

Introduction to R and RStudio

Instituto Nacional de Medicina Genómica

Daniel Rivas & Juan Manuel Mejia Arangure

August 25, 2023

Outline - Working with Data

- Working with Files
- Advanced Graphics

Working with Files

Data manipulation

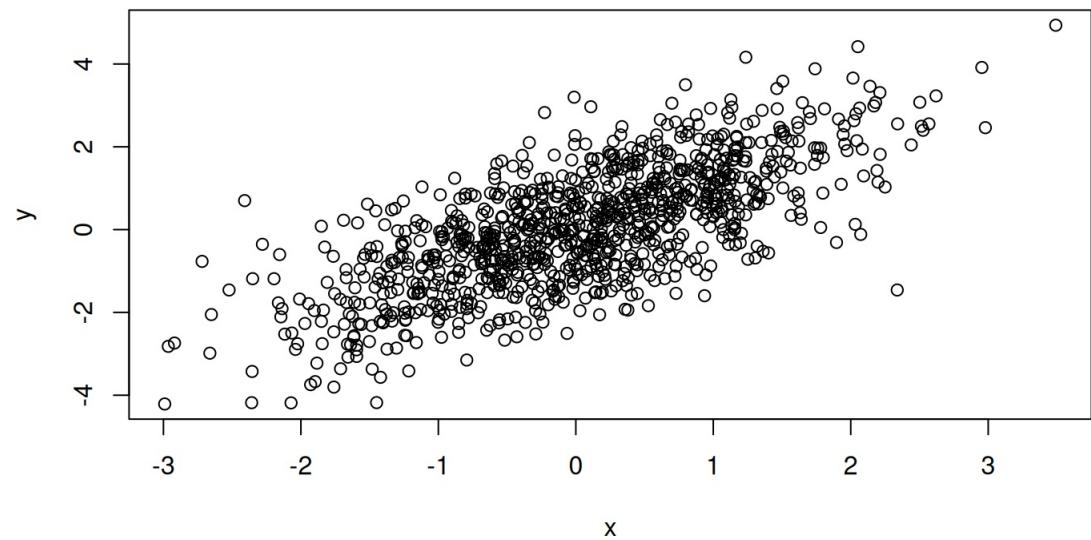
- `tidyverse` - <https://tidyverse.org>
- `dplyr` - <https://dplyr.tidyverse.org>
- `data.table` - <https://r-ddatatable.com>

Data import

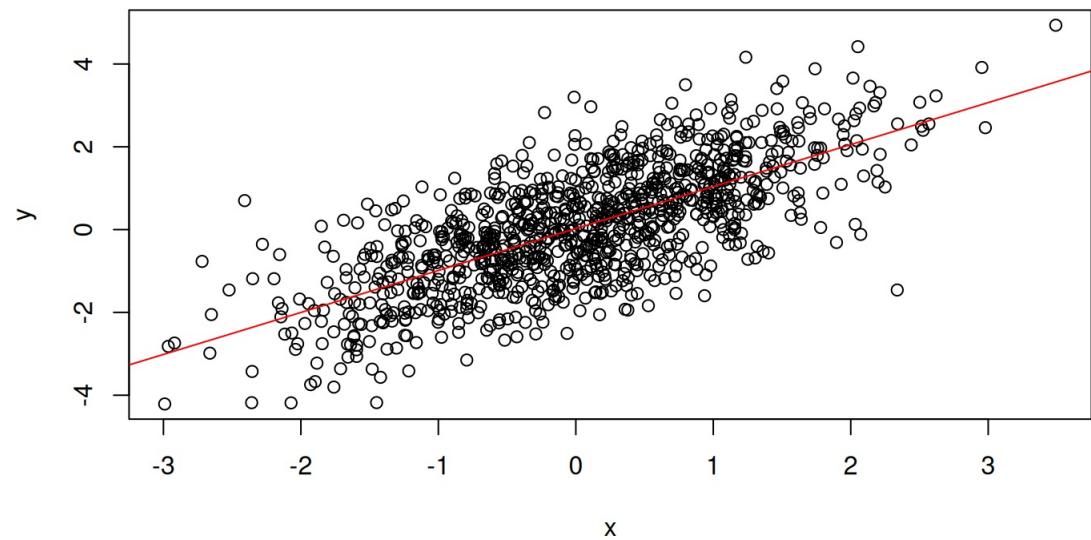
- `readr` - <https://readr.tidyverse.org>
- `readxl` - <https://readxl.tidyverse.org>
- `haven` - <https://haven.tidyverse.org>

Advanced Graphics

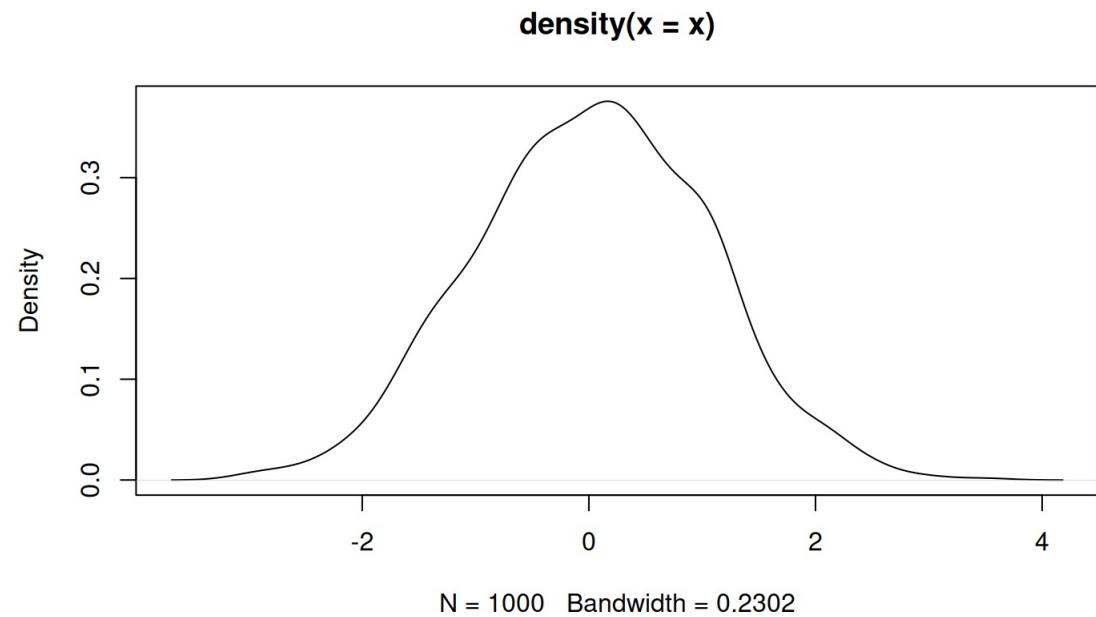
Graphics



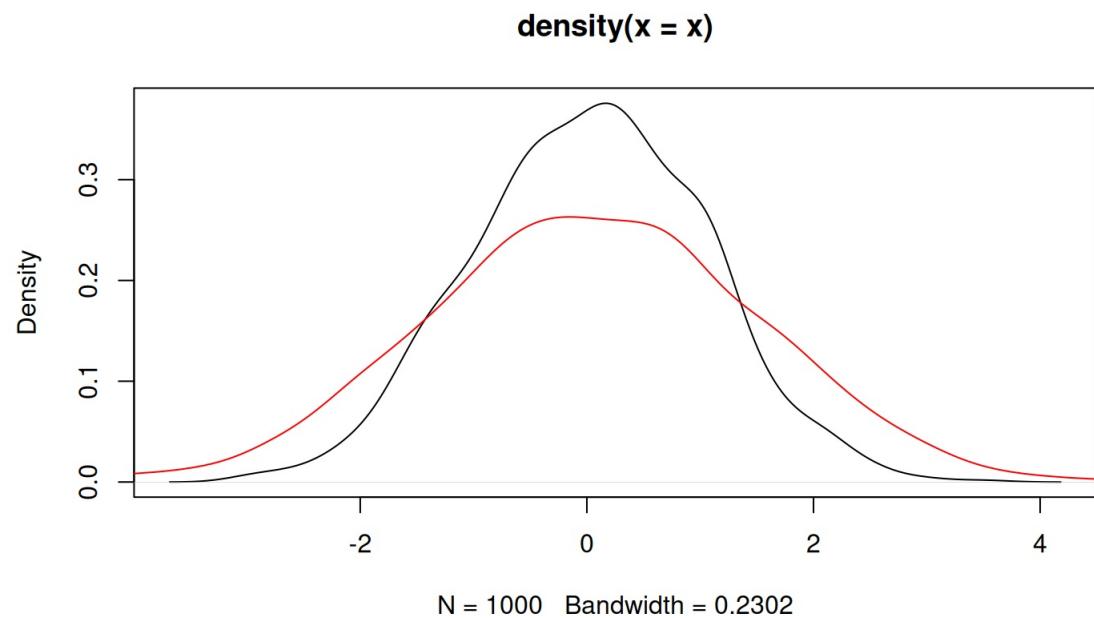
Graphics



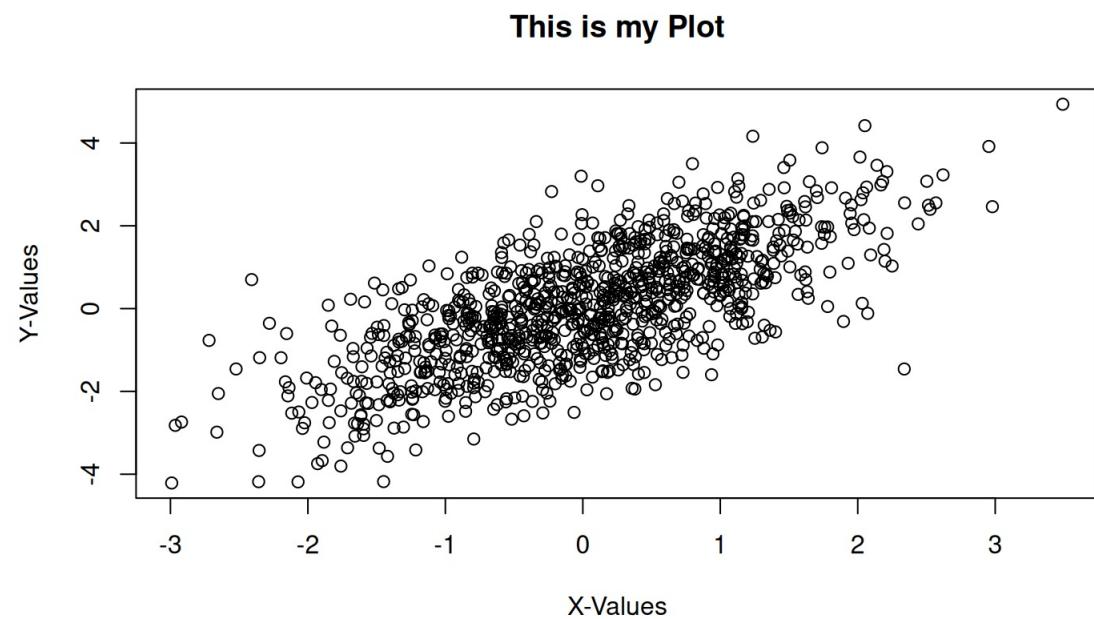
Graphics



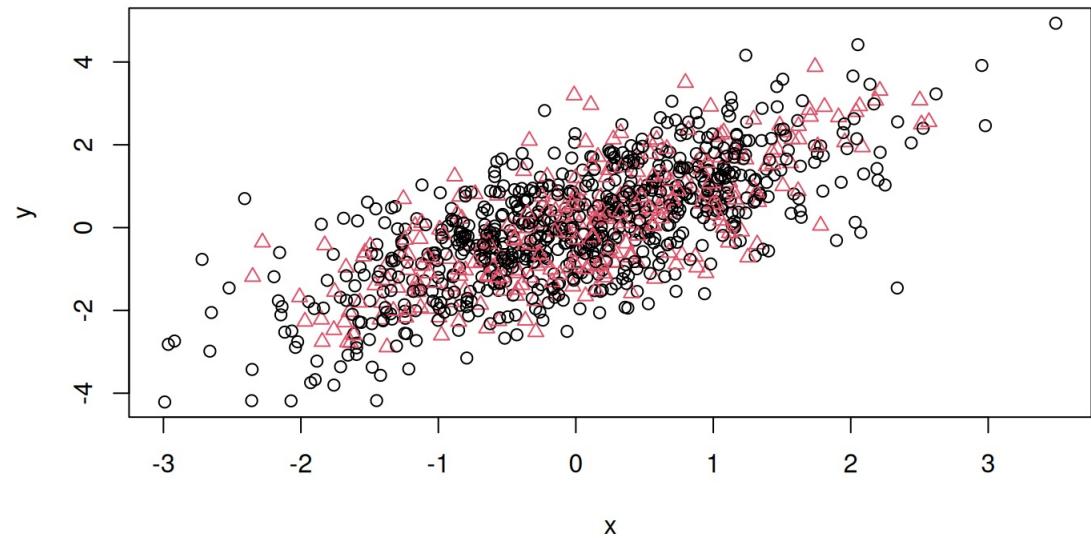
Graphics



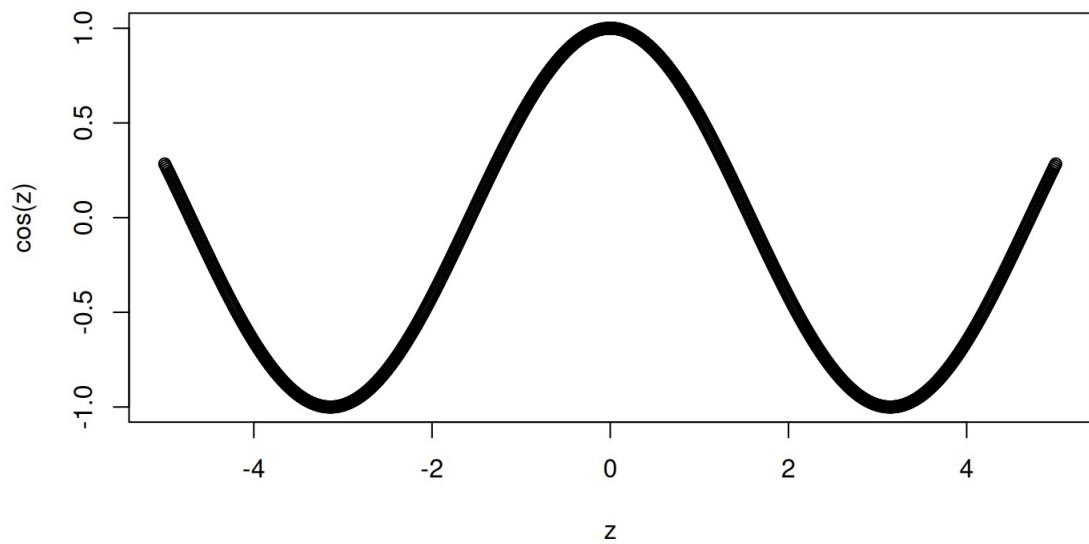
Graphics



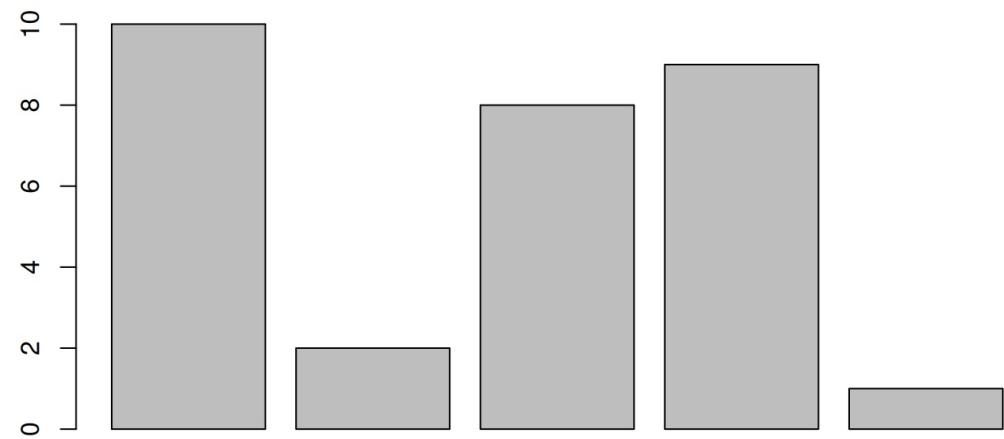
Graphics



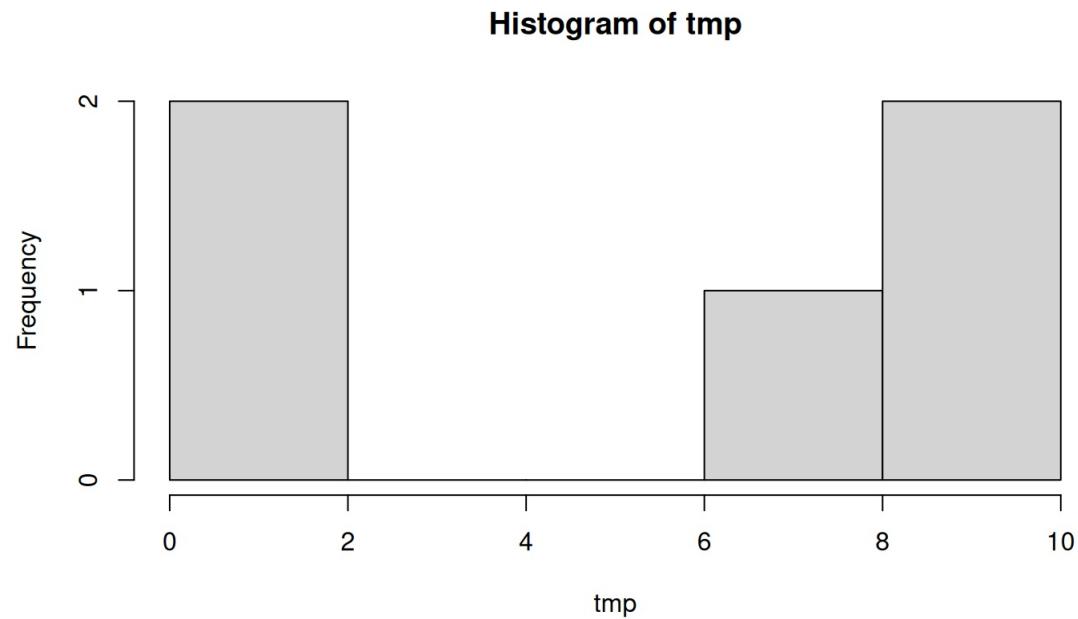
Graphics



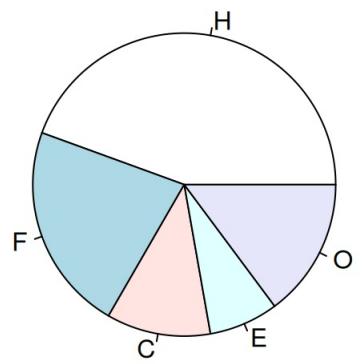
Graphics



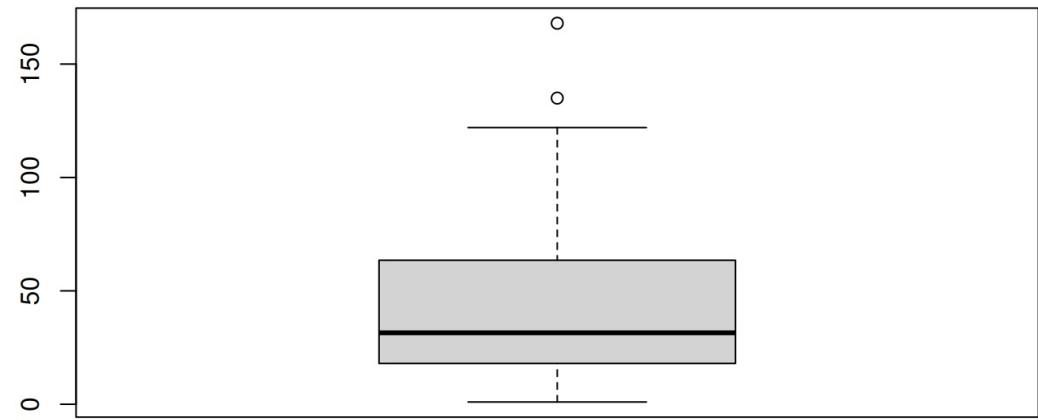
Graphics



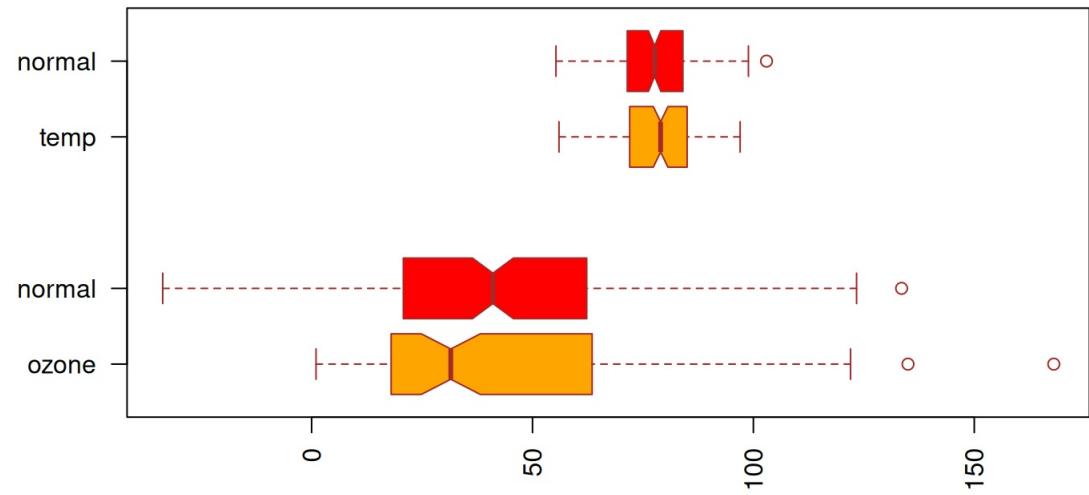
Graphics



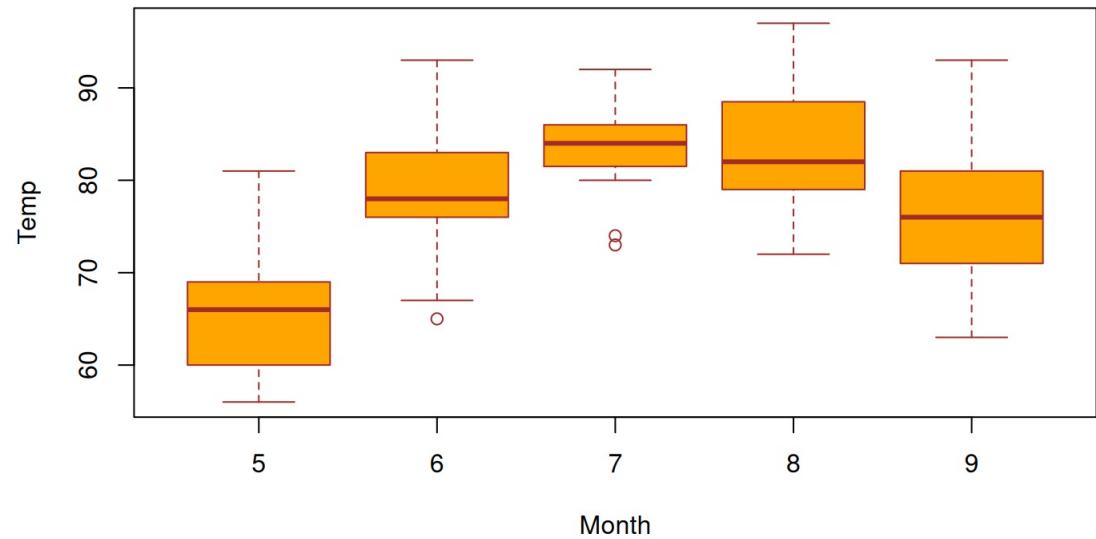
Graphics



Graphics



Graphics



Graphics

- Pie chart
- Bar chart
- Line chart
- Scatterplot

Graphics

Decompose graphics into its constituents

- theme
- coordinates
- facets
- geometries
- scales
- statistics
- mapping
- data

Graphics

- **ggplot2**

```
ggplot(  
  data = faithful,  
  mapping = aes(x = eruptions, y = waiting)  
) +  
  geom_point()
```

Which columns to use for x and y

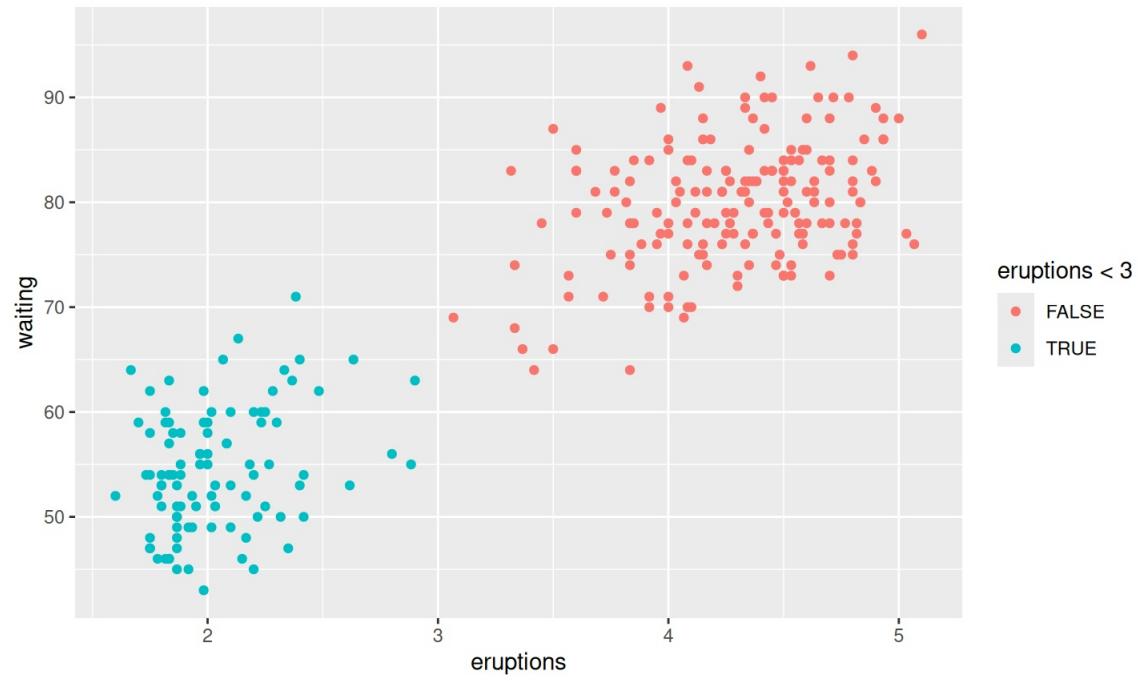
How to draw the plot

+ is used to combine **ggplot2** elements

Graphics

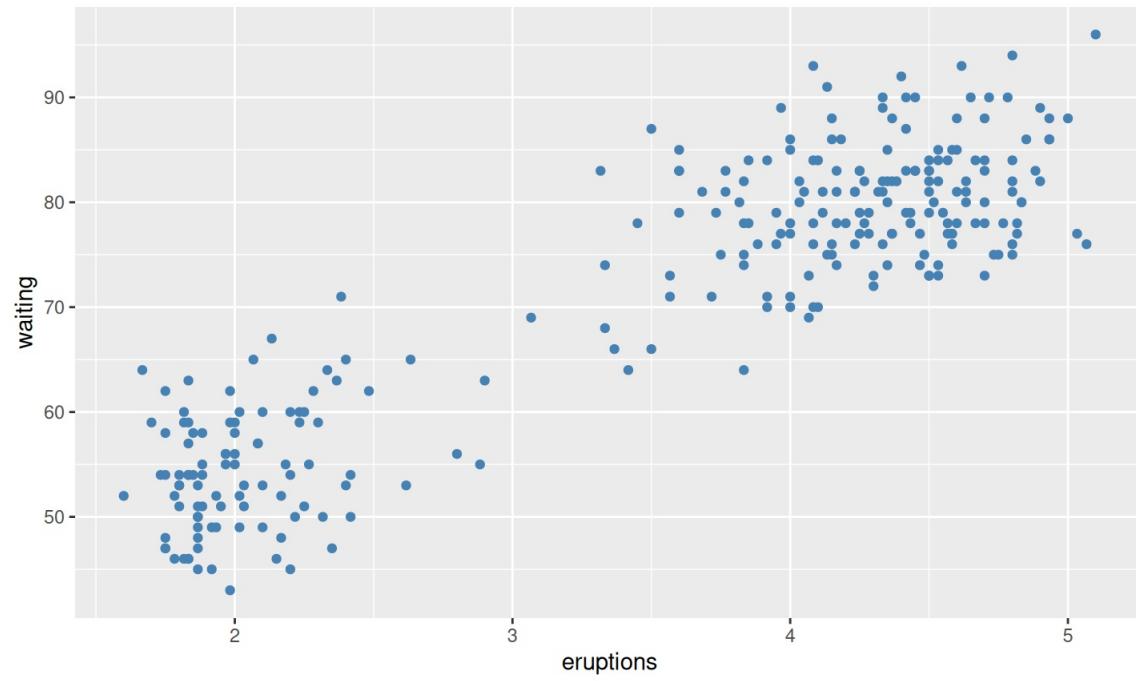
```
ggplot(  
  data = faithful,  
  mapping = aes(x = eruptions, y = waiting)  
) +  
  geom_point()  
  
ggplot(data = faithful) +  
  geom_point(  
    mapping = aes(x = eruptions, y = waiting)  
)  
  
ggplot() +  
  geom_point(  
    mapping = aes(x = eruptions, y = waiting),  
    data = faithful  
)
```

Graphics



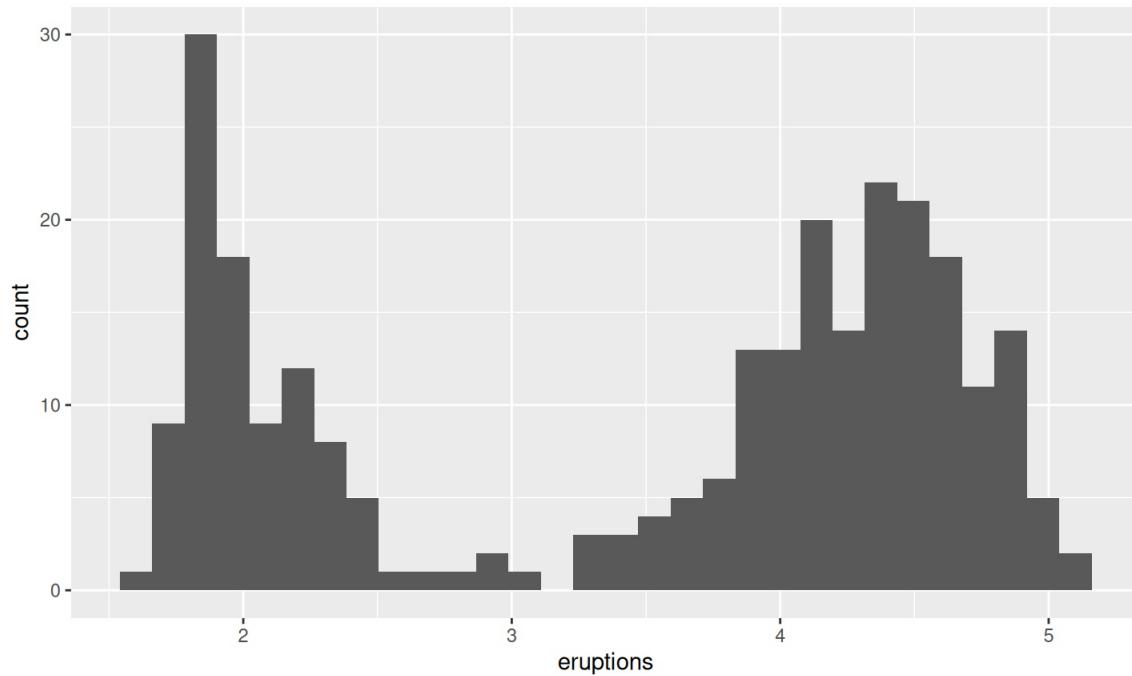
Mapping color

Graphics



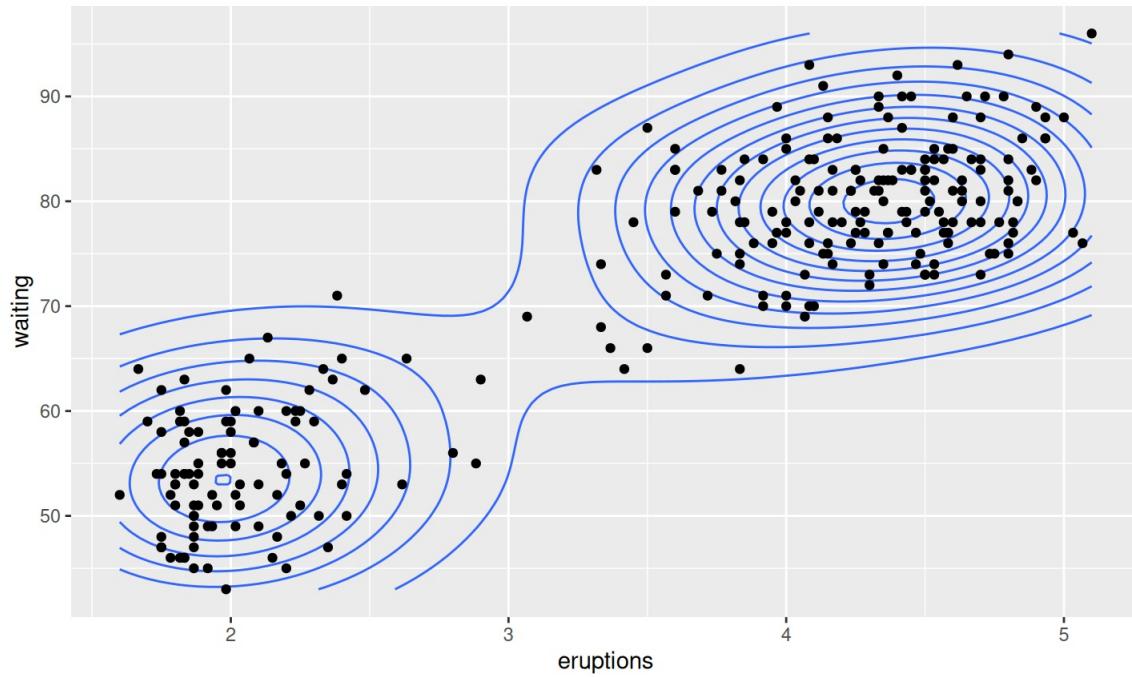
Setting color

Graphics



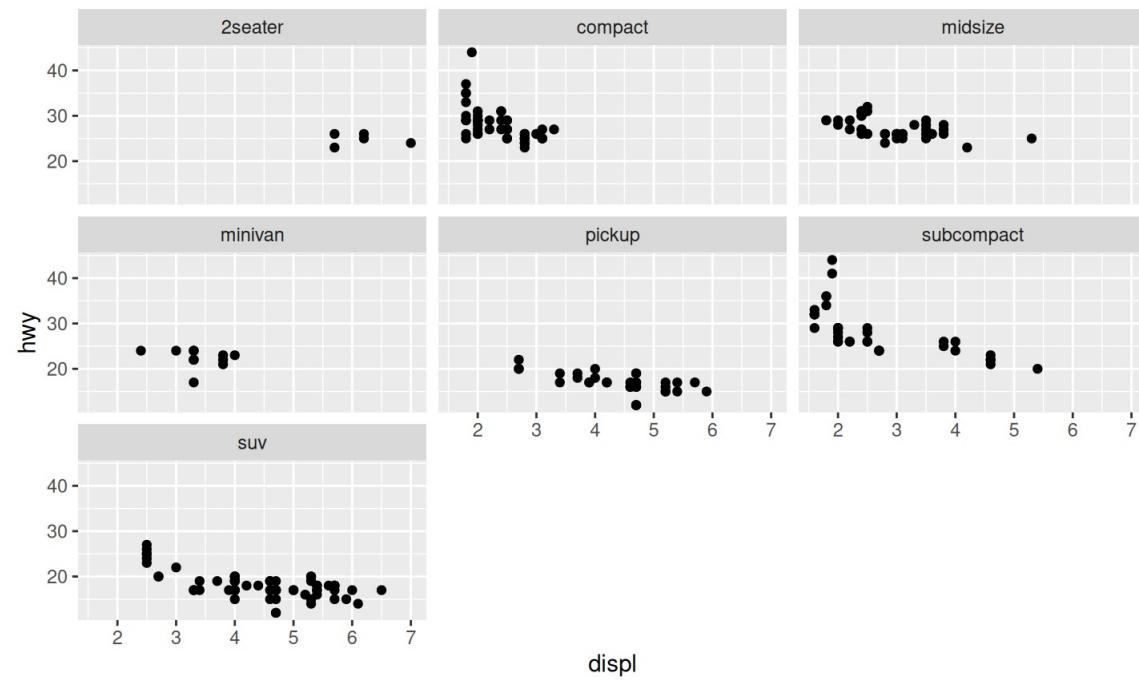
There are many types of **geoms**, but their mapping requirements differ

Graphics



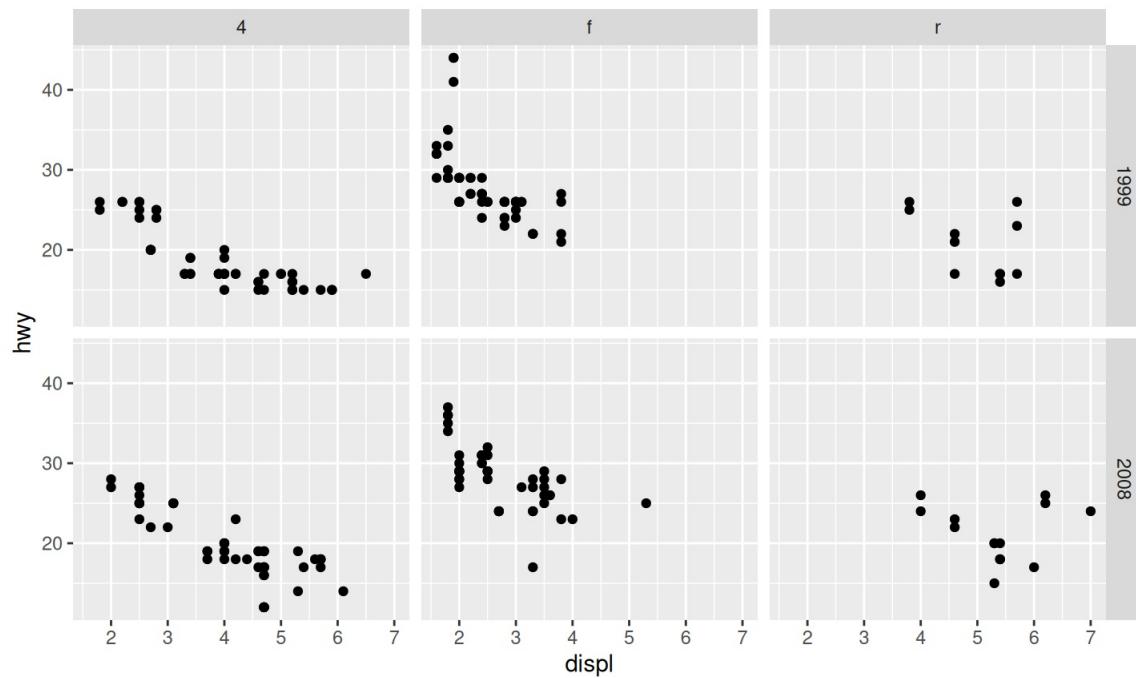
Layers are stacked in the order of code appearance

Graphics



Faceting is often the best way to avoid overplotting

Graphics



`facet_grid()` provides a way of doing graphic pivots

Beyond *ggplot2*

`ggplot2` is huge!

Maintenance is already a team effort

- 47 geom
- 25 stats
- 62 scales

Many extensions niche specific & better developed by filed experts

R Gallery

<https://r-graph-gallery.com/>

ggstatsplot

<https://github.com/IndrajeetPatil/ggstatsplot>

Daniel Rivas

danielrivasmd@gmail.com

Questions & Comments