Introduction au package **ggplot2**

Sébastien Déjean – math.univ-toulouse.fr/~sdejean

Préambule

- Pré-requis :
  - utiliser R régulièrement
  - oublier la façon dont vous produisiez les graphiques jusqu'à présent

- Objectifs :
  - Comprendre la logique de la ***g**rammar of **g**raphics* (**gg** de **gg**plot2)
  - S'initier à `ggplot2`
  - Apprendre à se débrouiller
  - Se faire plaisir !

Plan

- Grammar of Graphics
- `ggplot2`

Grammar of Graphics

Grammar of Graphics

- Grammaire : nom féminin, XIIe siècle. Dérivé du latin grammatica, emprunté du grec grammatikê, « grammaire, culture ». Ensemble des règles qui forment le système d'une langue et que l'on doit suivre pour parler ou écrire conformément à l'usage (Dictionnaire de l'académie Française, 9e édition)

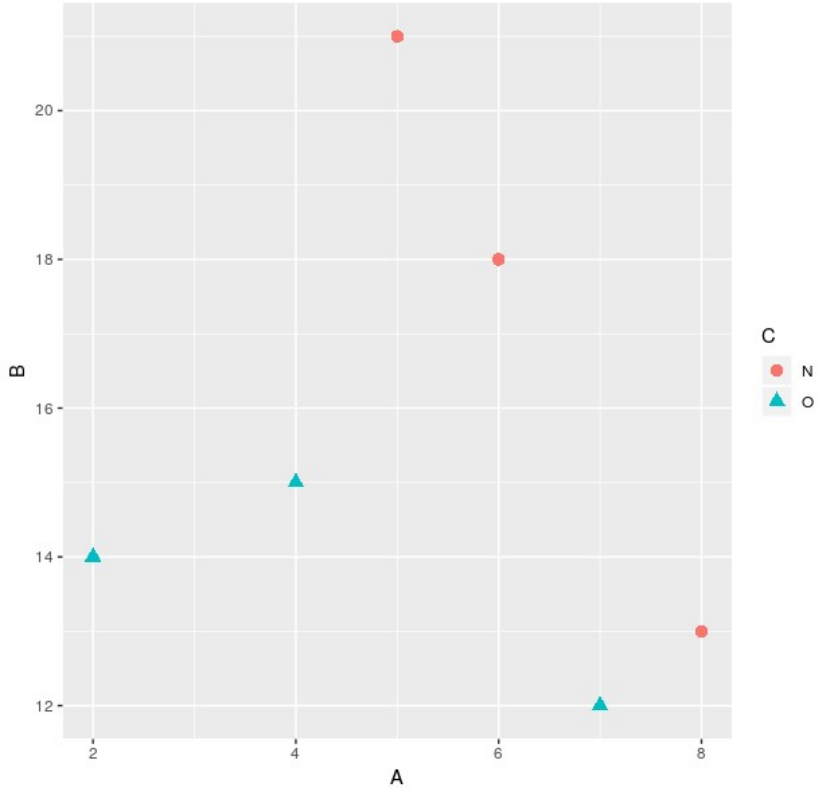
- *Grammar gives language rules*. Wilkinson 2005

L. Wilkinson. The Grammar of graphics. Statistics and Computing. Springer, 2nd edition, 2005.

Grammar of Graphics

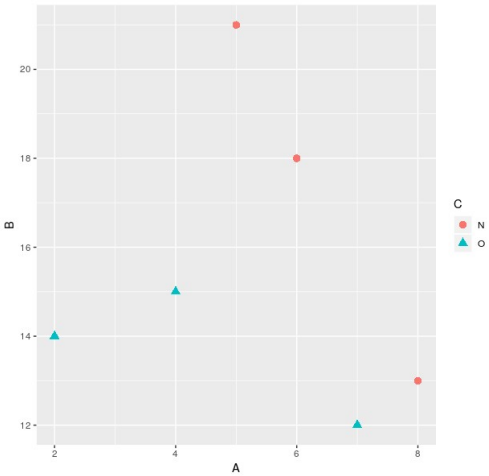
- Ensemble de règles qui vont permettre de passer d'un tableau de données à un graphique

A	B	C
2	14	0
4	15	0
5	21	N
6	18	N
8	13	N
7	12	0



Grammar of Graphics

A	B	C	x	y	Colour	Shape	Size
2	14	O	0	0.2	Blue	Circle	3
4	15	O	0.3	0.3	Blue	Circle	3
5	21	N	0.5	1	Pink	Triangle	3
6	18	N	0.6	0.6	Pink	Triangle	3
8	13	N	1	0.1	Pink	Triangle	3
7	12	O	0.8	0	Blue	Circle	3



Variable space

Aesthtetic space

The grammar tells us that a statistical graphic is a mapping from data to aesthetic attributes (colour, shape, size) of geometric objects (points, lines, bars).

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Use R!, Springer
https://ggplot2-book.org

Grammar of Graphics

- **Defaults**
  - Data
  - Mapping from variables to aesthetics
- **Layer**
  - Data (optional)
  - Mapping (optional)
  - **Geom**etric object
  - **Stat**istical transformation
  - **Position** adjustement
- **Scale**: controls the mapping from data to aesthetic attributes, one scale or each aesthetic property
- **Coord**inate system
- **Facet** specification : to build the same plot for different subsets of the dataset

Layered Grammar of Graphics

Wickham, Hadley. 2008. “Practical Tools for Exploring Data and Models.” PhD thesis, Iowa State University. `http://had.co.nz/thesis.`

Aesthetics

- Properties that can be perceived on the graphic
- Each aesthetic can be mapped to a variable or set to a constant value

Geoms

- Performs the actual rendering of the layer
- Control the type of plot that you create

Layer

Layers are responsible for creating the objects that we perceive on the plot.

The plot is not ready to be displayed until at least one layer is added.

- Data
- **Aes**thetic mapping
- **Geom**etric object
- **Stat**istical transformation
- **Position** adjustement

Minimal layer: specify a geom

Wickham, Hadley. 2008. “Practical Tools for Exploring Data and Models.” PhD thesis, Iowa State University. `http://had.co.nz/thesis`.

Scales

- Controls the mapping from data to aeshetics
- Tools to read the plot: axes and legends (guides)
- Scale : function from a region in data space to a region in aesthetic space.
- Without a scale, there is no way to go from the data to aesthetics so a scale is required for every aesthetics used on the plot
- For position aesthetics, the axes are the guides, for all other aesthetics, legends do the job
- All aspects of the guides are controlled by parameters of the scales

ggplot2

Input data

It must be a data.frame

Data structures

- A plot object **is a list** with:
  - Data (stored in the plot object as a copy not a reference)
  - Mapping
  - Layers
  - Scales
  - Coordinates
  - facet

ggplot2

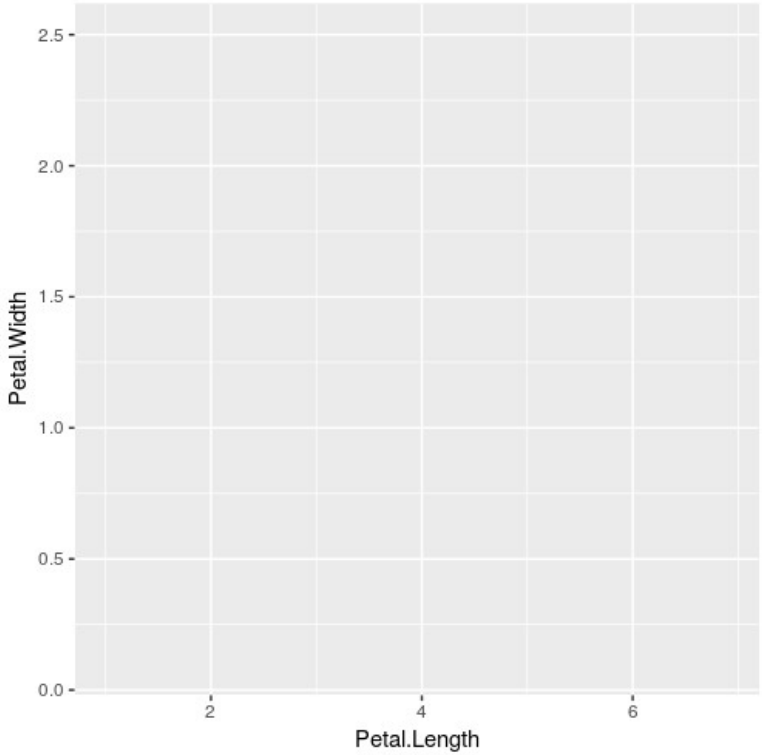
```
> ggplot(iris)
```

Data

ggplot2

```
> ggplot(data=iris,
         aes(x=Petal.Length,
             y=Petal.Width,
             colour=Species))
```

Aesthetics



ggplot2

```
> ggplot(data=iris,
         aes(x=Petal.Length,
             y=Petal.Width,
             colour=Species))
  + geom_point()
```

Layer with geom