Data acquisition and visualisation

Making maps online

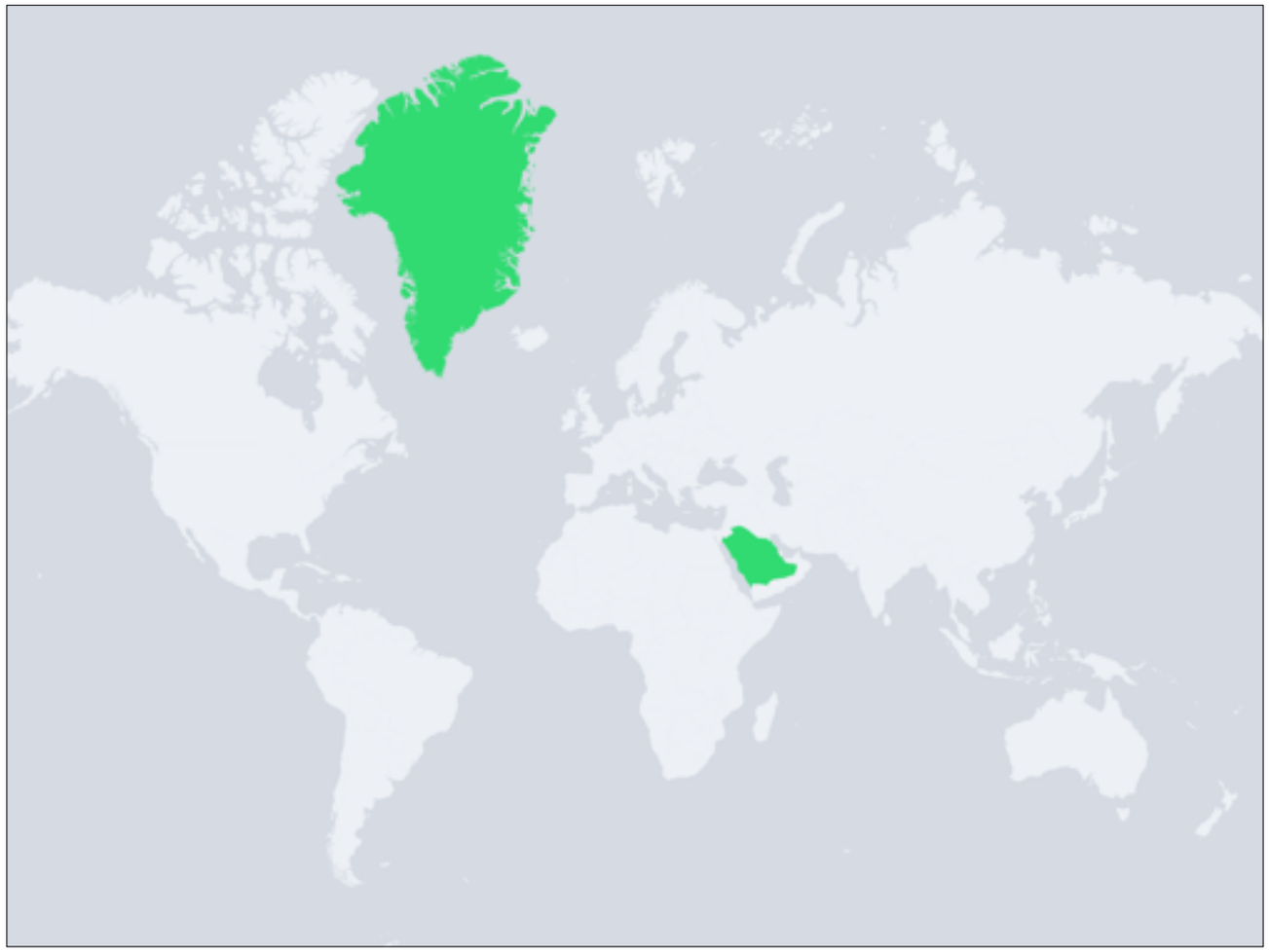
7 December 2015

How to make a map
in the 1960s

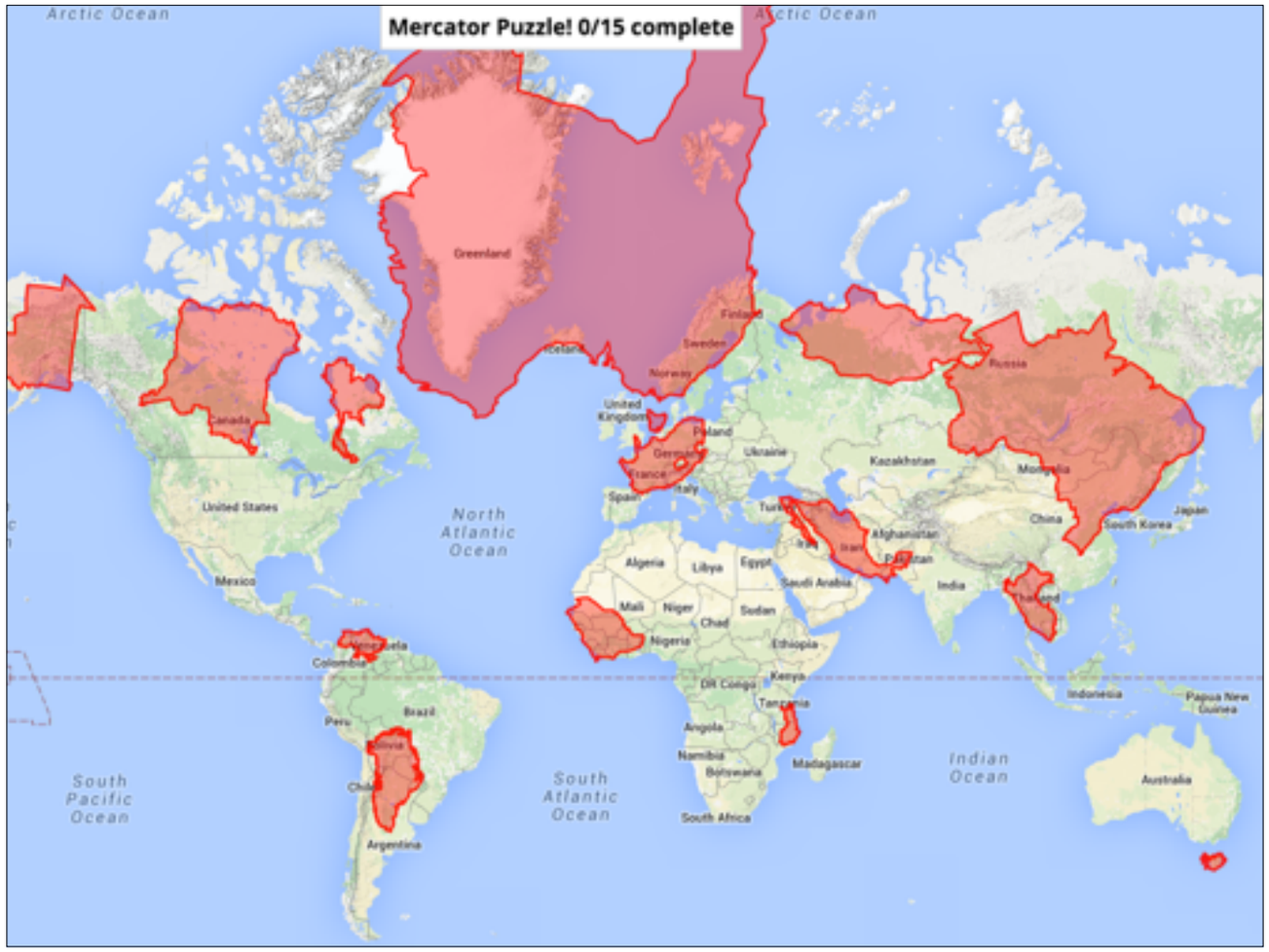
Fifty years ago map-making looked like this. Computers have altered cartography to the point it's hard to see any connection between now and then.

Computers — and the Internet — have also made map-making easier. But just because something's easy to do doesn't mean it's easy to do well. Today we're going to learn how to make a good digital map.

http://www.nls.uk/exhibitions/bartholomew/maps-engraver



Which is bigger: Greenland or Saudi Arabia? They’re almost exactly the same size.



https://gmaps-samples.googlecode.com/svn/trunk/poly/puzzledrag.html

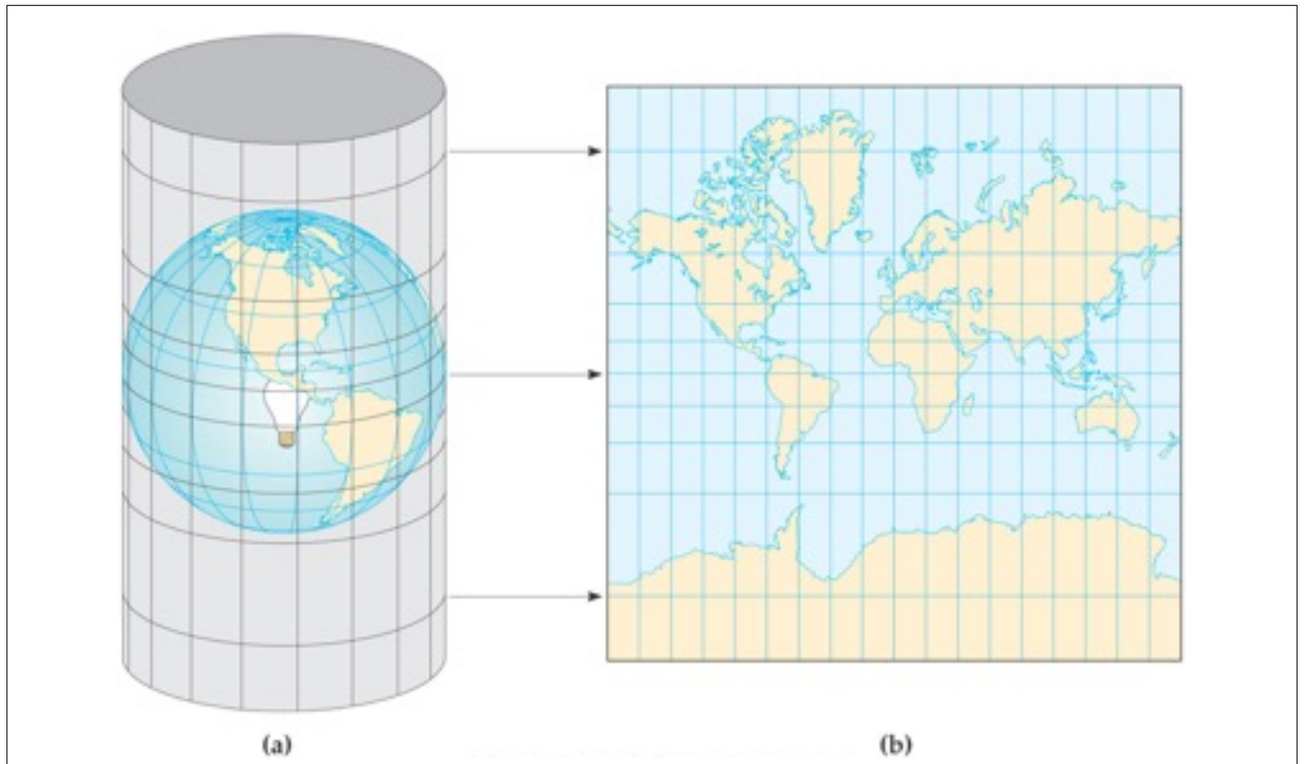
A globe can show size, shape, distance, and direction accurately, but it won't fit in your pocket. It also can’t show detail.

Map projections

The world is a sphere but a map is flat. Try to flatten orange peel: it splits and the shape changes. Same thing happens when flatten a planet onto a piece of paper.

By creating a map we're turning a three-dimensional object, Earth, into a two-dimensional drawing, and you can't do that without distortion.

If we flatten the globe we can carry it around but we need to *project* the globe onto a flat surface first: this is called a map projection.



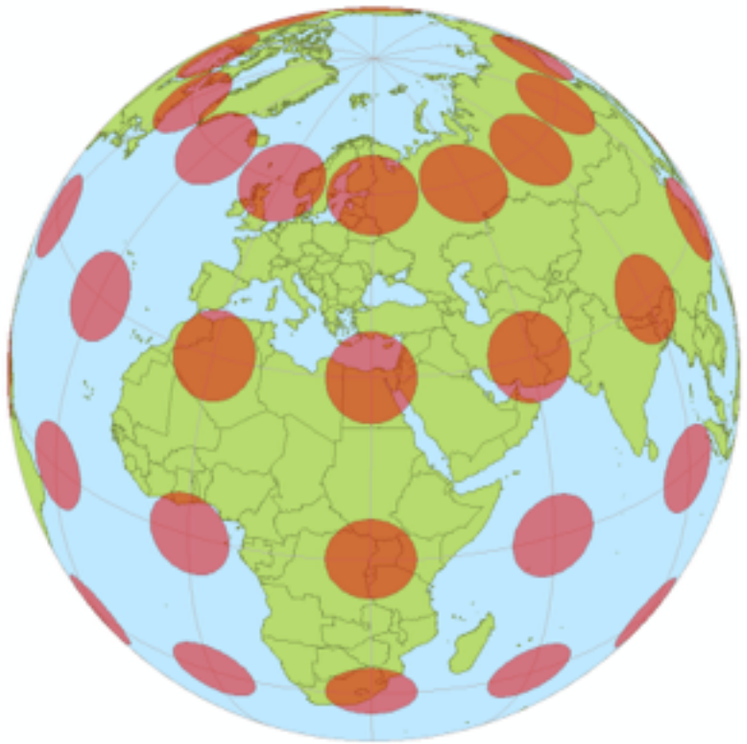
Mercator projection

There is no one way to project Earth onto a flat surface, each with different distortions

Area, direction, and distance will all be distorted when you create a two-dimensional map

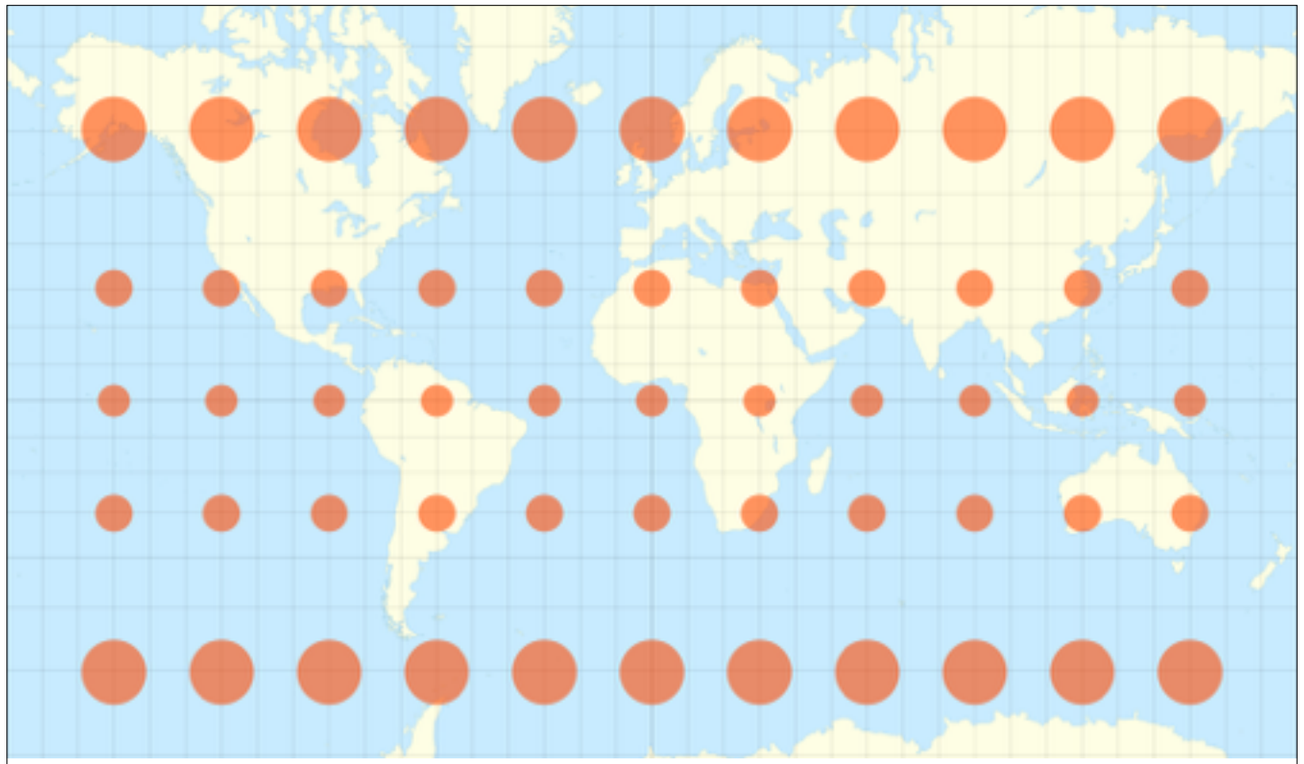
Gerardus Mercator’s 1569 map where lines of longitude and latitude crossed at right angles. Great for navigation at sea, and for local navigation.

But on a global scale it causes very bad distortions. Don’t use Mercator for global maps.



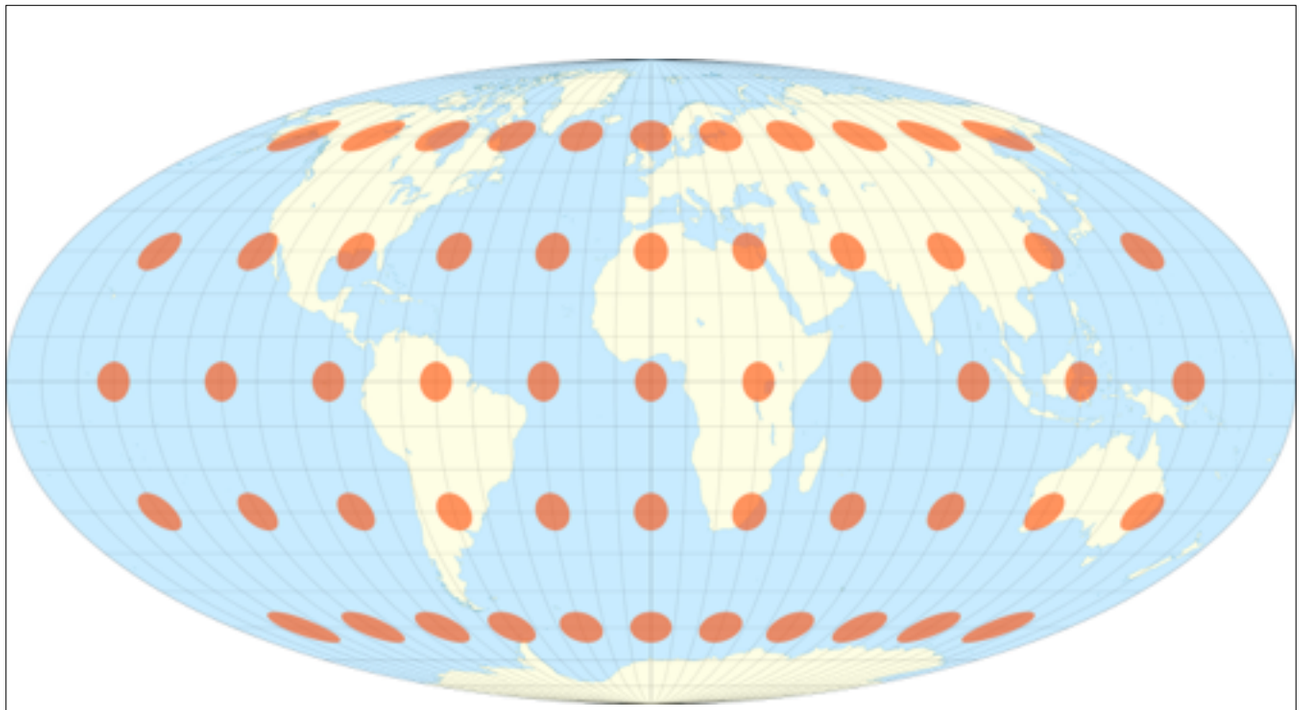
Tissot indicatrix
on a globe

On a sphere, the circles are all the same size.

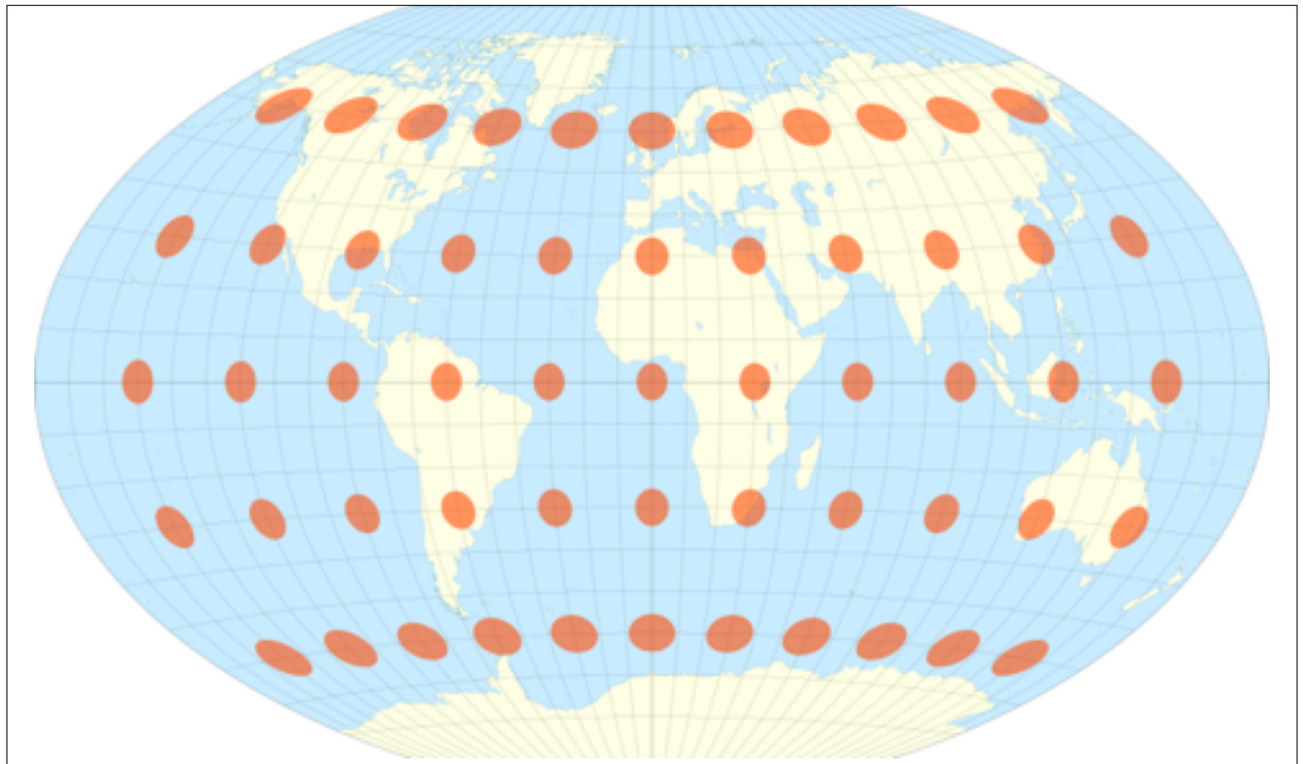


Tissot indicatrix for the
Mercator projection

Tissot's indicatrix displays local distortions on a map projection. Each circle covers an area of the same size but the projection causes those circles nearer the poles to appear to be larger.



Tissot indicatrix for the
Mollweide projection



Tissot indicatrix for the
Winkel Tripel projection

How to pick a projection

- What's important? Area, distance, direction?
- National map projections
  - Iceland's ISN2004
  - UK's National Grid
- Mercator for local, large scale, maps
  - But no comparisons

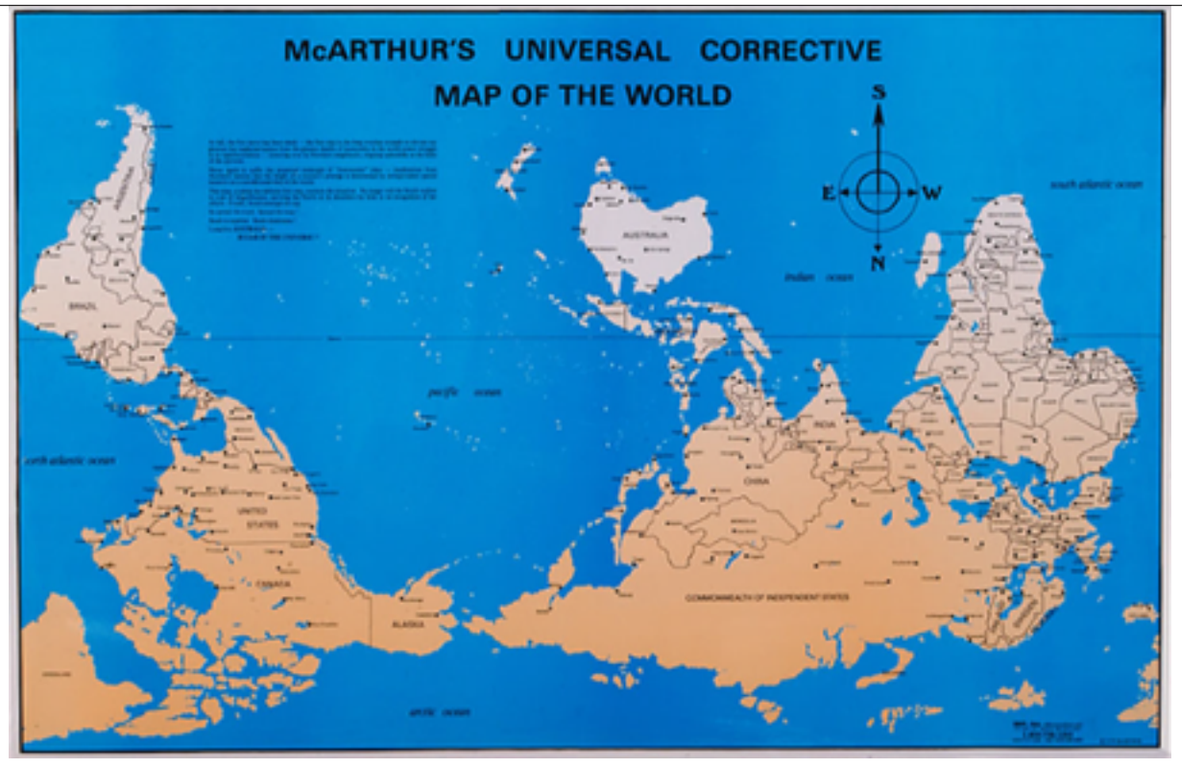
All projections result in distortions. Choose an appropriate projection so that the inherent distortions don't ruin the purpose of your map. A projection is always an algorithm of compromise.

Don’t use Mercator for comparisons

Oslo and Singapore are separated by 60 degrees of latitude. If we compare them using a Mercator projection they look roughly the same size. But if we switch to an equal-area projection you can see the two cities are very different in size.

Don’t use Mercator for comparisons

Oslo and Singapore are separated by 60 degrees of latitude. If we compare them using a Mercator projection they look roughly the same size. But if we switch to an equal-area projection you can see the two cities are very different in size.



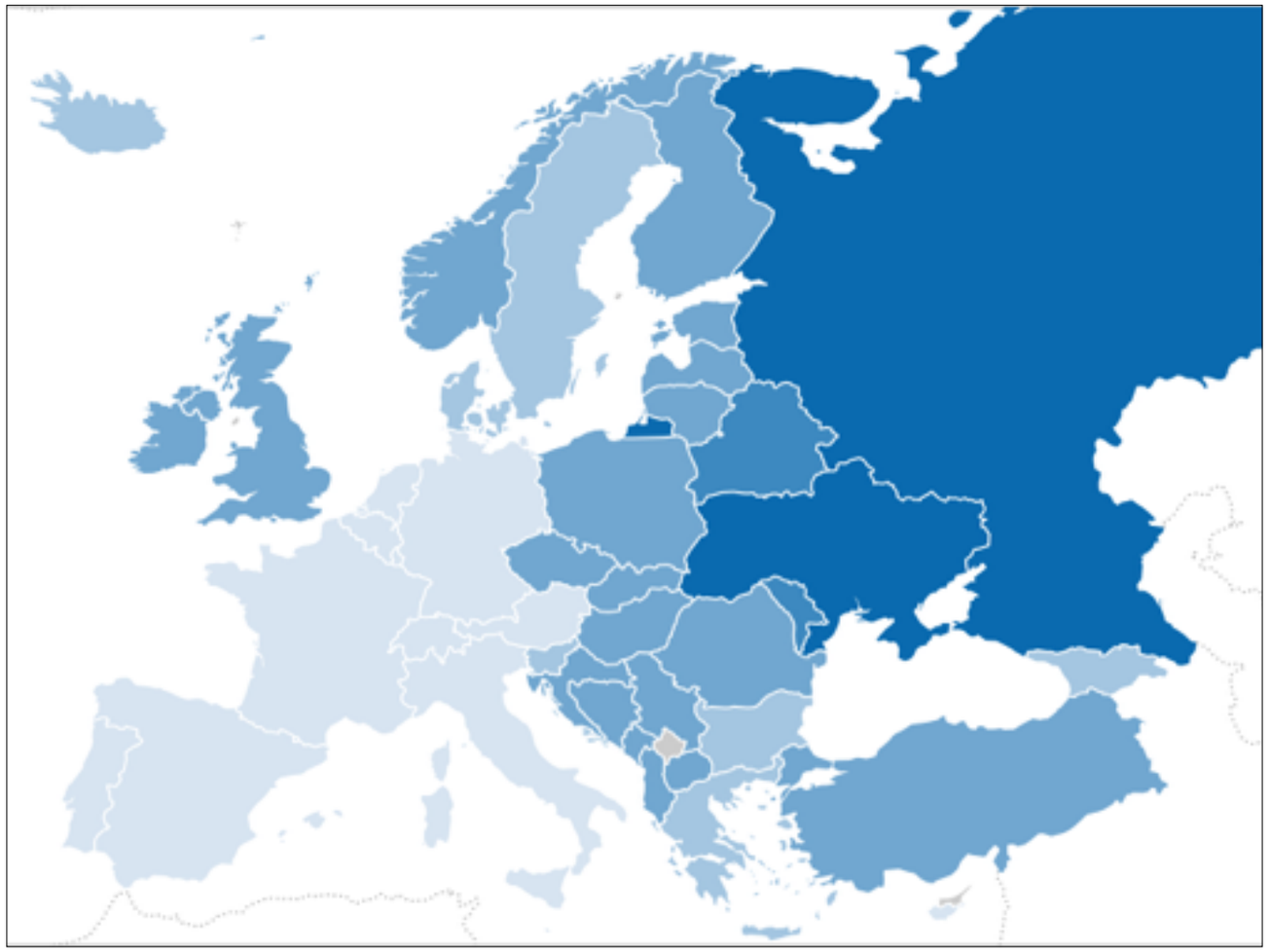
South is not up

Sometimes, people's perceptions are so rigid that they will think a map projected to look different is wrong. People expect north to be "up", for example.

Is this wrong?

Stuart McArthur drew his first “south is up” map when he was 12 years old. His geography teacher told him he’d fail unless he re-did his map the “correct” way up.

http://mapdesign.icaci.org/2014/02/mapcarte-38365-mcarthurs-universal-corrective-map-of-the-world-stuart-mcarthur-1979/



Með bestu drykkjumennina?

Í DAG

Pawel Bartoszek
stærðfræðingur

Grein Silju Daggar Gunnarsdóttur, þingkonu Framsóknarflokksins, „Áfengi er engin venjuleg söluvara“, var alls ekki slæm. Silja Dögg gerir til dæmis ekki tilraun til að halda því fram, eins og margir, að verðið muni stórhækka og dreifingin versna við afnám ríkiseinokunar. Ef hvort tveggja væri satt ættu áhugamenn um minni neyslu auðvitað að fagna þessum tillögum því hærra verð og verri dreifing ættu að minnka neyslu.

Nei, rök Silju Daggar eru skárri. „Frjáls sala mun þýða betri dreifingu. Betri dreifing mun þýða meiri neyslu. Meiri neysla mun þýða meira ofbeldi. Heildaráhrifin eru því skaðleg.“ Silja Dögg kemst svo að þeirri niðurstöðu að við eigum að vera öðrum þjóðum fyrirmynd og halda í einkaleyfi ríkisins á áfengissölu.

DRYKKJUMYNSTUR Í EVRÓPU

Heimild: Alþjóða-heilbrigðismálastofnunin

neyslumynstri fólks annað en löggjöf. En í engu þessara ríkja þar sem WHO telur að áfengisnotkun sé áhættuminnst sér ríkið um smásölu áfengis. Í mörgum þessara ríkja búa miklar vínþjóðir. Öll þessi ríki hafa

> Því er stundum réttilega haldið fram að heildaráfengisneysla

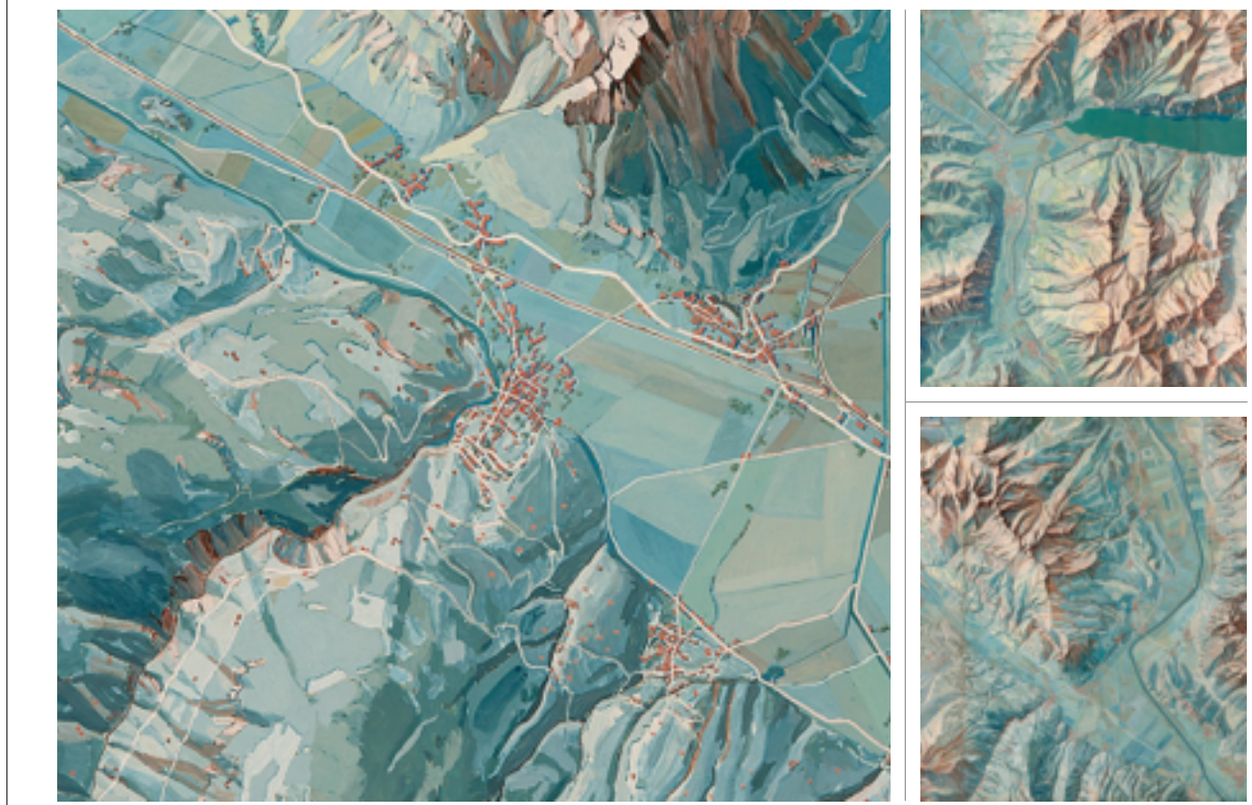
The two types of map

- Reference map
- Thematic map
- (Dynamic map)

Reference maps show you *places*, thematic maps show you *data*.

Some people say there's a third type, dynamic maps. These are just an interactive reference or thematic map.

Reference maps

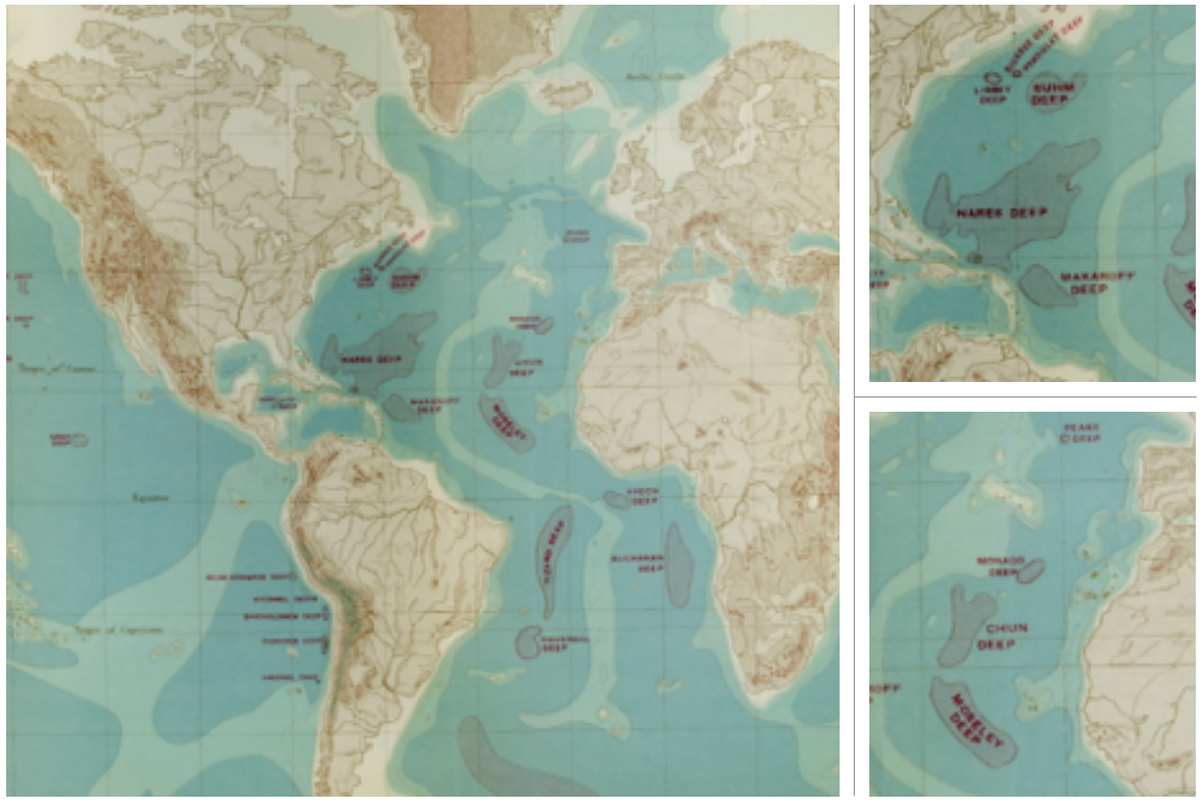


A reference map provides information on a location. This information might be rivers, buildings, topography, even satellite images.

Most paper maps you encounter are reference maps. They help you see where things are.

Eduard Imhof, 1938. Landscape painter as a mapmaker.

http://mapdesign.icaci.org/2014/02/mapcarte-58365-karte-der-gegend-um-den-walensee-by-eduard-imhof-1938/



Bathymetrical chart of the oceans, John Murray, 1899.

First map to focus on the ocean depths rather than the surface. Accurate survey of the ocean floor was paramount to the correct siting of telecommunications cables.

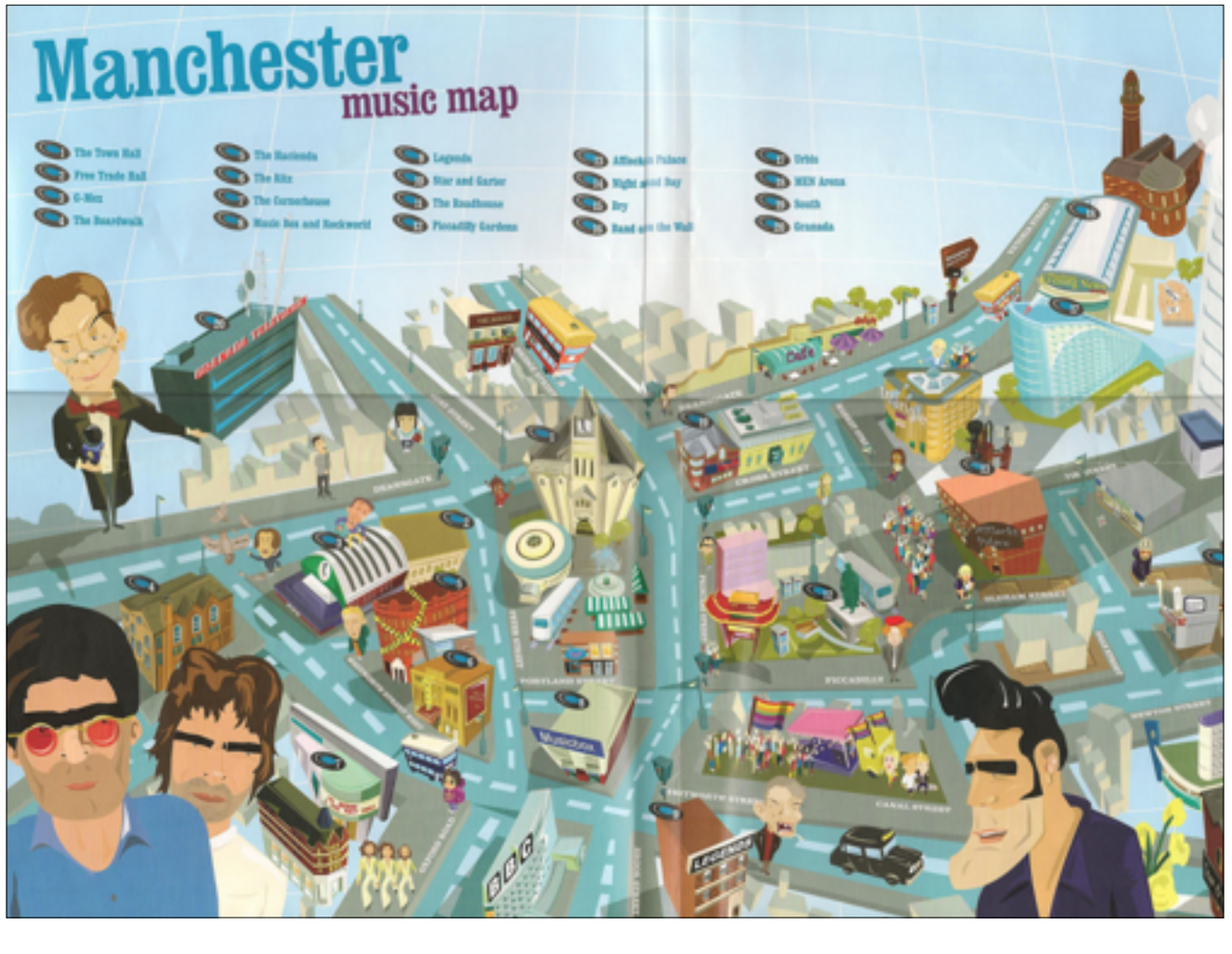
Accurate but simple.

http://mapdesign.icaci.org/2014/11/mapcarte-323365-bathymetrical-chart-of-the-oceans-by-john-murray-1899/

Bangkok, Nancy Chandler, 1974–present.

Hand-drawn version of TripAdvisor etc. You have to make a choice as to what to include, and what to ignore.

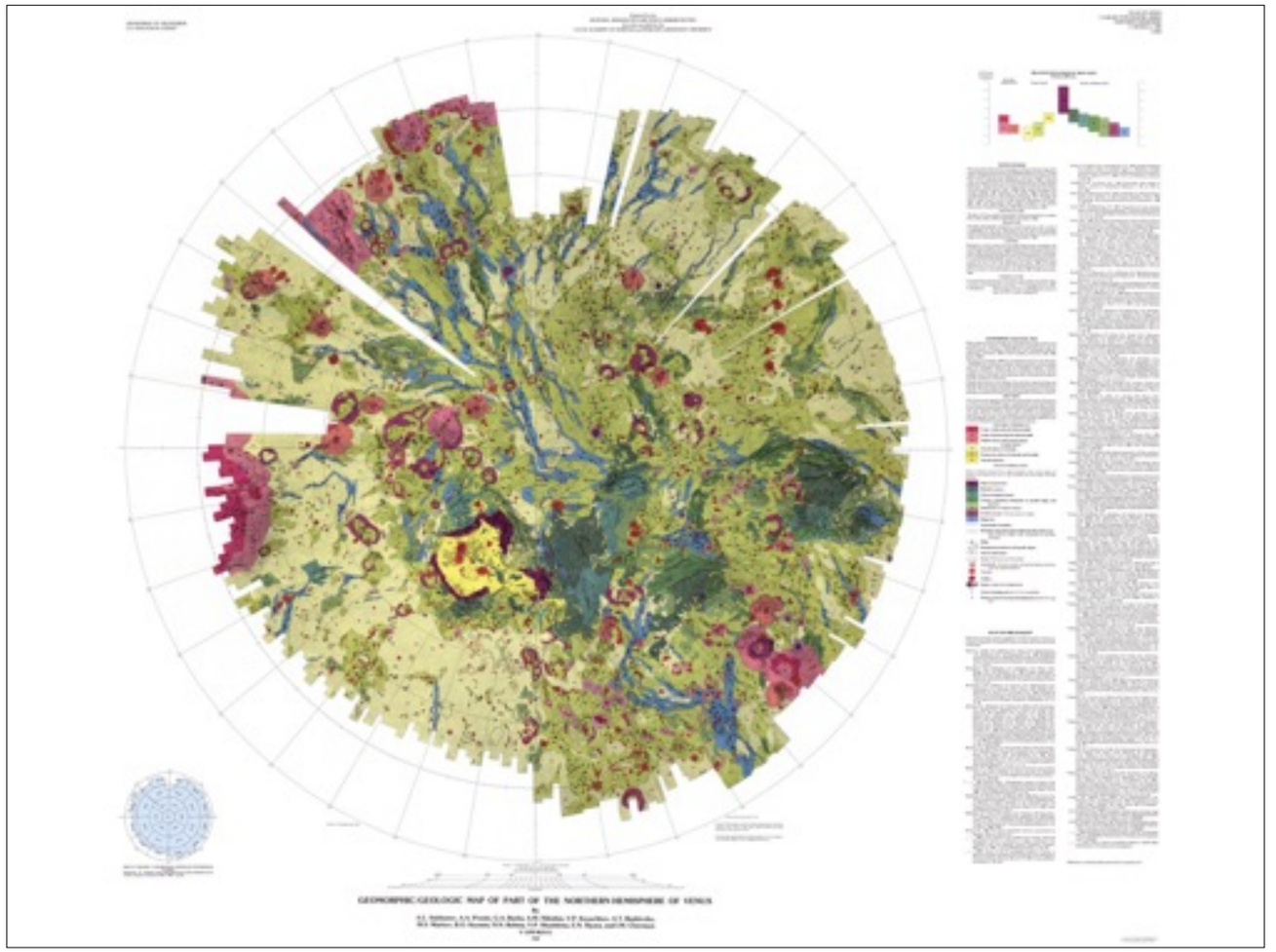
http://mapdesign.icaci.org/2014/08/mapcarte-224365-bangkok-by-nancy-chandler-1974-present/



Manchester Music Map, CreativeLynx, circa 2002.

Reference maps don’t always have to be useful for navigation. Focus on significance.

http://mapdesign.icaci.org/2014/07/mapcarte-205365-manchester-music-map-by-creativelynx-c-2002/

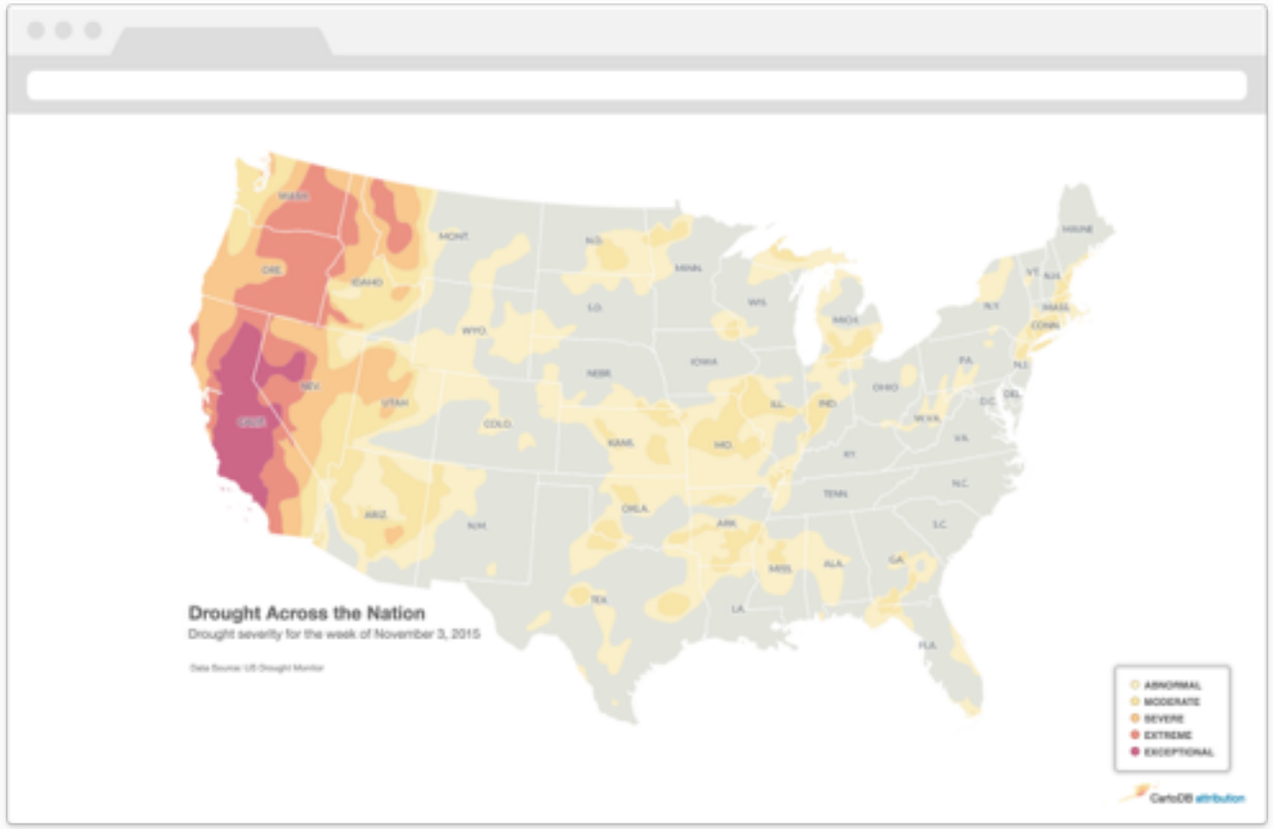


Geomorphic/geologic map of part of the northern hemisphere of Venus, USGS, 1989.

Reference maps don’t need to be of Earth. The choice of colour means it looks unearthly.

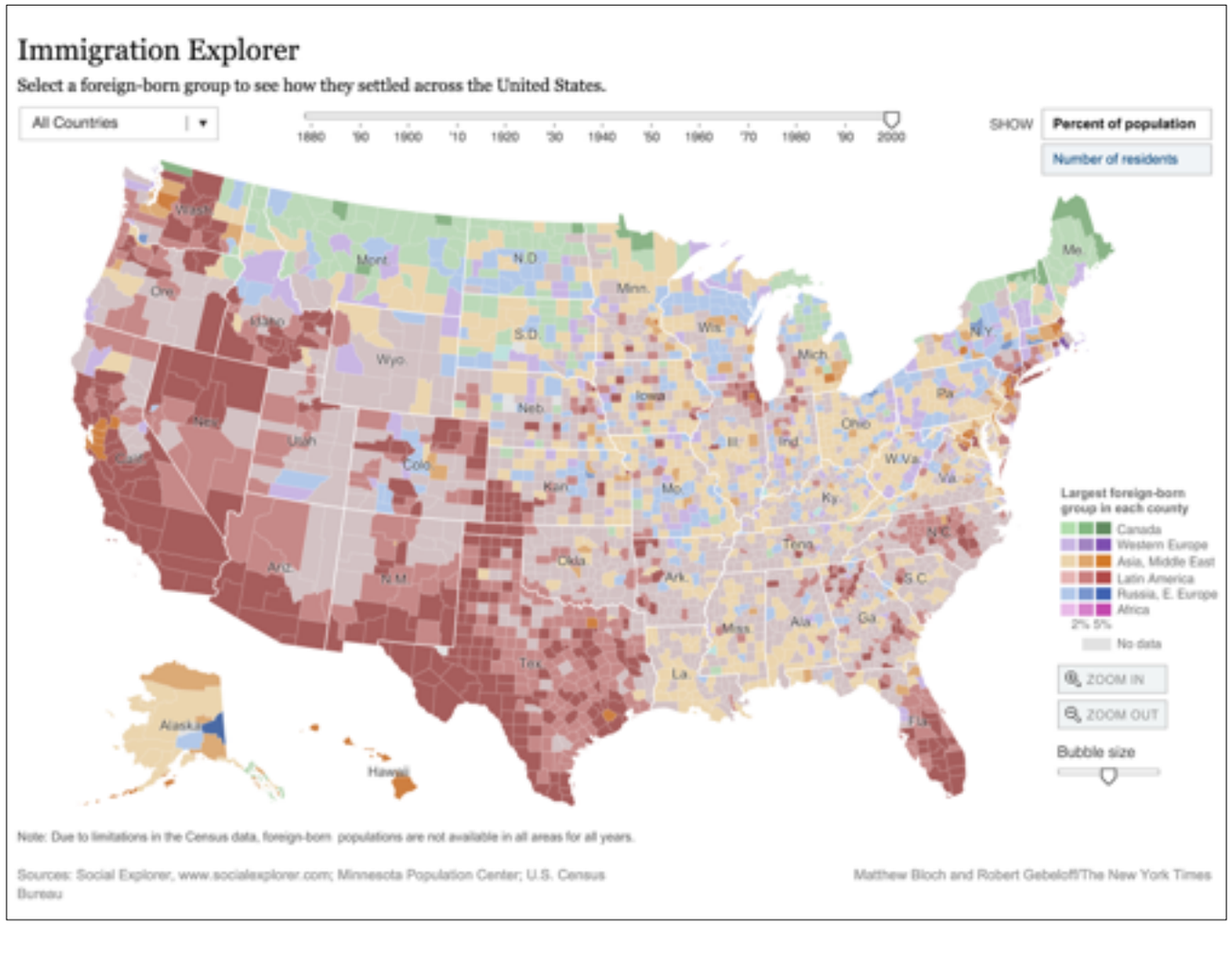
http://www.lpi.usra.edu/resources/venus_maps/2059/

Thematic maps



Many (the majority?) of online maps highlight a specific topic or theme of information. Typically, thematic maps have two components: one or more basemaps, and a thematic overlay.

http://blog.cartodb.com/thematic-map-drought-readings/

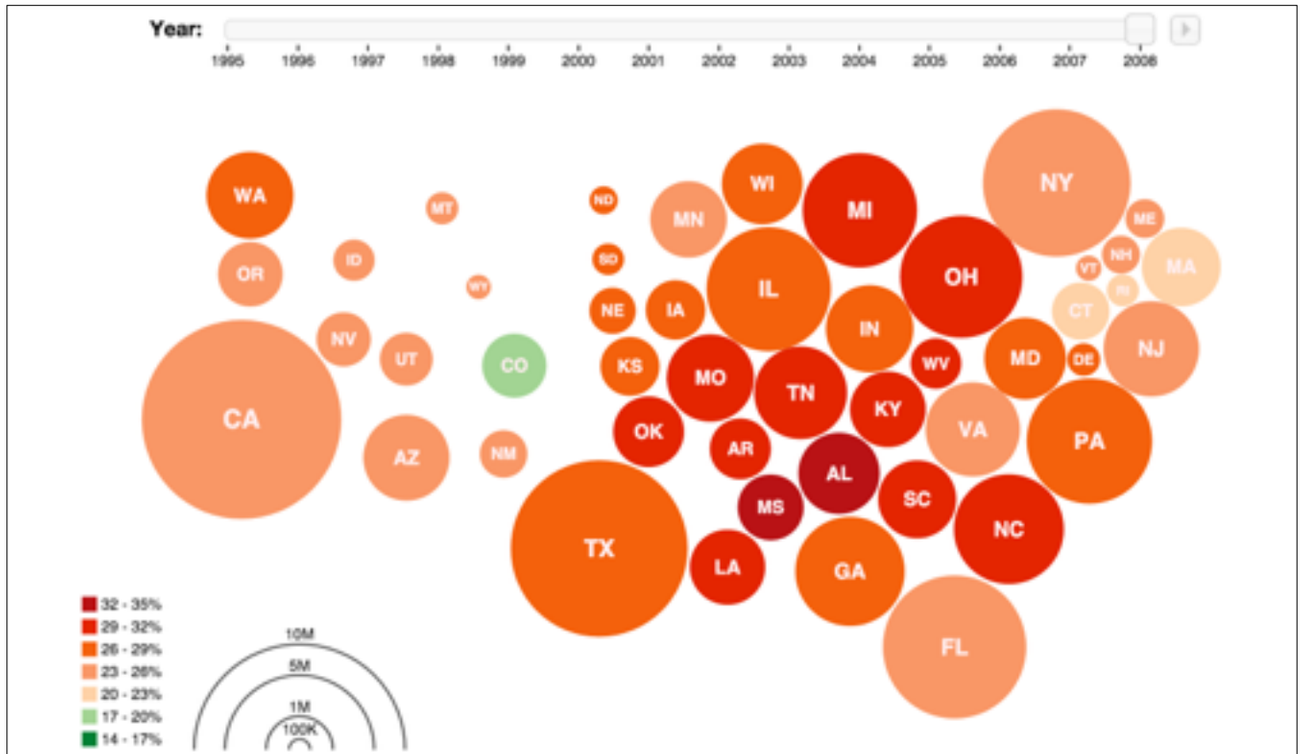


Immigration Explorer, New York Times, 2009.

A thematic map shows you attributes or statistics about a place. It's interested in a particular topic and it's distributed across the world: GDP or life expectancy, for example.

It’s important not to be too general. County level data makes this data. State level would be too broad.

http://mapdesign.icaci.org/2014/01/mapcarte-11365-immigration-explorer-by-new-york-times-2009/



Dorling cartogram

Here proportional symbols (where data is described using area) are used, instead of colour, to represent large numbers.

A Dorling cartogram is a proportional symbol map that doesn't allow symbols to overlap.

https://mbostock.github.io/protovis/ex/cartogram.html

How much does the big square represent?

Beware: humans are bad at estimating area.

If the small square represents 1,000 people, how much does the large square represent? 36,000.

You might want to classify your data in a few discrete classes to make it easier to compare and understand.

Combining (layering) maps

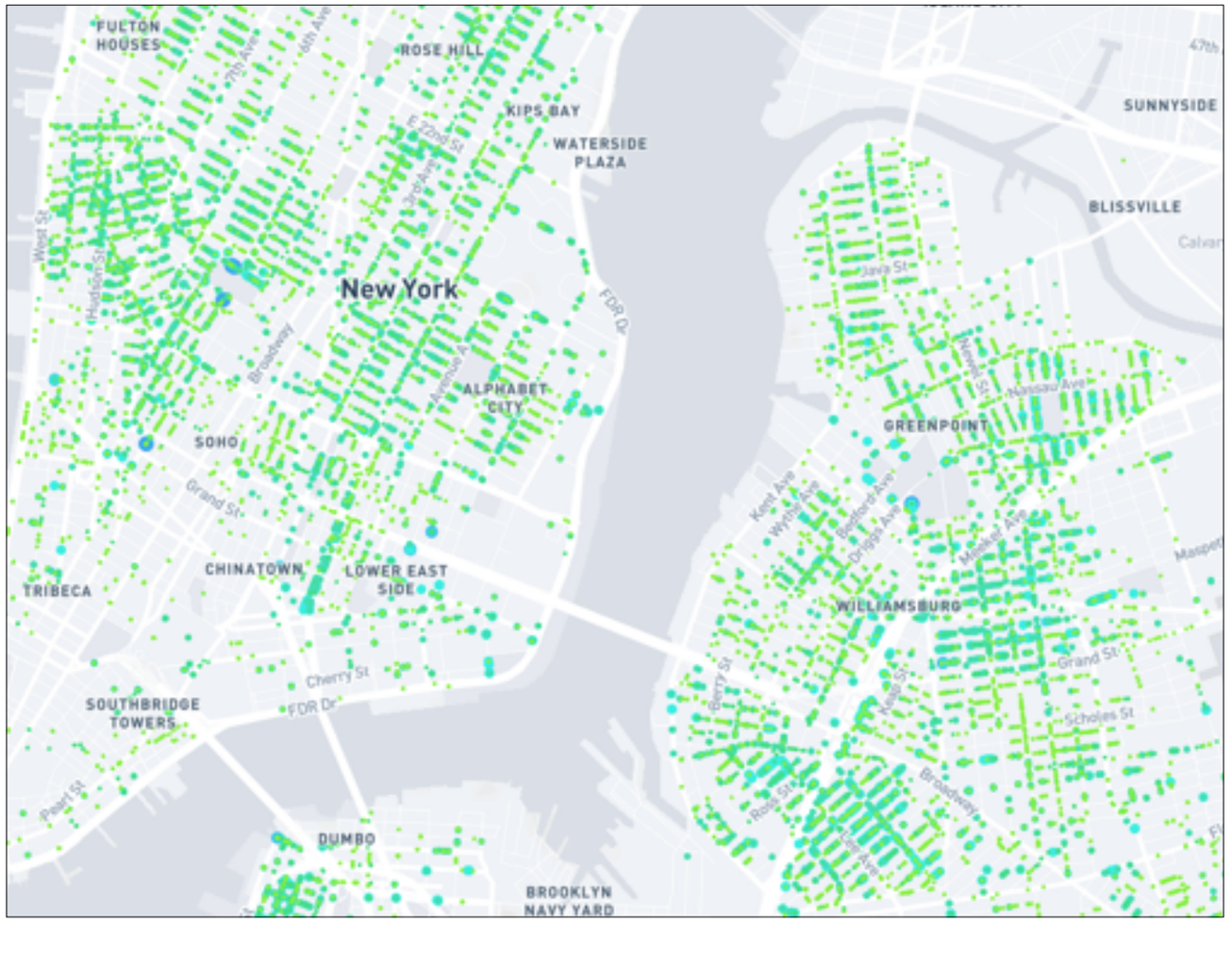
Reference maps and thematic maps are not mutually exclusive, one can be built on top of another and they can complement one another. When we combine maps we say we *overlay* them.

You might have a reference map of supermarkets in a city, overlaid on a thematic map showing neighbourhood population, to see if any neighbourhoods are underserved.

Maps descriptive of London poverty, Charles Booth, 1898–99.

Reference and thematic map combined. Data overlaid on a street map of London.

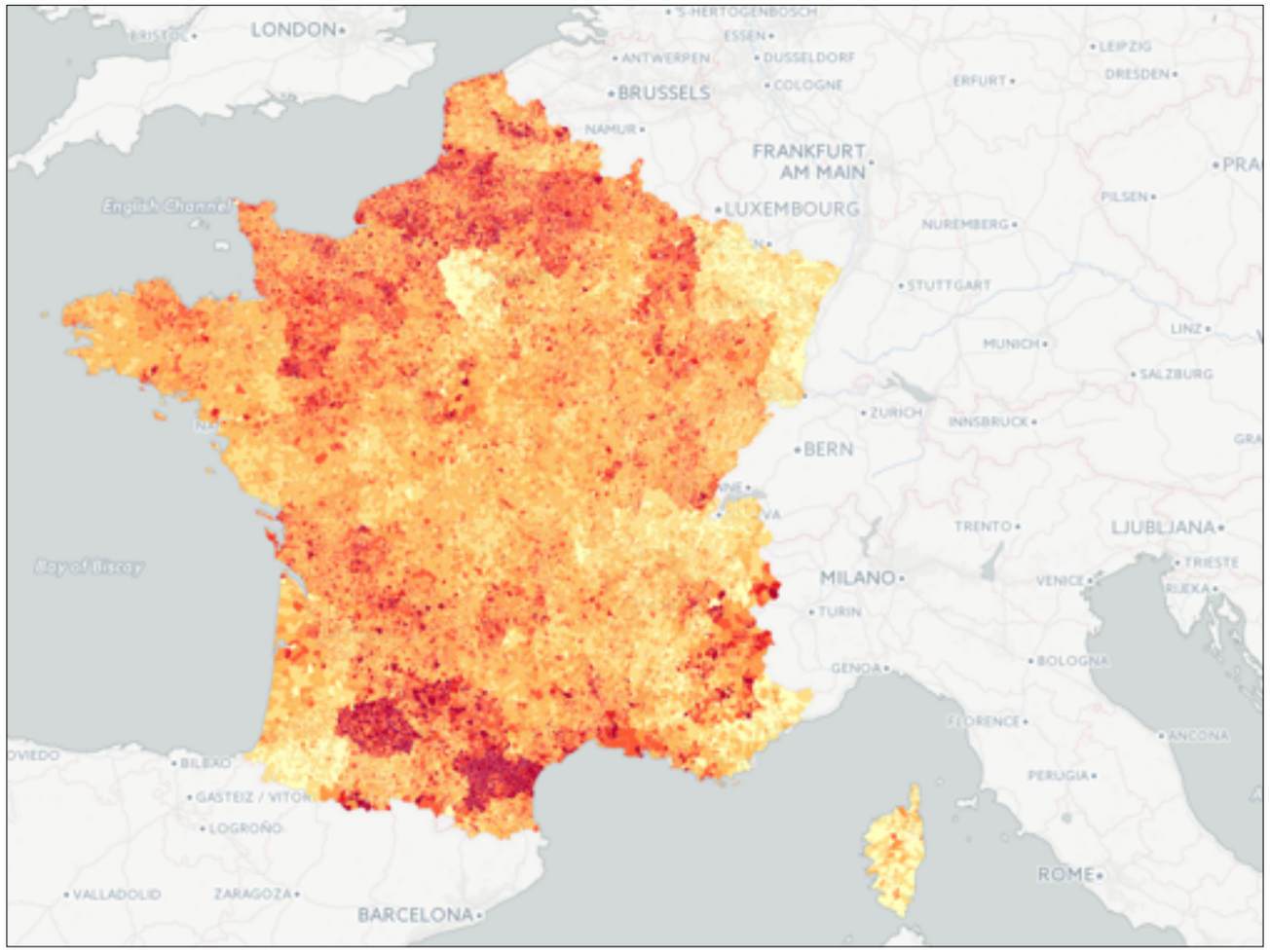
http://mapdesign.icaci.org/2014/06/mapcarte-161365-maps-descriptive-of-london-poverty-by-charles-booth-1898-9/



NY Trees, Eden Halperin, 2015.

MapBox's excellent Street reference map, overlaid with data on the city's trees.

https://www.mapbox.com/gallery/



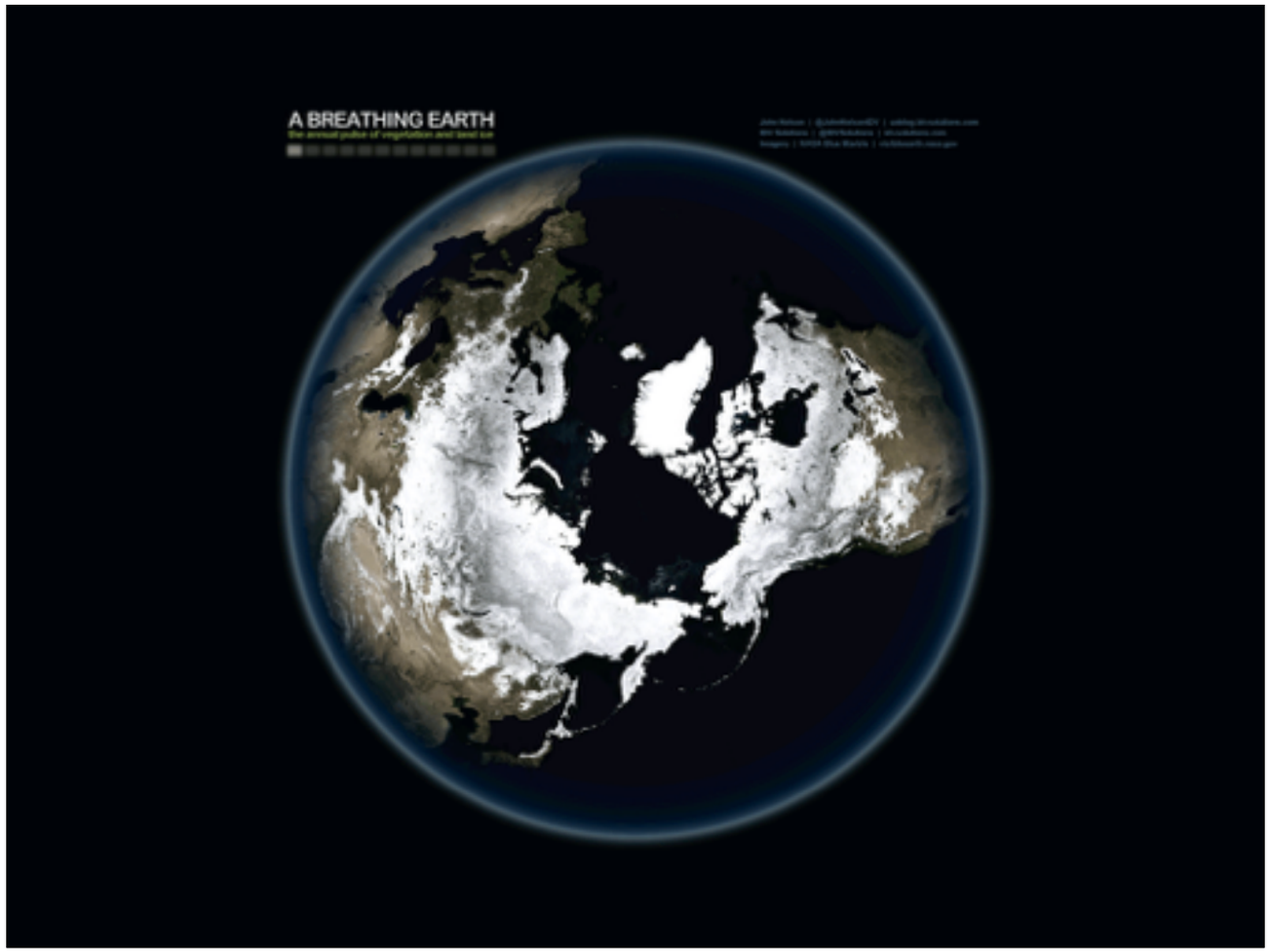
La carte des impôts locaux en France depuis 10 ans, Home’n’Go.

This is what many modern online maps look like: a reference map overlaid with a thematic map, where the reference map is of no use. Better to use a simpler base layer — land is useful but we don’t need the names of European cities.

https://homengo.com/immobilier/impots-locaux/carte/

Dynamic maps

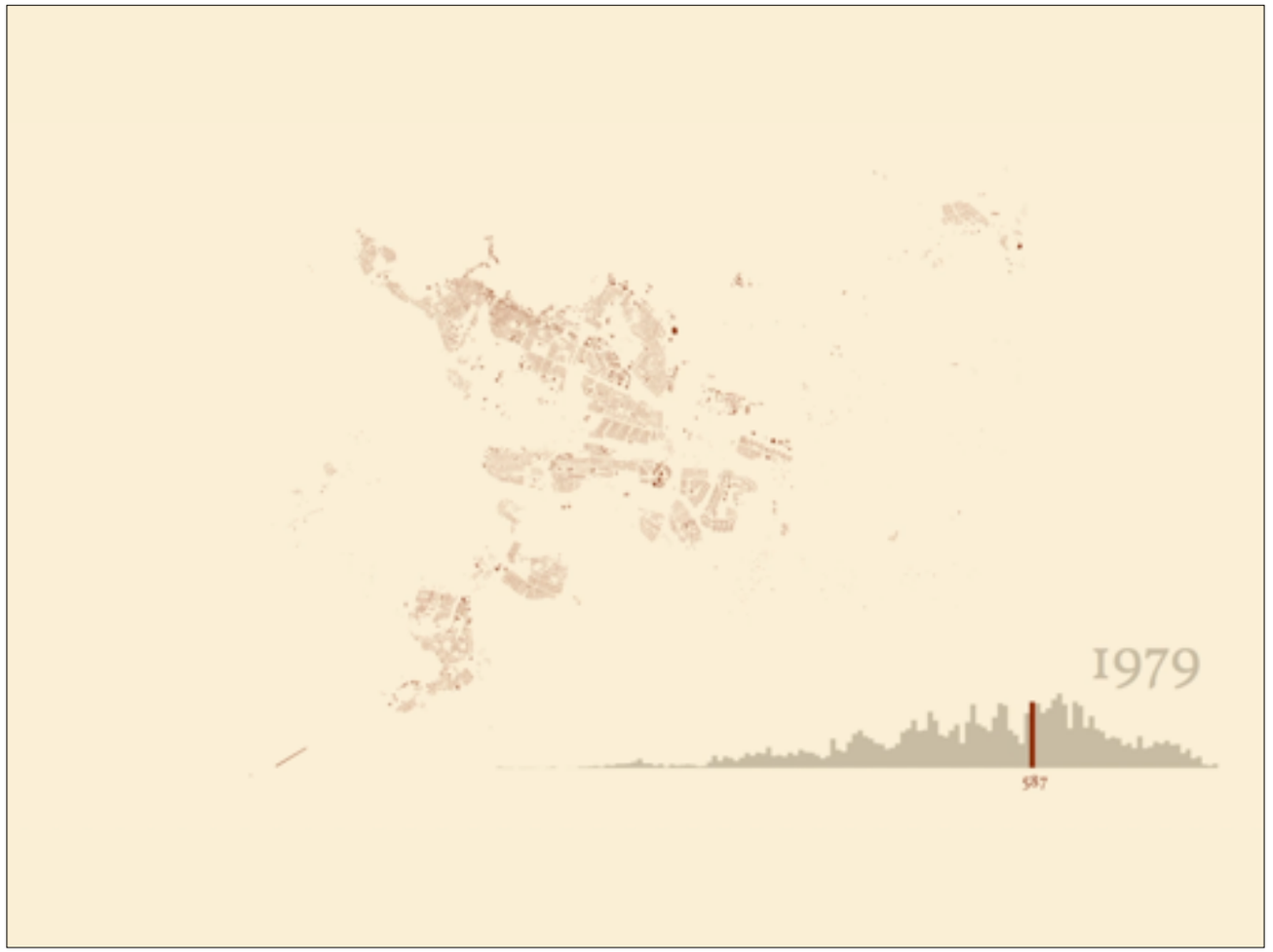
Some people say there are three types of map, the third being *dynamic*. But a dynamic map is simply either a reference or thematic map with interactivity and I don't consider it a separate type.



Breathing Earth, John Nelson, 2013.

Some people say there’s a third type of map, a dynamic map. But to me, this is just a subset of the other two types, reference or thematic. Here, this reference map is shown dynamically.

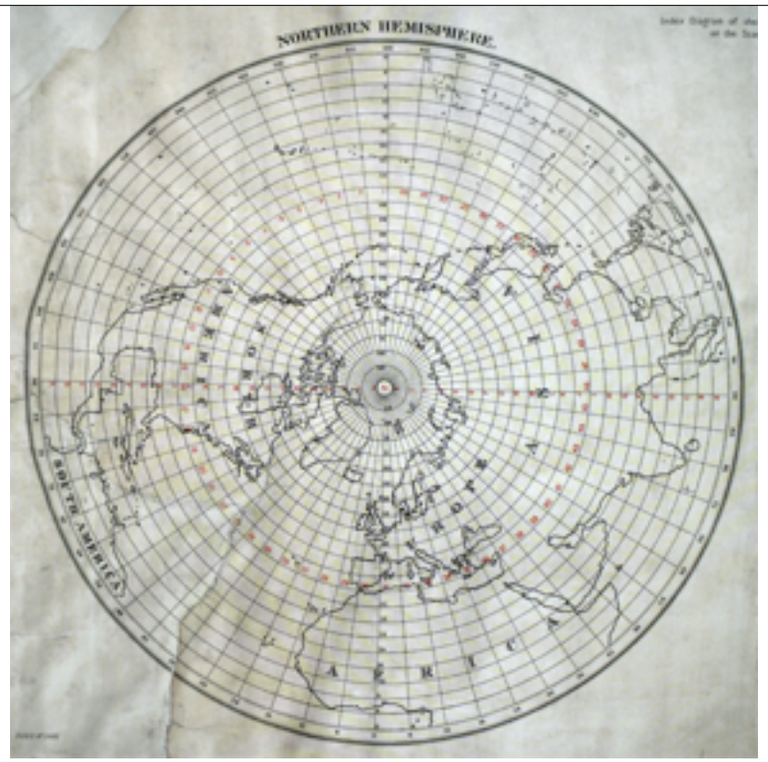
http://mapdesign.icaci.org/2014/08/mapcarte-216365-breathing-earth-by-john-nelson-2013/



Dynamic maps can seem terrifyingly complicated, but they’re really only a series of maps collected together. If you can make one map, you can make many maps.

This map I made by creating a map for each year and then stringing them together as frames in a movie.

Base maps



The International Map of the World

Late 19th century. Project started to create standard reference map. But eventually people realised different needs meant different maps. Same problem now with Google Maps as the base map for everything.

There would be no point using Google Maps as a base map here, for example. Different projects need different base maps.

Seafloor map of Hawai'i, Tom Patterson, 2012.

Here the base map includes flight lanes and flight information regions (FIRs). This is what Flight Radar 24 should do, and not just rely on Google Maps.

Airspace: the invisible infrastructure, NATS, 2014.

https://vimeo.com/110348926

Don't just use Google Maps as your base map.

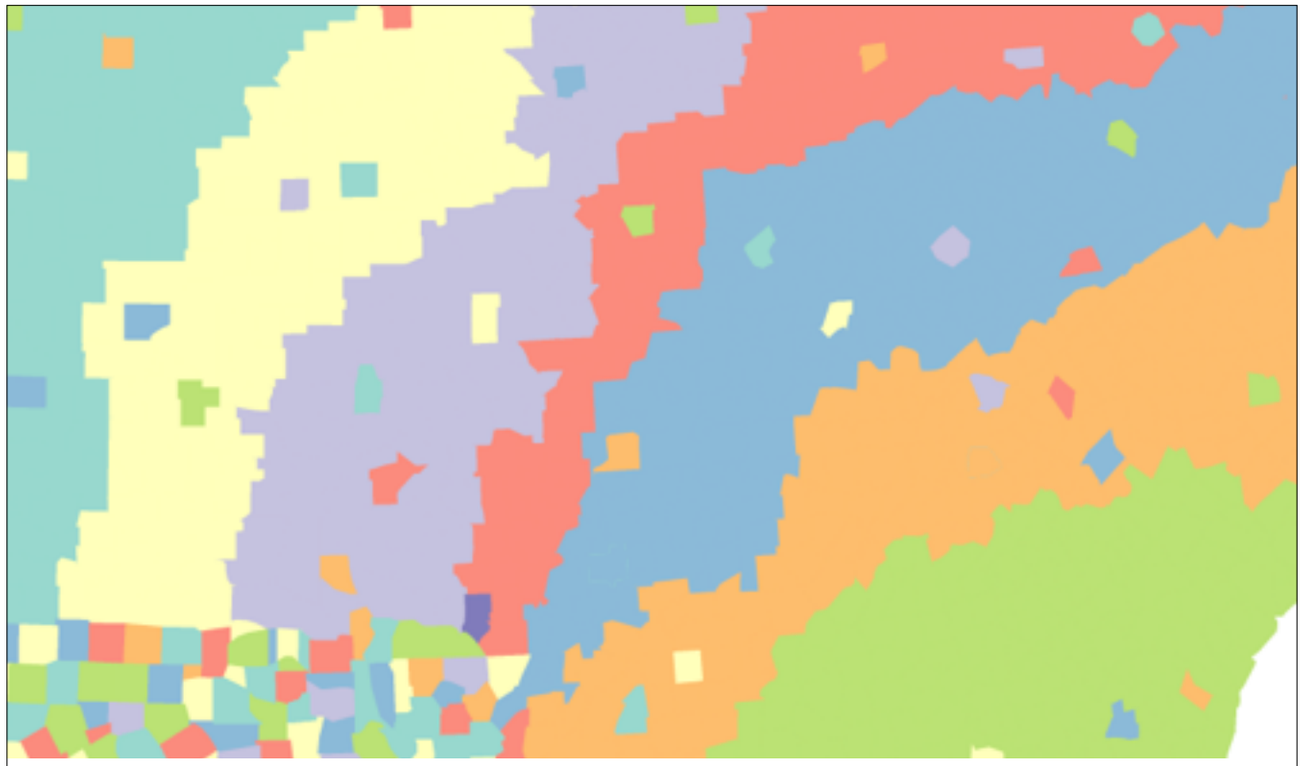
They make design decisions you might not want (Web Mercator projection, road map, human population centres, etc).

If you can, build your own.

If you can't, pick one that suits you as closely as possible. MapBox is good.

Colour schemes

- *Qualitative*: unordered categorical data
- *Sequential*: numerical data; ordered categories (low/medium/high)
- *Diverging*: data with natural mid-point (e.g. zero, positive/negative growth); comparing against national average

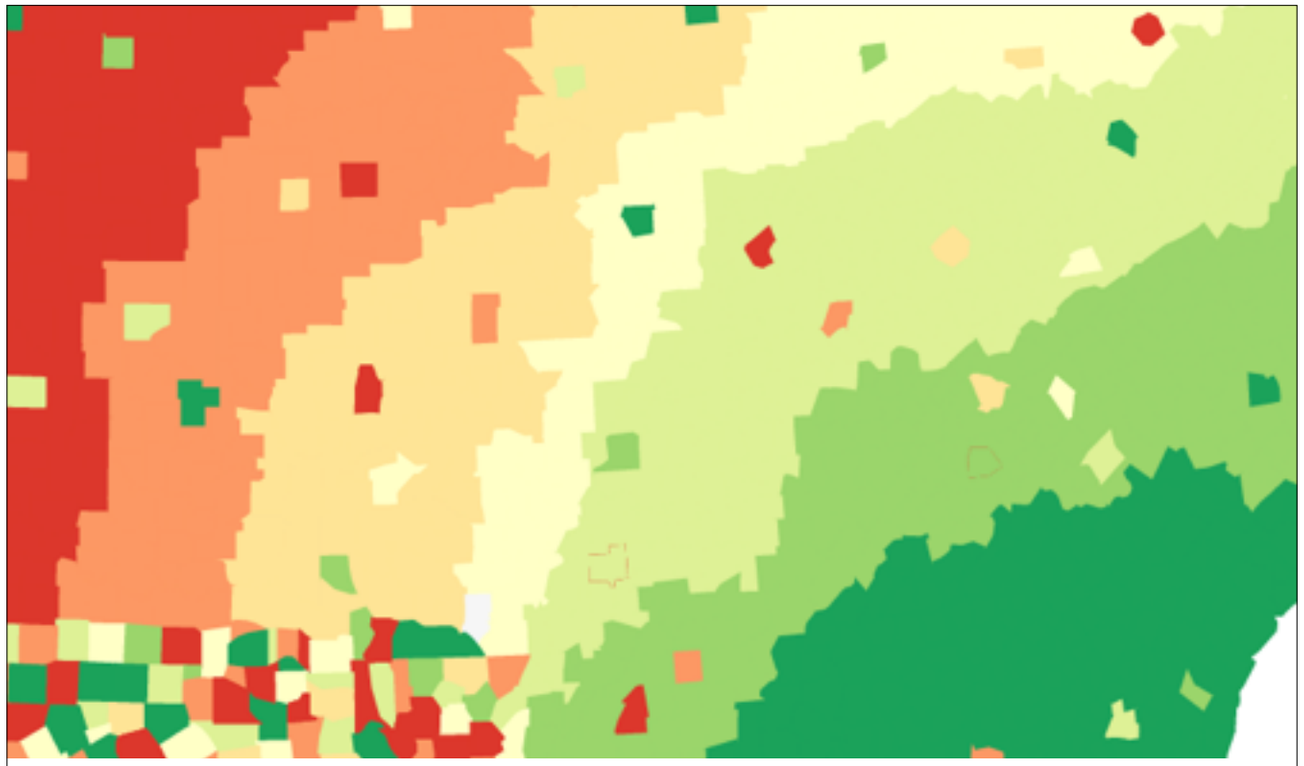


Qualitative

Unordered categorical data (ethnicity, alcohol type, etc).

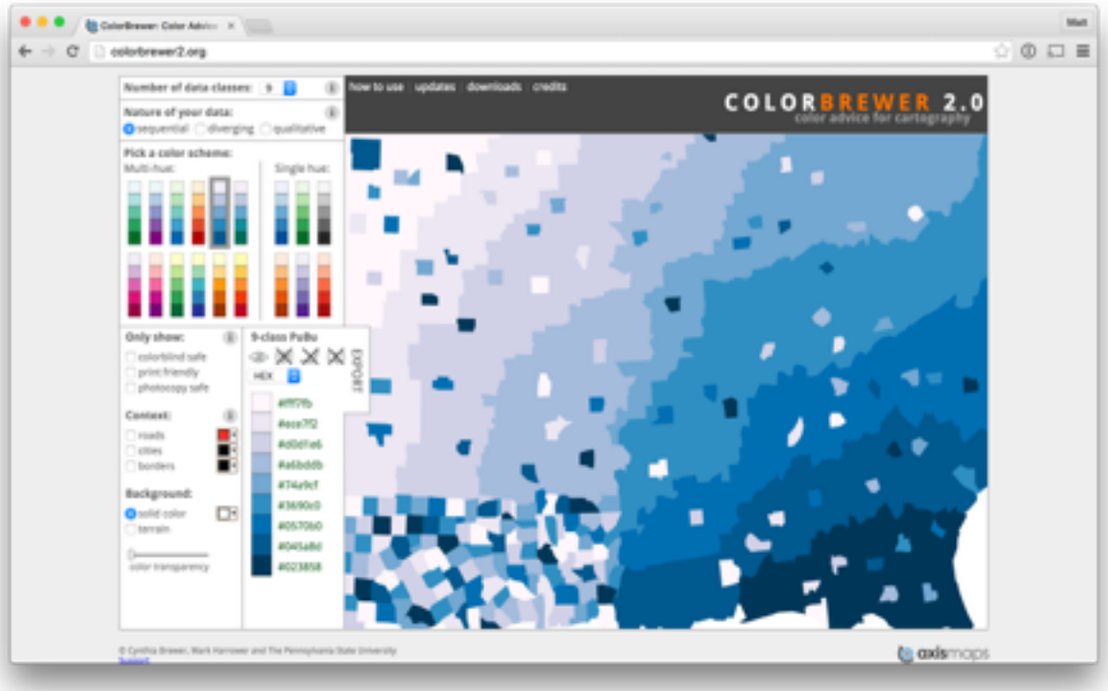
Sequential

Numerical data or ordered categorical data (e.g. high, medium, low).



Diverging

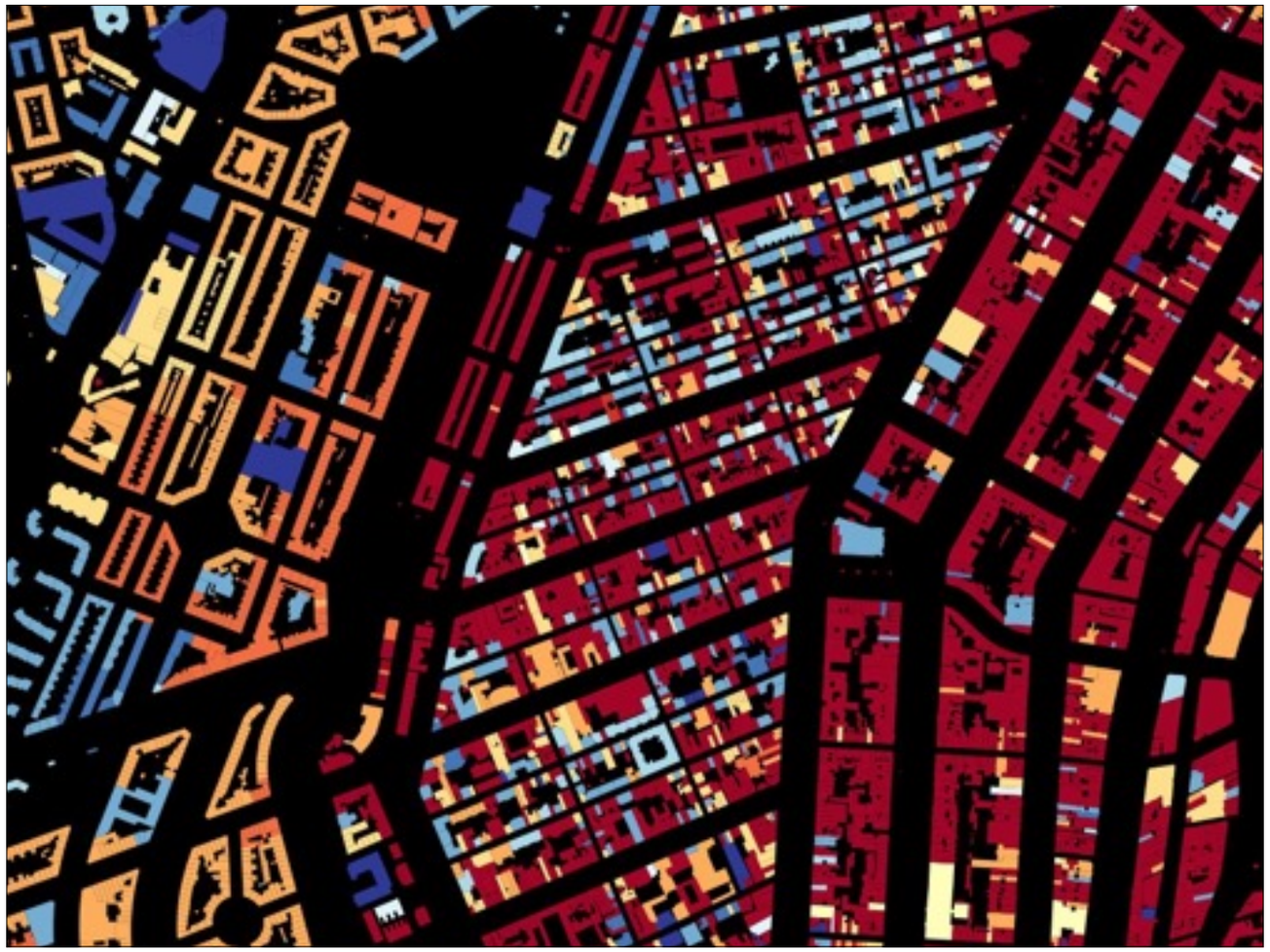
Data with a natural mid-point (divergence from national average).



Picking colours isn't subjective.
Use ColorBrewer.

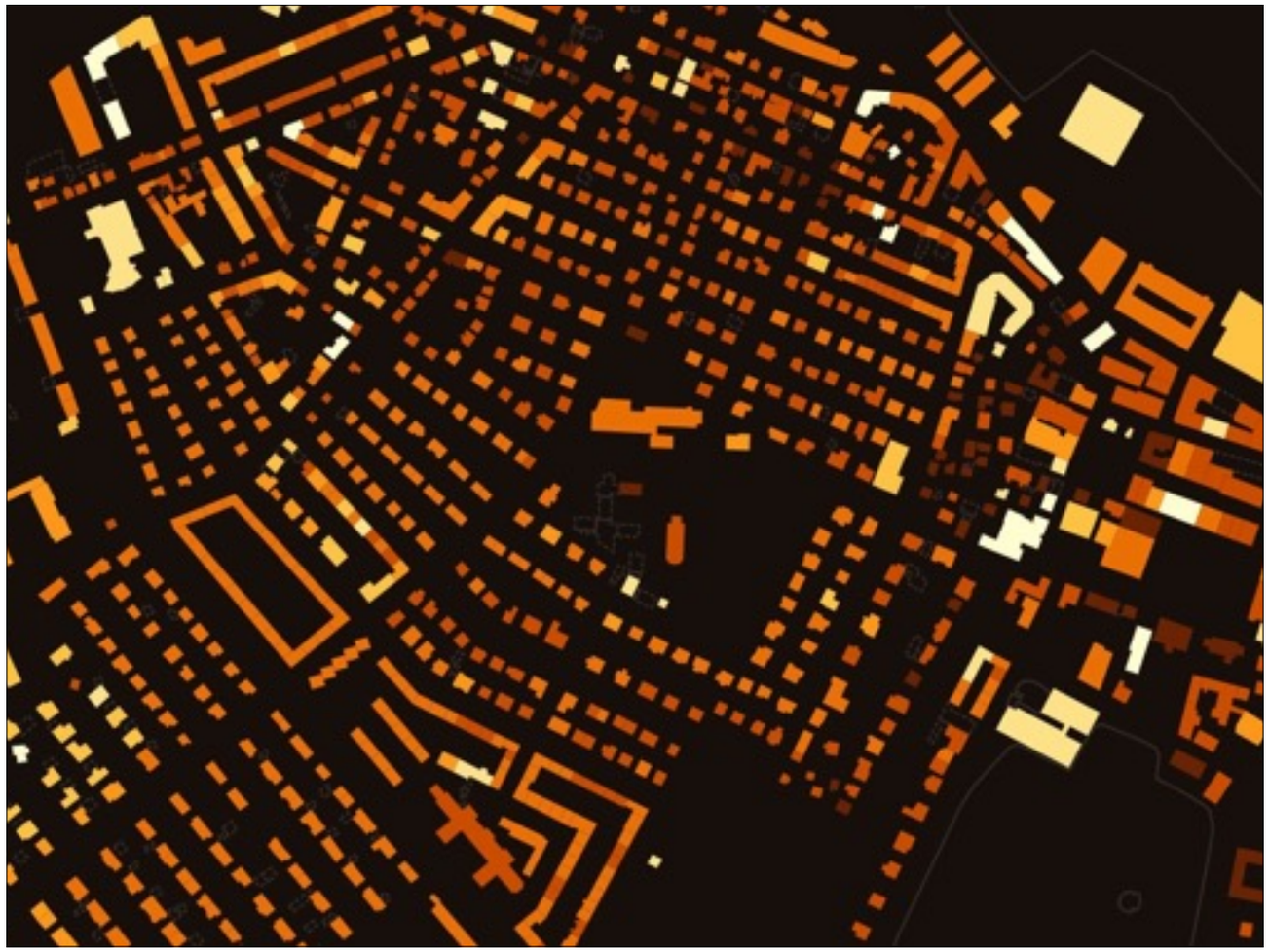
Don't pick colours because they look nice. Making professional colour schemes like this isn’t easy.

Colour schemes on ColorBrewer are perceptually graded so that the amount of colour change in each step looks consistent to our eyes. They're not, but they look like they are because humans do not perceive differences in all hues equally well.

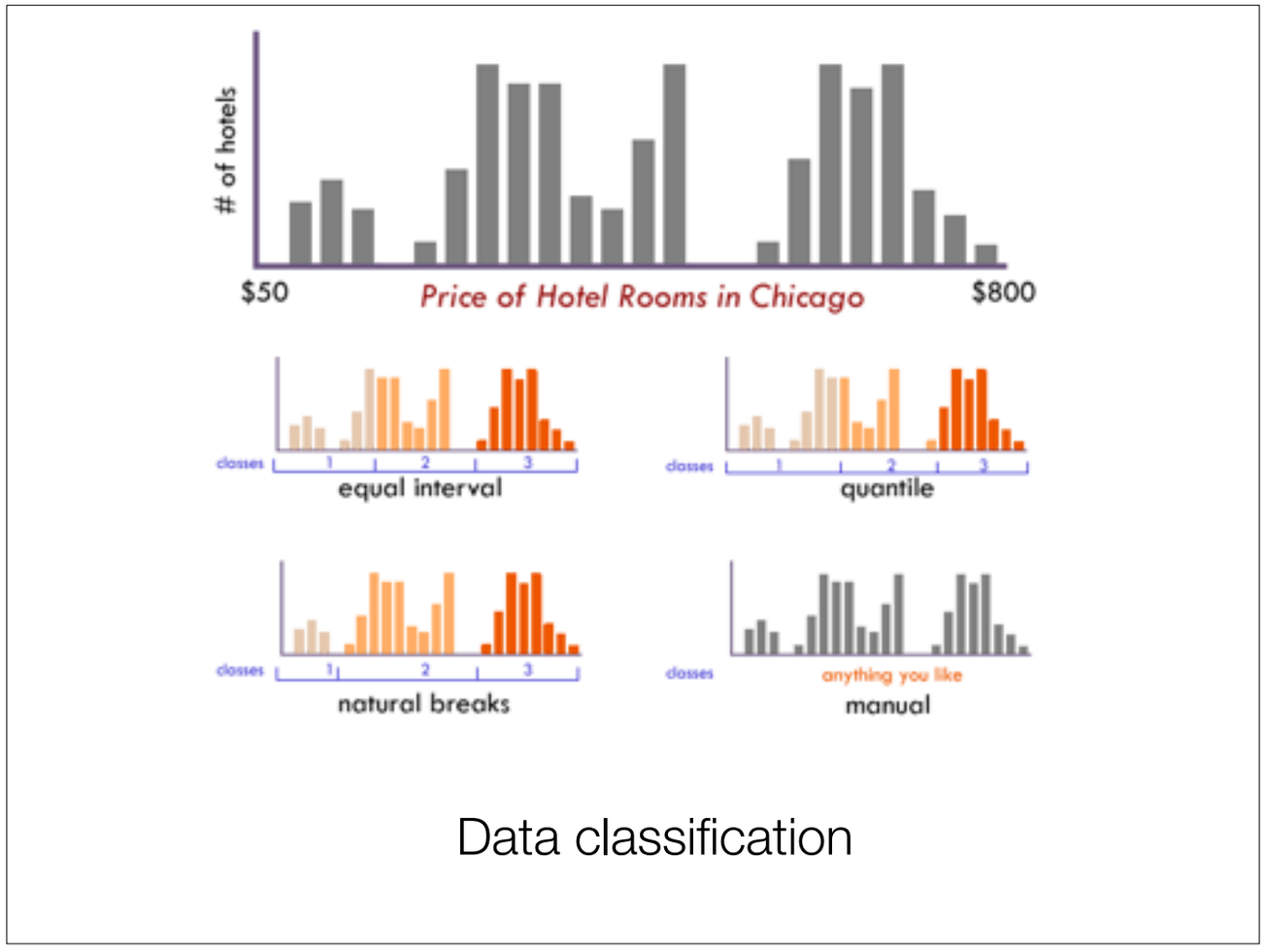


Use the wrong colour scheme for your map, and you make it harder to understand.

Are blue buildings older than yellow? Are yellow older than red? Here a qualitative colour scheme, one for unordered categorical data, has been used for numerical data. The sequence of colour is not understandable — even with a legend it’s hard.



This map uses a sequential colour scheme, moving from dark to light, allowing you to understand how the buildings compare to one another. You need a legend to know whether dark or light represents the oldest buildings, but once you know that it’s easy to memorise.



Data classification

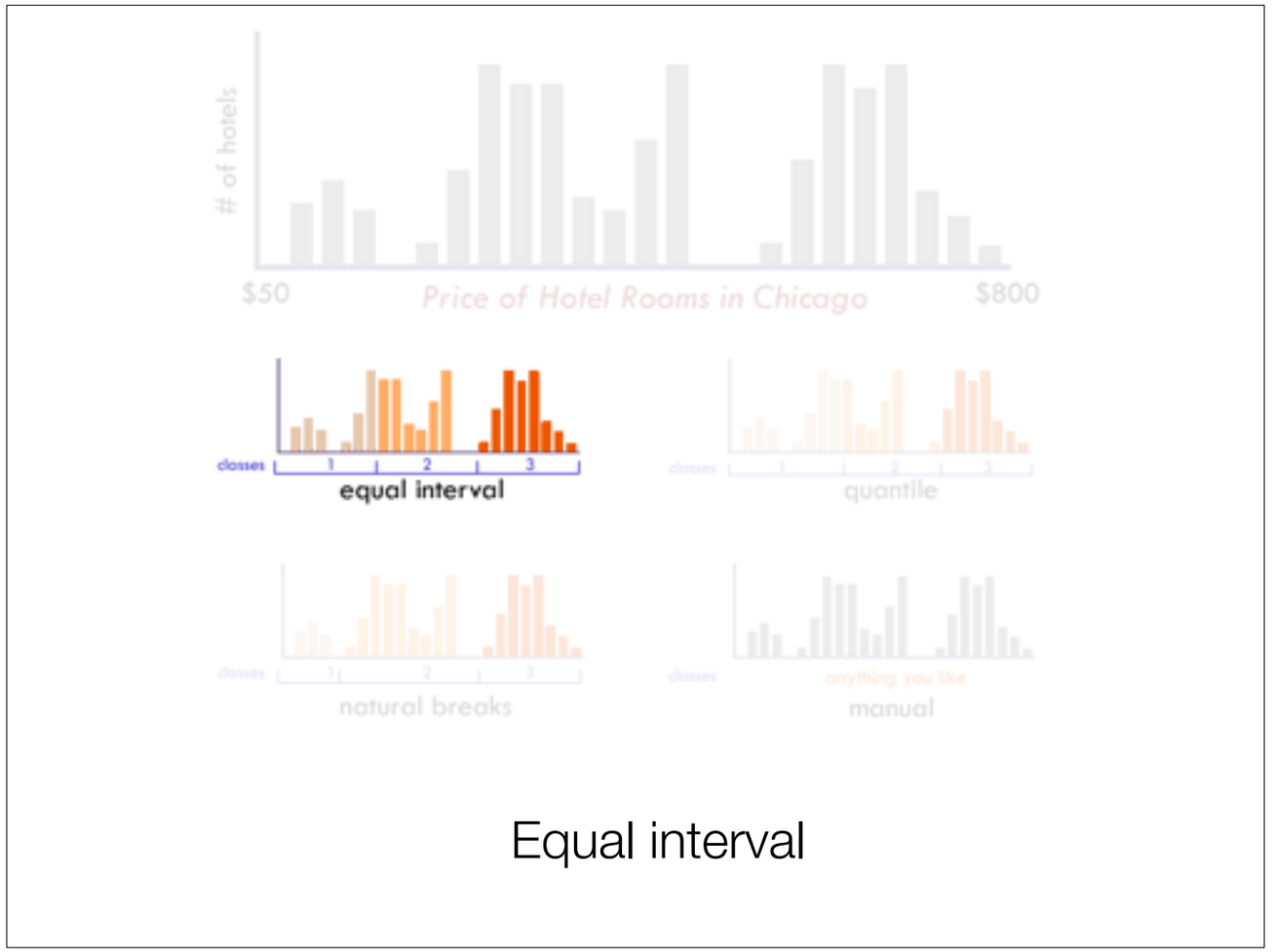
Decide: number of classes; and method for breaking into classes.

Why? It can be easier to read a map with a few well-defined classes, instead of raw data. They can simplify and clarify your message.

A bad choice of classification can create false patterns and result in a misleading map.

Three to seven classes works on maps. More classes means less generalisation (which is good) but less legibility (which is bad). How much generalisation do you want?

https://axismaps.github.io/thematic-cartography/articles/classification.html

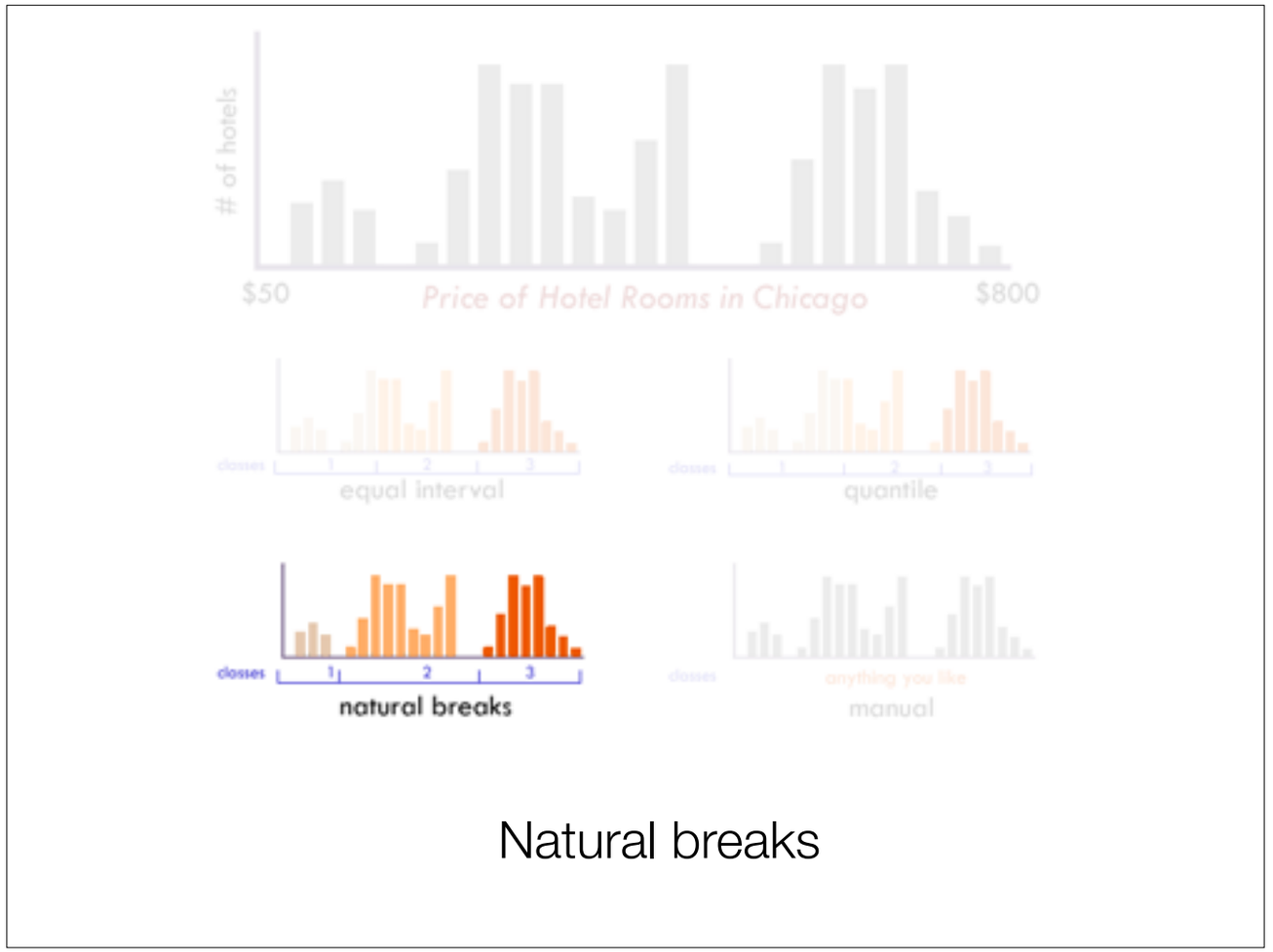


Equal interval

Equal interval gives you equally-sized classes (0–9, 10–19, 20–29) and works well on data that's spread equally.

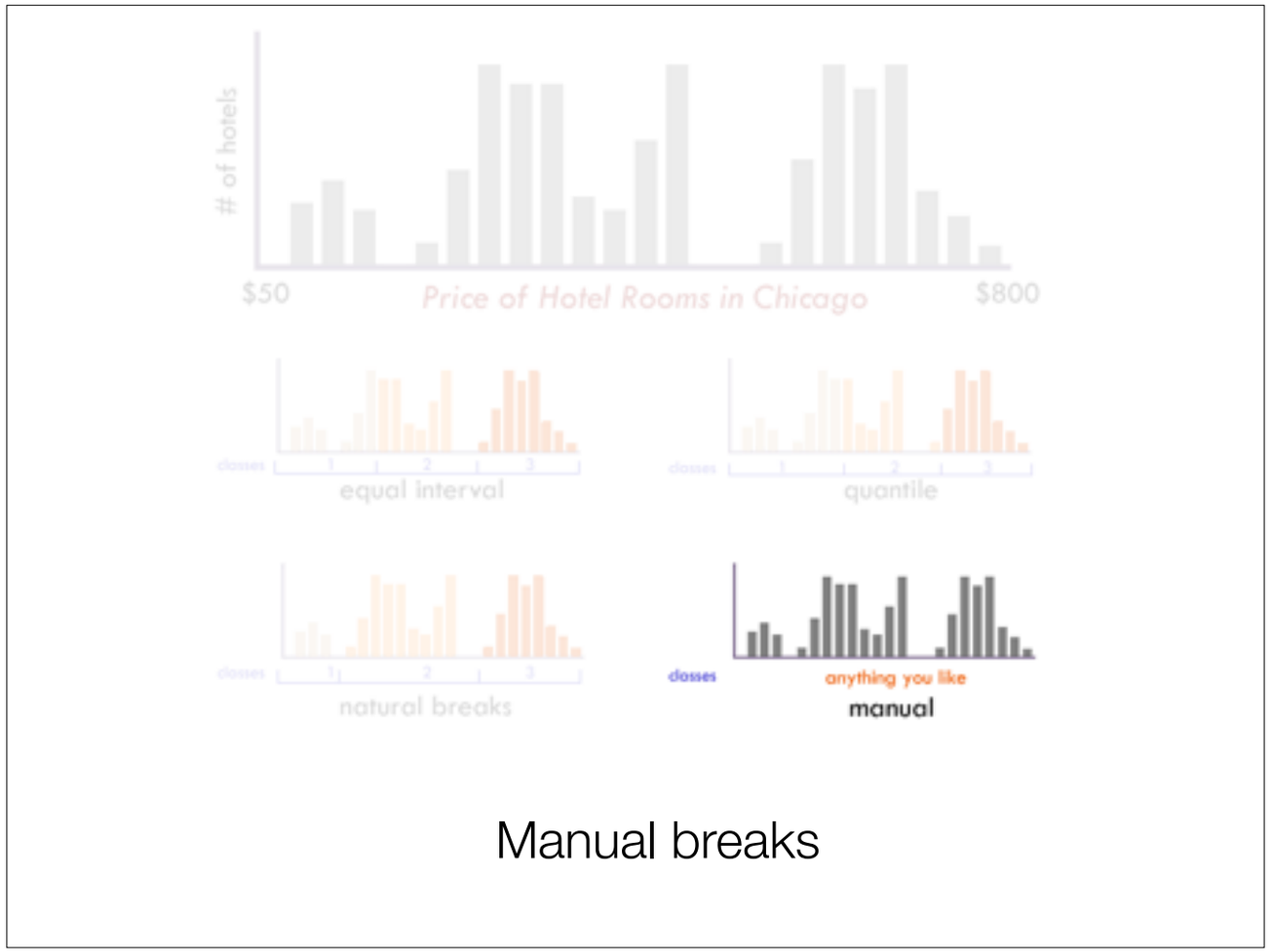
Quantiles

Quantiles give you classes that each contain the same number of observations. This gives you nice groupings (e.g. 50 countries in five classes gives you ten countries in each class) but can give you classes with very different numerical ranges (0–2, 3–6, 6–29). It can also classify close values separately and distant values together, which isn't great.



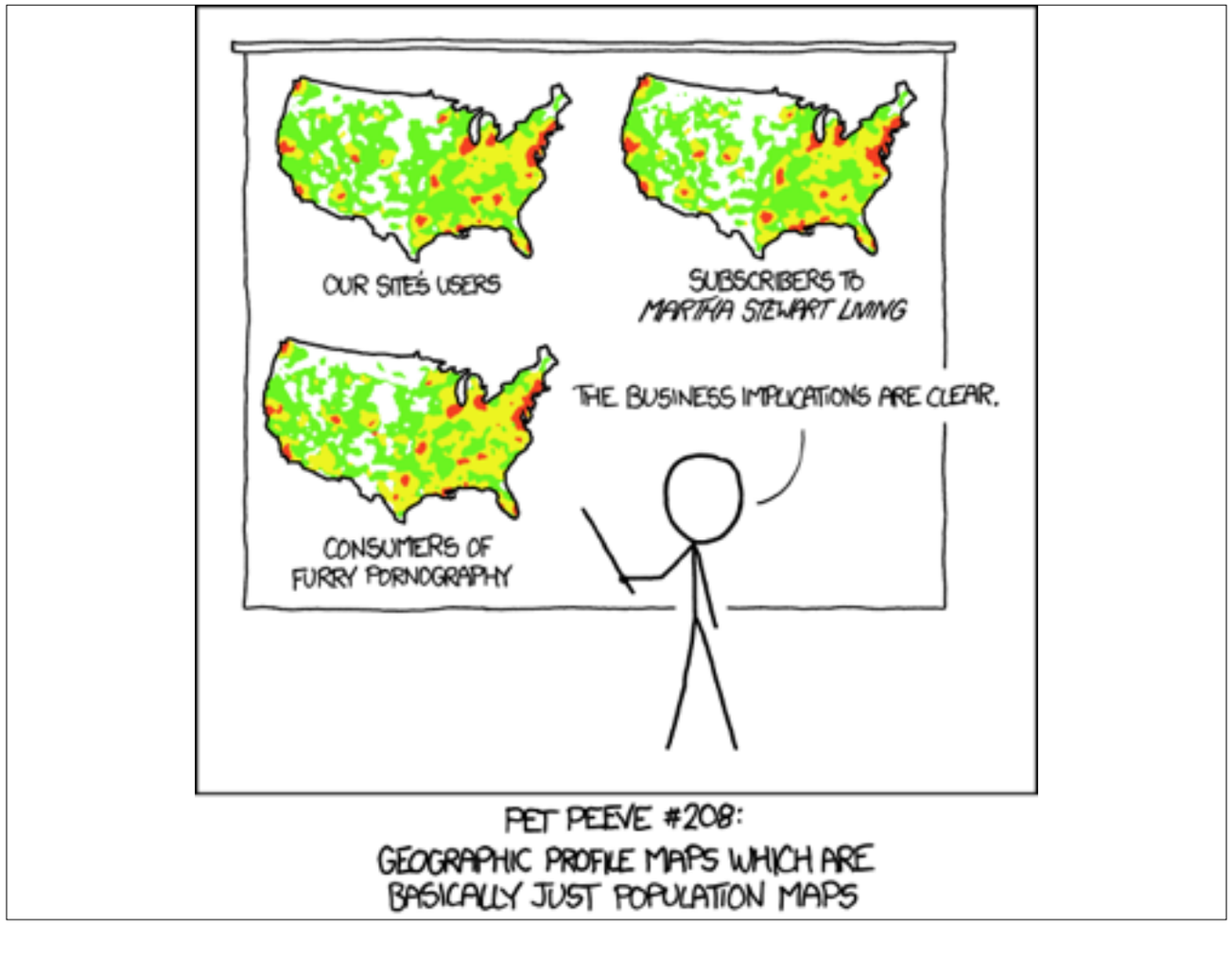
Natural breaks

Natural breaks tries to minimise the problems of these two approaches. One drawback is that each dataset generates a unique set of classes, making it hard to compare different maps.



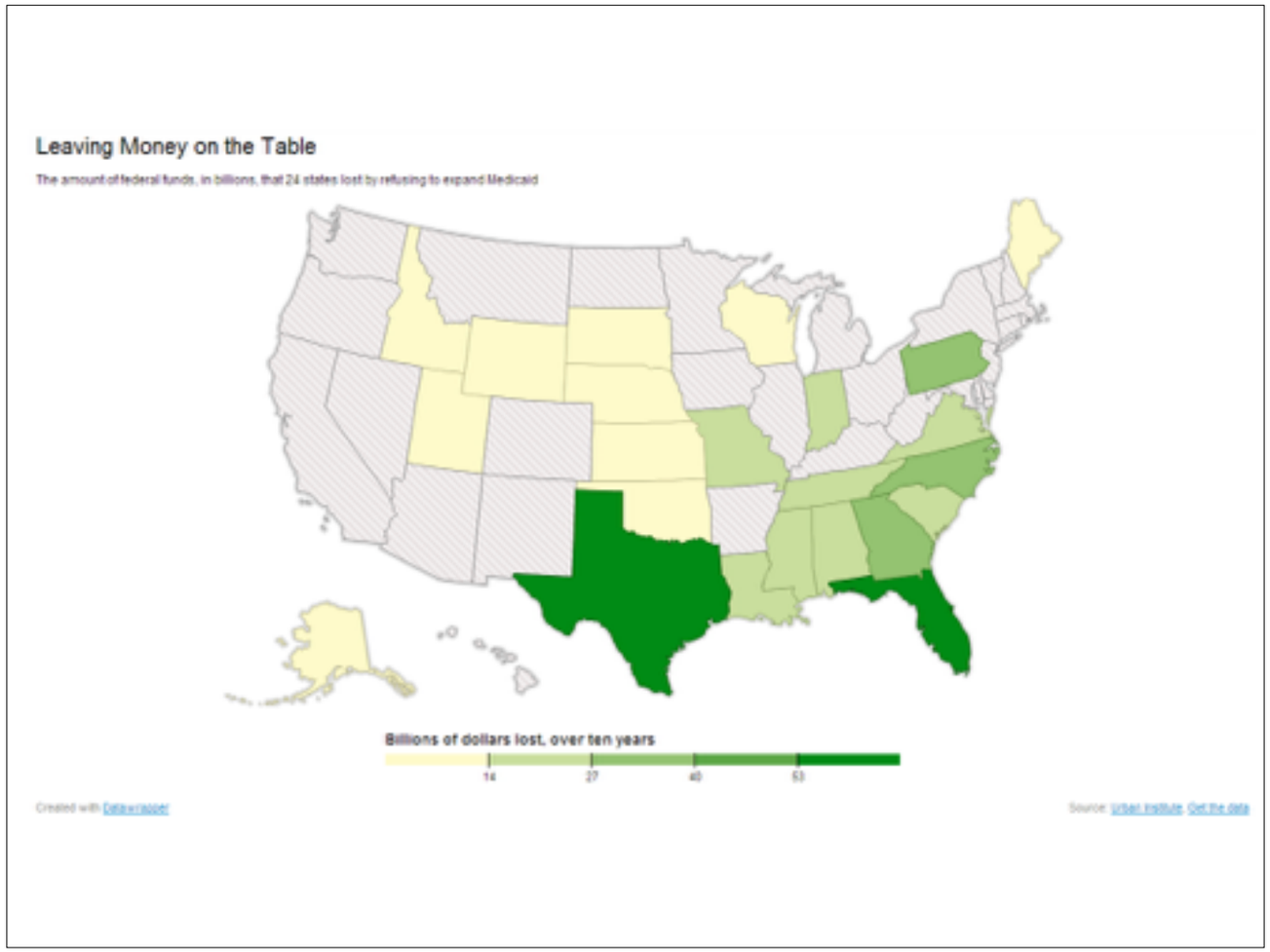
Manual breaks

You can always manually choose your classes. Perhaps you want to use the same classes across a series of maps, or if you want to tweak the classes created by an automatic method.

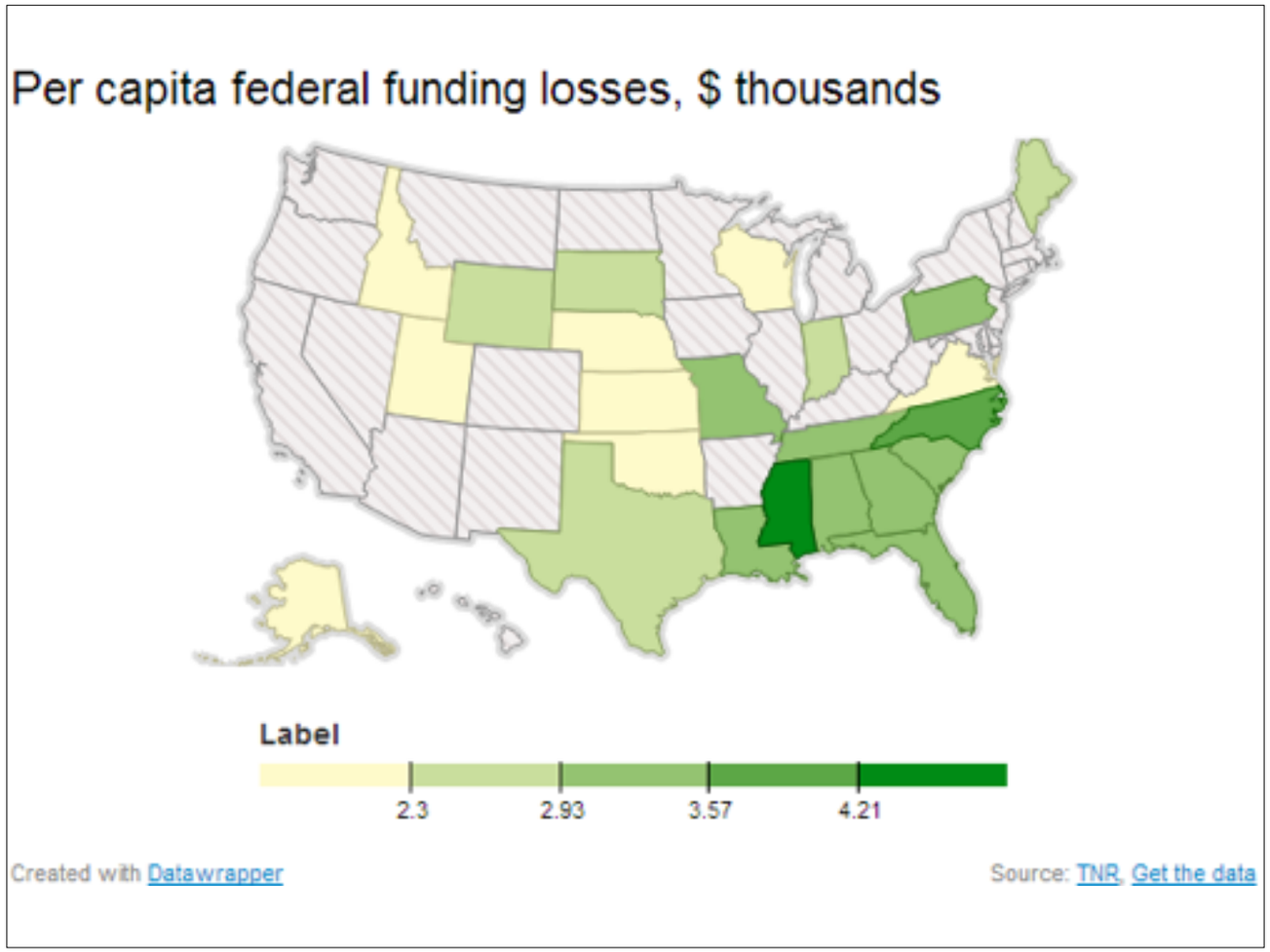


Normalise your data or you'll only end up with a map of populated places.

https://xkcd.com/1138/

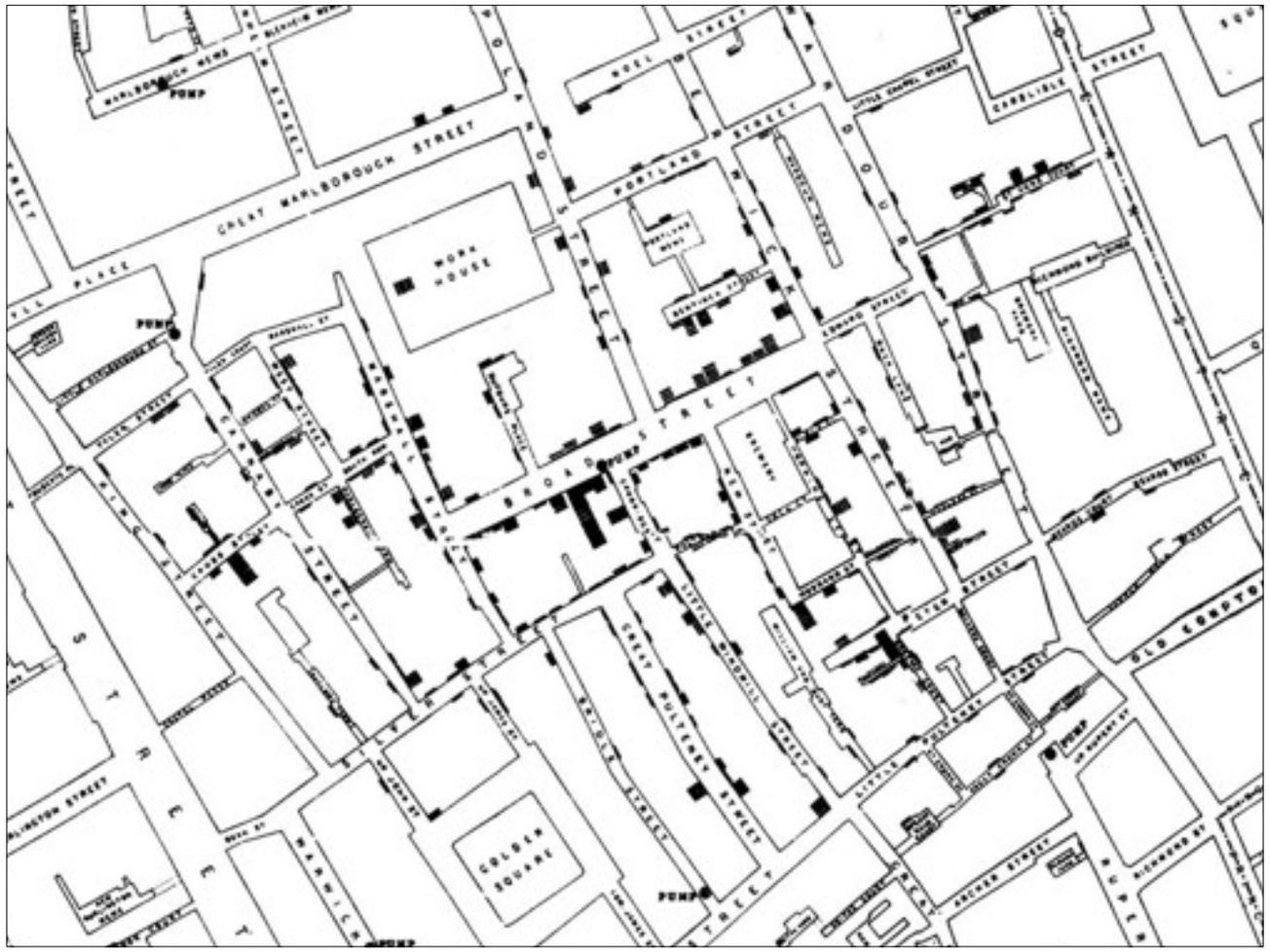


Unnormalised data show Texas and Florida as the biggest losers here. They happen to be the states with the largest population too.



Normalise the data and the real story appears.

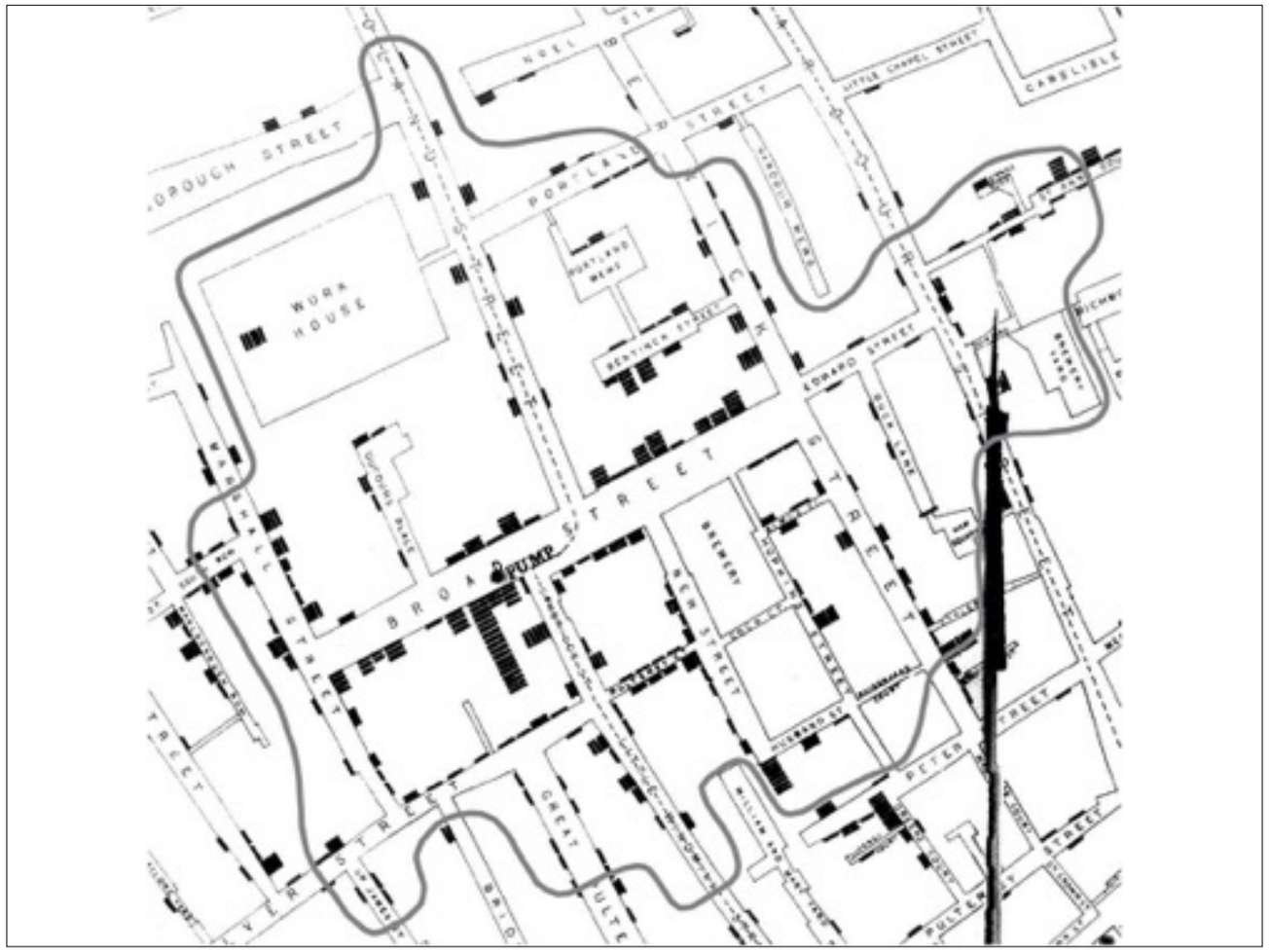
http://cartonerd.blogspot.is/2014/08/the-reason-for-cartography.html



In 1854 John Snow investigated an outbreak of cholera in London. The prevailing theory of the time was “bad air” caused cholera.

Snow had doubts, and decided to go door to door, asking at each house whether anyone had experienced cholera and from where they drew their drinking water. His map shows a clear clustering of cases around a particular well, the Broad Street pump.

But what if more people lived in the houses where more people died? The map doesn’t definitively show the pump was responsible, it’s missing normalisation.

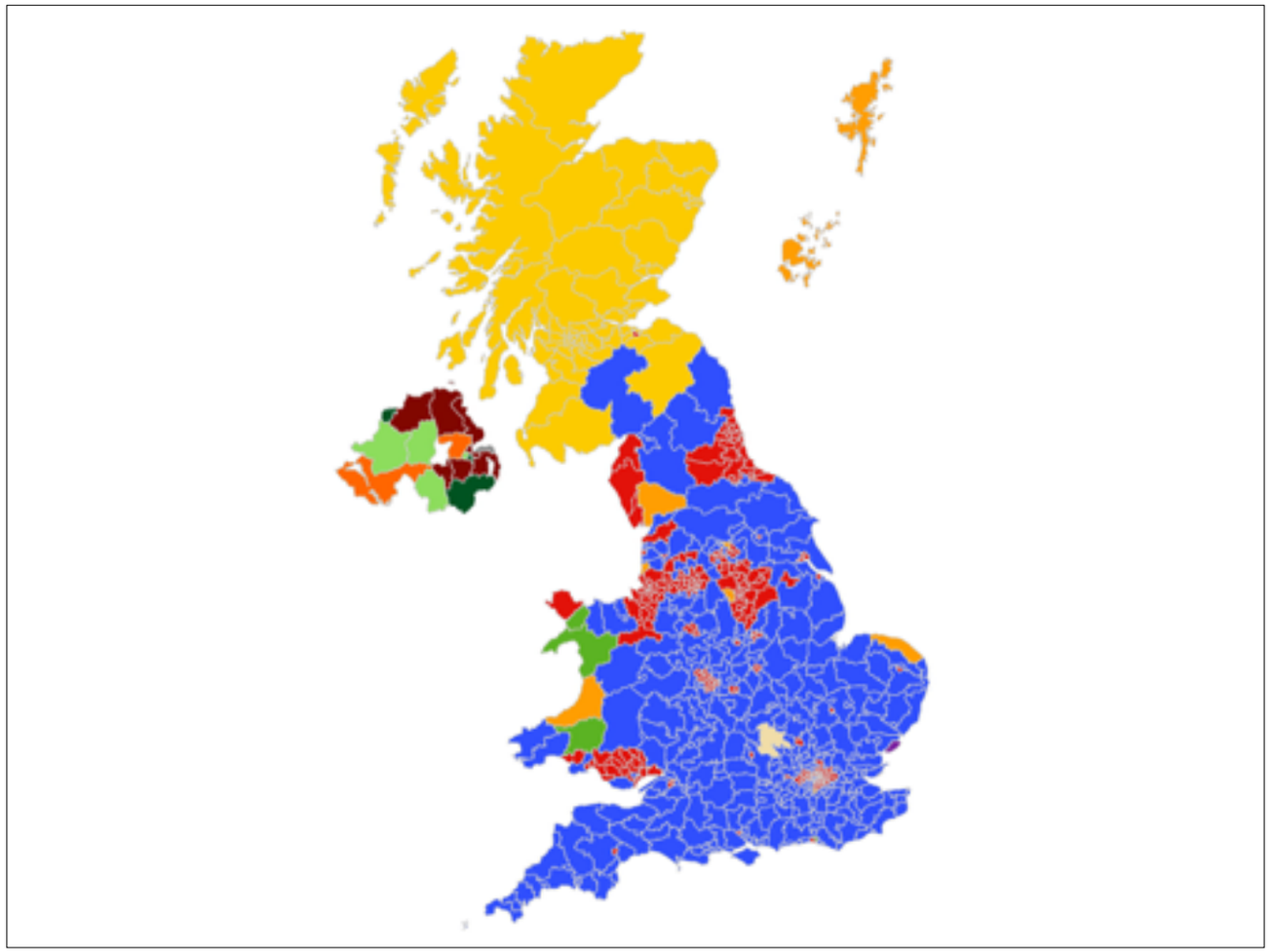


Aside: the map only proves that there were lots of deaths near the pump, not whether water was to blame --- perhaps the vapours emitted from the pump caused the deaths. A better map was needed.

Why are there clusters of deaths outside the walking distance? Those people used this pump instead of a closer pump because they thought the water tasted better. And what of the houses near the pump that had no deaths? One was a brewery where the workers drank enough malt liquor to mean they didn't bother drinking water.

Data normalisation

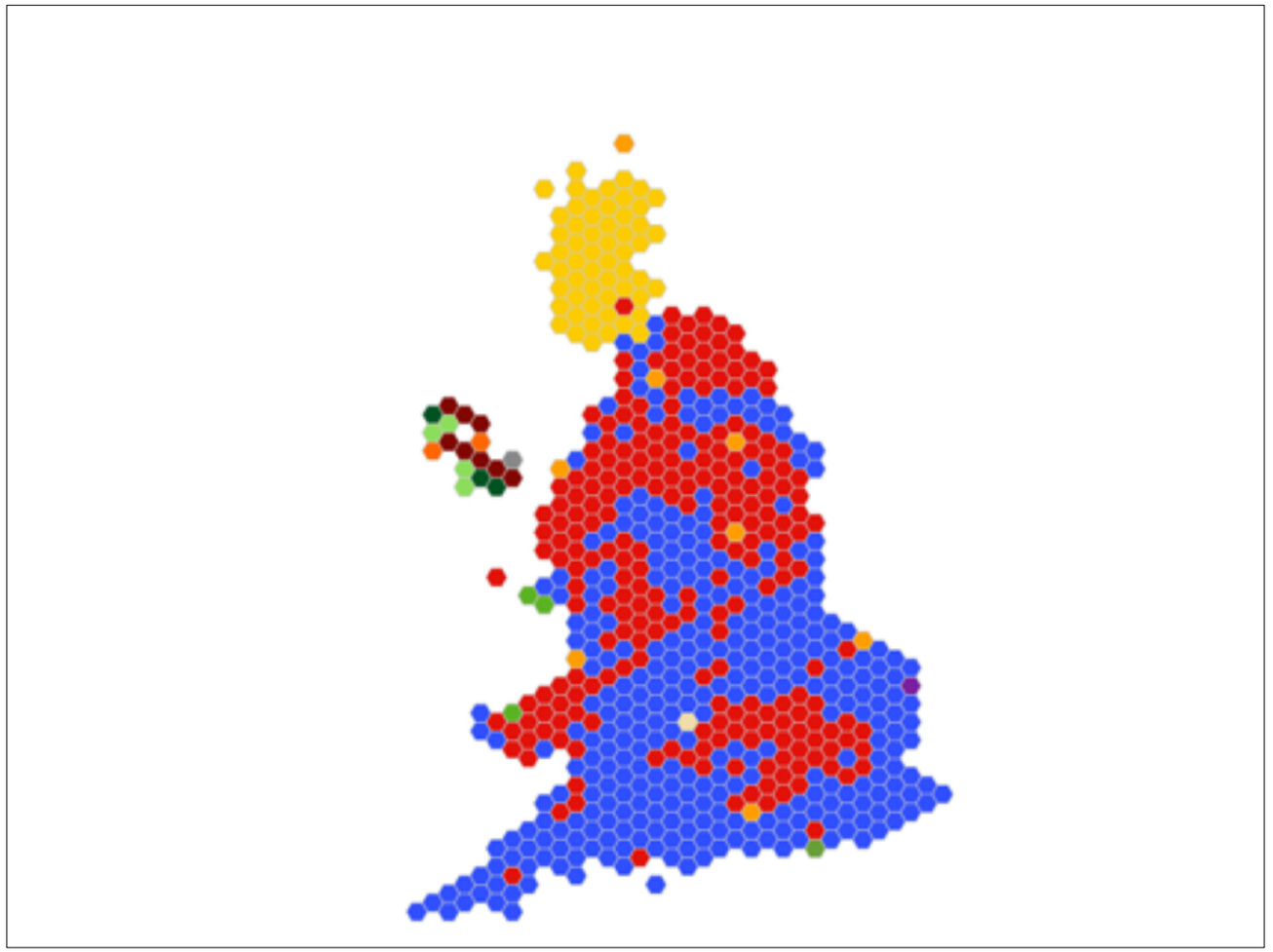
- Rate per 10,000 people
- Rate per 100,000 people
- Rate per 1,000,000 people
- People per square kilometre
- Per capita



Human geography is arbitrary: Two-thirds of the population of Iceland is on a tiny peninsula that covers 1% of the country.

Map of the 2015 UK general election: London has more parliamentary constituencies than Scotland but because London is only 1/50th the size of Scotland, Scotland looks more important. You can solve this by using a cartogram.

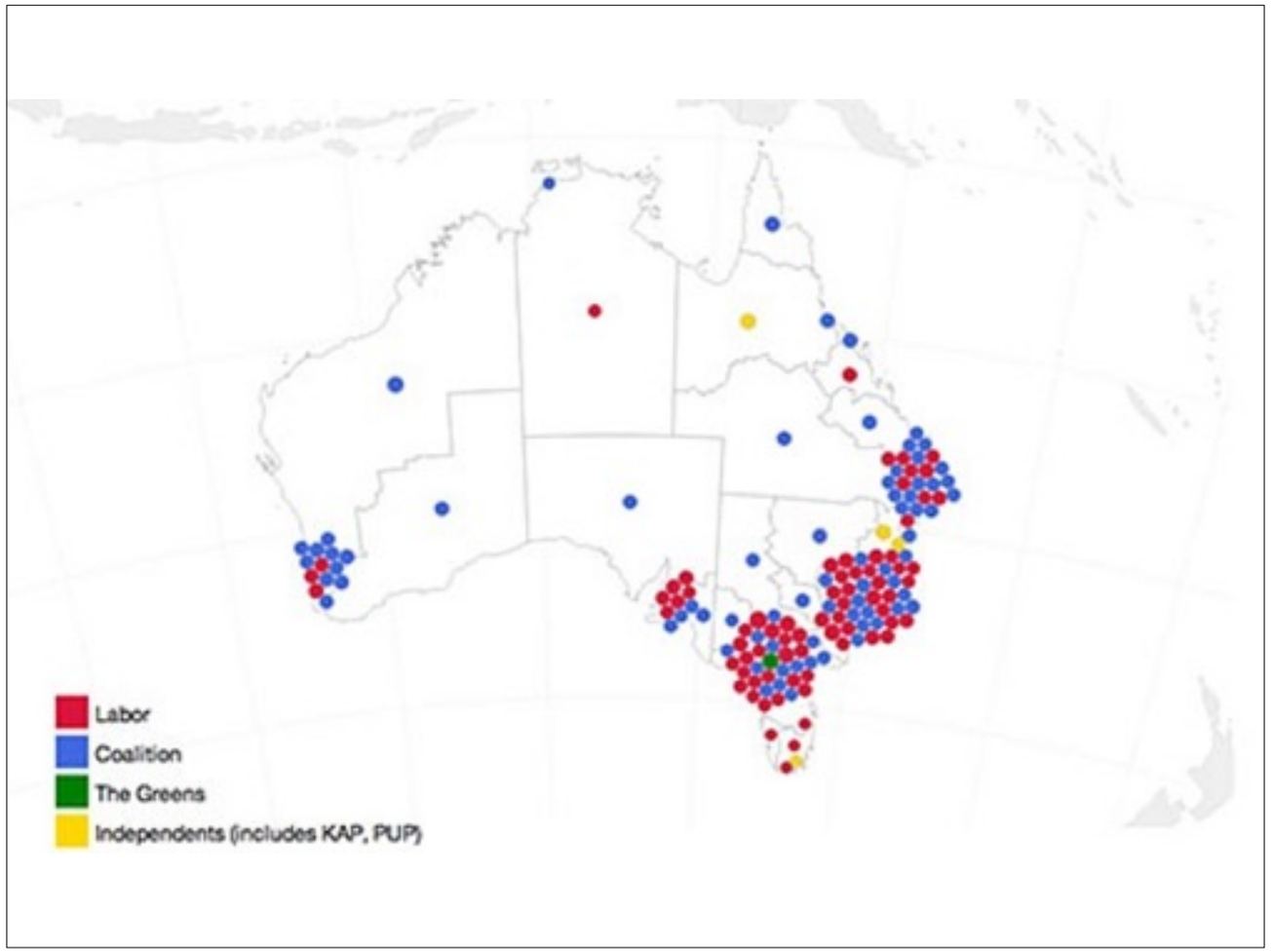
http://www.economist.com/uk2015data



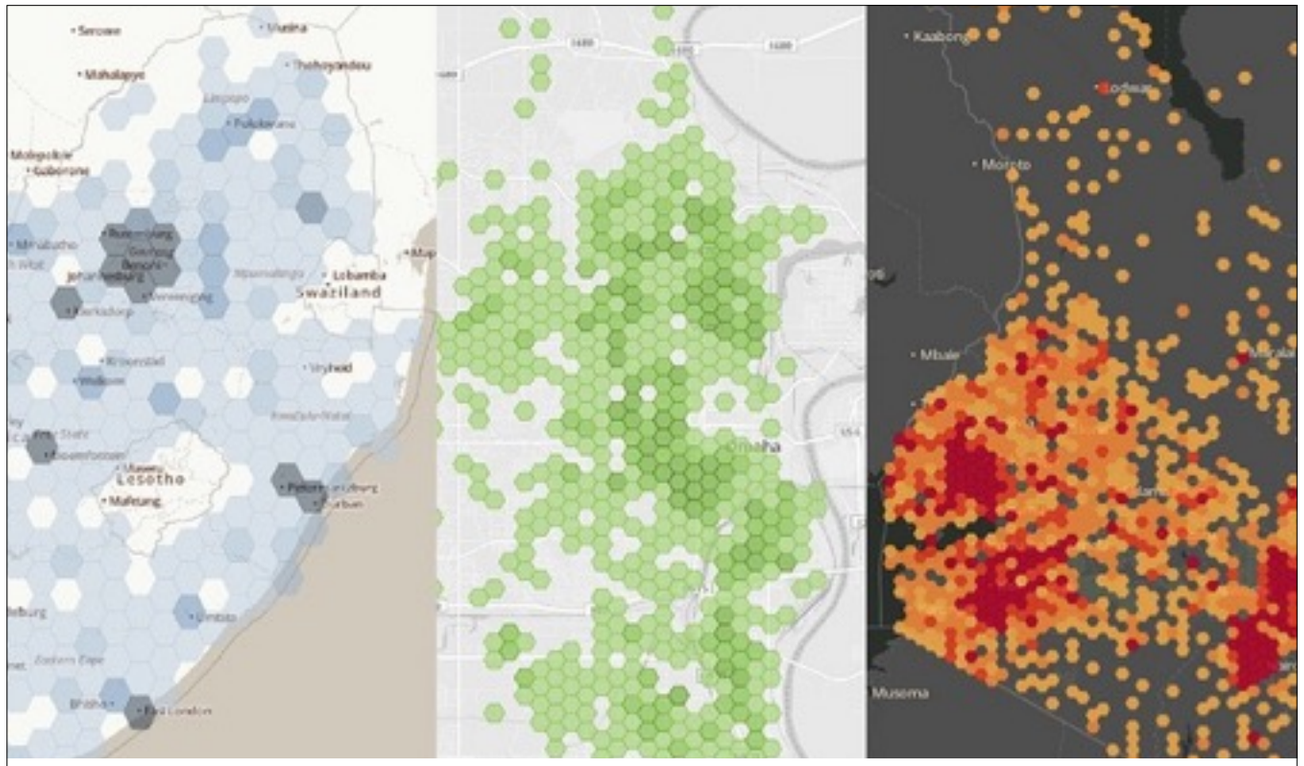
Scotland is much smaller, London much larger, and the north of England turns much redder.

Cartograms are useful for election maps because constituencies tend to have roughly an equal number of voters or because the outcome is equal (one member of parliament per constituency).

http://www.economist.com/uk2015data



Sometimes population is so sparse that something like the previous cartogram wouldn’t work, and you need to try something different. Australia’s population is sparse (as is Iceland’s) and so the cartogram needs to be spread across a geographically accurate reference map.

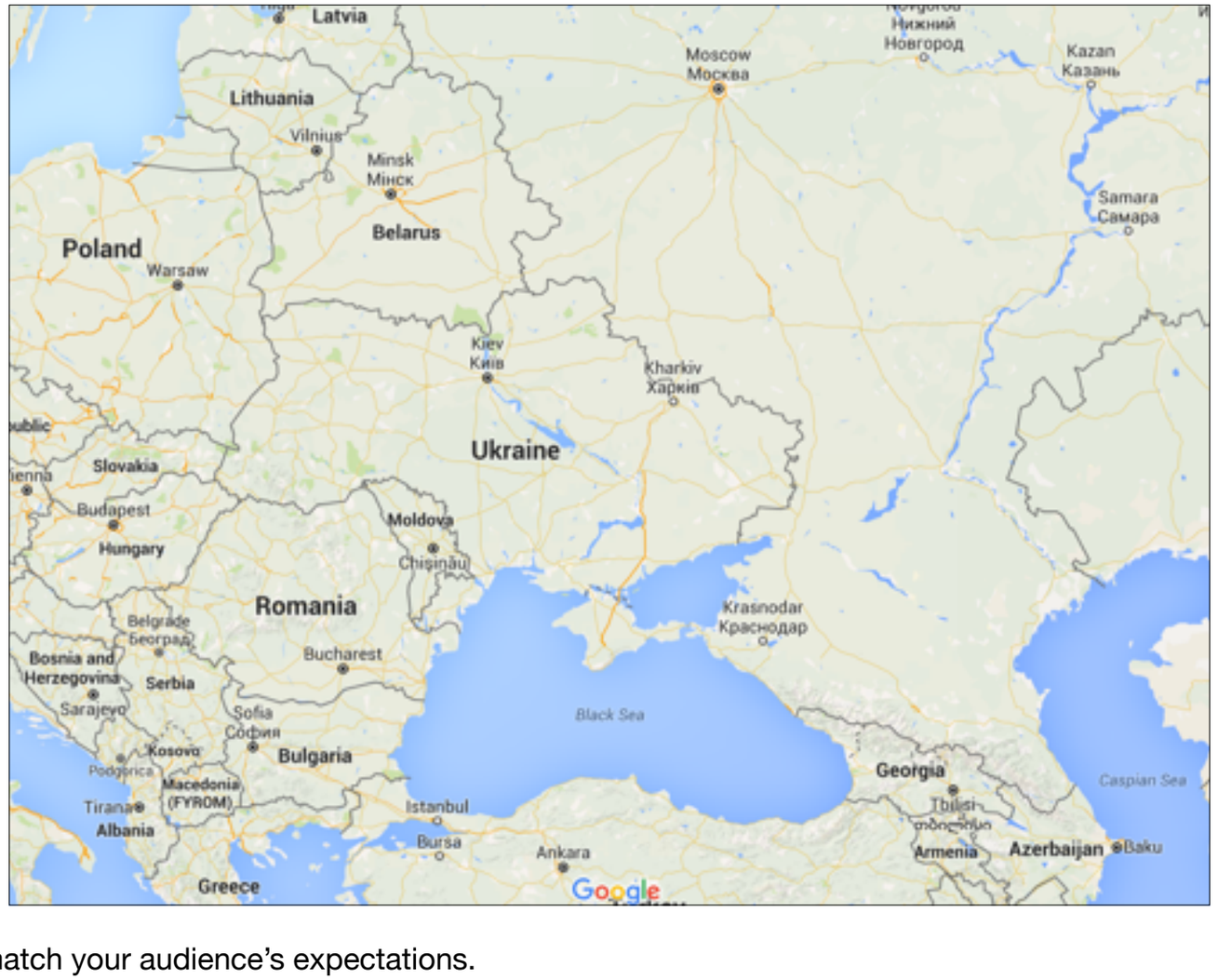


Hexagonal binning

When you want to plot point, but have a lot of data, hexagonal binning can do a better job of explaining things.

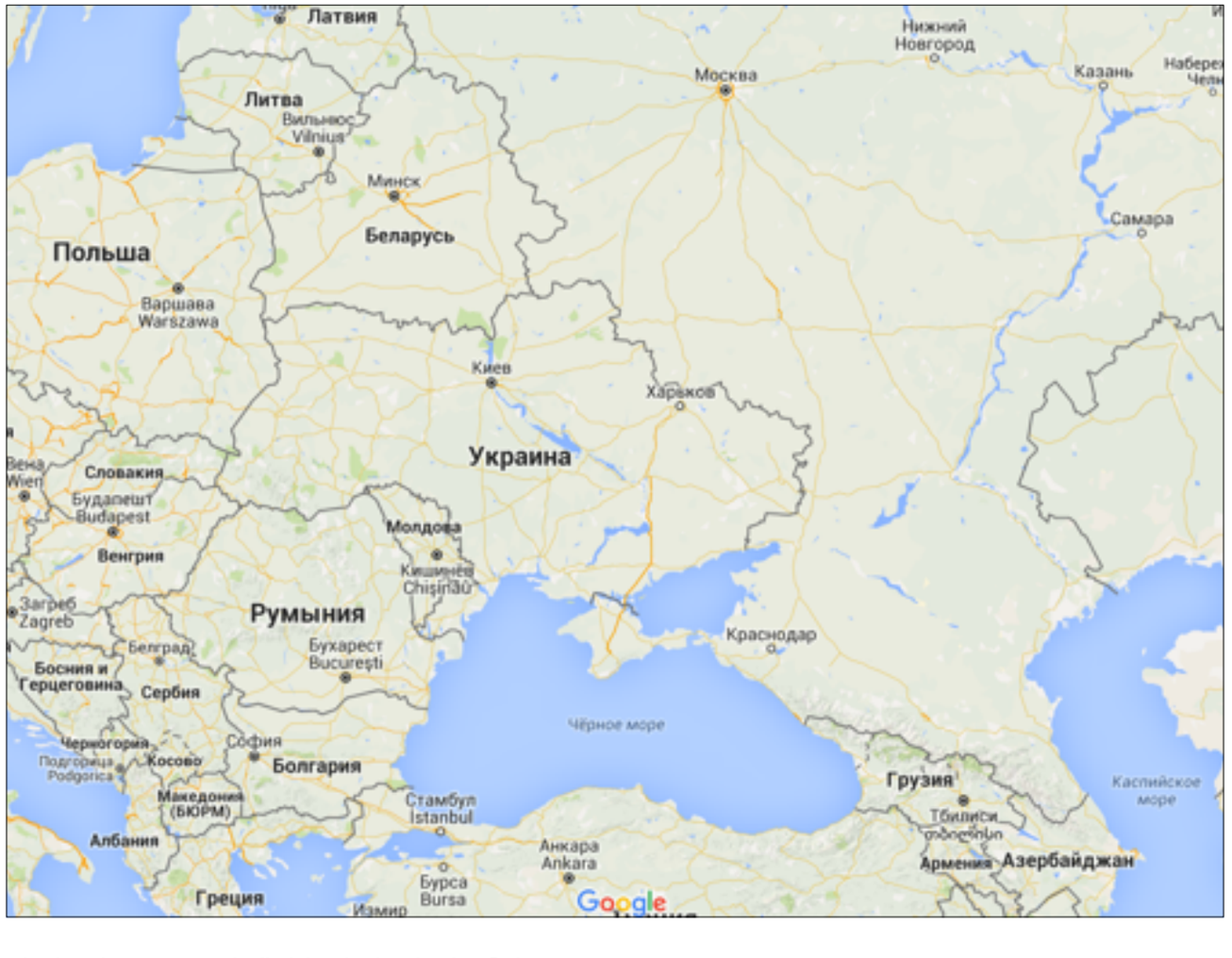
Binning means counting the number of points that fall within each hexagon (or rectangle) on a gridded surface, and using a sequential colour scheme to display the counts.

https://www.mapbox.com/blog/binning-alternative-point-maps/



You may need to localise your map to match your audience's expectations.

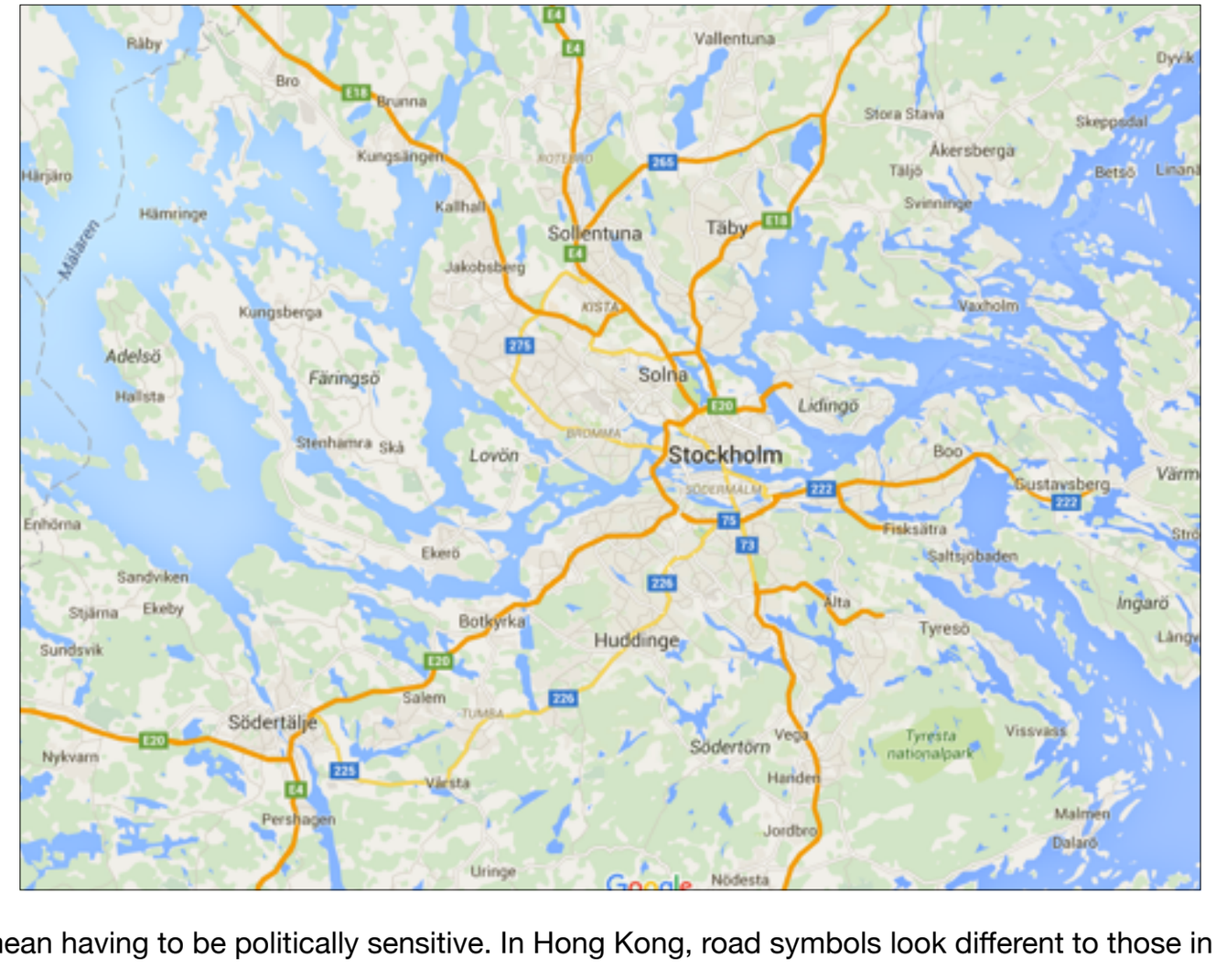
On the google.com world map, Crimea is shown as a disputed area (the border is dotted).



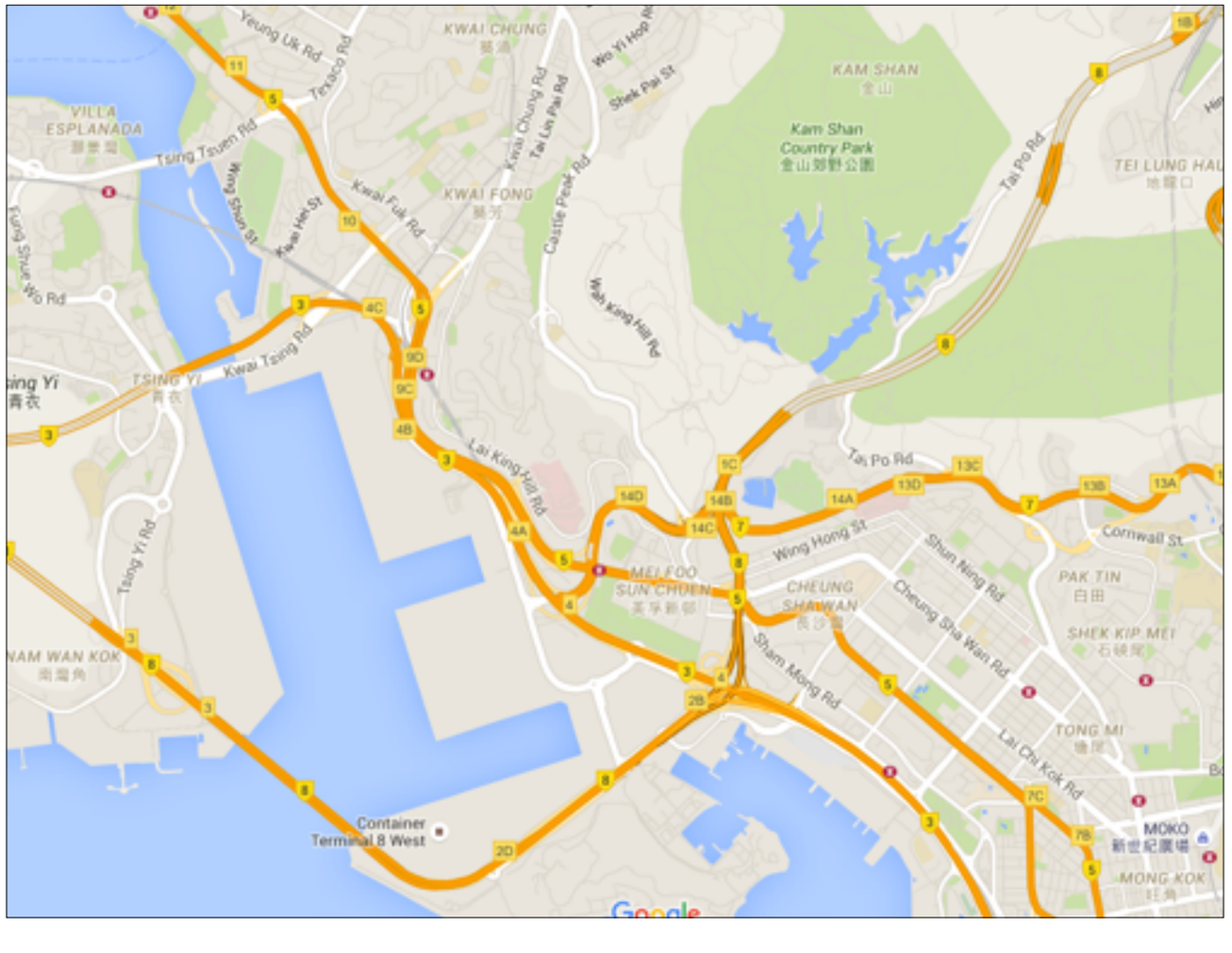
On the google.ru map, the border of Russia is shown to definitively include Crimea.

On google.in, India’s borders are shown definitively.

On google.com, the status of the north-west is shown to be very much less decided.



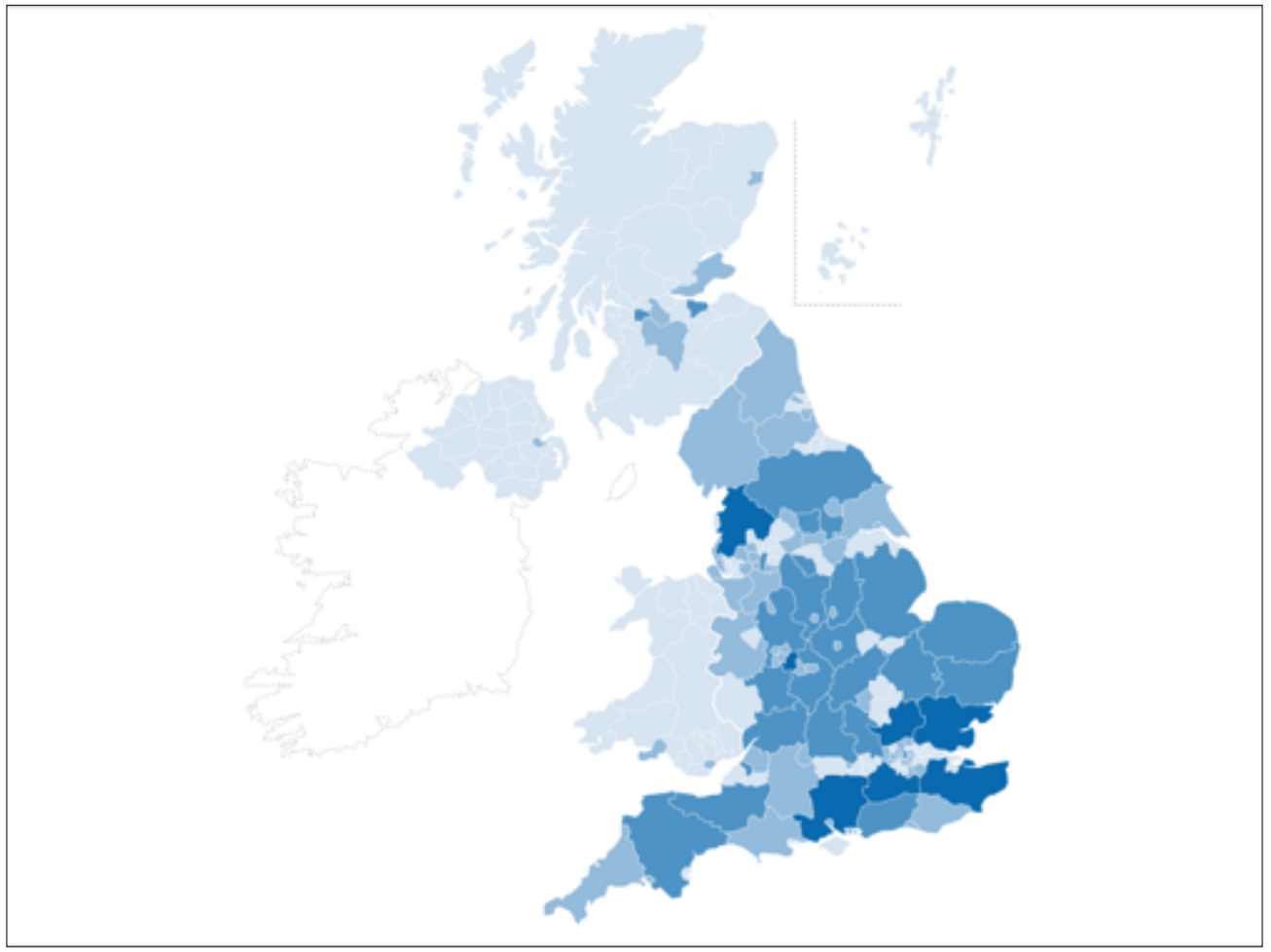
Localising a map doesn’t necessarily mean having to be politically sensitive. In Hong Kong, road symbols look different to those in Sweden.



Language is also important. Both Chinese and English appear in Hong Kong, but not in Sweden. The Russian map is also interesting: Cyrillic for Russia but Cyrillic and Latin for countries that use the Latin alphabet. There is lots to consider.

A country might be made up distant parts. The US is a famous example, with Hawaii and Alaska being so far from the rest of the country that they're often not shown in their real locations.

(Look at the border between the US and Canada there. The world's longest straight border is between those two countries, and yet on this National Geographic map it's shown as curved. On a Mercator map it's shown straight. That's an interesting difference between two projections.)



Maps of the UK do the same thing, moving the islands to the north of Scotland to reduce the size of the map that would be needed otherwise.

Notice also that although no other countries are shown on this map (the coast of France should be visible at the bottom right) but the dotted outline of Ireland is still required to stop Northern Ireland looking like an island.

Data sources

- Natural Earth
- OpenStreetMap
    - Overpass Turbo API
    - MapZen
- Google search (for shapefiles, GeoJSON, etc)
- National & local governments (open data sites)

http://www.naturalearthdata.com/
http://www.openstreetmap.org/
http://overpass-turbo.eu/
https://mapzen.com/data/metro-extracts

Tools

- MapBox Studio
- MapBox Studio Classic
- CartoDB
- QGIS
- ArcGIS
- D3
- Leaflet

https://www.mapbox.com/studio
https://www.mapbox.com/mapbox-studio-classic/
https://cartodb.com/
http://www.qgis.org/
https://www.arcgis.com/
http://d3js.org/
http://leafletjs.com/