

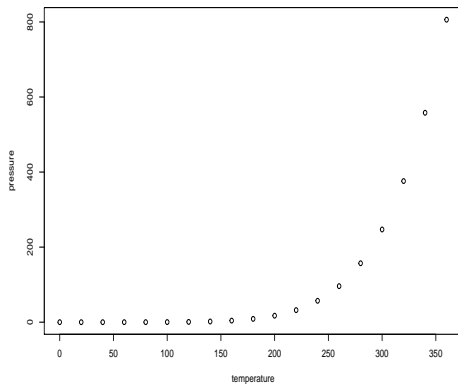
Faire des figures avec ggplot2

Matthias Grenié

3 mars 2016

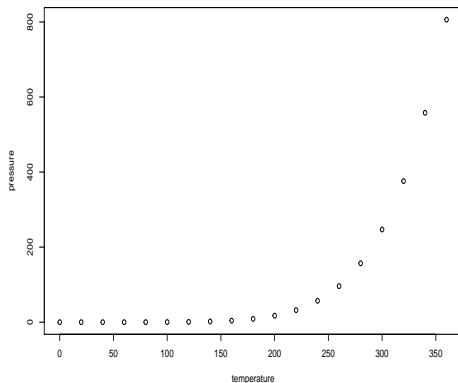
Statistiques Graphiques

```
plot(pressure)
```



Statistiques Graphiques

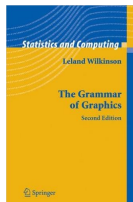
```
plot(pressure)
```



De quoi est constitué un graphique ?

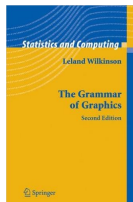
Réponse

Figure 1: Wilkinson, *Grammar of Graphics*, 2005



Réponse

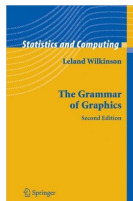
Figure 1: Wilkinson, *Grammar of Graphics*, 2005



Grammaire ?

Réponse

Figure 1: Wilkinson, *Grammar of Graphics*, 2005

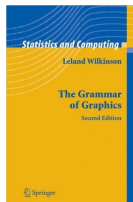


Grammaire ?

A grammar provides a strong foundation for understanding a diverse range of graphics.

Réponse

Figure 1: Wilkinson, *Grammar of Graphics*, 2005



Grammaire ?

A grammar provides a strong foundation for understanding a diverse range of graphics.

Décomposition graphique en éléments simples

- ▶ Données
- ▶ Éléments (Points, Lignes, etc.)
- ▶ Échelles (en X ? en Y ? en couleur ?)
- ▶ Coordonnées (cartésiennes ? polaires ? autres ?)

ggplot2

Hadley Wickham



ggplot2



Hadley Wickham

Package **ggplot2** (**g**rammar of **g**raphics)

ggplot2



Hadley Wickham

Package **ggplot2** (**g**rammar of **g**raphics)

A Layered Grammar of Graphics en 2010 inspiration Wilkinson

Préambule

Installer le package !

```
install.packages("ggplot2")
```

Le charger

```
library(ggplot2)
```

Exemple avec qplot()

Jeu de données “trees”

```
# On charge les données
```

```
data(trees)
```

```
head(trees, n = 3)
```

```
##      Girth Height Volume  
## 1      8.3      70   10.3  
## 2      8.6      65   10.3  
## 3      8.8      63   10.2
```

Exemple avec qplot()

Jeu de données “trees”

```
# On charge les données
```

```
data(trees)
```

```
head(trees, n = 3)
```

```
##      Girth Height Volume
```

```
## 1      8.3      70   10.3
```

```
## 2      8.6      65   10.3
```

```
## 3      8.8      63   10.2
```

```
qplot()
```

Exemple avec qplot()

Jeu de données “trees”

```
# On charge les données
```

```
data(trees)
```

```
head(trees, n = 3)
```

```
##      Girth Height Volume
```

```
## 1    8.3      70   10.3
```

```
## 2    8.6      65   10.3
```

```
## 3    8.8      63   10.2
```

```
qplot() = quickplot()
```

Exemple avec qplot()

Jeu de données “trees”

```
# On charge les données
```

```
data(trees)
```

```
head(trees, n = 3)
```

```
##      Girth Height Volume
```

```
## 1      8.3      70   10.3
```

```
## 2      8.6      65   10.3
```

```
## 3      8.8      63   10.2
```

qplot() = quickplot() = analogue de la fonction plot() pour ggplot2

Exemple avec qplot()

Jeu de données “trees”

```
# On charge les données
```

```
data(trees)
```

```
head(trees, n = 3)
```

```
##      Girth Height Volume
```

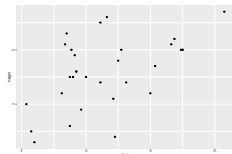
```
## 1    8.3      70   10.3
```

```
## 2    8.6      65   10.3
```

```
## 3    8.8      63   10.2
```

qplot() = quickplot() = analogue de la fonction plot() pour ggplot2

```
qplot(Girth, Height, data = trees)
```



Avec la fonction ggplot()

Données \Leftrightarrow variable **esthétique** du plot (**aesthetics**)

```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_point()
```

Avec la fonction ggplot()

Données \Leftrightarrow variable **esthétique** du plot (**aesthetics**)

```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_point()
```

Types de variable esthétique :

- ▶ x
- ▶ y
- ▶ taille
- ▶ couleur
- ▶ type de ligne
- ▶ forme

aes()

aes() autre exemple : la **couleur**

```
ggplot(data = trees, aes(x = Girth, y = Height,  
                          color = Volume)) +  
  geom_point()
```

Ou la taille des points :

```
ggplot(data = trees, aes(x = Girth, y = Height,  
                          size = Volume)) +  
  geom_point()
```

Objets géométriques

Même données, représentations **différentes** :

```
# Nuage de points
```

```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_point()
```

```
# Ligne
```

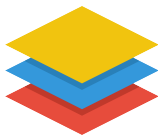
```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_line()
```

```
# Histogramme
```

```
ggplot(data = trees, aes(x = Girth)) +  
  geom_histogram()
```

Calques

Système de Calques de ggplot2 :



Calques

Système de Calques de ggplot2 :



Exemple : ajout régression linéaire

```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_point() +  
  geom_smooth(method = "lm")
```

Question : Que se serait-il passé si on avait inversé l'ordre ?

Facetting

Subdivise les graphiques en fonction des sous-jeux de données:

```
ggplot(InsectSprays, aes(x = count, color = spray)) +  
  geom_density()
```

```
ggplot(InsectSprays, aes(x = count, color = spray)) +  
  facet_grid(. ~ spray) +  
  geom_density()
```

Échelles

Facile de changer une échelle

```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_point() +  
  scale_x_log10()
```

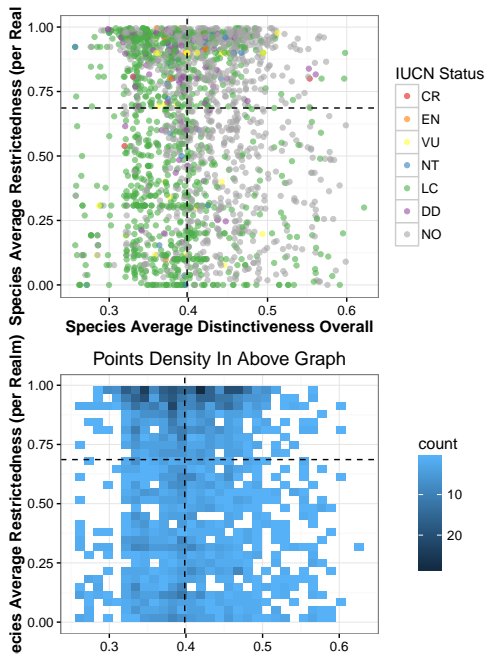

Échelles

Facile de changer une échelle

```
ggplot(data = trees, aes(x = Girth, y = Height)) +  
  geom_point() +  
  scale_x_log10()
```

Ou `scale_x_sqrt()` ou `scale_x_continuous(trans = 'log10')`

Exemples de graphiques (1)



Exemples de graphiques (2)

Acanthuridae



Balistidae



Caesionidae



Carangidae



Chaetodontidae



Haemulidae



Holocentridae



Labridae



Lethrinidae



Lutjanidae



Mullidae



Muraenidae



Pomacentridae



Scaridae



Serranidae



Siganidae



Sparidae



Tetraodontidae

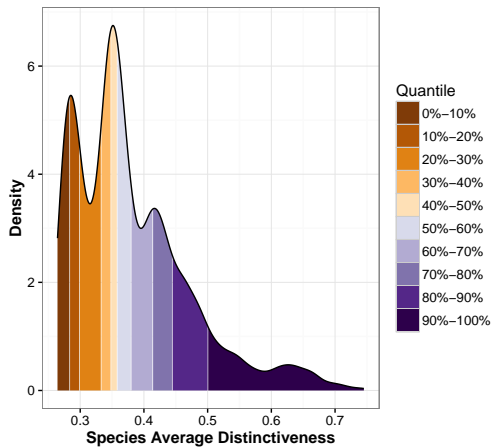


Top 100 Fur
Distinct

• FALSE

• TRUE

Exemples de graphiques (3)



Intérêts de ggplot2

- ▶ Grosse Communauté
- ▶ Facilement Extensible (version $\geq 2.0.0$)
- ▶ Facilement Paramétrable
- ▶ Beau graphiques par défaut

Et après ?

- ▶ Documentation officielle de `ggplot2`
- ▶ R graph catalog source d'inspiration
- ▶ O'Reilly R Cookbook
- ▶ Questions sur StackOverflow (~12000 questions)
- ▶ Packages annexes (GGally, ggtree, ggmap, etc.)