Data visualisation in R using crime data

Sam Langton

11 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_

Material: https://github.com/langtonhugh/data_viz_R_workshop

Contents

10.00-10.15

- Intro

10.15-11.00

- Presentation: data visualisation

11.00-11.15

- Short break

11.15-13.00

Practical

13.00-14.00

- Long break

14.00-14.30

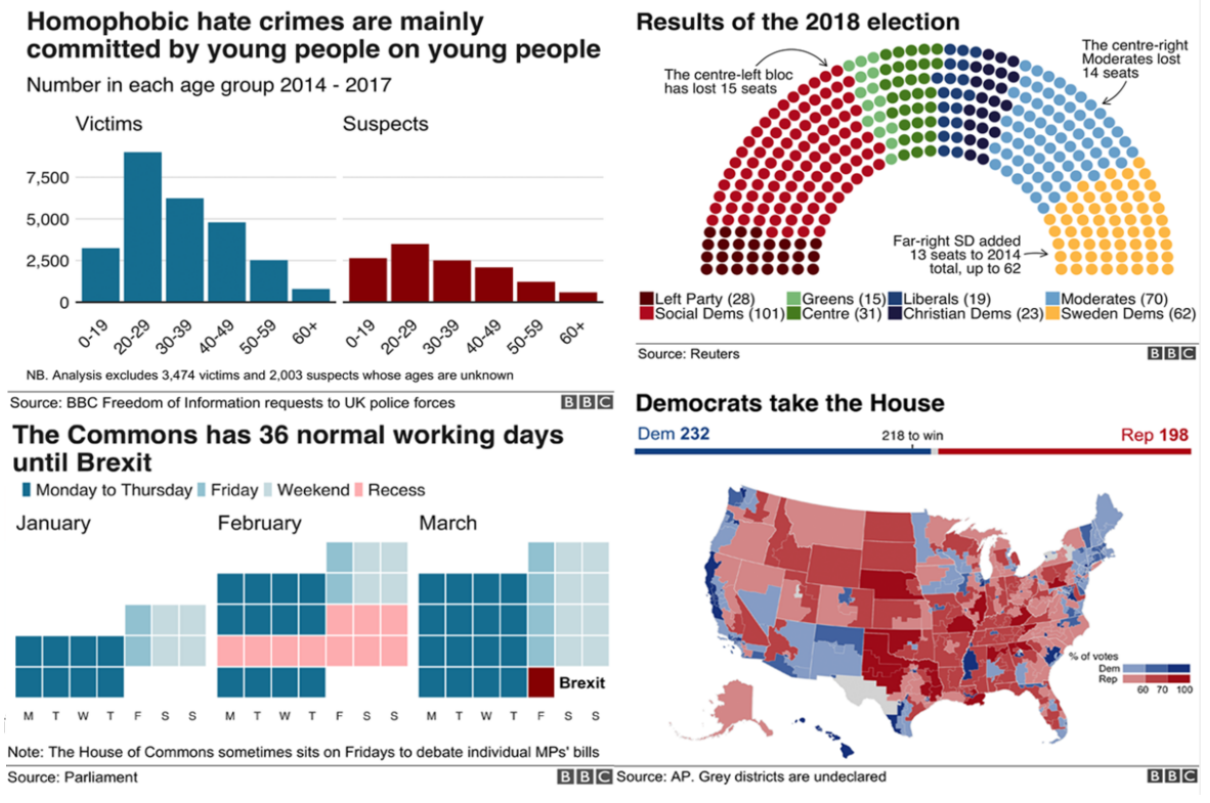
- Presentation: applied research using data visualisation

Data viz

"The visual representation and presentation of data to facilitate understanding" (Kirk, 2019)

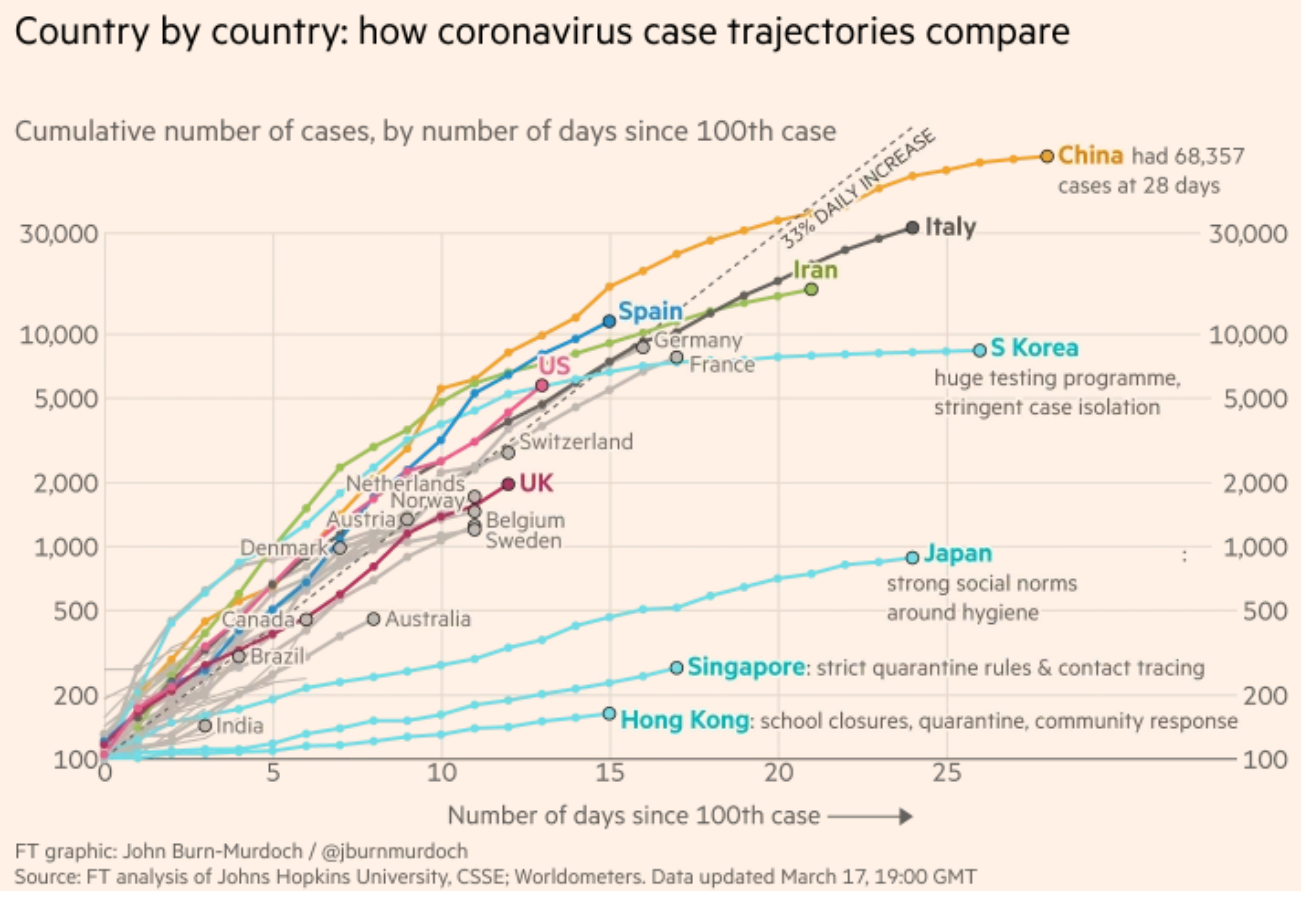
Data viz in R

"The visual representation and presentation of data to facilitate understanding" (Kirk, 2019)



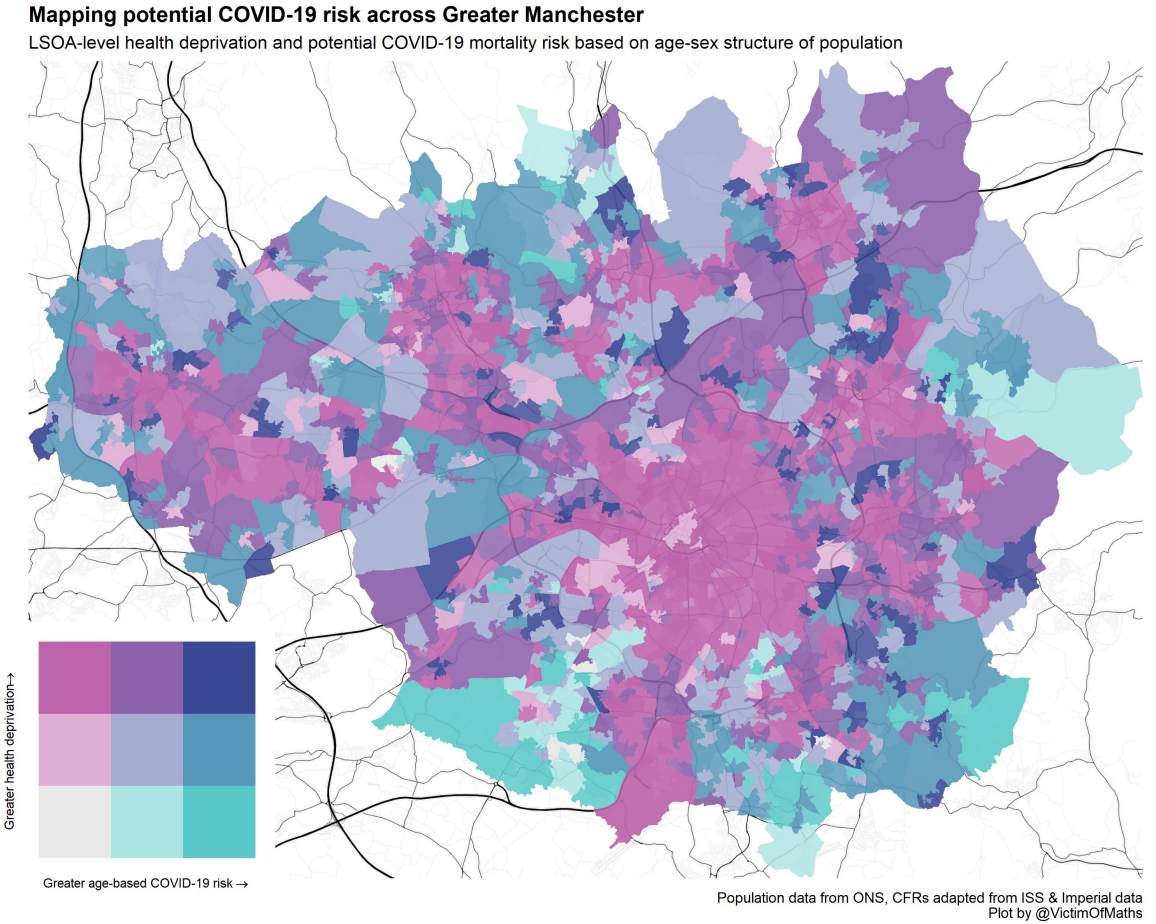
Source: Medium

Data viz in R



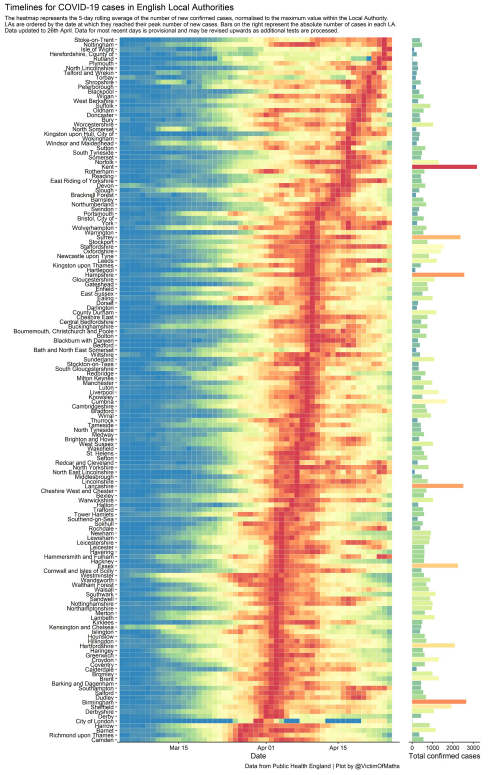
Source: John Burn-Murdoch

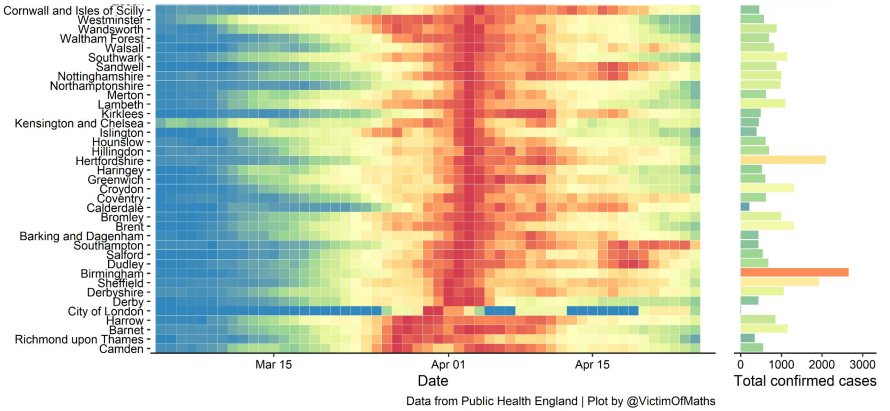
Data viz in R



Source: Colin Angus

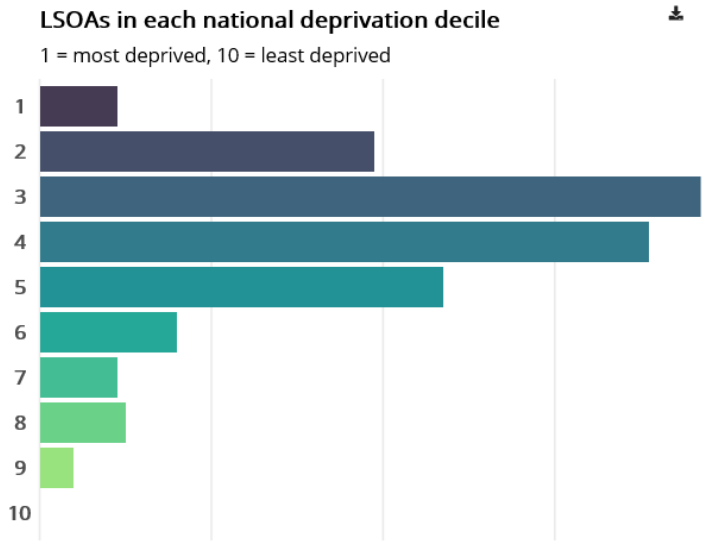
Data viz in R

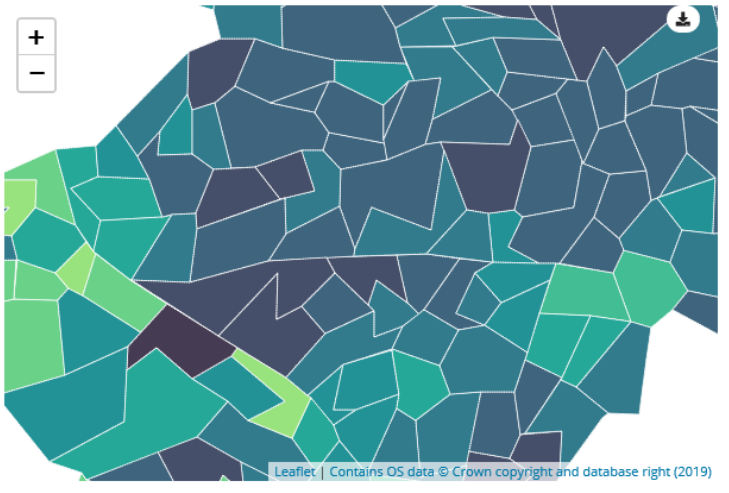




Source: Colin Angus

Data viz in R

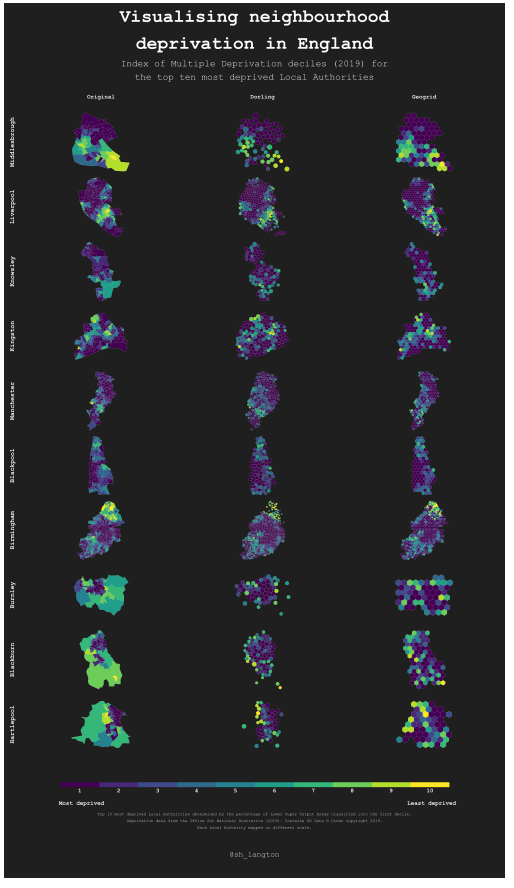




TRAFFORD COUNCIL Indices of Deprivation, 2019

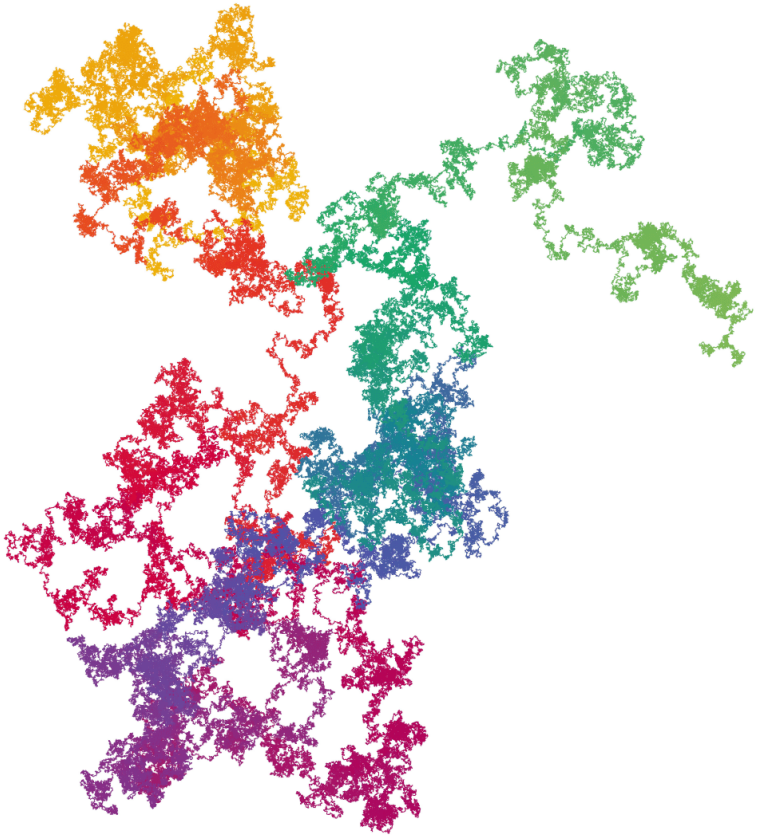
Source: Trafford Data Lab

Data viz in R



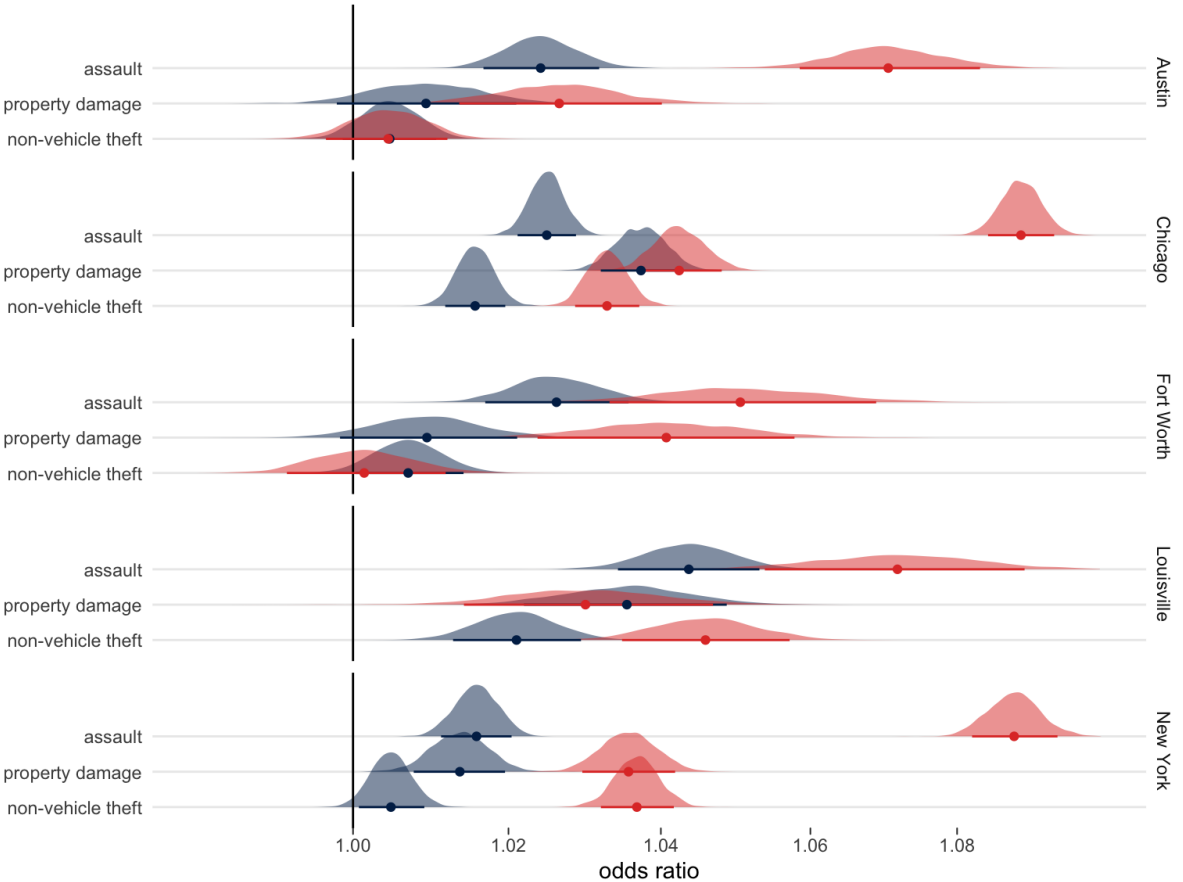
Source: Sam Langton

Data viz in R



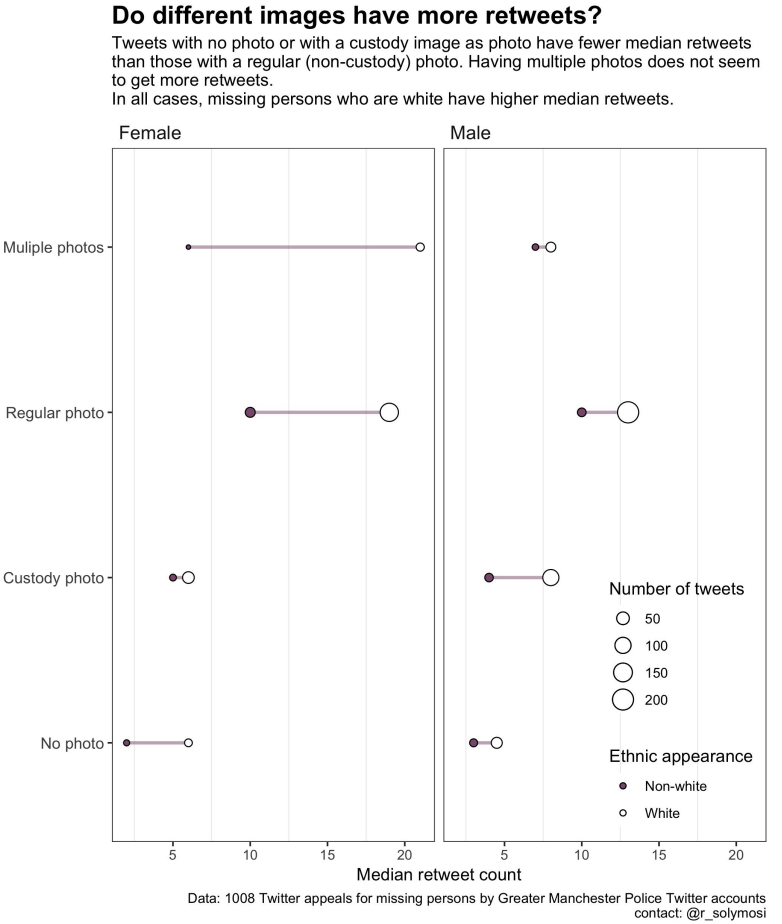
Source: Nadleh Bremer

Data viz in R using crime data



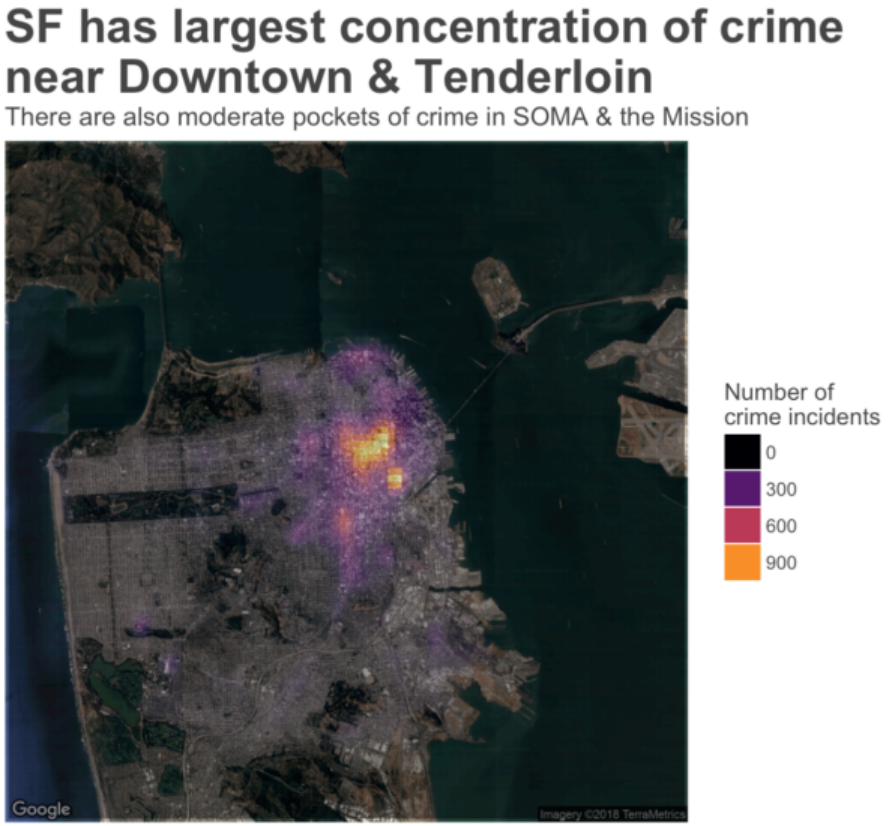
Source: Matt Ashby

Data viz in R using crime data



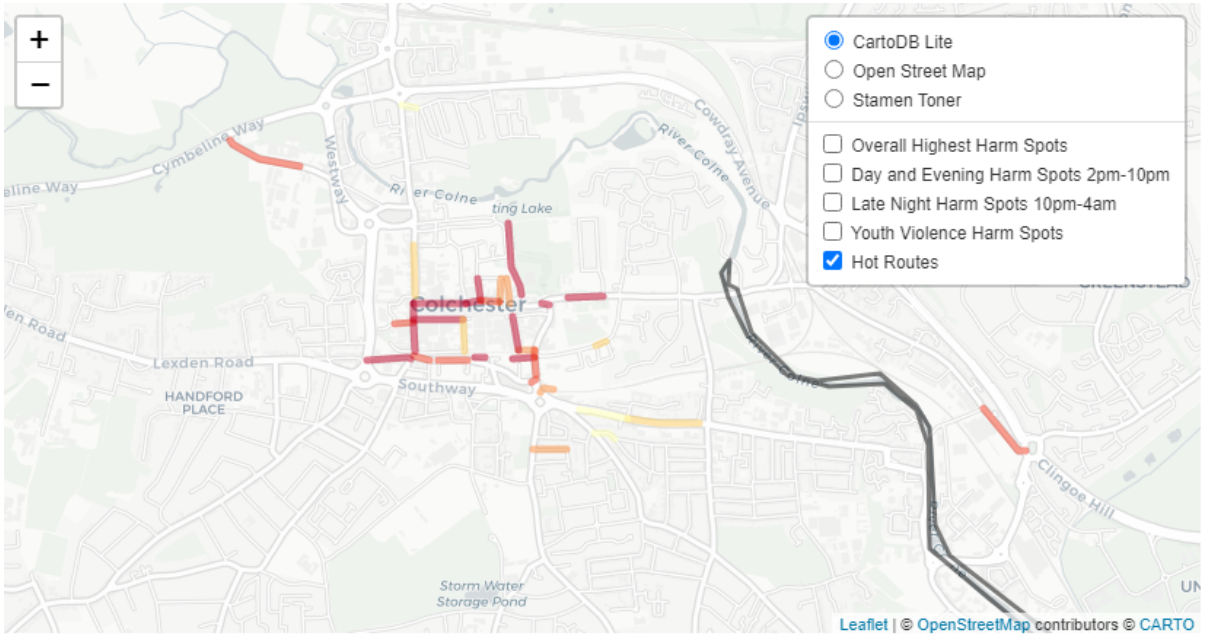
Source: Reka Solymosi

Data viz in R using crime data



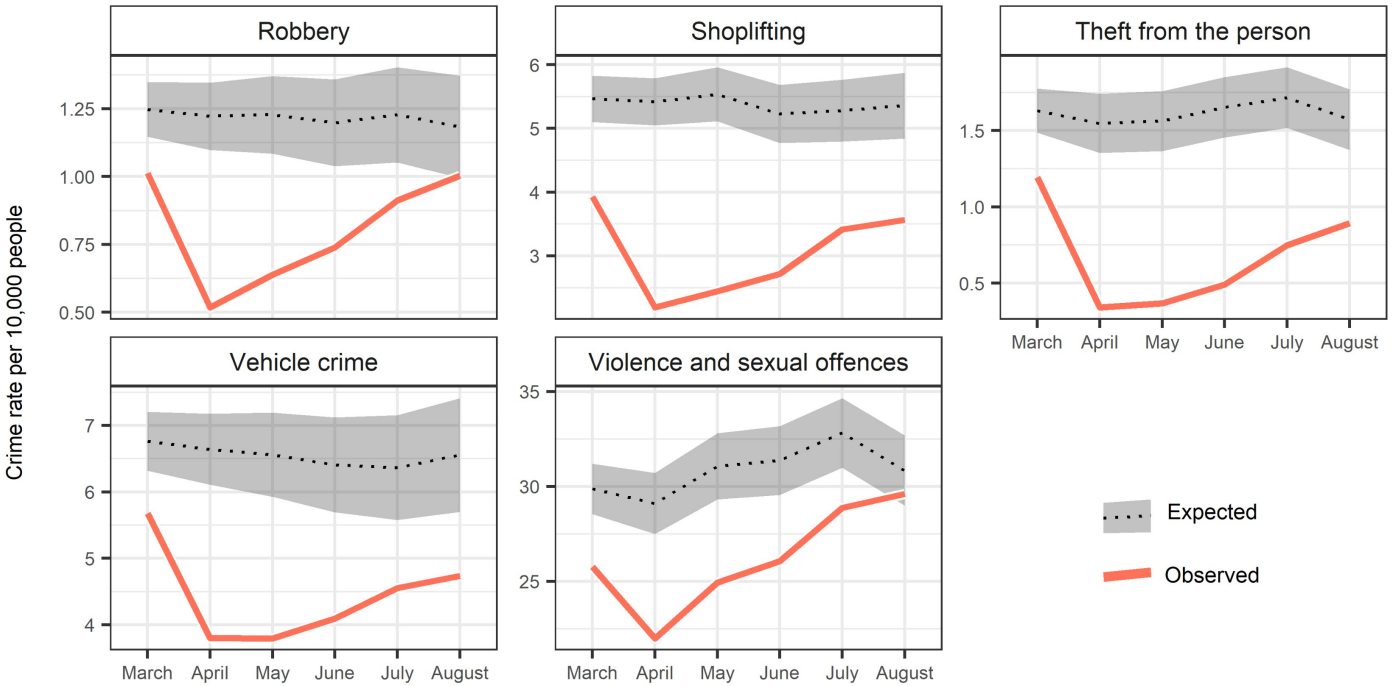
Source: Sharp Insight

Data viz in R using crime data



Source: Iain Agar

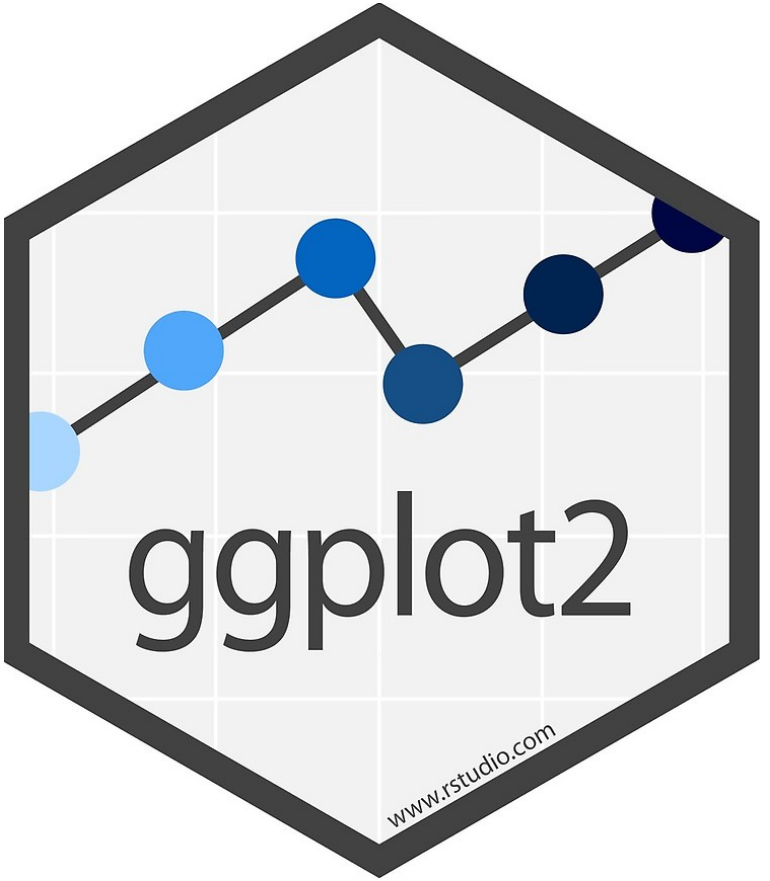
Data viz in R using crime data



Source: Langton, Dixon and Farrell (2020)

Common thread?

Common thread?



ggplot2

- ggplot2 is a package for creating graphics in R based on the grammar of graphics.
- A fundamental component of this is that graphics are made up **layers**.
- This way of thinking is reflected in how we write ggplot2 code.

Image source: Skill Gaze

ggplot2

df1

var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

ggplot2

df1

var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

What is the relationship between var1 and var2?

ggplot2: data

Image source: Skill Gaze

ggplot2: data

```
ggplot(data = df1)
```


df1

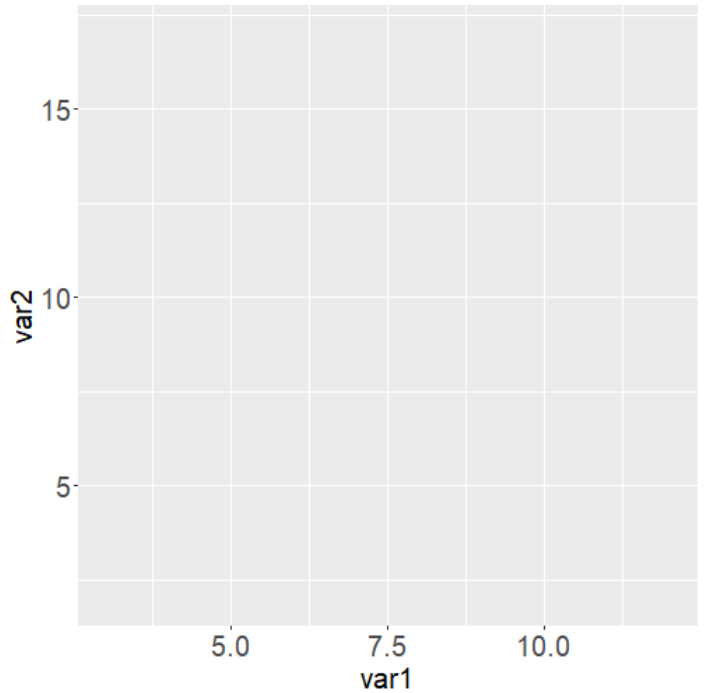
var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

ggplot2: aesthetics

Image source: Skill Gaze

ggplot2: aesthetics

```
ggplot(data = df1, mapping = aes(x = var1, y = var2))
```



df1

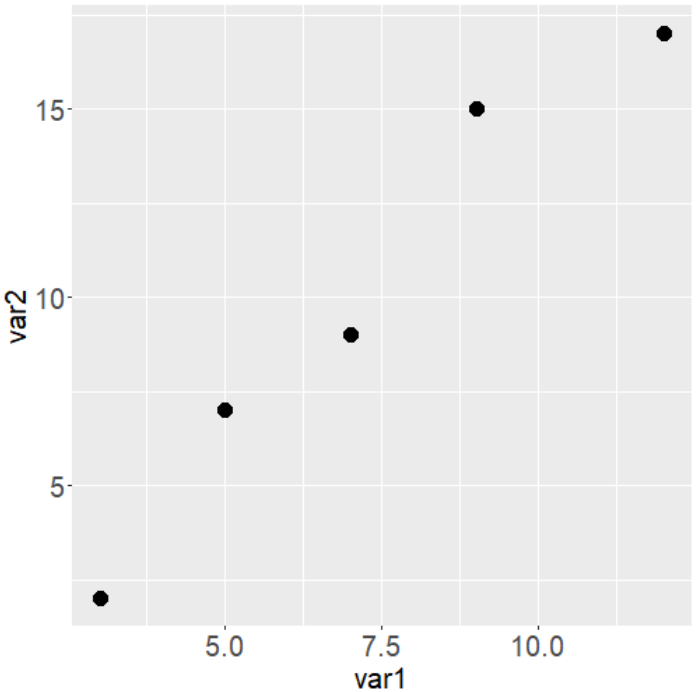
var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

ggplot2: geometries

Image source: Skill Gaze

ggplot2: geometries

```
ggplot(data = df1, mapping = aes(x = var1, y = var2)) +
  geom_point()
```



df1

var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

ggplot2: different aesthetics

df1

var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

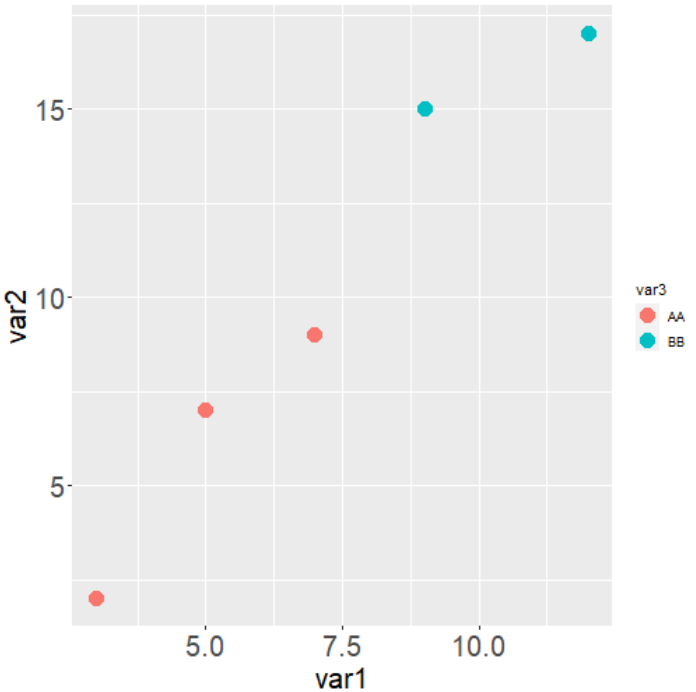
How does var3 factor into this relationship?

ggplot2: different aesthetics

- x
- y
- colour
- fill
- shape
- size
- alpha
- linetype
- ...

ggplot2: different aesthetics

```
ggplot(data = df1, mapping = aes(x = var1, y = var2, colour = var3)) +
  geom_point()
```

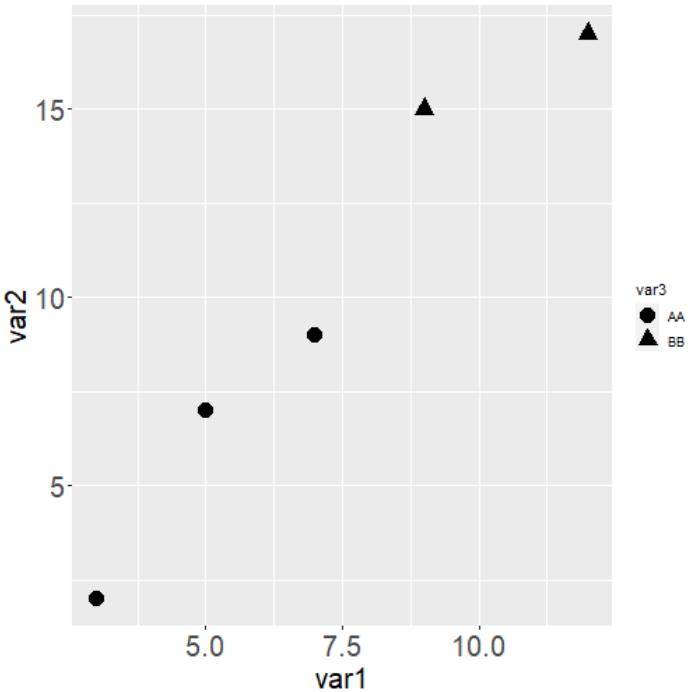


df1

var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

ggplot2: different aesthetics

```
ggplot(data = df1, mapping = aes(x = var1, y = var2, shape = var3)) +
  geom_point()
```

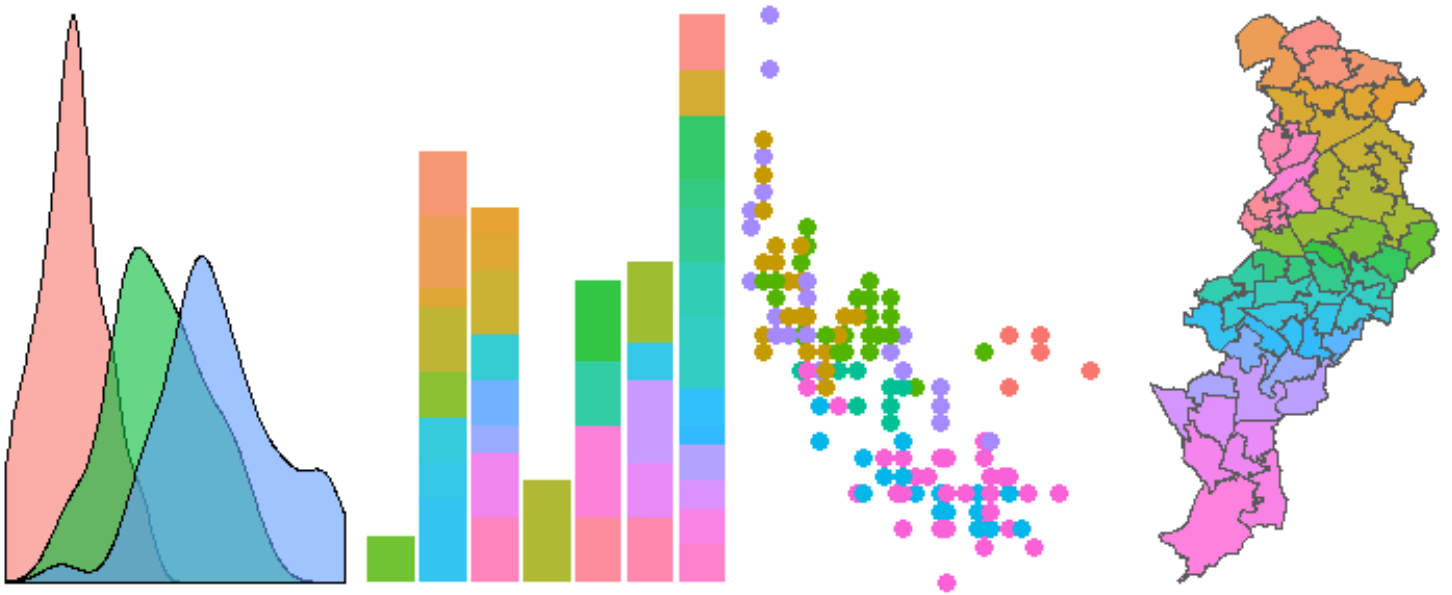


df1

var1	var2	var3
5	7	AA
3	2	AA
7	9	AA
9	15	BB
12	17	BB

ggplot2: different geometries

- geom_point()
- geom_bar()
- geom_density()
- geom_smooth()
- ...



Crime demo