

# Introduction to GIS

Raphaëlle ROFFO

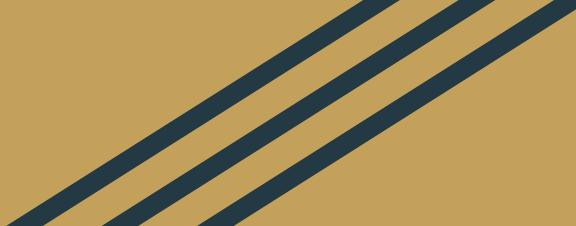
Sciences Po - GETEC Masters  
Fall Semester 2020

# Session 4

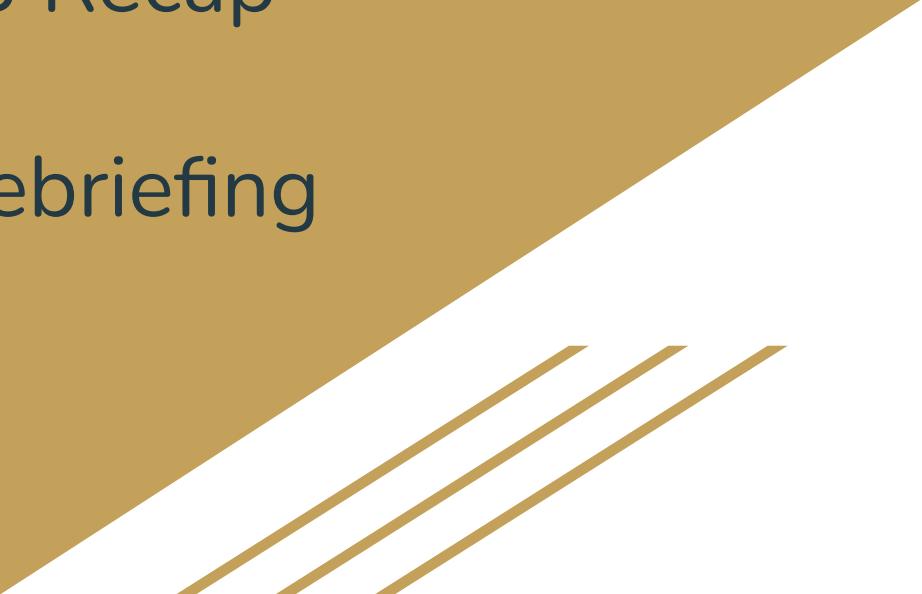
## Cartographic Design

# Today's plan

1. Tutorial debriefing
2. Cartographic Design Principles
3. Accessibility
4. Map layouts
5. Choropleths
6. Coursework & homework



# Session 3 Recap & Tutorial Debriefing



# Session 3 learning objectives

After completing Session 3 & tutorial you should now be comfortable with:

- **Connecting to basemaps** using XYZ tiles
- Accessing **summary statistics** and basic information about a layer
- How **records and features** are related, and how to access them
- Creating a **point layer from a \*.csv file**
- **Joining two tables** using a common attribute field
- **Creating a query** to access features based on their attributes
- Creating a **definition query** to only display a subset of a given dataset

# Remarks

- **QGIS syntax:** this is the hardest part of working with queries and expressions in QGIS. As always, [refer to the documentation](#) to explore all the possibilities. [Section 14.2.1.2. Writing an expression](#) is especially useful and will likely help you debug most of your expression issues.
- **Table joins:** to carry out a table join, the key column you're using to perform your join working with **must be of the same data type** (typically, string). For a computer, the string `123` is different from the integer number **123**. You can use the *Refactor Fields* tool or the *Field calculator* in the Fields tab of your *Layer Properties* to change a field's data type.

# Data types in QGIS

- Integer (whole numbers)
- Real number (decimal numbers)
- String (text)
- Date / time
- Binary / boolean (1 or 0)

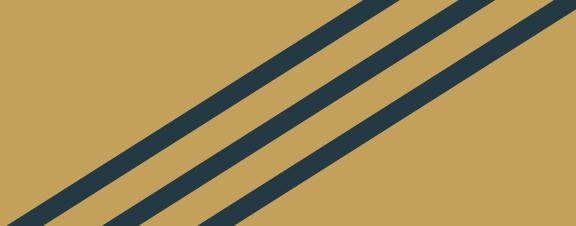
The screenshot shows the QGIS 'Layer Properties' dialog for a layer named 'Session3-London LondonBoroughs'. The 'Fields' tab is selected, displaying a table of existing fields:

ID	Name	Alias	Type	Type name	Length	Precision	Comment	Configuration
123	0 fid		qulonglong	Integer64	0	0		
123	1 objectid		int	Integer	0	0		
abc	2 name		QString	String	22	0		
abc	3 gss_code		QString	String	9	0		
1.2	4 hectares		double	Real	0	0		
1.2	5 nonld_a							
abc	6 ons_inn							
abc	7 sub_20							

An 'Add Field' dialog is open, prompting for a field name ('Name') and type ('Type'). The 'Type' dropdown menu is expanded, showing the following options:

- Whole number (integer)
- Whole number (integer 64 bit)
- Decimal number (real)
- Text (string)
- JSON (string)
- Date
- Date & Time
- Binary object (BLOB)
- Boolean

Questions?



# Cartographic Design Principles (a non-exhaustive list)

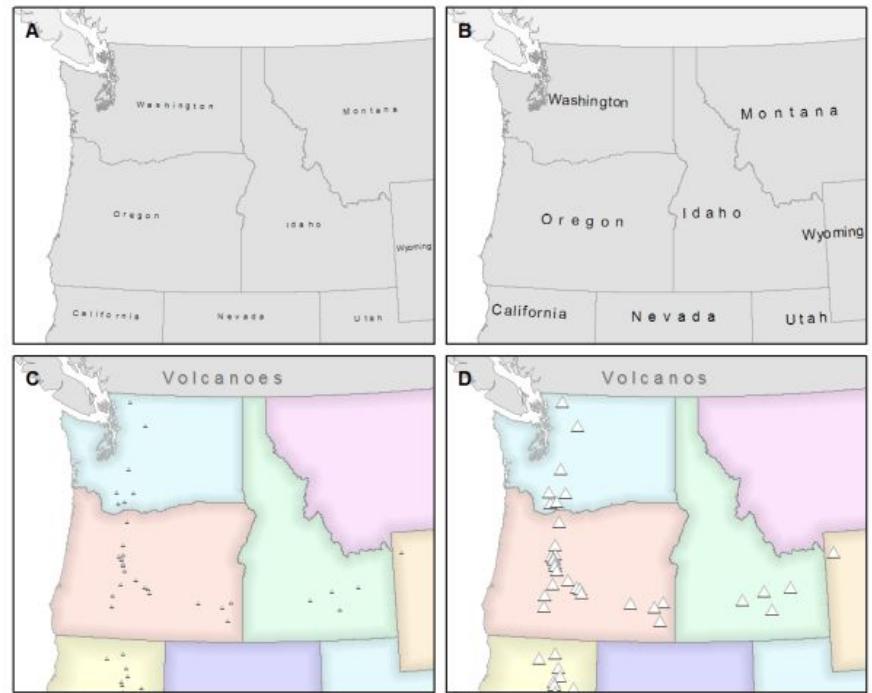


# Cartographic Design Principles

## 1. Legibility

Legibility is “the ability to be seen and understood”. Making a **legible map** depends on you selecting appropriate symbols and sizes so that anyone can easily read and understand the various elements of your map.

If you must include small elements, bear in mind that simple symbols (dots, triangles etc) and minimalist fonts are easier to read than more complex symbols.



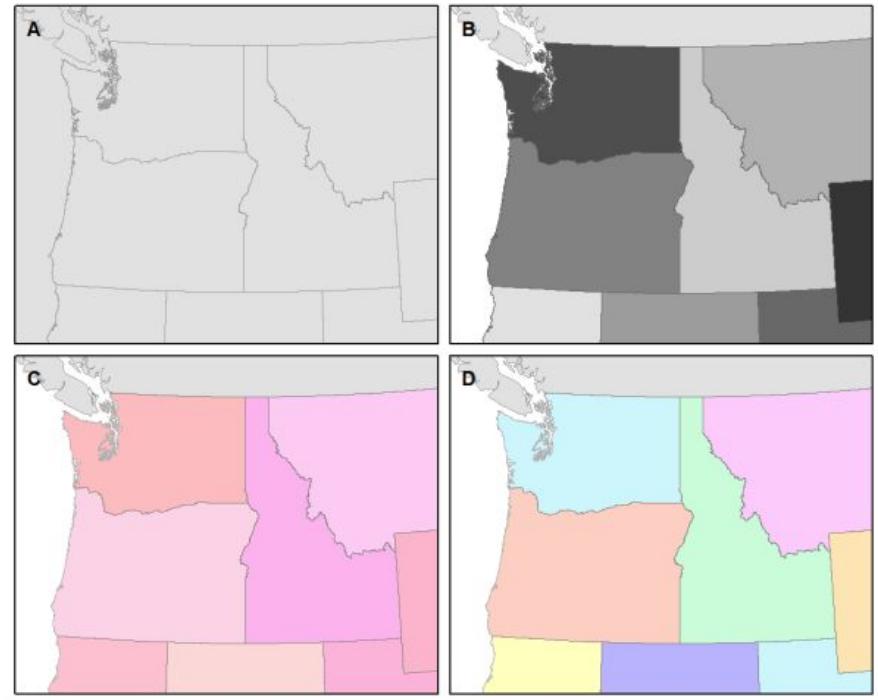
Source: [ESRI](#)

# Cartographic Design Principles

## 2. Visual Contrast

When there is little **contrast between your features**, they appear to belong together. To make something *stand out* in your map, you need to use *high contrast* between features. Usually the darker/brighter feature will automatically attract the eye.

If you're trying to convey a more subtle impression, you can *choose* to use low visual contrast.

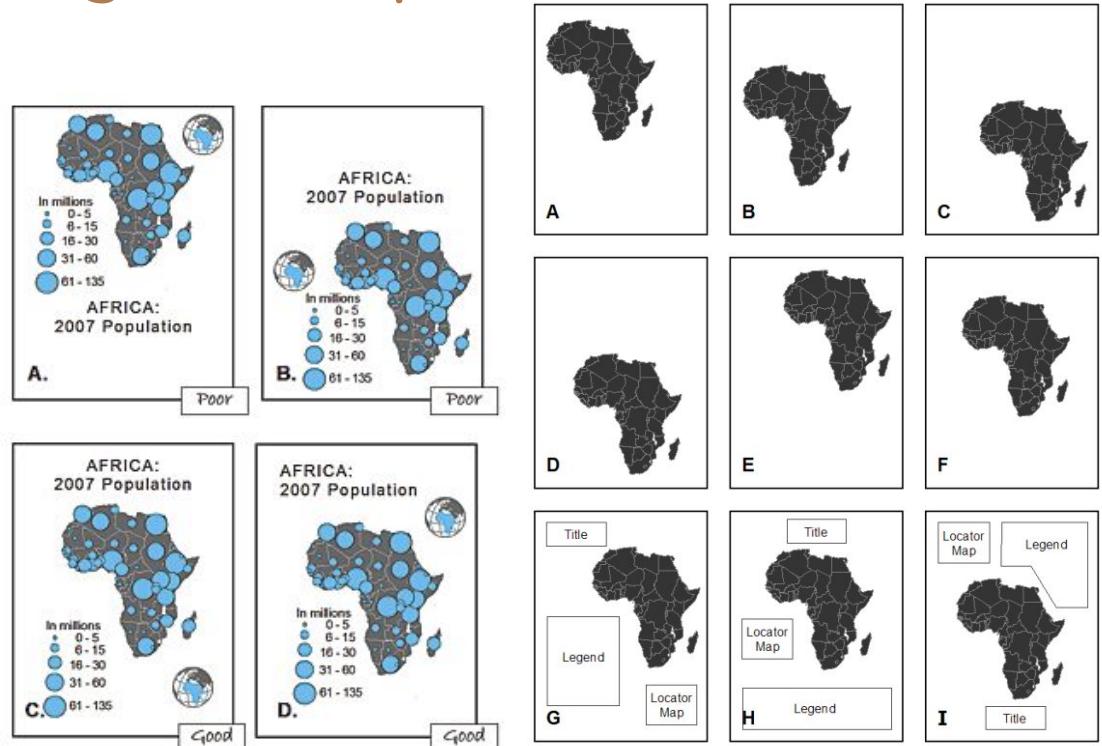


Source: [ESRI](#)

# Cartographic Design Principles

## 3. Balance

When you look at a map, you don't want to get a "tilted" feeling. Visual weight and visual direction are two key elements in making a **balanced** map. You want the various elements in your map to be relatively centered and not crammed in one side of your layout.

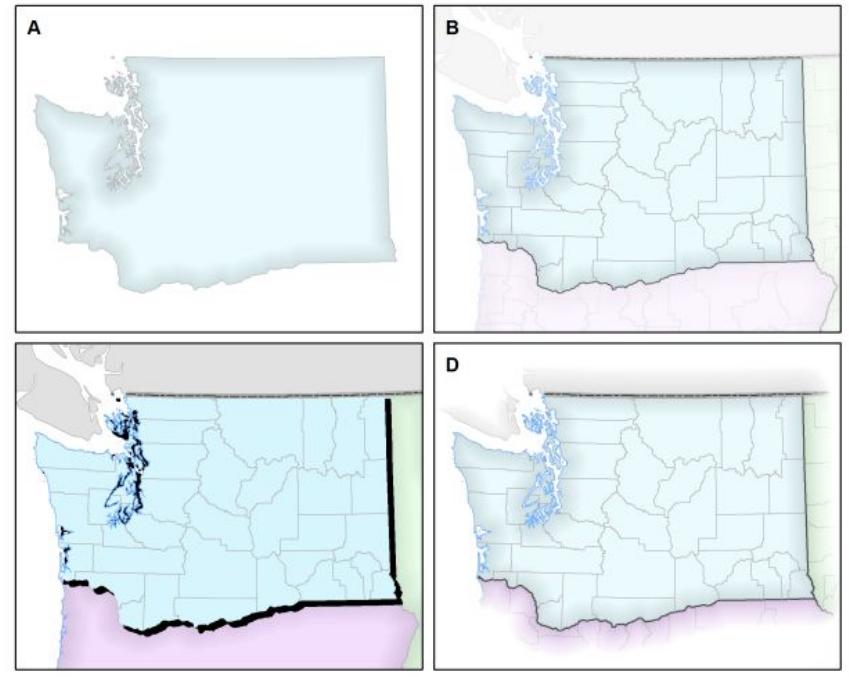


Source: [ESRI](#)

# Cartographic Design Principles

## 4. Figure-ground

When your audience reads your map, how easy is it for them to figure out the elements they should focus on? **Figure-ground organization** refers to how elements in the foreground and background can be distinguished and spontaneously separated. If you're focusing on one administrative region, how can you make sure your audience knows right away which region to look at and which to ignore? White wash, drop shadow or feathering are usually helpful in improving figure ground organization.



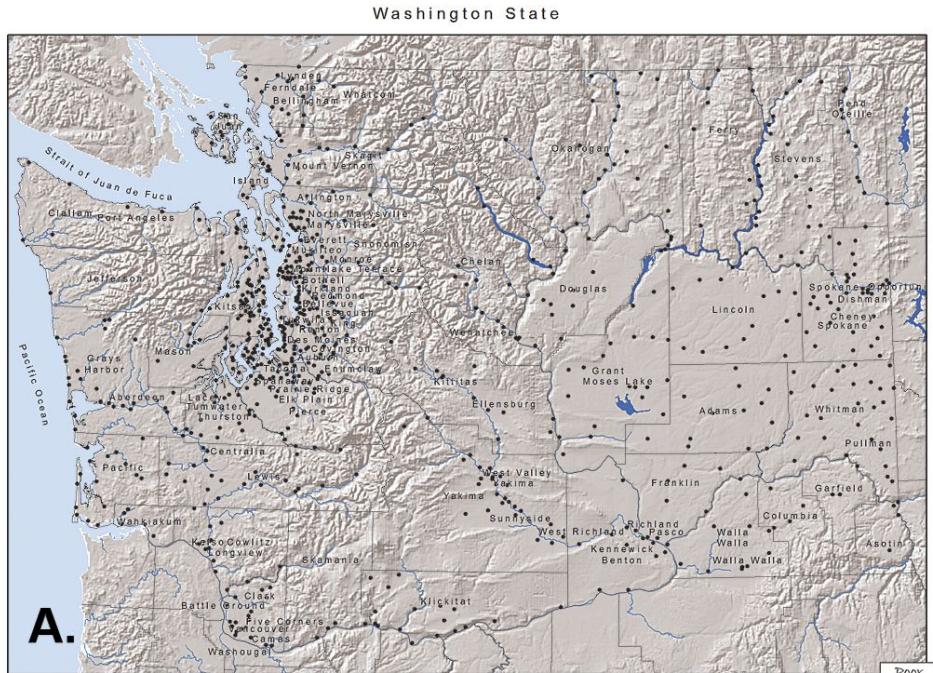
Source: [ESRI](#)

# Cartographic Design Principles

## 5. Hierarchical Organisation

Your map usually contains multiple layers of information, of which some are more important than others (e.g. your basemap only gives you background, not key information).

Your map should be the centre of attention, with a title and legend that must be easy to read without squinting. It should be obvious to the reader what they should focus on and what patterns emerge.



Source: [ESRI](#)

# Cartographic Design Principles

## 5. Hierarchical Organisation

In a general reference map, you can play with the boundary line styles, line widths, and the contrast between your basemap and the areas you're referencing to ensure your readers know which region to focus on and which to ignore / only use as context.

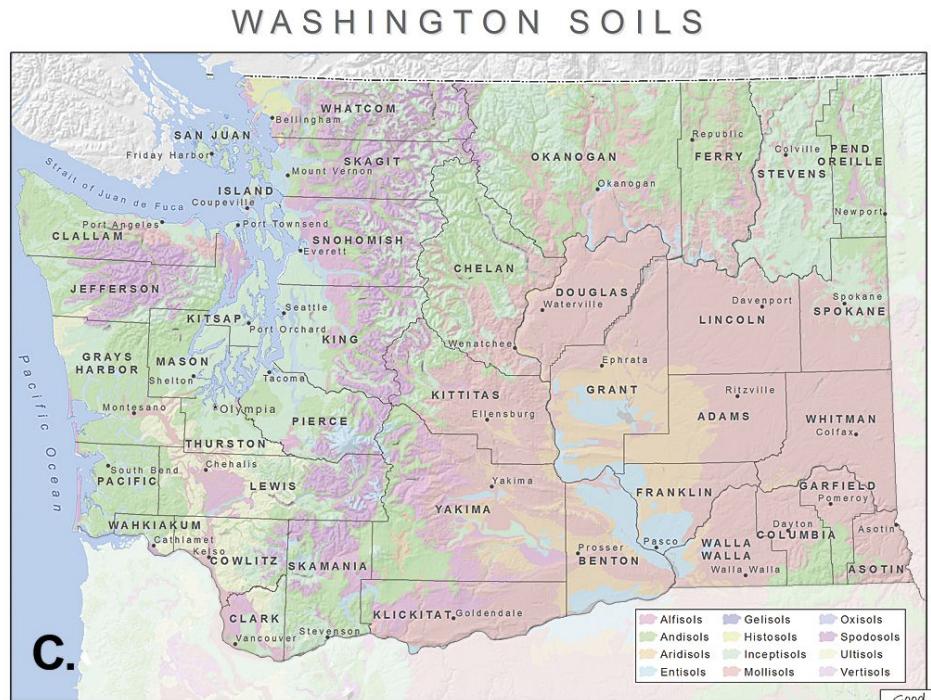


Source: [ESRI](#)

# Cartographic Design Principles

## 5. Hierarchical Organisation

When you are mapping thematic data, keep the basemap information to a minimum so that the actual theme (in this example: the nature of soils) is at the highest visual level in the hierarchy.



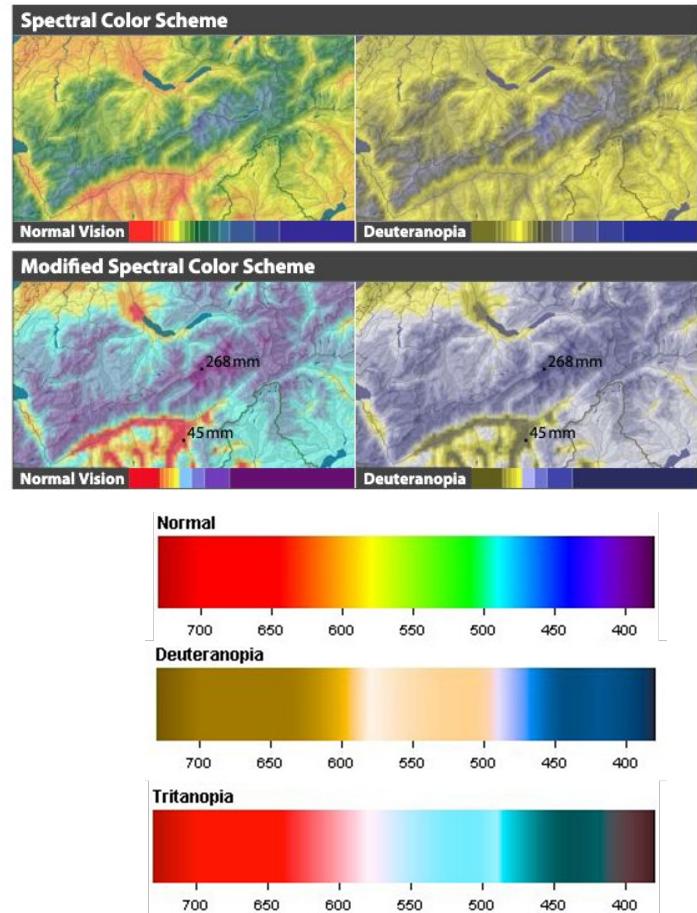
Source: [ESRI](#)

# Accessibility

# Accessibility

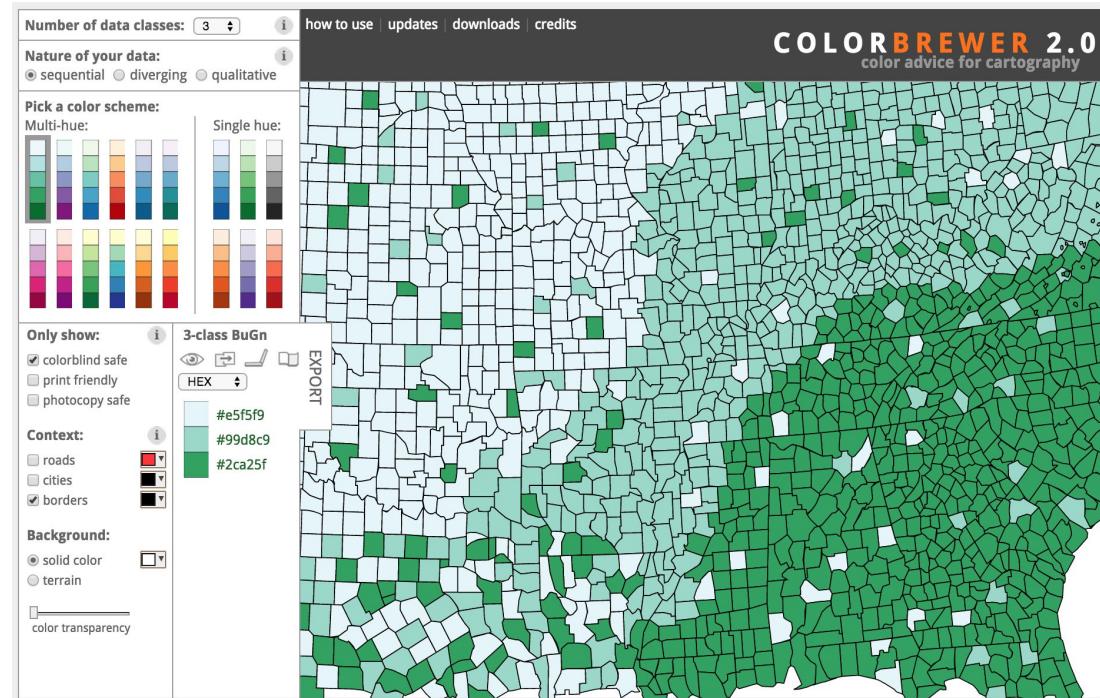
There are many forms of visual impairments. Things to consider to create a map that's accessible:

- Make it colour blind safe
- Use contrasted colour ramps
- Less is more - don't try to cram too much in a single map (you can make two instead!)
- Consider combining cartographic elements (e.g. size *and* colour of your points / lines)



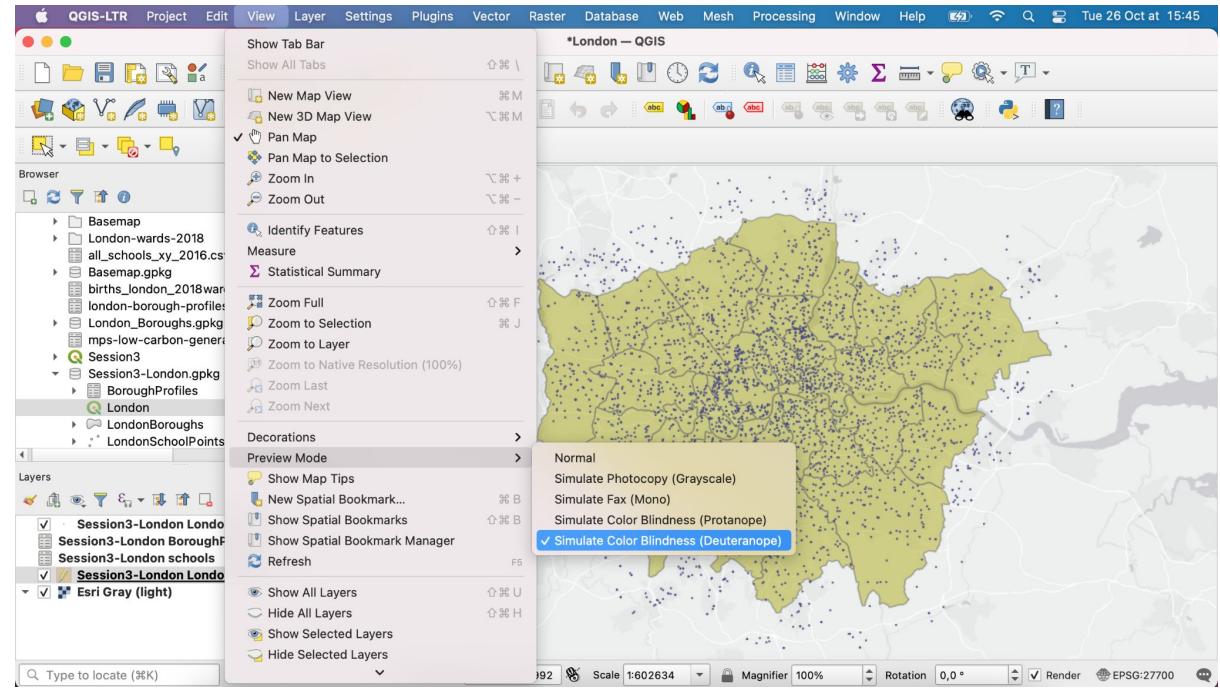
# Accessibility: colourblind safe maps

ColorBrewer ramp tools



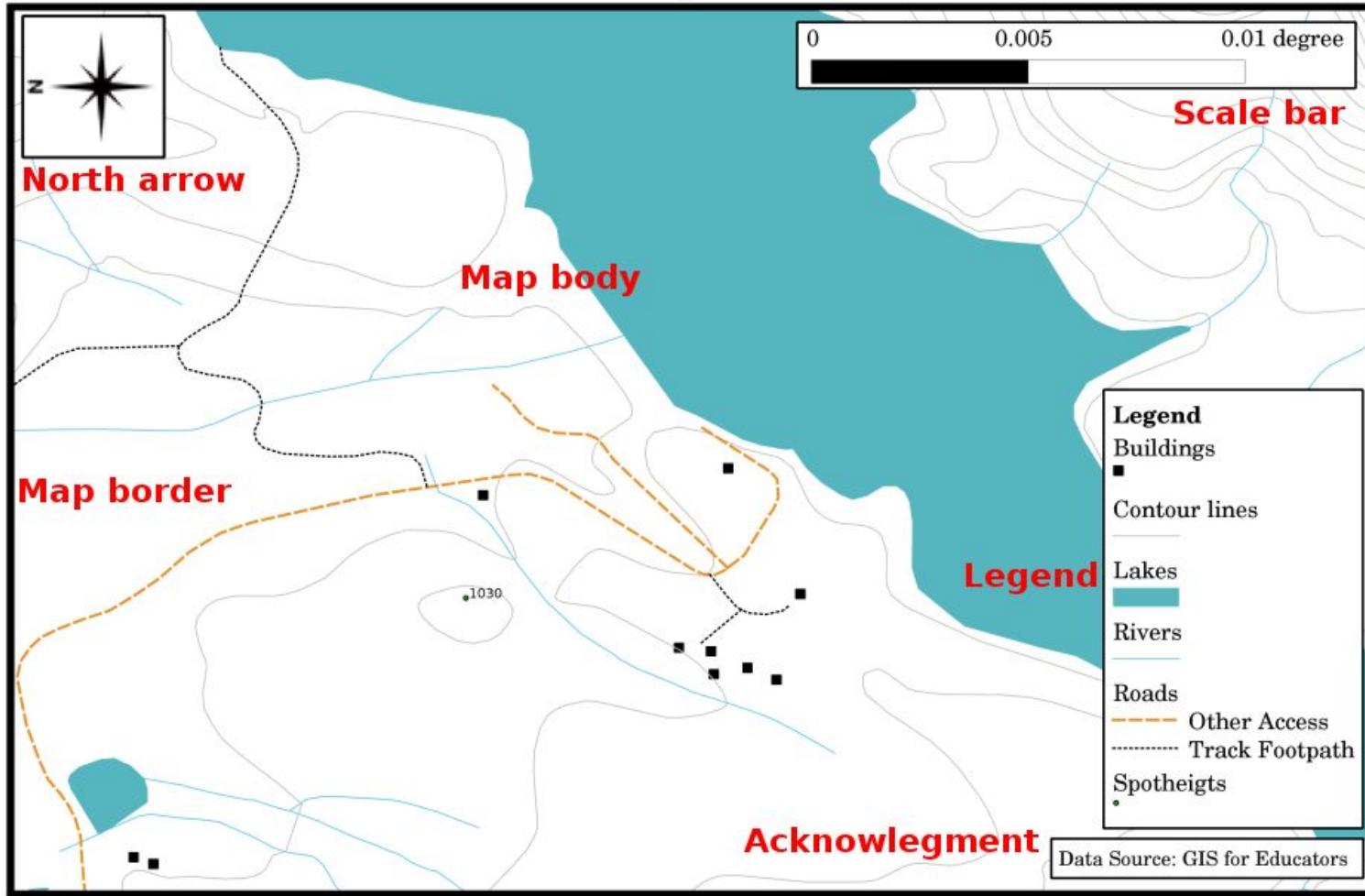
# Accessibility: colourblind safe maps

QGIS previews for colour blindness and b&w printer safe maps.



# Map Layouts

# Title Common map elements



# Cartographic elements

1. Map frame
2. Title
3. Legend
4. Directional element (North arrow)
5. Scale (scale bar or RF)
6. Source

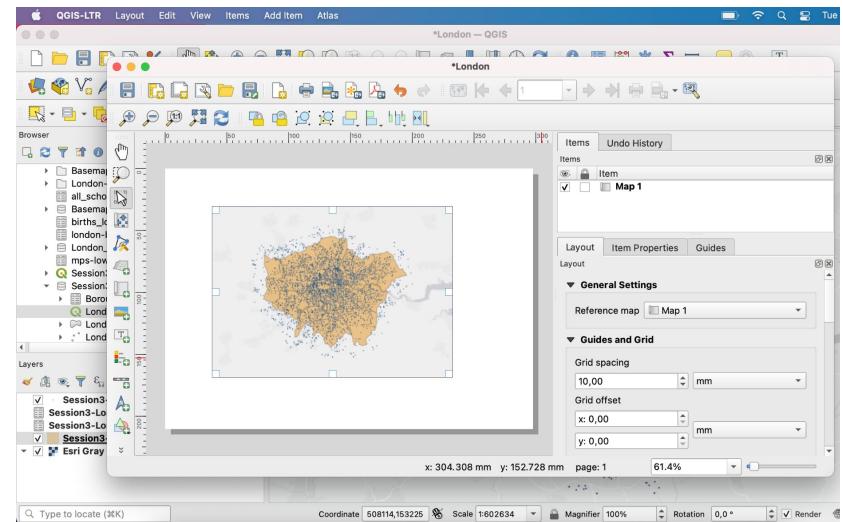
*Optional:*

7. *Border to delineate your map layout area*
8. *Locational references (inset of the broader context)*
9. *Projection*
10. *Date*

# Exporting maps: Print Layout Composer

## How do you share your map?

1. You can use a \*.qgz file or save your project in a geopackage if you're sharing with someone who uses QGIS and will need to manipulate your layers.
2. You can export your map as a \*.jpeg, \*.png, \*.pdf or simply print it. Both exporting and printing are handled via the **Print Layout Composer**.

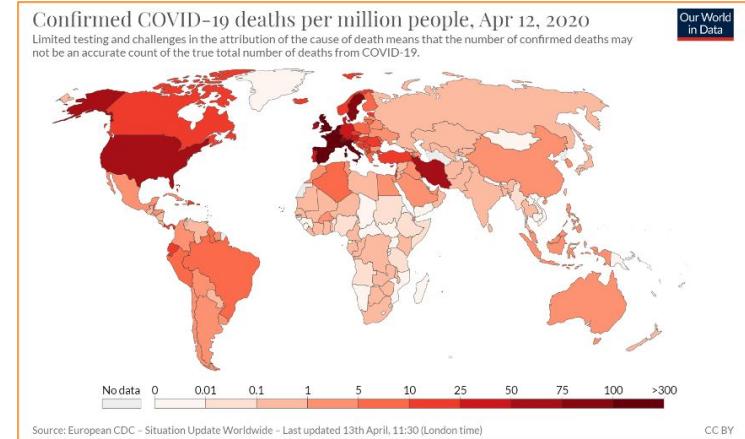
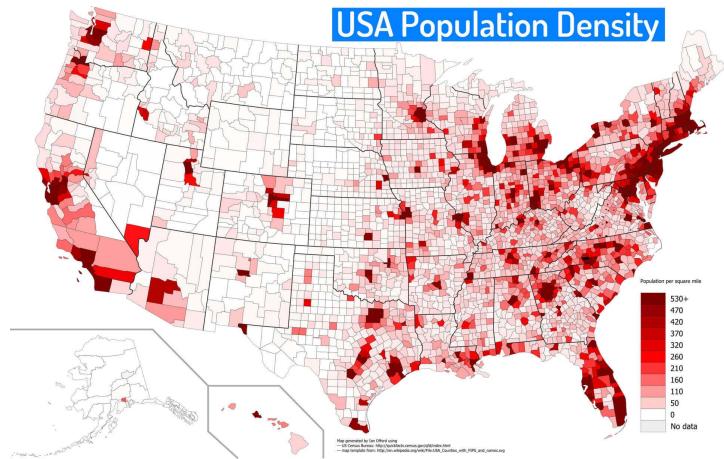


# Choropleths

# 1. What's a choropleth?

A choropleth is a type of thematic map that uses **graded differences in colours** of some areas (polygons), such that the difference in colour is in proportion to a statistical variable aggregated value at that area level.

For example: a map of the US population density per county, or a map of the world and covid contamination rates per country, etc.



## 2. Making a choropleth

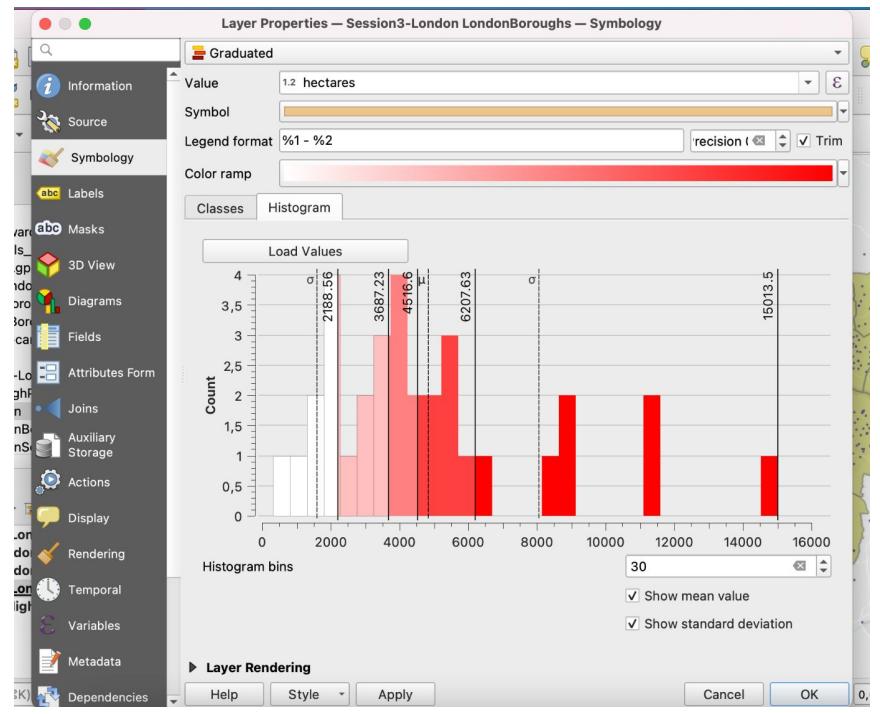
To build a choropleth map:

1. You need a vector layer representing **polygons** (*regions / sub-regions / countries / some kind of administrative, census or electoral boundary ...*)...
2. ... that contains one **numerical field** (*data type = integer or real number*) of interest.  
Note that this can be tabular data (census etc) that you joined to the vector layer.
3. Have a look at that variable's **distribution** (is it normal? Skewed? Can you identify clusters or gaps?)
4. Pick a **number of classes** (based on domain knowledge and the data distribution)
5. Pick a **method** for determining your **class breaks**
6. Pick a **colour ramp** (think about accessibility, and implicit meaning of certain colours)

### 3. How many classes should you pick?

This is an art and a science! There is no magic rule - you have to proceed by trial and error. Usually anywhere between 3 to 7 data classes is a good range.

Look at the data distribution in the histogram to identify gaps or clusters that already suggest possible data breaks around those clusters.



### 3. How many classes should you pick?

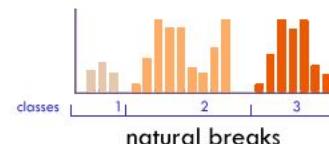
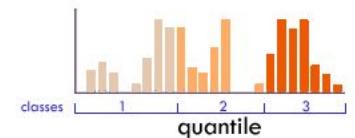
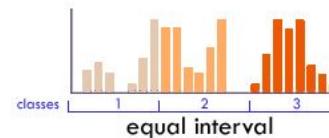
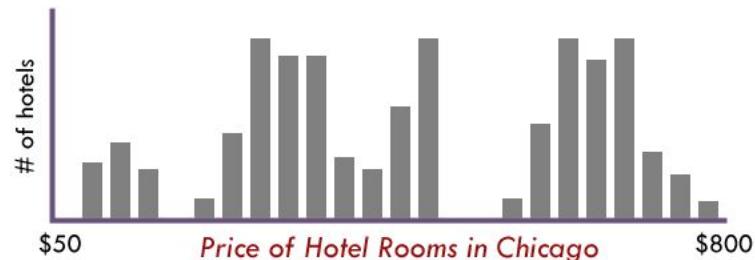
Please note that:

- If you use too many classes, you increase the risk of **confusion** when your audience reads the map (they can't read properly the statistical value for a given area as the colours are too close).
- When you use less classes, you need to do some **data generalization** (you will lump a larger range of data together), so you're automatically reducing precision. For instance a map with only 3 levels is easy to interpret (low, medium, high) but might **over-simplify** the reality of the phenomenon you're looking at, and maybe group together regions that are actually quite different.

## 4. Which method should you use?

Please read [this GISGeography.com article](#) and [this ESRI article](#) for an overview of the main data classification caveats, including choosing the number of breaks and a classification method.

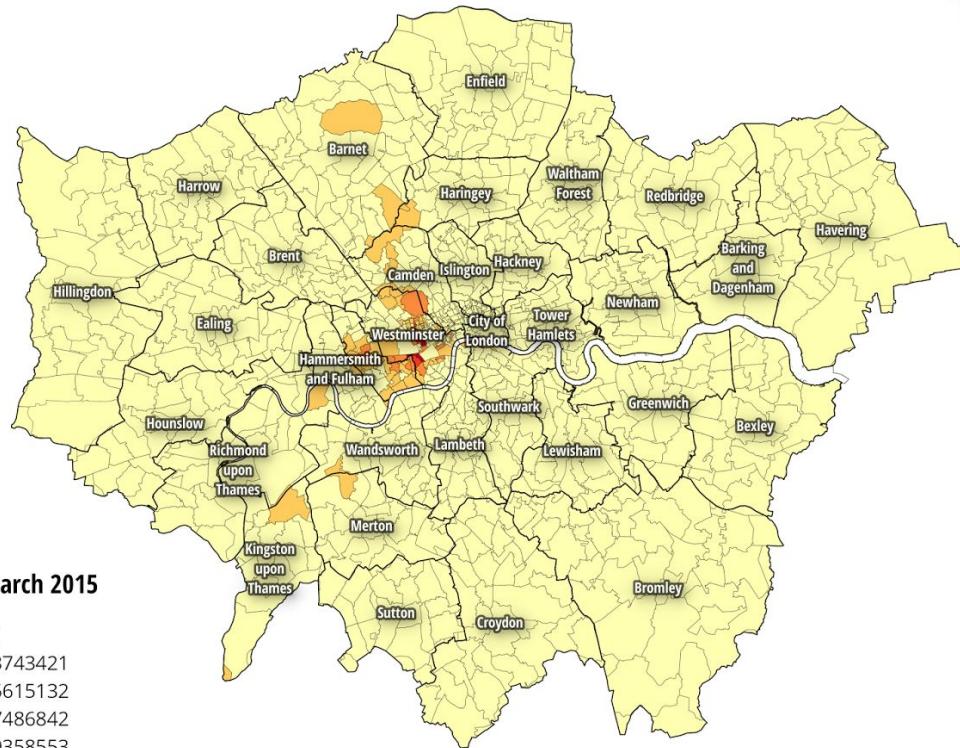
Read [this StatsMapsNpix article](#) and see maps on the next slides for a comparison of classification methods on a same dataset (London house prices).



Source: <https://www.axismaps.com/guide/data-classification>

# London House Prices

Equal Interval



Average Price, March 2015

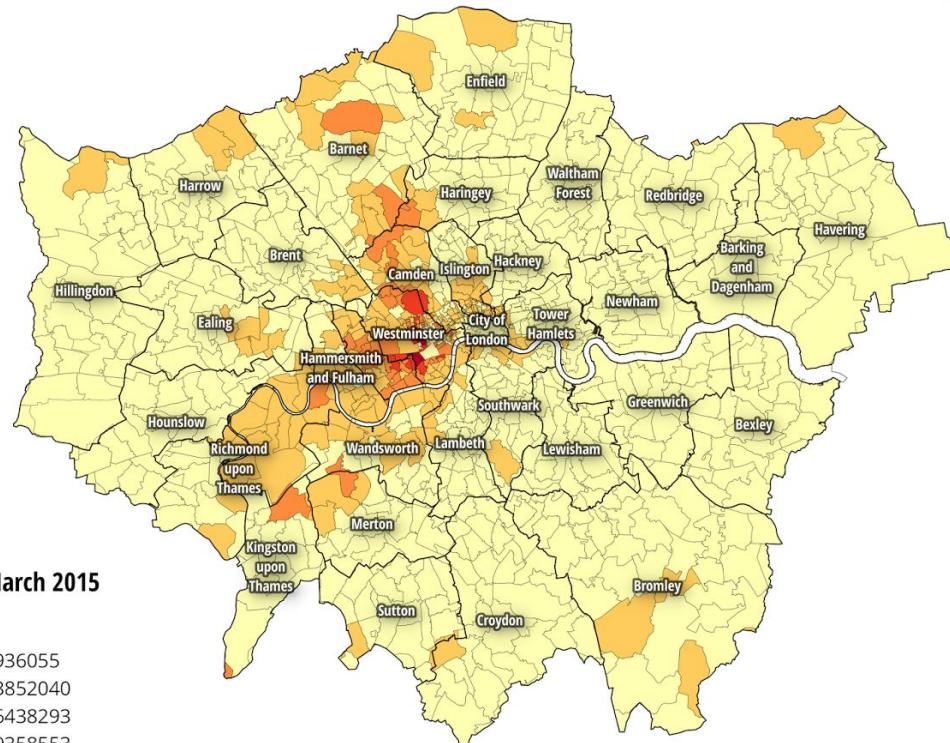
- £0 - £1871711
- £1871711 - £3743421
- £3743421 - £5615132
- £5615132 - £7486842
- £7486842 - £9358553

Data Source: HM Land Registry, 2015

Source: <http://www.statsmapsnpix.com/2016/01/from-anscombes-quartet-to-choropleth.html>

# London House Prices

Natural Breaks



Average Price, March 2015

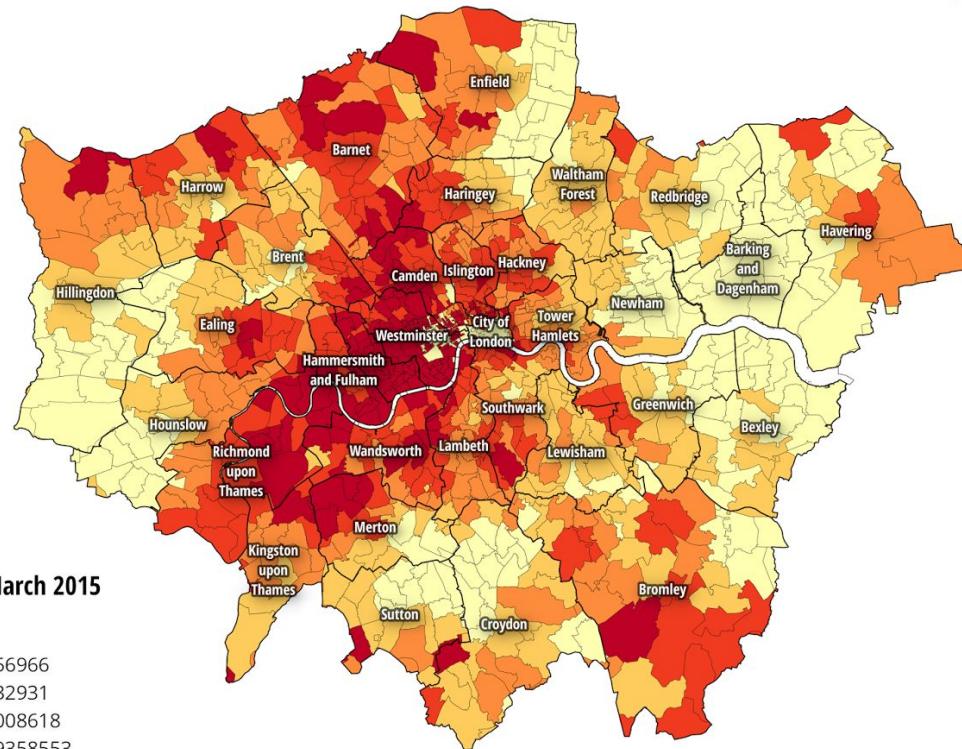
- £0 - £817200
- £817200 - £1936055
- £1936055 - £3852040
- £3852040 - £6438293
- £6438293 - £9358553

Data Source: HM Land Registry, 2015

Source: <http://www.statsmapsnpix.com/2016/01/from-anscombes-quartet-to-choropleth.html>

# London House Prices

Quantile

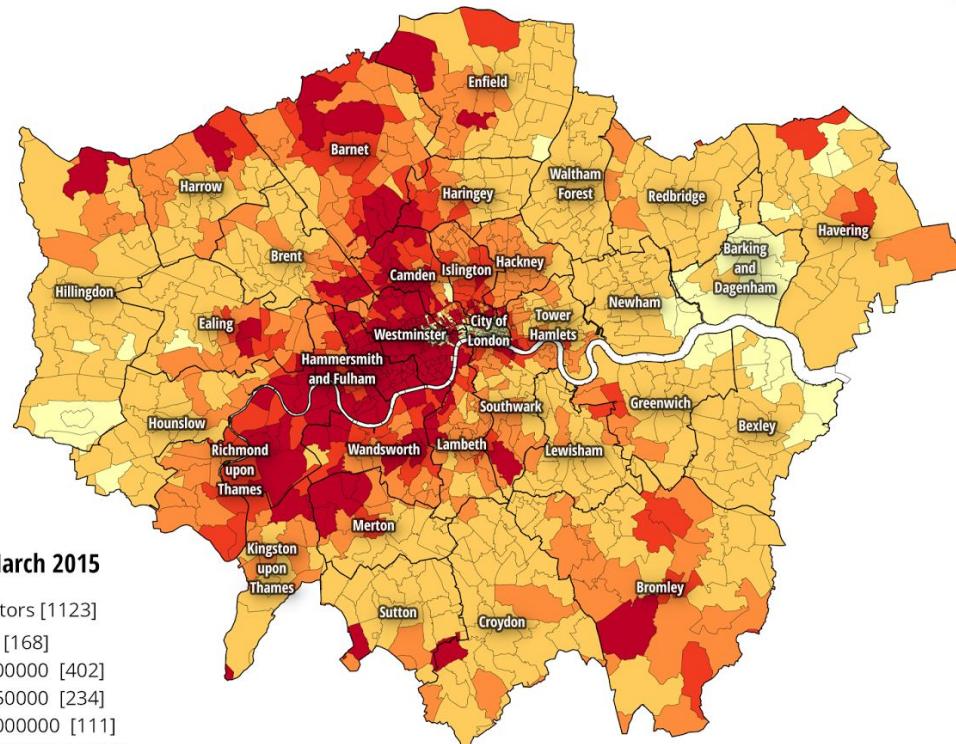


Data Source: HM Land Registry, 2015

Source: <http://www.statsmapsnpix.com/2016/01/from-anscombes-quartet-to-choropleth.html>

# London House Prices

Manual



Average Price, March 2015

Total Postcode Sectors [1123]

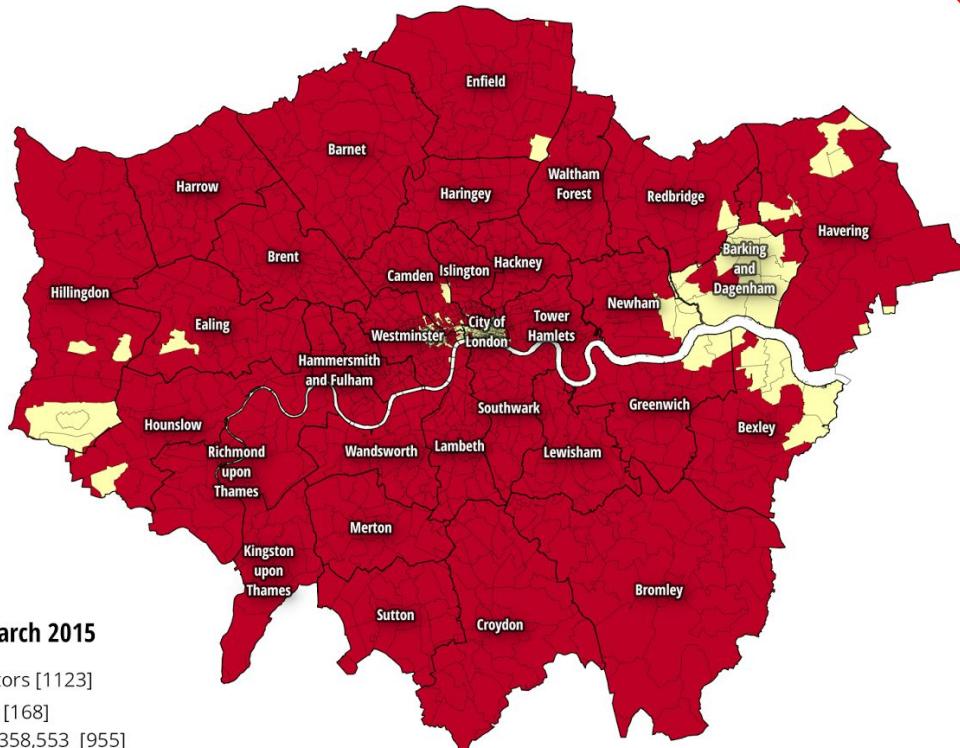
- £0 - £250000 [168]
- £250000 - £500000 [402]
- £500000 - £750000 [234]
- £750000 - £1000000 [111]
- £1000000 - £9358553 [208]

Data Source: HM Land Registry, 2015

Source: <http://www.statsmapsnpix.com/2016/01/from-anscombes-quartet-to-choropleth.html>

# London House Prices

Above/Below UK average



Data Source: HM Land Registry, 2015

Source: <http://www.statsmapsnpix.com/2016/01/from-anscombes-quartet-to-choropleth.html>

# 5. How to pick your colour ramp?

We associate **meaning to certain colours**; for example we're used to associating water and blue, fire and red, or to automatically interpreting darker colours as the higher values in a distribution. Be aware of these conventions and only go against them if you have a very good reason to!

**Sequential palettes** display the progression of a continuous variable. If you're interested in the variance below and above a central value (for example: 0 or a mean value), then you will need a **divergent palette**. Save **qualitative palettes** for categorical variables.

## Sequential Palette

- Continuous Data
- Ordered Data
- Numerical Data

## Pick a color scheme:

Multi-hue:



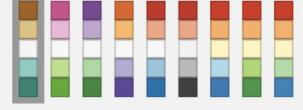
Single hue:



## Diverging Palette

- Continuous Data with 0
- Ordered Data with 0
- +/- Numerical Data

## Pick a color scheme:



## Qualitative Palette

- Categorical Data
- Nominal Data

## Pick a color scheme:

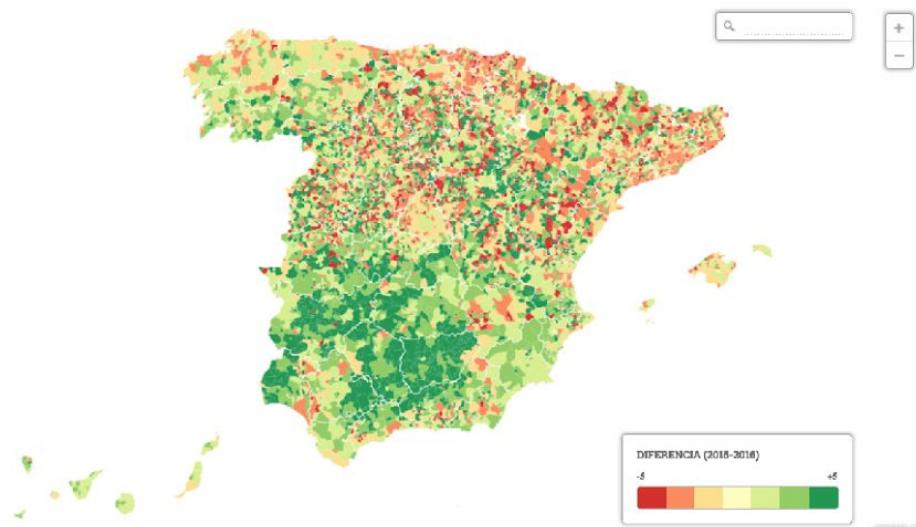


# 5. How to pick your colour ramp?

/!\ Please DO NOT use a red to green colour ramp ever (not colourblind safe!) :)

THIS IS WHAT YOU WANT READERS TO SEE...

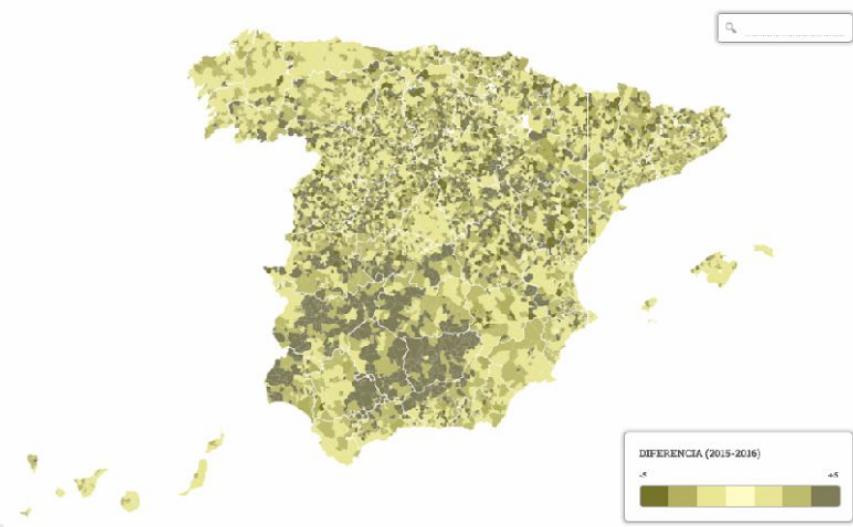
Variación de la participación a las 14h entre el 20D y 26J, por municipios



...BUT THIS IS WHAT PEOPLE WITH DEUTERANOPIA\* SEE

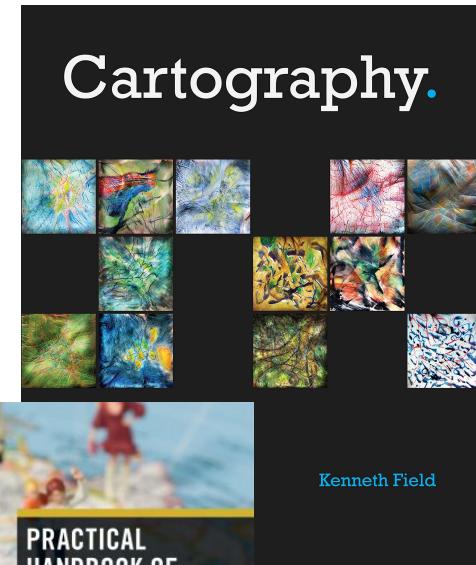
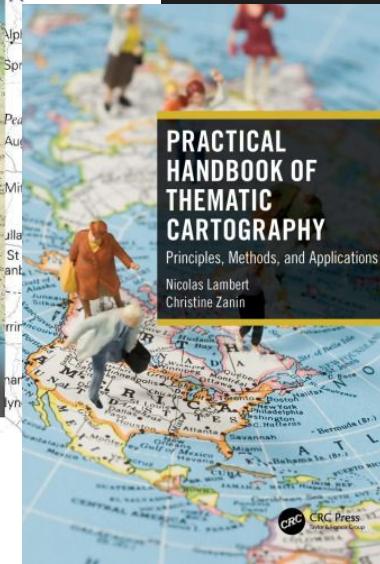
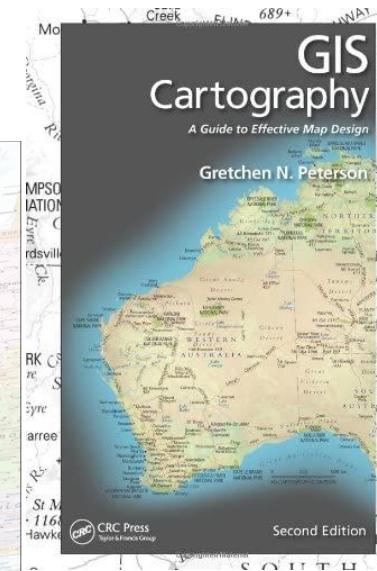
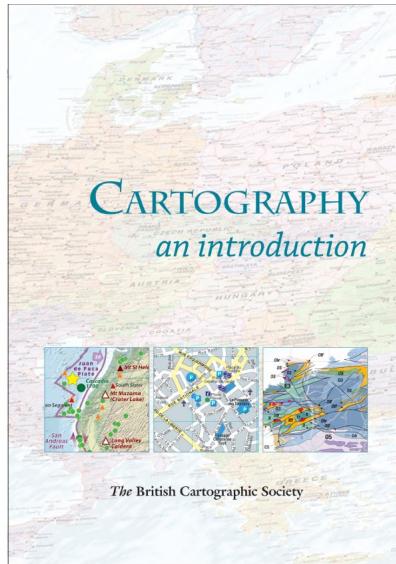
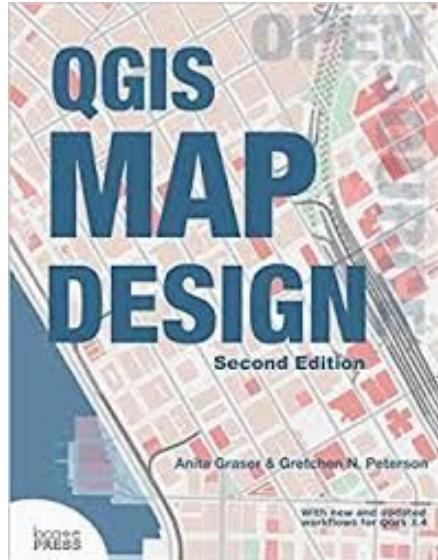
\*the most common kind of color blindness

Variación de la participación a las 14h entre el 20D y 26J, por municipios



Source: <http://www.thefunctionalart.com/2016/06/red-green-color-schemes-in.html>

# To go further on map design



Cartography.

# Coursework

# Final Coursework

The final coursework is a **map production exercise**. Find a research question, carry out a simple map production workflow, and write a report summarizing your findings. You will be assessed on your capacity to:

- Frame your research question in a clear and concise manner, and ensure a few maps can provide interesting insights,
- Identify datasets that are relevant to answering your question (*technically here you may want to work backwards and use the data you already have to find your research question*),
- When appropriate, use table joins to “enrich” your vector data,
- Carefully choose your symbology, and ensure your map is accessible and colour-blind safe,
- Create map exports complete with all key cartographic elements (title, legend, data source etc).
- Analyse and interpret the patterns that emerge from your maps, explaining what this might mean in terms of policy or research outlooks.

# Final Coursework

## Some practical points:

- Deadline: Wednesday 24th November, 23.59 Paris time. *If you're late, minus 1 point for each day behind the deadline.*
- Work in groups of 2-3 students
- Work on a **European city** of your choice (*cities in the UK and other non-EU european countries are ok*)
- The report must be 3 pages minimum, 5 pages *maximum*
- You are encouraged to use the data provided for the tutorials, however if you want to challenge yourselves, you will get **+2 points bonus for working with data you have sourced yourselves.**

# Final coursework: Proposed outline

You are strongly encouraged to follow this outline:

1. Executive summary (maximum ½ page, bullet points are fine)
2. Introduction / Problem / Context
3. Data sources in a table
4. High-level methodology. Keep it very short but use precise terminology. Explain your symbology choices including your choice of class breaks if you built a choropleth.
5. A minimum of 2 maps. Careful, you only have 5 pages maximum in this report so these maps must be relevant to answering your policy question
6. Analysis of the findings (half a page)
7. Conclusion / next steps for policy makers or future research outlook (half a page)

In this exercise, concision and precision are key!

# Homework

(In 2 weeks)

# Homework for 10/11/2021

1. Do the [QGIS tutorial](#) on the attribute table
2. Use Slack if you have questions (#help).
3. Make sure you have formed groups of 2 or 3 for the final coursework