HDAT9800 Visualisation and Communication of Health Data

Chapter 7

Mapping

We have looked at how *leaflet* makes map-based visualisations easy

Let’s have a look at some different types of maps

Different map types

Common types

- Choropleths
- Isopleths
- Proportional symbols
- Cartograms
- Heat maps

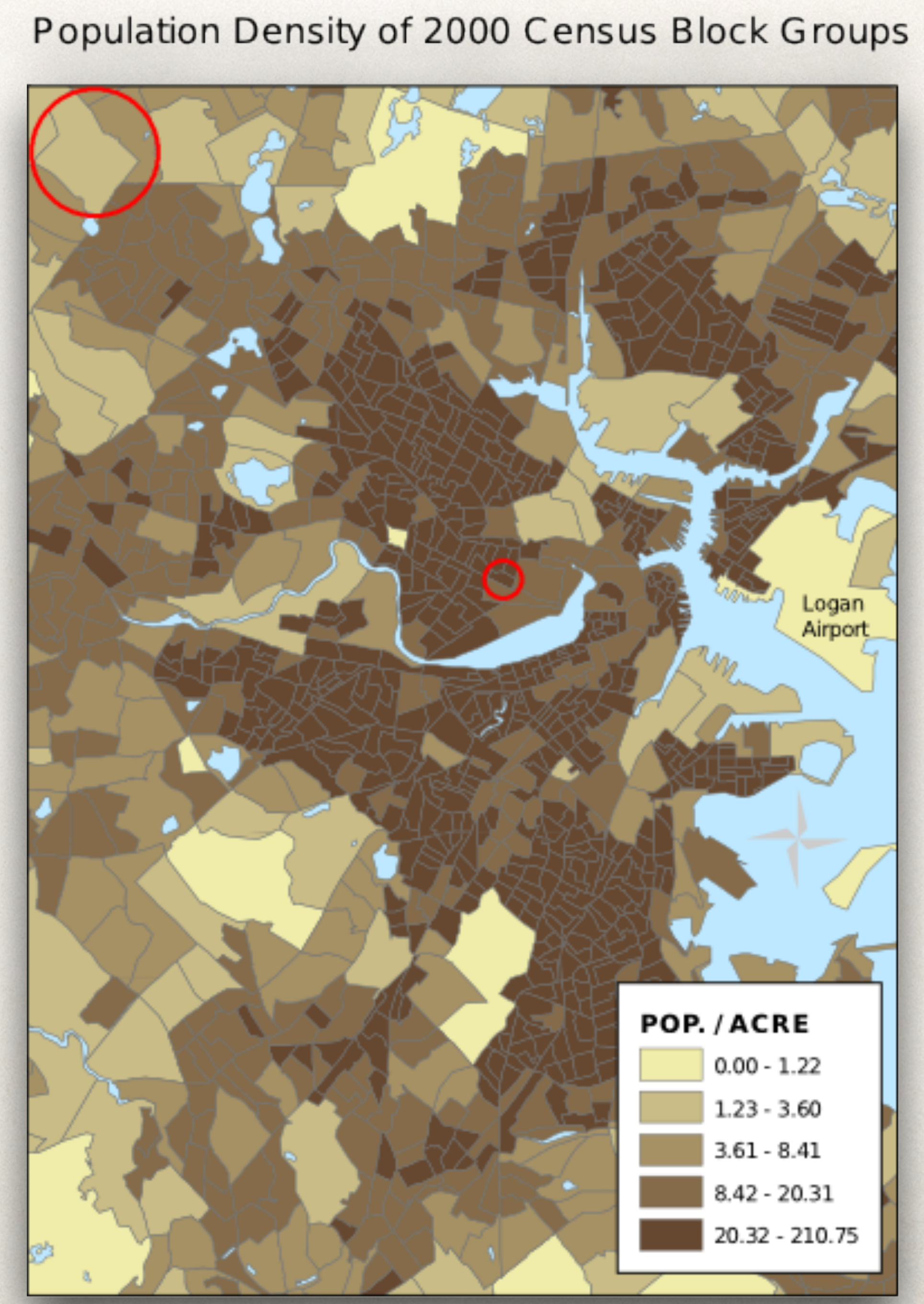
These are not the only ones

Choropleths

Choropleths are maps with areas shaded according to a key

Each shading or colour type represents a range or category

Choropleths



Choropleth drawbacks

Choropleths can give a false indication of abrupt changes at a boundary

Choropleths are often not suitable for showing total values

- symbol/mini-graph overlays can help with this

It can be difficult to distinguish between different shades

Variations within map units are hidden

Uniform colour regions hide reliability of estimates

- area of unit not proportional to number of cases

Choropleth caveats

Values should be normalised to allow comparison

Progressive colour gradients should be used with quantitative data

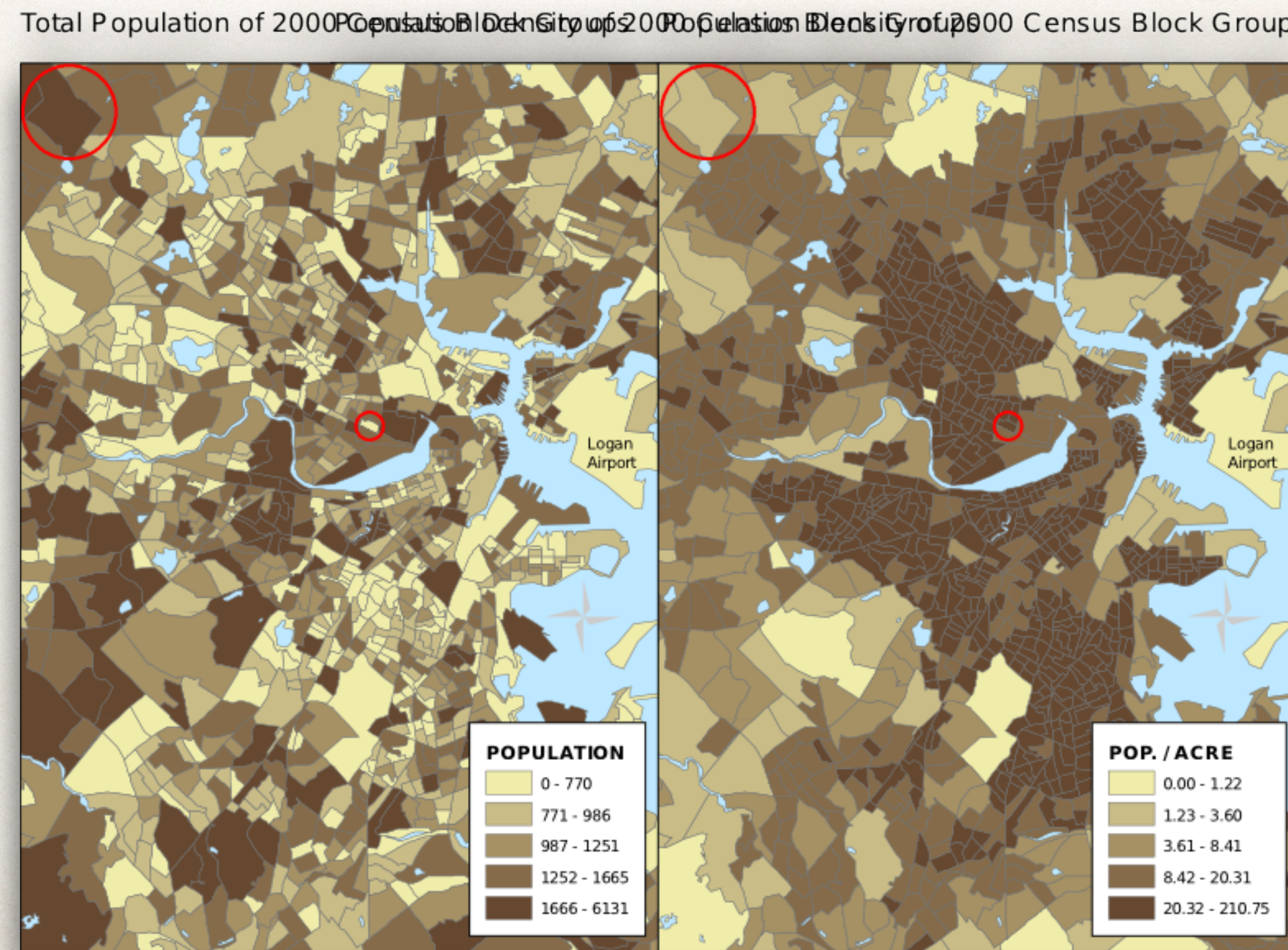
Bipolar (two-colour) gradients should be used with ± data

Distinct colours should be used with categorical data

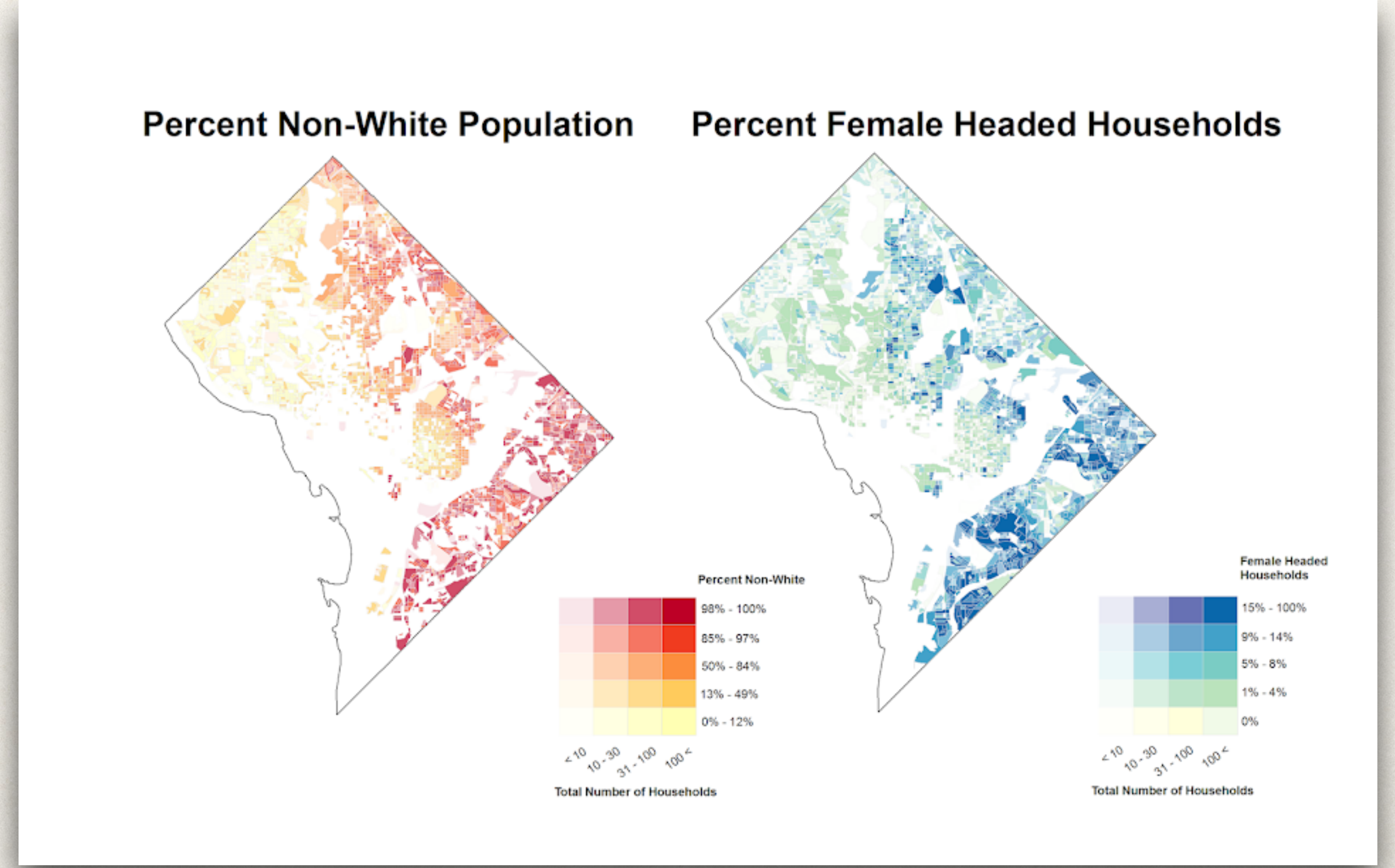
As with all visualisations be aware of colour-blind users

Smaller units show more detail than larger units

Unnormalised choropleths



Using the alpha channel



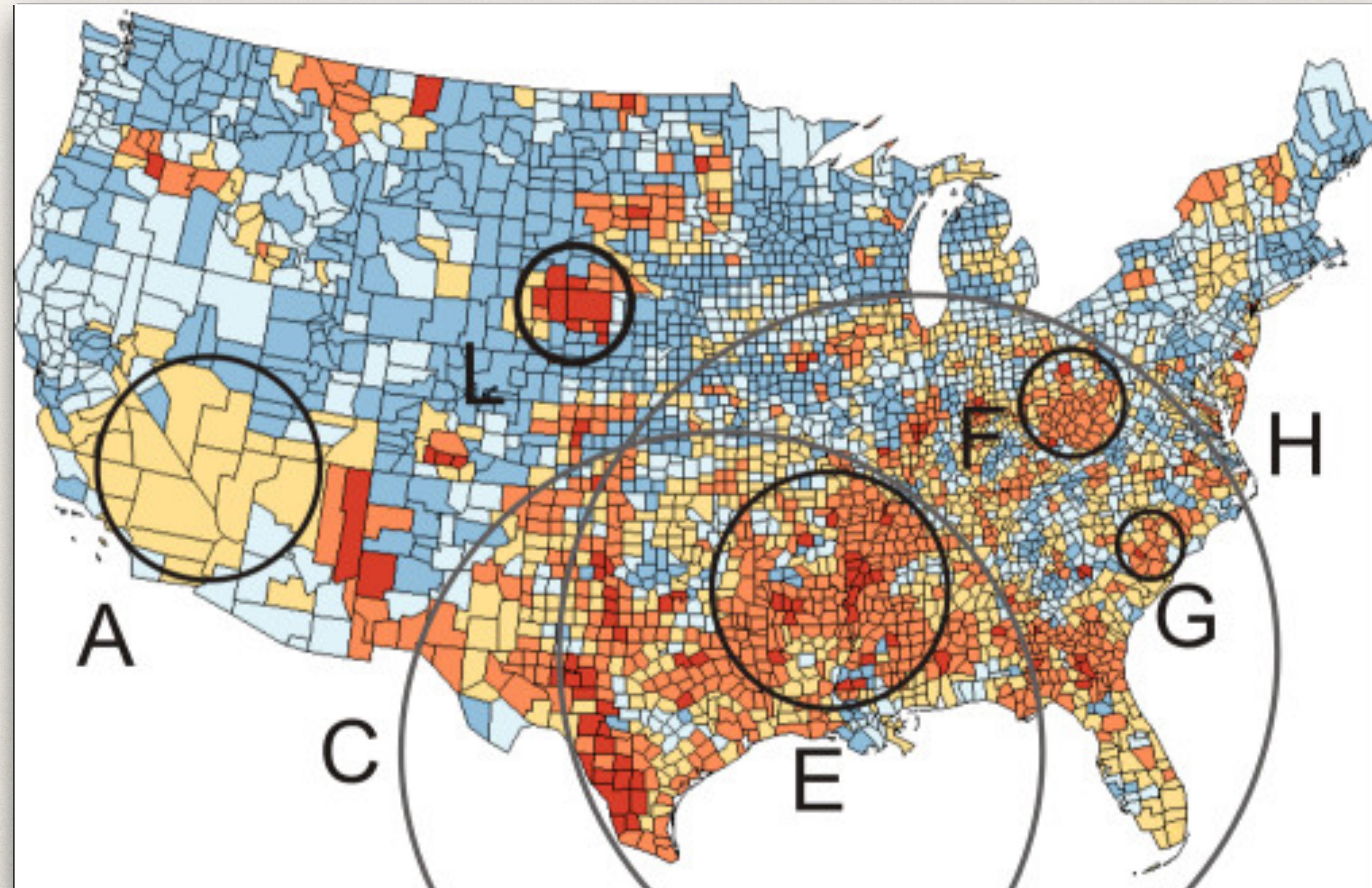
Empirical Bayes smoothing

Empirical Bayes Smoothing uses the population in a region as a measure of the confidence in the data

Higher populations in a given area lend a higher confidence to the estimated number of events in that location.

Empirical Bayesian Smoothing leaves estimates for areas with low margins of error alone, but nudges estimates in regions with high margins of error closer to the global average of the event rate.

Empirical Bayes smoothing



Modifiable unit area problem

A statistical bias that can significantly impact the results of statistical hypothesis tests

MAUP affects results when point-based measures of spatial phenomena are aggregated into districts, *e.g.*, population density or illness rates

The resulting summary values (*e.g.*, totals, rates, proportions, densities) are influenced by both the shape and scale of the aggregation

Data aggregation can obscure a strong correlation between variables, making the relationship appear weak or even negative.

MAUP can cause random variables to appear as if there is a significant association where there is not

MUAP: scale (aggregation)

Aggregation data at different levels gives different results

The cancer rate for all of NSW is different to that of individual regions and different yet again from those for all of Australia

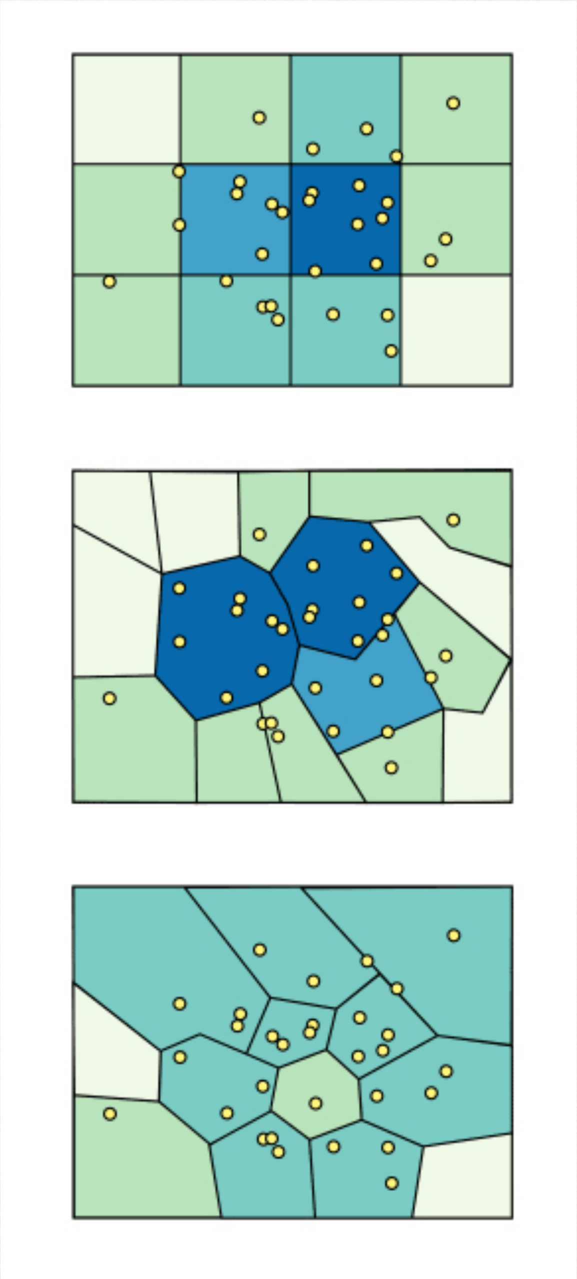
The level to which we aggregate affects results

MUAP: zone (grouping)

How one groups data can affect interpretation

Here we see the effects of different boundary choices

The same underlying data gives rise to three very different pictures



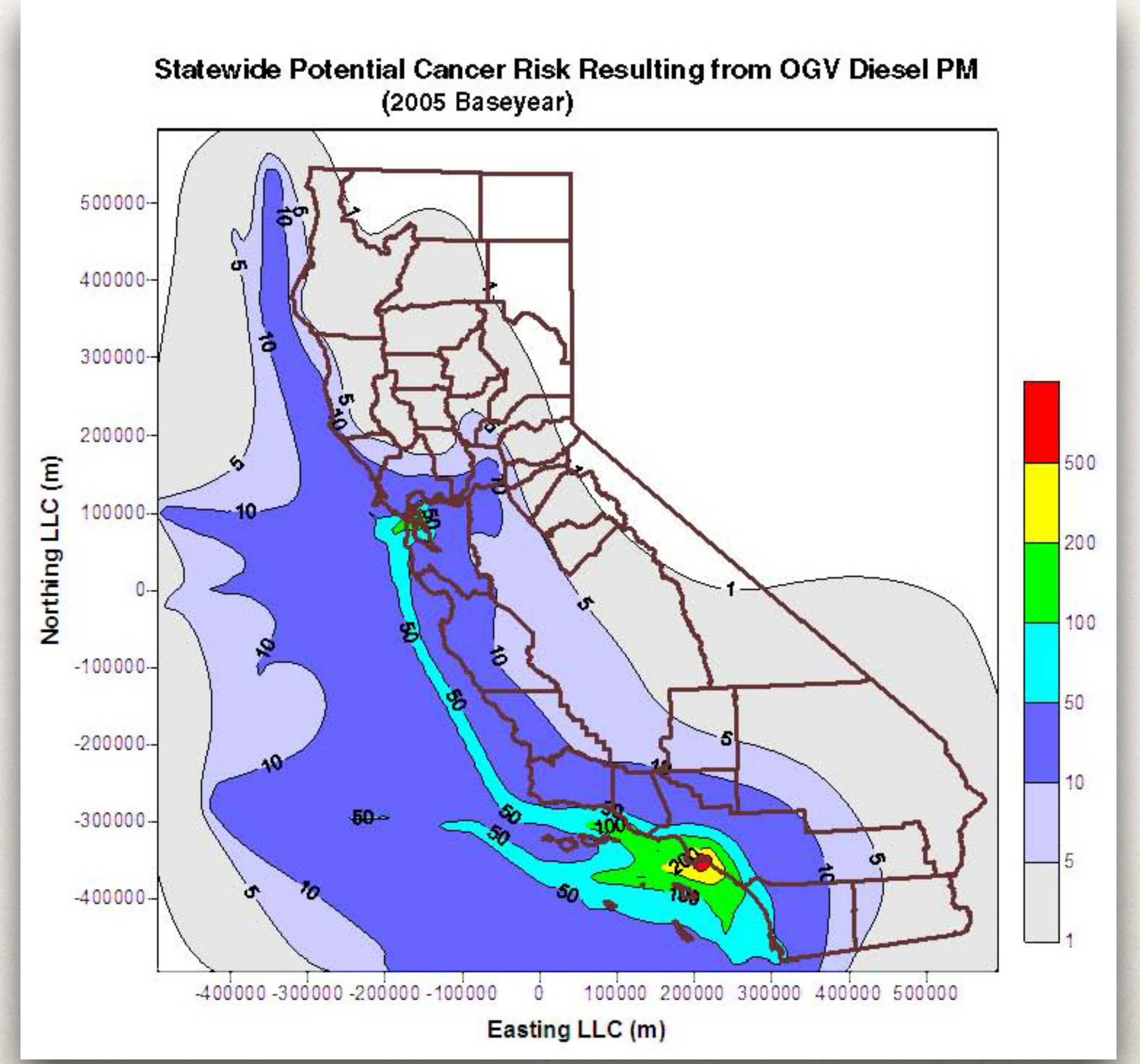
Isopleths

Isopleths are maps with lines of equal value drawn (possibly filled)

The lines on some common isopleths have specific names

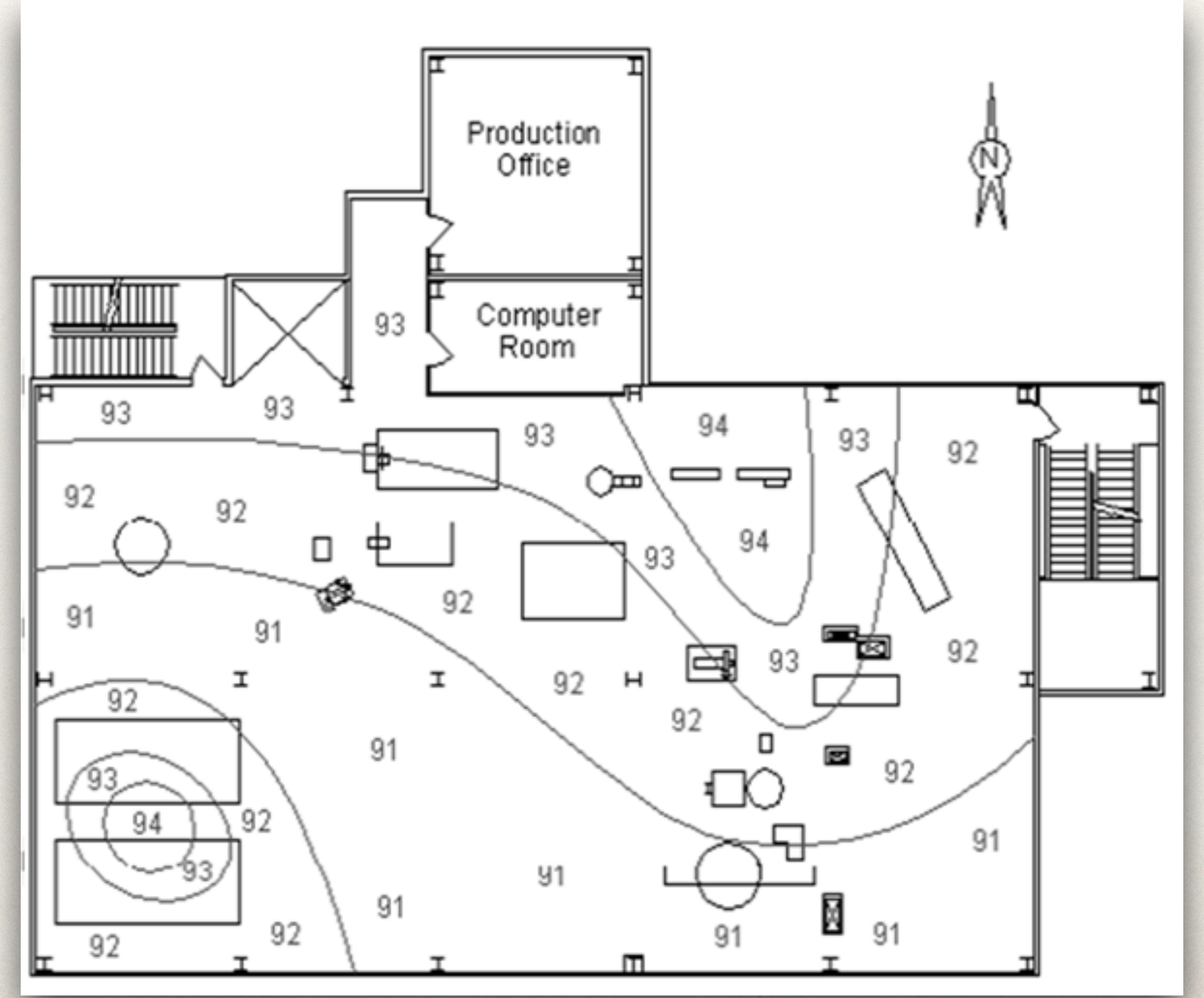
- isobars on a weather (barometric) map showing pressure
- contour lines on a topographic map showing equal height
- isobaths on a bathymetric map showing places of equal depth

Isopleths



Statewide Potential Cancer Risk Resulting from OGV Diesel PM
(2005 Baseyear)
Northing LLC (m)
Easting LLC (m)
500000
400000
300000
200000
100000
0
-100000
-200000
-300000
-400000
-400000 -300000 -200000 -100000 0 100000 200000 300000 400000 500000
500
200
100
50
10
5
1

Isopleths


Production Office
Computer Room
N
93
93
93
93
94
93
92
92
92
92
94
93
91
91
92
92
93
92
92
92
91
93
94
92
93
92
91
91
91
92
92
91
91

Isopleths drawbacks

Isopleths are unsuitable for showing discontinuous or 'patchy' distributions

A large amount of data are needed for accurate representation

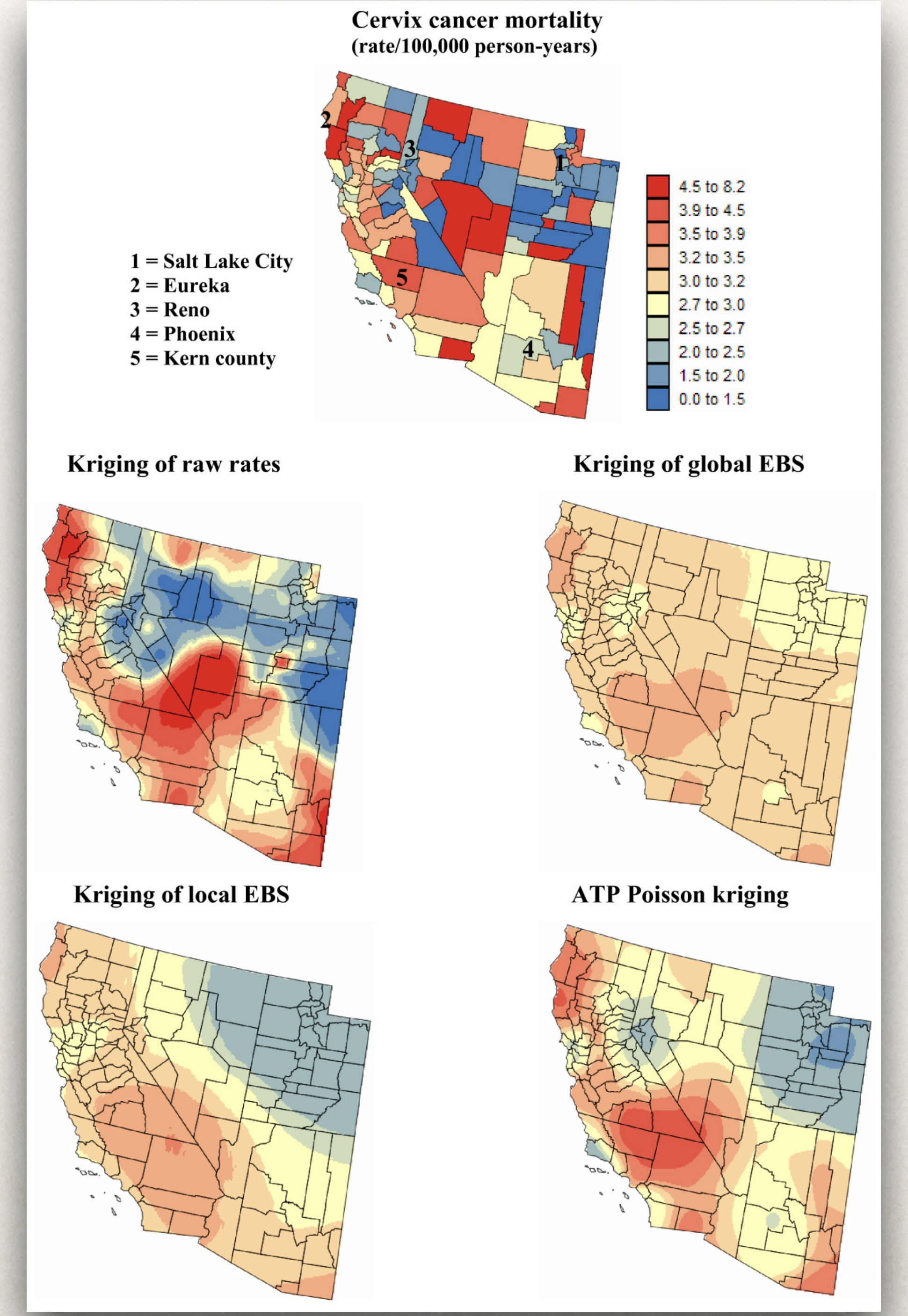
Isopleths caveats

This type of map is ideal for showing gradual change over space and avoids the abrupt changes which boundary lines produce on choropleth maps

Temperature, for example, is a phenomenon that should be mapped using isoplething, since temperature exists at every point (is continuous), yet does not change abruptly at any point

Relief maps should always be in isopleth form for this reason.

Isopleths caveats

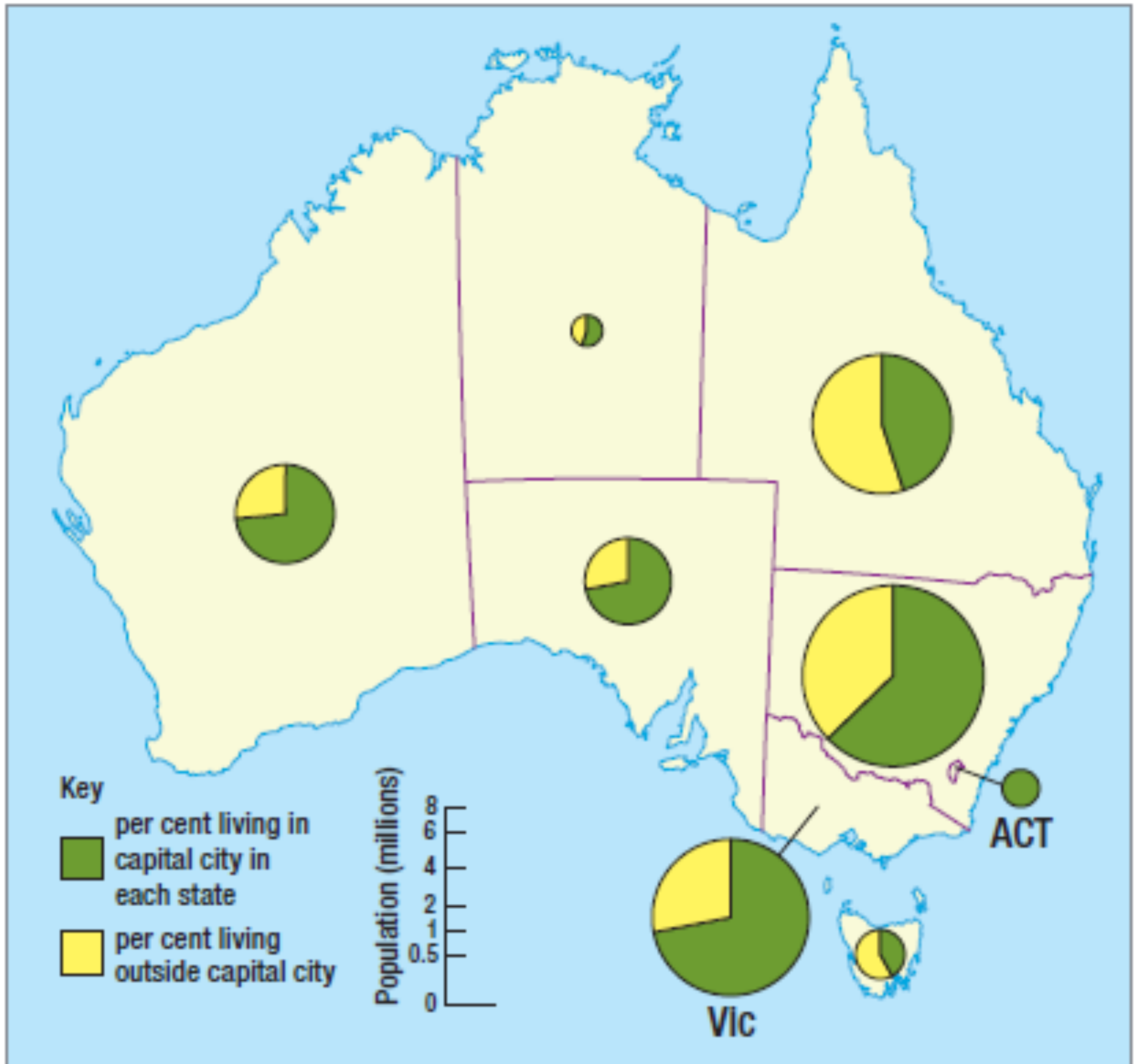


Proportional symbols

As the name implies, symbols (often circles) are drawn proportional in size to the size of the variable

Proportional symbol maps are not dependent on the size of the area associated with the variable.

Proportional symbols



10.19 The proportion of the population of each of the Australian states living in the capital city in that state, in 2006—an example of divided proportional circles

Proportional symbol drawbacks

It can be hard to judge the different sizes of the symbols

Too many symbols can result in a confusing mess

Proportional symbol caveats

Much research has gone into the optimal scaling for proportional symbols.

Make sure that the *area*, rather than linear proportions like radius or length of a side, is the scaled parameter.

- If there are four times as many businesses in Site 1 than in Site 3, the area of the symbol should be four times greater for Site 1
- If the symbol choice is a circle, the radius of the Site 1 symbol should thus be only twice as great (since area scales with the square of the radius)

Symbol size should be proportional to the underlying denominator not the value (areas with more people should be bigger)

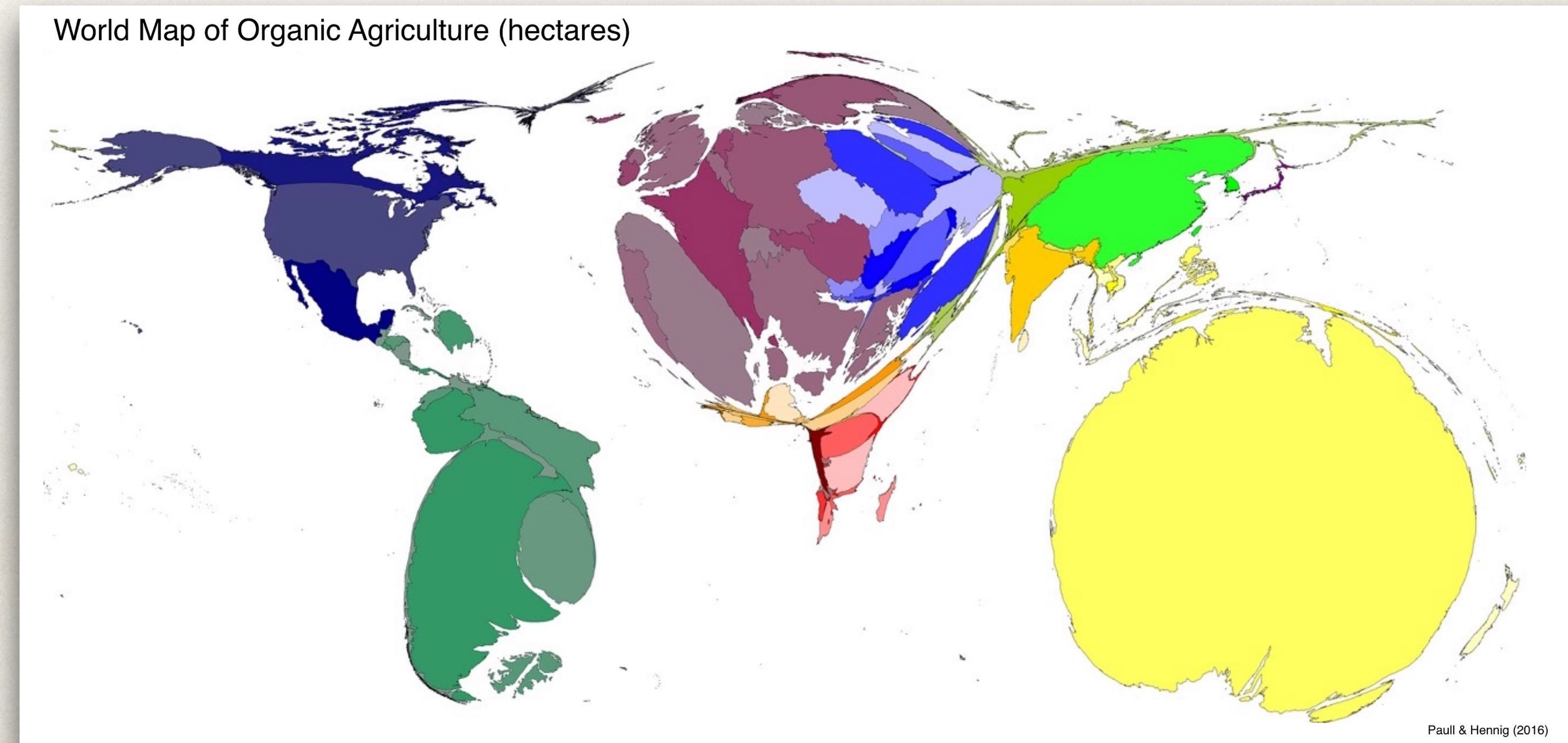
Cartograms

A *cartogram* is a map in which some thematic mapping variable, such as travel time, population, or GNP, is substituted for land area or distance

The geometry or space of the map is distorted, sometimes extremely, in order to convey the information of this alternate variable

They are primarily used to display emphasis

Cartograms


World Map of Organic Agriculture (hectares)
Paull & Hennig (2016)

Heat maps

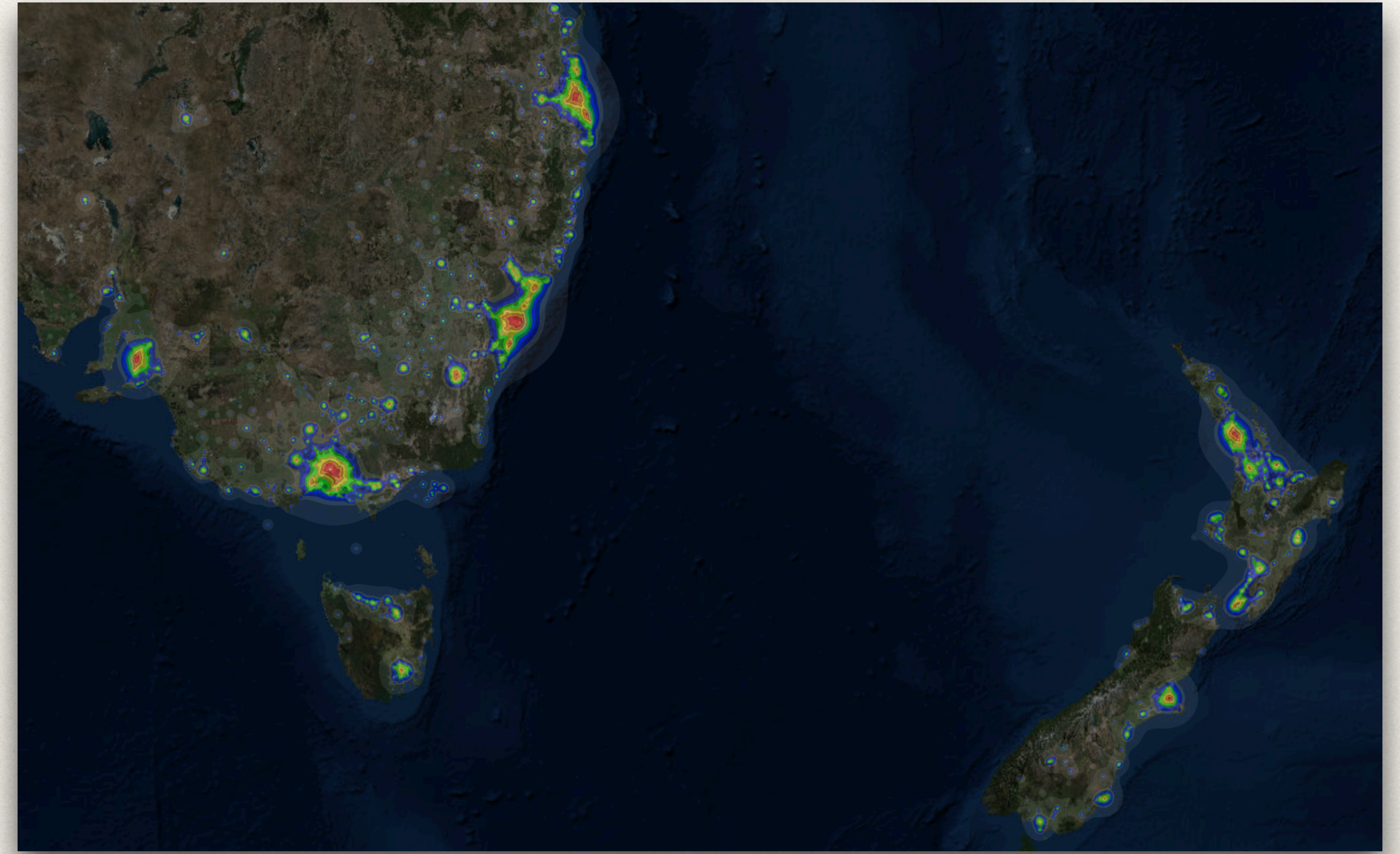
Heat maps originated in 2D displays of the values in a data matrix

- larger values were represented by small dark grey or black pixels
- smaller values by lighter pixels

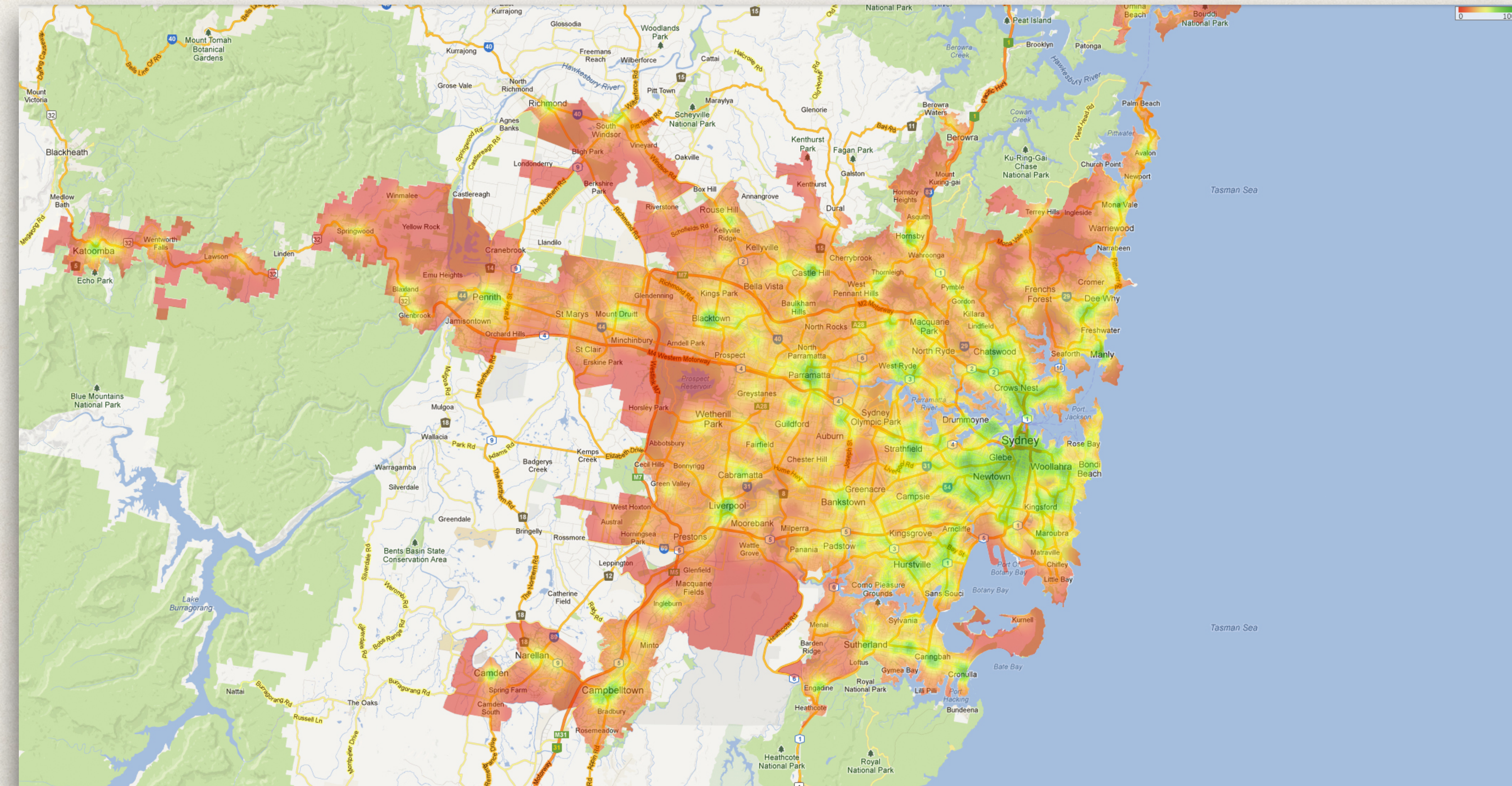
Different colour schemes can be used to illustrate a heat map, with perceptual advantages and disadvantages for each

- rainbow colour maps can be used but these have drawbacks

Heat maps



Heat maps


Tasman Sea
Sydney
Parramatta
Penrith
Blacktown
Liverpool
Campbelltown
Camden
Narellan
Katoomba
Richmond
Bankstown
Sutherland
Hornsby
Chatswood
Manly
Newtown
Glebe
Woollahra
Bondi Beach
Blue Mountains National Park
0
100

Heat map drawbacks and caveats

The use of rainbow maps is discouraged by many in the scientific community, for the following reasons:

- they lack the natural perceptual ordering found in grayscale or blackbody spectra
- some colormaps have uncontrolled changes in luminance that prevent meaningful conversion to grayscale for display or printing
- luminance changes distract from the actual data, arbitrarily making yellow and cyan regions appear more prominent than the regions of the data that are actually most important.
- colour changes lead to perception of gradients that aren't actually present, making actual gradients less prominent, meaning that rainbow colormaps can actually obscure detail in many cases rather than enhancing it

The group assignment

Need to be in groups of three or four

The assignment will be larger than the weekly assignments but not impossibly huge

You will have plenty of time to complete it

Must be done collaboratively using GitHub and GitHub Classroom