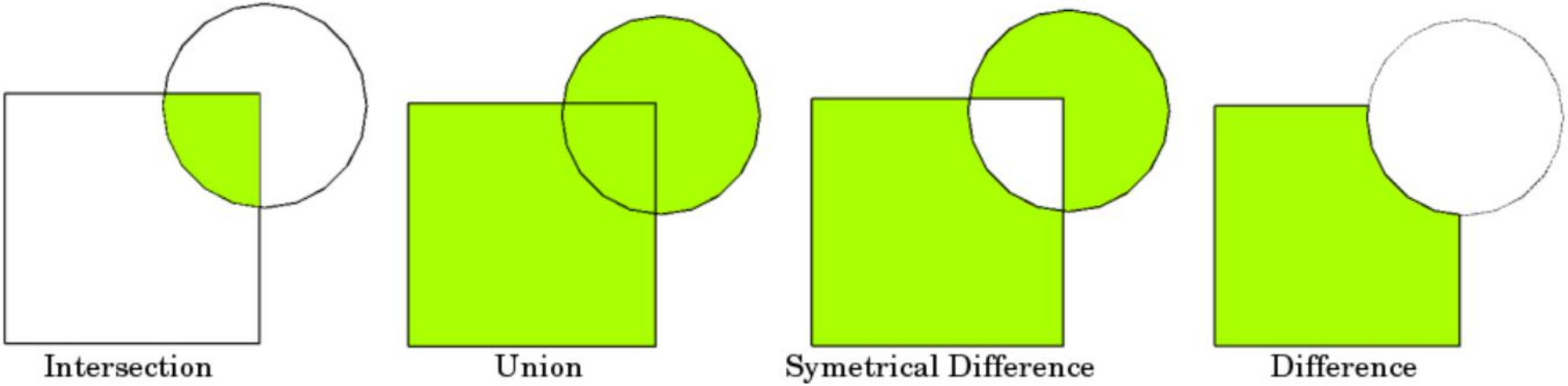
Within distance of
Which schools are within 10km radius to my home?



Touch the boundary of
Which country shares a border with Germany?

https://saylordotorg.github.io/text_essentials-of-geographic-information-systems/s10-data-characteristics-and-visua.html

Set / Overlay Operations



Source: https://geopandas.org/en/stable/docs/user_guide/set_operations.html