

P4a : Complexité

M. Zimmermann

IUT Robert Schuman, Département Informatique

2021/22

Chapitre 2

A large blue rounded rectangle with a subtle drop shadow, centered on the slide. It contains the white letter 'R' in a serif font.

Section 1

csv et data frame

On utilisera essentiellement un type de données particulier en R : la dataframe.

```
  x  y  name
1 0  0 carré
2 1  1 carré
3 2  4 carré
4 3  9 carré
5 4 16 carré
6 5 25 carré
...
```

Récupérer un csv

Si tout se passe bien, vous avez produit des fichiers csv. Comment les importer sous forme de dataframe en R? EZ!

```
read.csv("mon_joli_fichier_de_donnees.csv",  
         header=TRUE,  
         sep="\t")
```

TIP : Pour les exports excel français (lol), `dec=","` peut être utile.

Sous ensemble de données

```
d[d$name == "x",]  
d[d$name == "x" | d$name == "carré",]  
d[d$x > 5,]  
d[c("x", "y")]  
d[1:10,]
```

Rassembler des données

```
cbind(df1, df2)  
rbind(df1, df2)  
merge(df1, df1, by = "nom.col")  
merge(df1, df1, by = c("nom.col.1", "nom.col.2"))
```

Section 2

ggplot2

ggplot

```
ggplot(data, aes(x="col.x", y="col.y", color="col.diff"))
```

met en place les limites du graphiques

geom_point

```
+ geom_point()
```

place des points

geom_smooth

```
+ geom_smooth(method="loess", se=TRUE)
```

trace des courbes avec ou sans intervalle de confiance, avec méthode au choix

labs

```
+ labs(x = "nom des abscisses",  
      y = "nom des ordonnées",  
      colours = "nom des différentes courbes",  
      title = "meilleur titre du monde pour le  
              meilleur graphique du monde")
```

theme

```
+ theme(...)
```

ajuster plein de réglages

facets

```
+ facet_wrap(~ nom.col, ncol=3)
```

ou

```
+ facet_grid(nom.col.1 ~ nom.col.2)
```

pour produire aisément un tableau de plusieurs graphes sur des sous-ensembles de données définis selon une variables.