

Exemple 2 : Création d'un document .pdf

Amandine Blin
UAR 2700 2AD, Service Analyse de Données

16/11/2023

Table des matières

1	Chargement du jeu de données iris	2
2	Description du jeu de données <i>iris</i>	2

1 Chargement du jeu de données iris

```
library(tidyverse)
```

- Chargez le jeu de données iris disponible dans les packages de base.

```
data(iris)
head(iris)
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa

2 Description du jeu de données *iris*

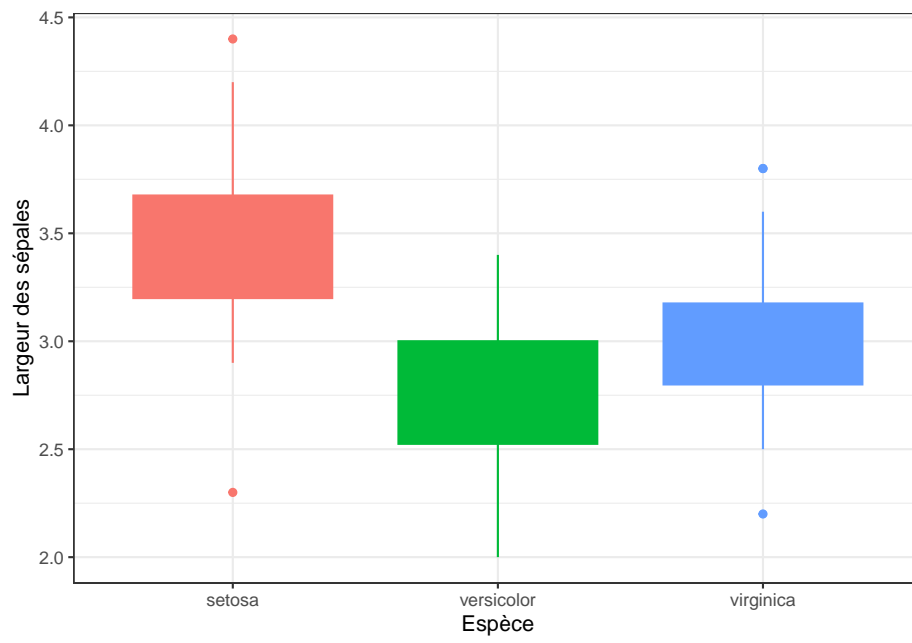
- Calculons dans un tableau la moyenne, la variance, l'écart-type, la médiane de la longueur des sépales pour chaque espèce.

```
iris %>% group_by(Species) %>%
  summarise(Moyenne_longueur_sepale = mean(Sepal.Length),
            SD_longueur_sepale=sd(Sepal.Length),
            mediane_longueur_sepale=median(Sepal.Length))
```

Species	Moyenne_longueur_sepale	SD_longueur_sepale	mediane_longueur_sepale
setosa	5.006	0.3524897	5.0
versicolor	5.936	0.5161711	5.9
virginica	6.588	0.6358796	6.5

- Effectuons le boxplot de la largeur des sépales en fonction des espèces.

```
iris %>% ggplot(aes(x=Species, y=Sepal.Width)) +
  geom_boxplot(aes(colour=Species, fill=Species), show.legend = FALSE) +
  theme_bw() + xlab("Espèce") + ylab("Largeur des sépales")
```



Références

- [1] Hadley WICKHAM et al. « Welcome to the tidyverse ». In : *Journal of Open Source Software* 4.43 (2019), p. 1686. DOI : 10.21105/joss.01686.