Data Visualisation
Chapter 7: Spatial Data

Dr James Baglin

How to use these slides

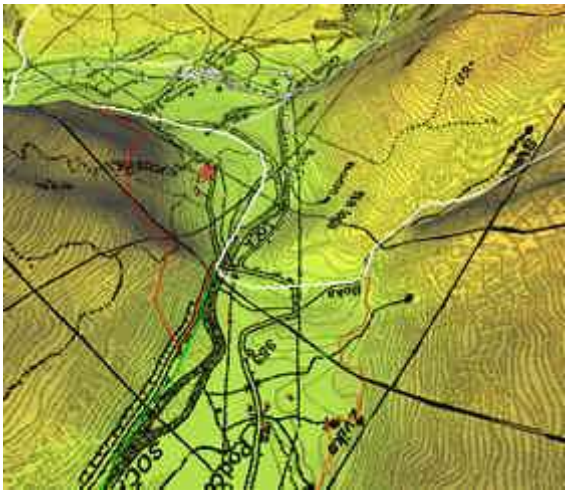
Viewing slides…

- Press ‘f’ enable fullscreen mode
- Press ‘o’ or ‘Esc’ to enable overview mode
- Pressing ‘Esc’ exits all of these modes.
- Hold down ‘alt’ and click on any element to zoom in. ‘Alt’ + click anywhere to zoom back out.
- Use the Search box (top right) to search keywords in presentation

Printing slides…

- Click here to open a printable version of these slides.
- Right click and print from browser or save as PDF (e.g. Chrome)

Spatial Data Visualisation



- Everything happens somewhere - variables can be linked with location.
- The spatial dimension of the data must be used to convey meaning or a spatial trend.
- Spatial autocorrelation:

> *Everything is related to everything else, but near things are more related than distant things. – Tolber (1970)*

Spatial Data Visualisation Cont.

- The most common spatial data visualisations map data to a geographical coordinate system.
- Other spatial systems might include the following:
  - sporting arenas
  - the universe/galaxy
  - a building
  - a virtual world in a computer game

Spatial Referencing Systems

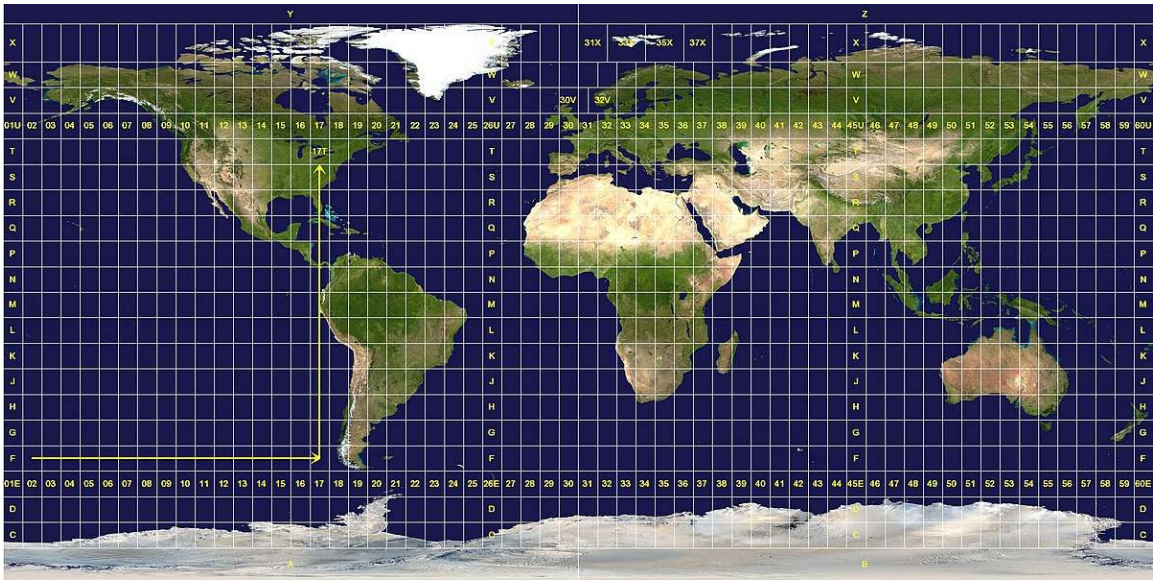
- **Spatial referencing system** refers to a system for defining a spatial location on Earth using coordinates.
- Australia's official system is the **Geocentric Datum of Australia 2020** (GDA2020)
- Referencing systems need to be updated:
  - continental drift (plate tectonics)
  - Improved measurements of the Earth's shape
  - changes to Australia's crust (Geoscience Australia 2018).
- This system projects latitude and longitude to a coordinate system composed of Eastings and Northings.

Map Projection Cont.

- **Map projection** is the process of transforming geographic coordinates from a three-dimensional surface (latitude and longitude) to a two-dimensional plane (Cartesian x and y coordinates).
- The Map Grid of Australia (MGA) coordinate system is based upon the **Universal Transverse Mercator (UTM)** projection.
- The UTM projection splits Earth into 60 zones of 6 degrees (6 * 60 = 360 degrees).
- The main advantage of the UTM system is that it minimises distortions to maps due to projection.

Map Projection Cont.

- Projections distort area, particularly at the edges of maps - The True Size of



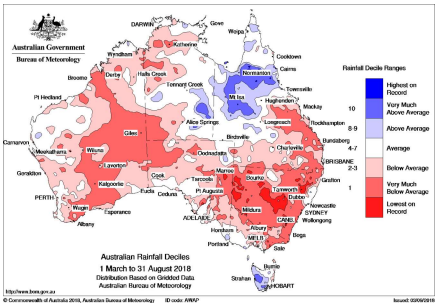
UTM Zones - Wikimedia Commons, Link

People and Organisational Systems

- Street addresses, electorates, local government areas, postcodes, and suburbs are examples of **people or organisational reference systems**.
- These systems are more practical but sacrifice precision.
- Often, spatial data visualisation will require People/Organisational reference system to be converted to coordinates.
- PSMA Geocoded National Address File (G-NAF) which provides longitude and latitude coordinates for over 13 million Australian addresses.

Thematic Mapping - Isarithmic Map

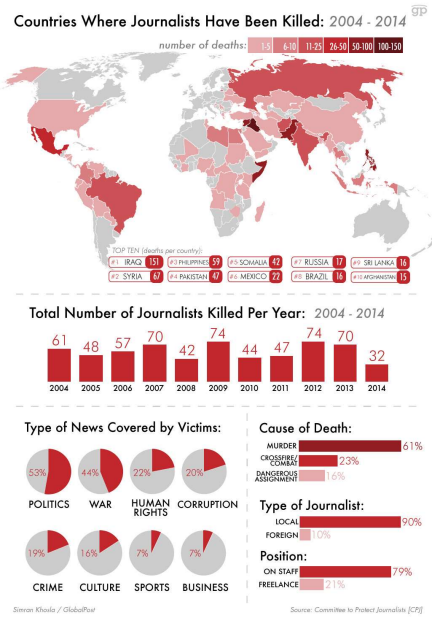
- **Thematic maps** link an area with a variable of interest, i.e. theme.



- The five main types of thematic maps include isarithmic maps, choropleth maps, points/dot density maps, proportional symbol/bubble maps and cartograms.
- **Isarithmic maps**, also known as a contour or topological maps, use colour or lines to convey the spatial variability in a statistical quantity across the map
- Example: Australia in winter 2018 -BOM

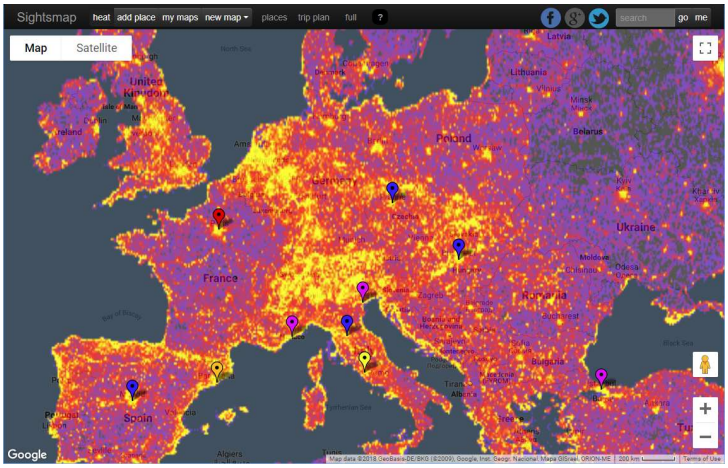
Thematic Mapping - Choropleth Map

- **Choropleth maps** use predefined regions of a map and a colour scale to visualise the variability in a statistical quantity
- Example: Journalist Killing Map



Thematic Mapping - Point/Dot Density Map

- **Point maps**, also known as dot density maps, use location points to visualise the spatial distribution of a statistical frequency.
- Example: Sightsmap

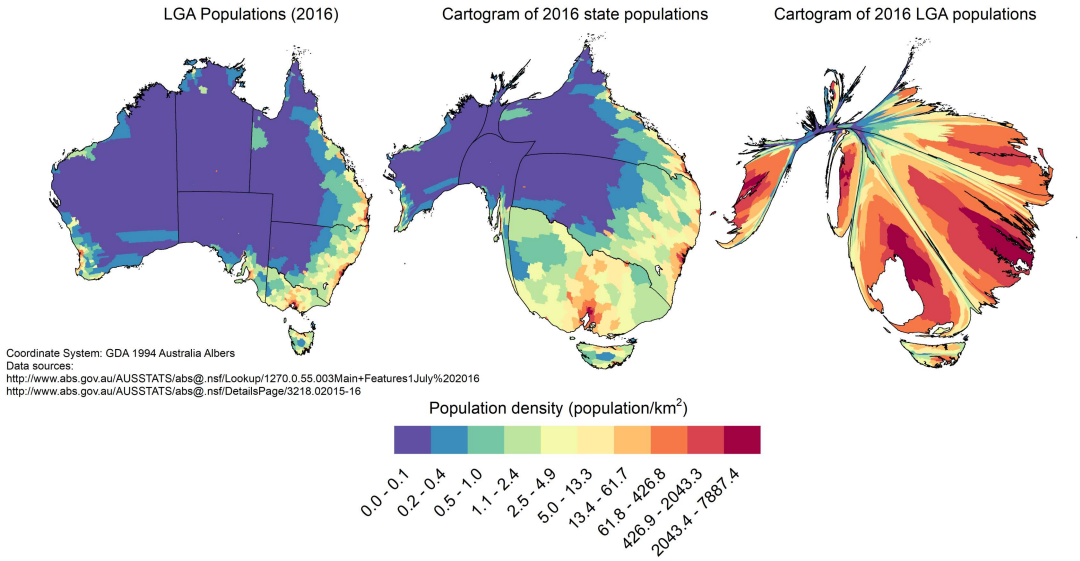


Thematic Mapping - Proportional Symbol/Bubble Maps

- **Proportional symbol/bubble maps** are point maps with the added advantage of being able to map an additional variable to the size of a point.
- Example: Earthquakes - PTWC

Thematic Mapping - Cartograms

- **Cartograms** take a map and then distort the distance or area of the map to represent a variable
- Example: Australia's Population 2016



Choropleth Maps - Packages

- Let's create a choropleth map using R.
- You will need to install and load the following packages.
- Order is important!

```
library(ggplot2)
library(dplyr)
library(rgeos)
library(maptools)
library(ggmap)
library(broom)
```

Leaflet

- Leaflet is an open-source JavaScript library for interactive maps
- Leaflet for R allows you to easily create Leaflet maps using R.
- Ensure you install and load Leaflet in R.

```
install.packages("leaflet")
library(leaflet)
```

Australian Marriage Law Postal Vote Survey 2017

- We will demonstrate how to create a choropleth map of the 2017 Australian Marriage Law Postal Vote Survey
- We are interested in the spatial distribution of the percentage of “Yes” votes based on Commonwealth Electorate Divisions (CED)
- The ABS reports the data which can be downloaded as a `.csv` here.
- Load the data into R

```
vote <-
  read.csv("../data/australian_marriage_law_postal_sur
```

Federal Electorate Boundaries

- In order to draw the Federal Electorate Divisions we will need an Esri `.shp` file.
- The ABS also hosts the official boundaries file for 2016 here.
- We will download the **Commonwealth Electoral Divisions ASGS Ed 2016 Digital Boundaries in ESRI Shapefile**.
- Extract the files from the .zip archive and save them into a folder.

Import .shp file

- Now we can import the `.shp` file into R.

```
national<-readShapeSpatial("1270055003_ced_2016_aust_s
                           delete_null_obj=TRUE)
```

- Confirming…

```
class(national)
```

```
## [1] "SpatialPolygonsDataFrame"
## attr(,"package")
## [1] "sp"
```

Import .shp file Cont.

- Checking variables…

```
head(national@data)
```

```
##   CED_CODE16 CED_NAME16 AREASQKM16
## 0        101      Banks    49.4460
## 1        102     Barton    39.6466
## 2        103  Bennelong    58.6052
## 3        104    Berowra   749.6360
## 4        105   Blaxland    61.1166
## 5        106  Bradfield    98.3974
```

Simplify

- Polygons in shape files might be too accurate and slow down visualisation.
- The `gSimplify` function from the `rgeos` package can help speed things up.

```
print(object.size(national), units = "MB")
```

```
## 50.3 Mb
```

```
national_simp <- gSimplify(national, tol = .001,
  topologyPreserve=TRUE)
national_simp <-
  SpatialPolygonsDataFrame(national_simp,

  data=national@data)
print(object.size(national_simp), units = "MB")
```

```
## 17.2 Mb
```

Merge .shp with ABS data

- We need to ensure both data objects have an electorate column named `CED_NAME16` for matching

```
names(vote)
```

```
##  [1] "CED_NAME16"         "STATE"              "yes"
##  [4] "yes_perc"           "no"                 "no_pe
##  [7] "n"                  "n_clear"            "n_cle
## [10] "not_clear"          "not_clear_perc"     "non_r
## [13] "non_response_perc" "n_total"
```

```
names(national_simp)
```

```
## [1] "CED_CODE16" "CED_NAME16" "AREASQKM16"
```

Merge .shp with ABS data Cont.

- Perfect. Now for the merge.

```
electorates_vote<-sp::merge(national_simp, vote,
  by="CED_NAME16")
```

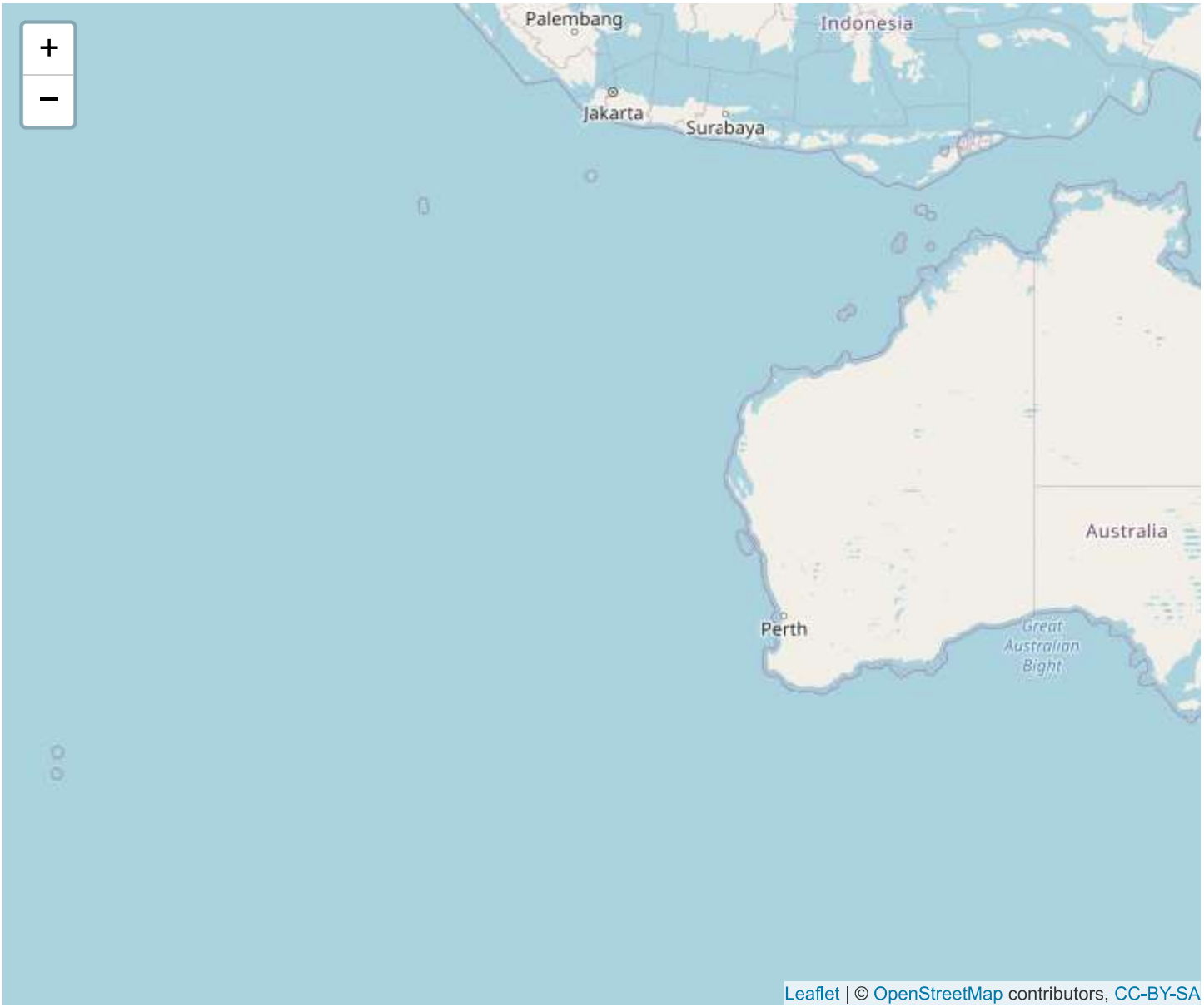
Basemap and View

- Let's start creating the choropleth map.
- We begin by setting a default view and zoom level.
- We can quickly check our choices using `addTiles()`

```
map <- leaflet(national_simp) %>%
  setView(lng = 134, lat = -28, zoom = 4)
map %>% addTiles()
```

- The result…

Basemap and View Cont.



Add Polygons

- We can check the electorial boundaries using `addPolygons()`
- This function automatically detects and projects the polygons in the `.shp` file.

```
map %>% addPolygons()
```

- The result…

Add Polygons Cont.

Continuous Colour Scale

- We need to create a continuous colour scale to represent the yes vote percentage
- We will use the `colorNumeric()` function from the Leaflet package and the `RbBu` palette from the colour Brewer system.

```
pal <- colorNumeric(
  "RdBu",
  domain = electorates_vote$yes_perc
  )
```

Add Colour Scale to Polygons

- The colour scale can now be added to the polygons.

```
map %>% addPolygons(
  fillColor = ~pal(yes_perc),
  weight = 2,
  opacity = 1,
  color = "white",
  dashArray = "3",
  fillOpacity = 0.7)
```

- The result…

Add Colour Scale to Polygons Cont.

Adding a Title, Highlights, Labels and Legend

```
labels <- sprintf(
  "<strong>%s</strong><br/>%g%% Yes",
  electorates_vote$CED_NAME16,
  electorates_vote$yes_perc
) %>% lapply(htmltools::HTML)

library(htmlwidgets)
library(htmltools)

title <- tags$div(
   HTML('<h3>Australian Marriage Law Postal Vote
  Survey 2017</h3>')
 )

map %>% addPolygons(
  fillColor = ~pal(yes_perc),
```

Adding ... Cont.

Choropleth Caution

- Choropleth maps have a few key limitations
  - Large regions can dominate
  - Small regions can be hard to see
  - Rely on colour - low visual comparison accuracy
  - Changing colour scales can drastically change appearance

Cartograms

- Can cartograms fix some of the issues with choropleths?

```
library(cartogram)

# Use detailed shp file
electorates_vote<-sp::merge(national, vote,
  by="CED_NAME16")

# Filter VIC divisions

electorates_vote@data <-
  na.omit(electorates_vote@data)
vic <- electorates_vote[electorates_vote$STATE ==
  "VIC",]

# Convert .shp to cartogram .shp

vic_cart <- cartogram_cont(vic, "n", itermax = 30
```

Cartograms Cont.

Chapter 7 - Further Reading

- Further details in Chapter 7:
  - Understanding spatial data
  - Choropleth maps in ggplot2 and Leaflet
  - Point maps in ggplot2
  - Case study - The City of Melbourne's Urban Forest
- Enjoy

References

Geoscience Australia. 2018. “Geocentric Datum of Australia 2020 (GDA2020).” http://www.ga.gov.au/scientific-topics/positioning-navigation/geodesy/datums-projections/gda2020.