

Un puissant package de visualisation

Benjamin Louis

15/10/2019 (MàJ: 18/10/2019)

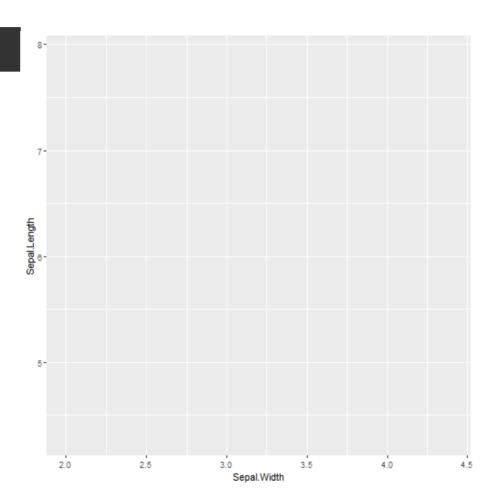
ggplot2

- Développé par Hadley Wickham
- Package du *Tidyverse*
- Se veut élégant et polyvalent
- Applique le *grammar of graphics* sytème
- Template:

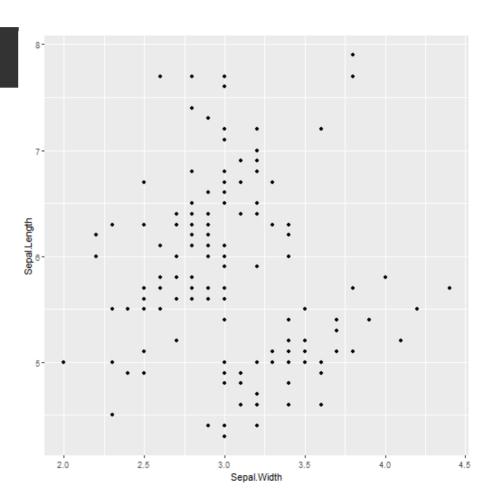
```
ggplot(data = mon_data) +
  aes(x = var_x, y = var_y) +
  geom_***() +
  ...
```

ggplot(iris)

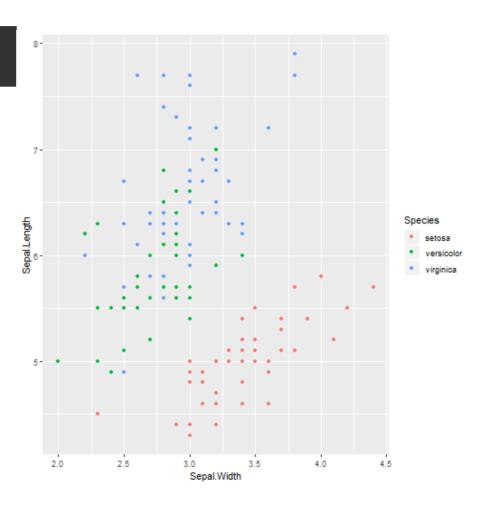
```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length)
```



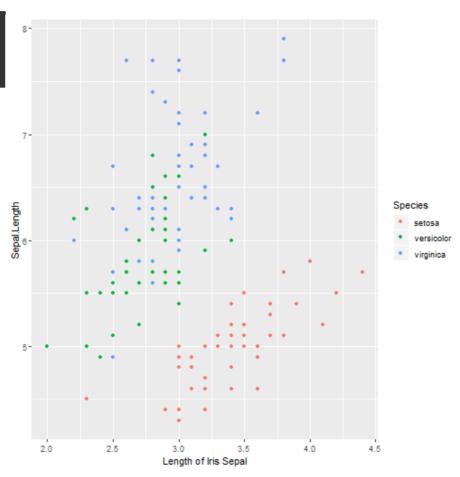
```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length) +
  geom_point()
```



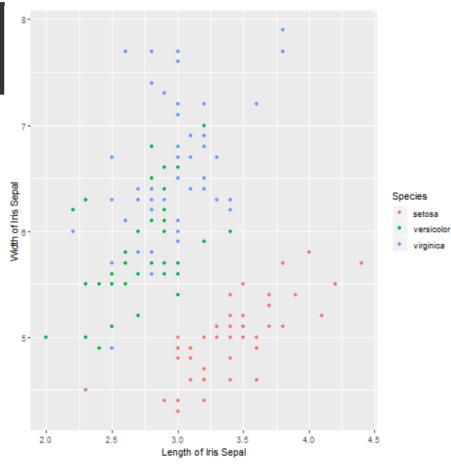
```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length) +
  geom_point(aes(color = Species))
```

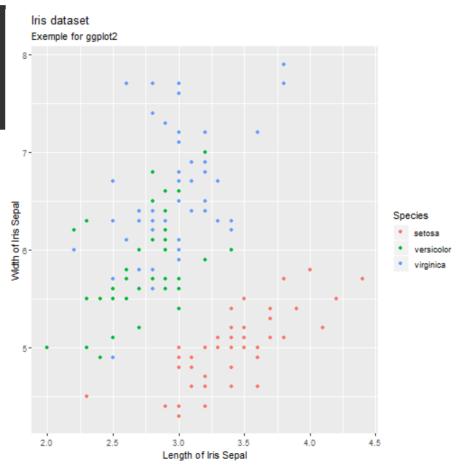


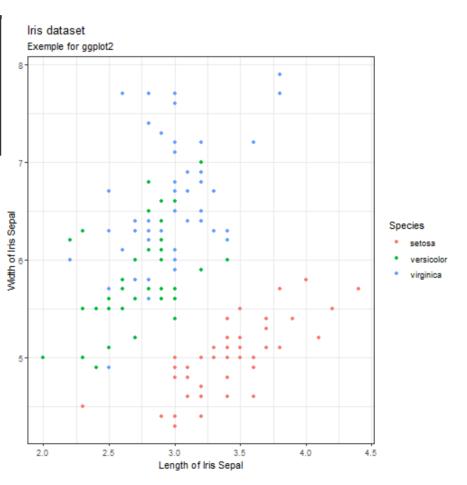
```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length) +
  geom_point(aes(color = Species)) +
  labs(x = "Length of Iris Sepal")
```

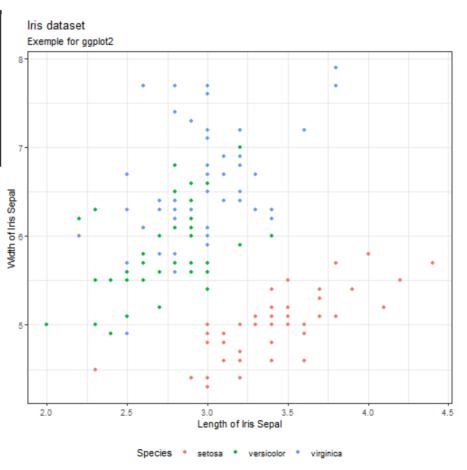


```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length) +
  geom_point(aes(color = Species)) +
  labs(x = "Length of Iris Sepal") +
  labs(y = "Width of Iris Sepal")
```

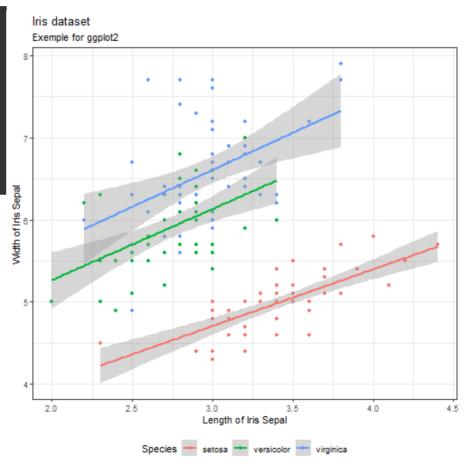








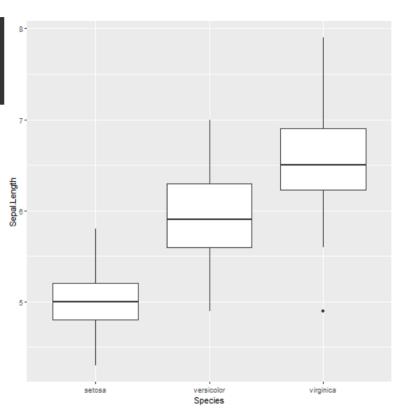
```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length,
      color = Species) +
  geom_point() +
  labs(x = "Length of Iris Sepal") +
  labs(y = "Width of Iris Sepal") +
  labs(title = "Iris dataset",
      subtitle = "Exemple for ggplot2") +
  theme_bw() +
  theme(legend.position = "bottom") +
  geom_smooth(method = "lm")
```



Autres geom_*

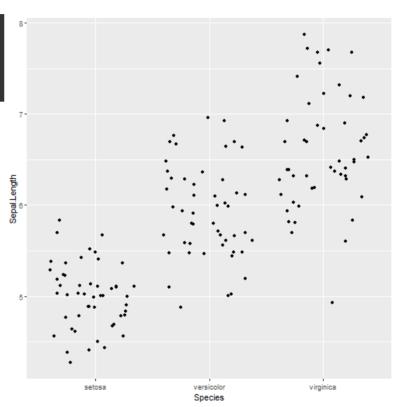
geom_boxplot()

```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_boxplot()
```



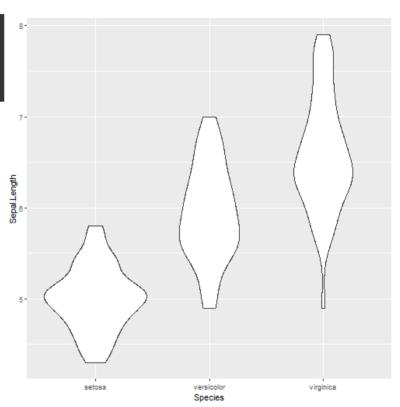
geom_jitter()

```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_jitter()
```



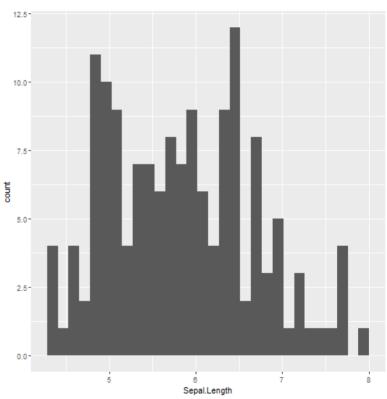
geom_violin()

```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_violin()
```



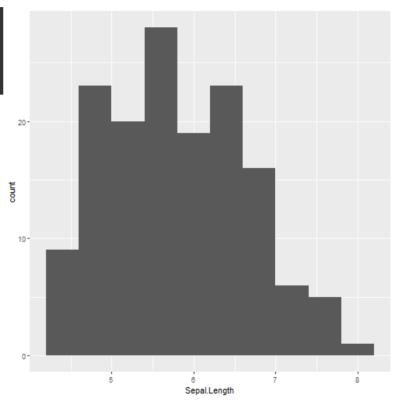
geom_histogram()

```
ggplot(iris) +
  aes(x = Sepal.Length) +
  geom_histogram()
```



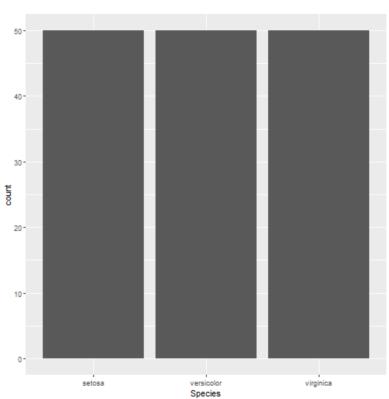
geom_histogram()

```
ggplot(iris) +
  aes(x = Sepal.Length) +
  geom_histogram(bins = 10)
```



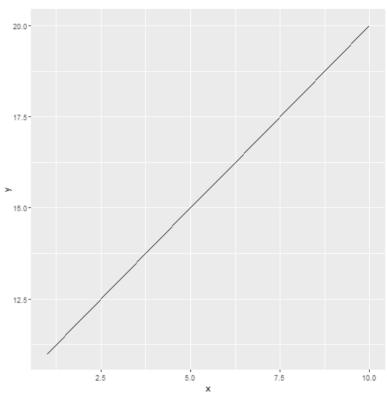
geom_bar()

```
ggplot(iris) +
  aes(x = Species) +
  geom_bar()
```



geom_line()

```
dt <- tibble(x = 1:10, y = 11:20)
ggplot(dt) +
  aes(x = x, y = y) +
  geom_line()</pre>
```



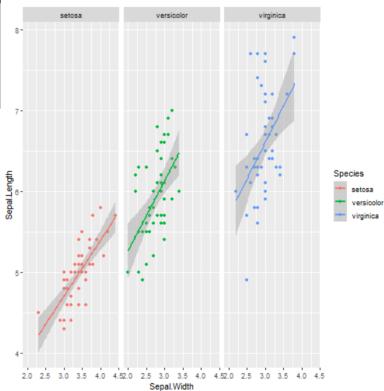
Autres

```
geom_smooth()
• geom_freqpoly()
• geom_abline()
• geom_hline()
• geom_vline()
• geom_polygon()
• geom_text()
• geom_label()
```

Facetting

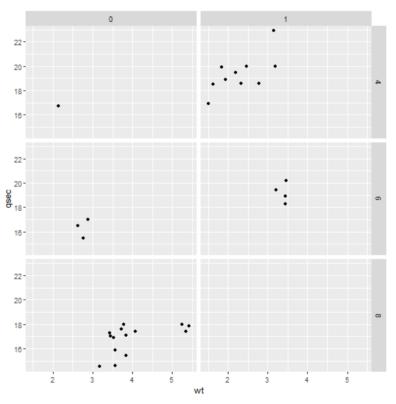
facet_wrap()

```
ggplot(iris) +
  aes(x = Sepal.Width, y = Sepal.Length,
      color = Species) +
  geom_point() +
  geom_smooth(method = "lm") +
  facet_wrap(~Species)
```



facet_grid()

```
ggplot(mtcars) +
  aes(x = wt, y = qsec) +
  geom_point() +
  facet_grid(cyl ~ vs)
```

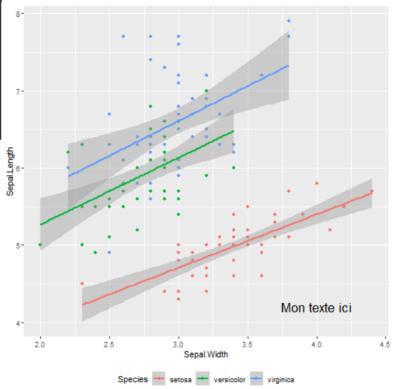


Controler les échelles

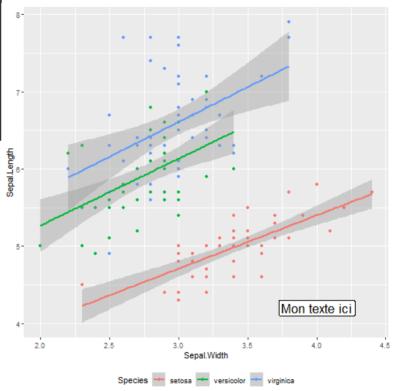
- scales = "fixed: échelles fixes pour tous les graphiques
- scales = "free_x: abscisses libres, ordonnées fixes
- scales = "free_y: abscisses fixes, ordonnées libres
- scales = "free : échelles libres pour tous les graphiques

Annotation

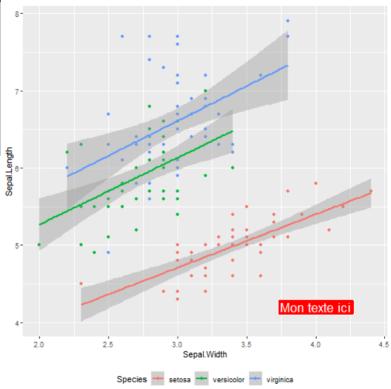
annotate()



annotate()

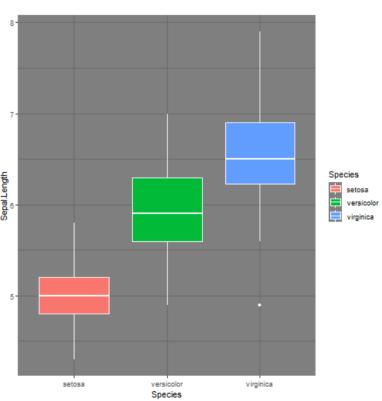


annotate()

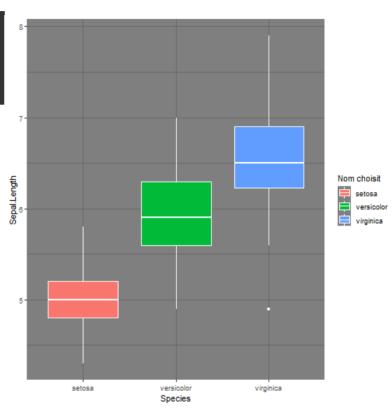


Couleurs & légendes

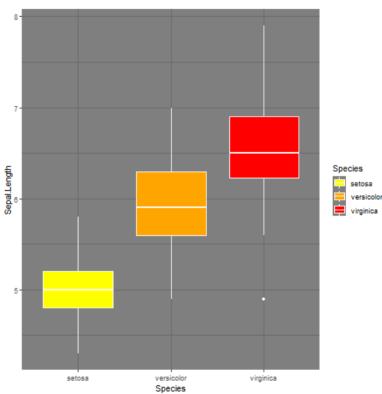
```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_boxplot(aes(fill = Species), cold
  theme_dark()
```



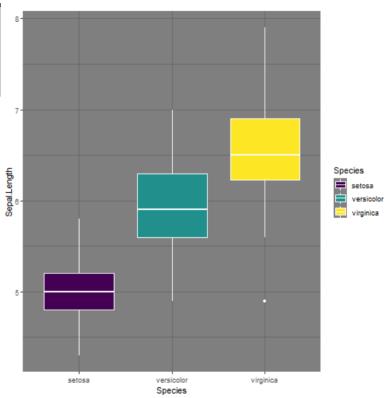
```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_boxplot(aes(fill = Species), color = "white") +
  theme_dark() +
  scale_fill_discrete(name = "Nom choisit")
```



```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_boxplot(aes(fill = Species), color = "white") +
  theme_dark() +
  scale_fill_manual(values = c("yellow", "orange", "red")
```



```
ggplot(iris) +
  aes(x = Species, y = Sepal.Length) +
  geom_boxplot(aes(fill = Species), color = "white") +
  theme_dark() +
  scale_fill_viridis_d()
```



^{*}https://rtask.thinkr.fr/fr/ggplot2-welcome-viridis/

Extensions

R graph gallery

https://www.r-graph-gallery.com/

ggplot2 extensions gallery

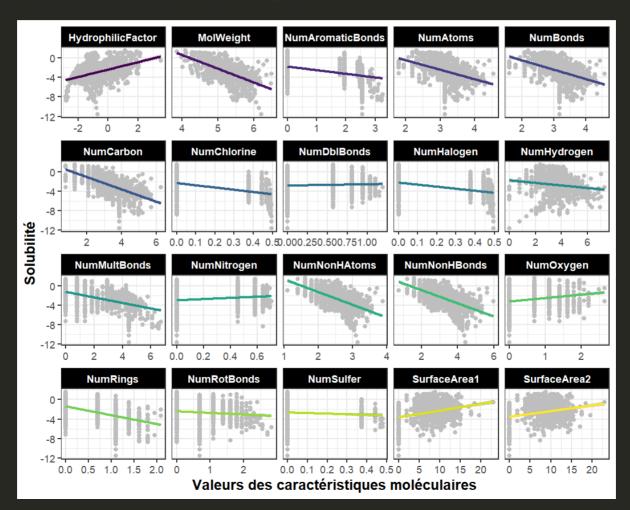
https://www.ggplot2-exts.org/gallery/

Exercice 1

À partir de *tidywoo*, faire différents graphiques pour prendre ne main ggplot2

Exercice 2

À partir du jeu de données solubility.csv



Correcion exercice 2

```
# Importation
dataset <- read_delim(here::here("static/data/solubility.csv"), ";")</pre>
# manipulation
dataset %>%
  pivot_longer(MolWeight:SurfaceArea2, names_to = "caract",
               values to = "value") %>%
  # Graphique
  ggplot() +
  aes(x = value, v = solubility) +
  geom_point(color = "grey") +
  facet_wrap(~caract, ncol = 5, scales = "free_x") +
  geom_smooth(aes(color = caract), show.legend = FALSE, method = "lm") +
  theme_bw() +
  theme(strip.background = element_rect(fill = "black"),
        strip.text = element_text(color = "white", face = "bold"),
        axis.title = element_text(size = 12, face = "bold")) +
  scale_color_viridis_d(option = "D") +
  labs(y = "Solubilité", x = "Valeurs des caractéristiques moléculaires")
```