

# Data visualisation in R using crime data

Sam Langton

12 February 2021

# Welcome



@sh\_langton

s.langton@leeds.ac.uk

# Welcome



@sh\_langton

s.langton@leeds.ac.uk

All materials for today are available online.

Web link: [https://rpubs.com/langton\\_](https://rpubs.com/langton_)

Material: [https://github.com/langtonhugh/data\\_viz\\_R\\_workshop](https://github.com/langtonhugh/data_viz_R_workshop)

# Contents

## **10.00-10.15**

- Intro

## **10.15-11.00**

- Presentation: data mapping

## **11.00-11.15**

- Short break

## **11.15-13.00**

- Practical

## **13.00-14.00**

- Long break

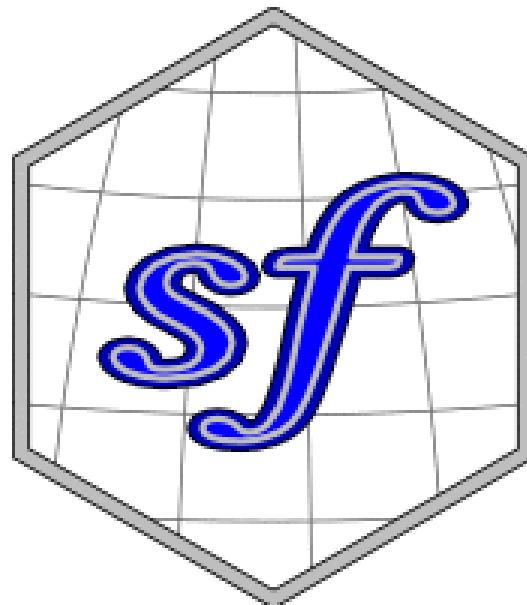
## **14.00-14.30**

- Presentation: applied research using data mapping

# Maps in R

# The good news...

- Making a map in R is *very similar* to what we have covered already.
- Making use of an additional package (`sf`) we can create beautiful maps using `ggplot2`.



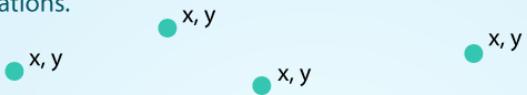
# The bad news...

- You might be entering uncharted territory!
- Spatial visualisation (and the prep needed to get to that point) necessitates some additional level of understanding in geography and Geographic Information Systems (GIS), including:
  - Spatial data
  - Projection
  - Visualisation issues

# The bad news: spatial data

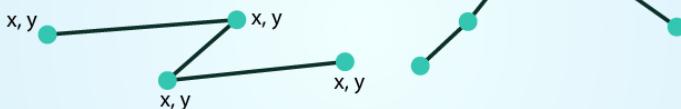
- Vector data.

**POINTS:** Individual **x, y** locations.  
ex: Center point of plot locations, tower locations, sampling locations.



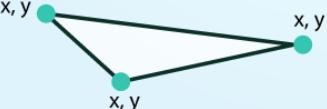
---

**LINES:** Composed of many (at least 2) vertices, or points, that are connected.  
ex: Roads and streams.



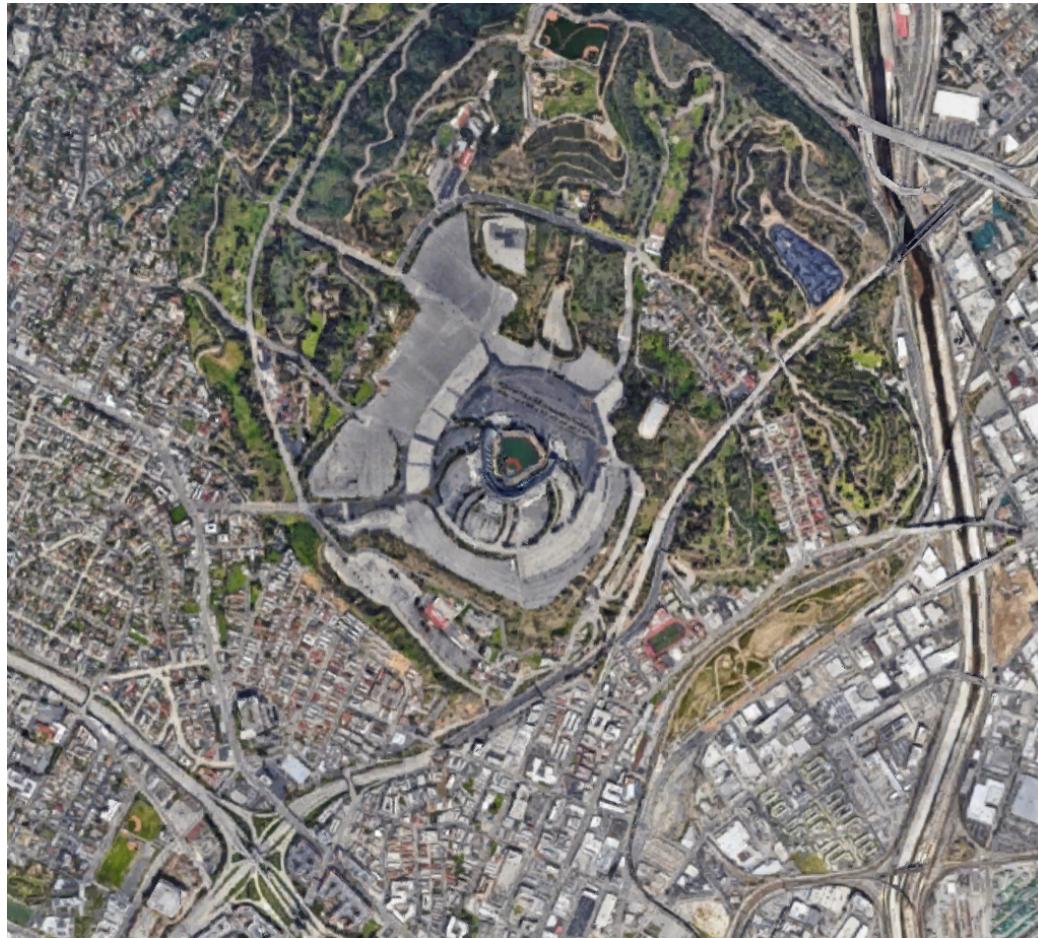
---

**POLYGONS:** 3 or more vertices that are connected and **closed**.  
ex: Building boundaries and lakes.



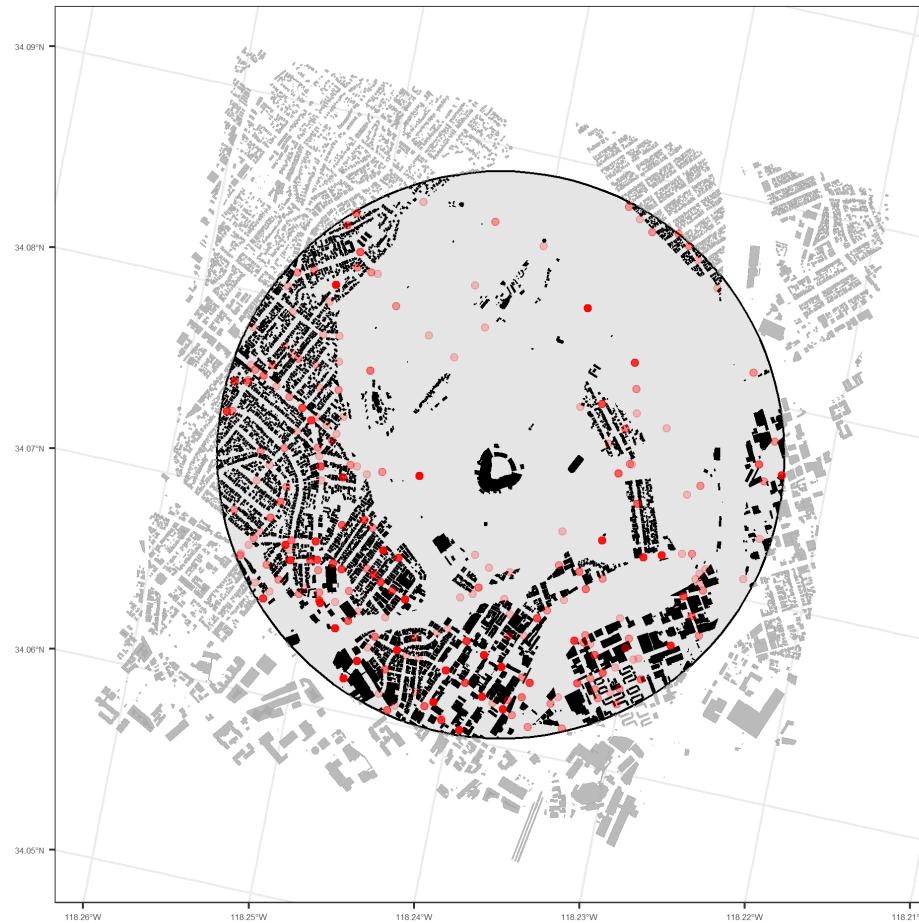
Source: NEON via Data Carpentry.

# The bad news: spatial data



Source: [Google Maps](#).

# The bad news: spatial data



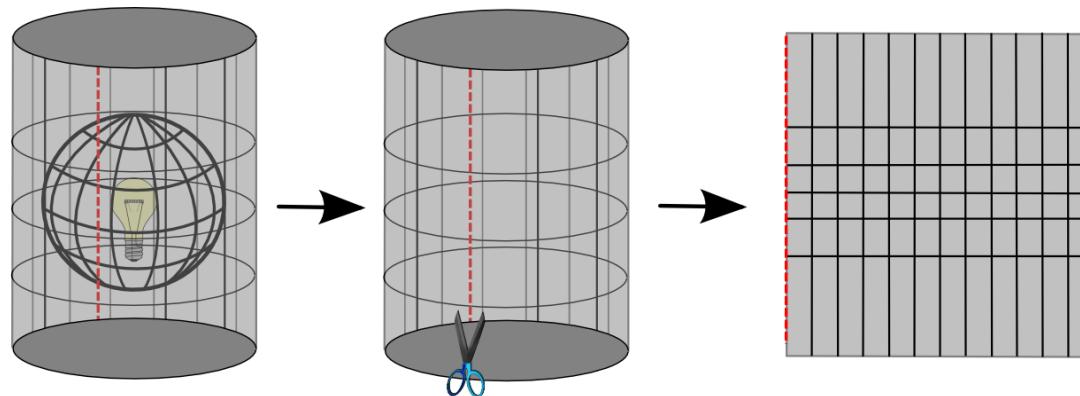
Source: [Langton & Solymosi](#).

# The bad news: spatial data

- Today, we will use shapefiles (`.shp`).
- Shapefiles are a popular format to store geospatial vector data.
- Unlike standard 2D data frames, such as `.csv` files, shapefiles contains multiple components.
  - `.shp`
  - `.shx`
  - `.dbf`
  - `.prj`
- But there are other formats, such as `.geojson` or `kml`, which you might come across.
- There is some **debate** over the most appropriate, with each having their own advantages and disadvantages.

# The bad news: projections

- Maps tend to represent the (~spherical) earth on flat surfaces (e.g. paper, screens).
- This transformation is known as a **projection**.
- Be aware of the Coordinate Reference System (CRS).
- Geographic CRS (e.g. **WGS 84**).
- Projected CRS (e.g. **British National Grid**).



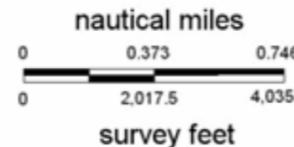
Source: [QGIS documentation](#).

# The bad news: visualisation issues

## Ten tips to make your maps more ‘interesting’

### Tip 1: Do not include a scale bar

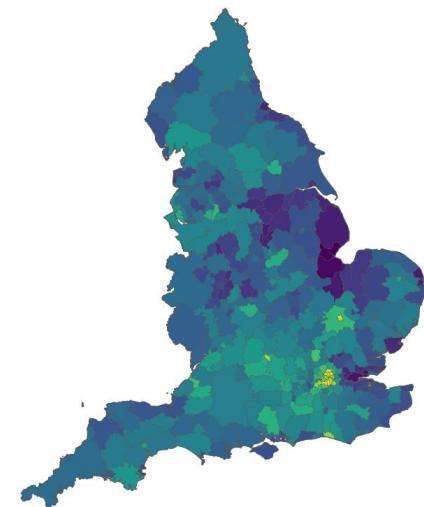
This will make it much more interesting as your map readers have to guess the distance between objects. One of the main purposes of mapping crime is to compare areas and examine the proximity of objects, so why make it easy for the uninitiated to understand your map? Without a scale bar nobody will have a clue how far things are apart and this gives you the opportunity to have impromptu quizzes or make things up as you are presenting. If you accidentally include a scale bar use a scale that goes; "0 ----- 6.75 ----- 13.25 kilometers" instead of the usual "0 ---5 ---10" or similar. Big complex numbers really impress audiences, and you want to look your most clever.



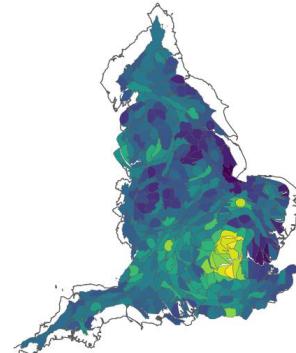
Source: Jerry Ratcliffe

# The bad news: visualisation issues

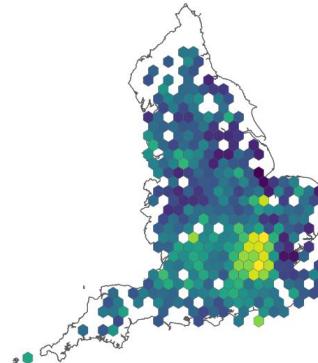
Original boundaries



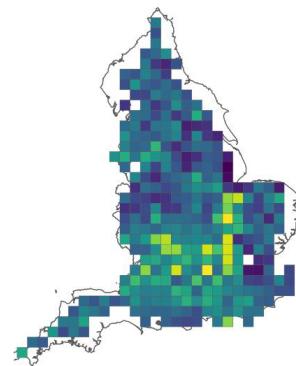
(a) Balanced cartogram



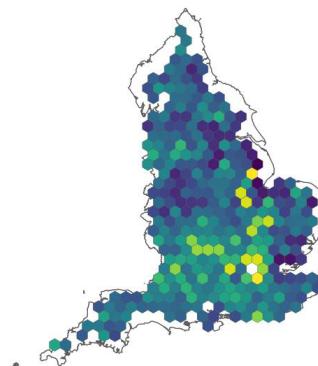
(b) Hexogram



(c) Square grid



(d) Hexagonal grid



Source: Langton & Solymosi (2020)

# Crime demo