

TP2

Manipulation d'objets - Outils graphiques

Exercice I : Création de data.frame

1. Créer la data.frame suivante à partir des vecteurs 'name', 'taille', 'poids', 'nat', 'niveau', 'ave':

```
> etudiants
  name age taille poids nat niveau ave
1 Giorgio 21 190 90 Ita Lic2 15.0
2 Laura 18 167 56 Ita Lic3 16.4
3 Marie 18 164 68 Fra Mr1 17.1
4 Luc 24 166 78 Fra Mr2 14.5
5 John 20 172 82 USA Mr1 16.0
```

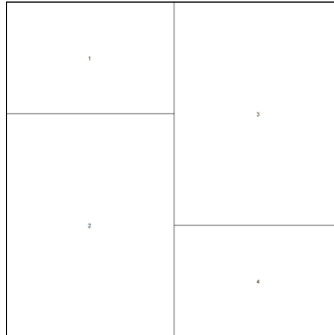
2. À partir des vecteurs 'name' et 'niveau', créer un tableau de contingence nommé tab_nn
3. Afficher les noms des dimensions de tab_nn
4. Changer 'name' en 'nom', 'Giorgio' en 'George' et 'John' en 'Jean' dans tab_nn
5. Créer un array de dimensions 4 x 3 x 2 avec les premier 24 nombres entiers. Nommer le dimensions 'a1'..'a4', 'b1'..'b3' et 'c1', 'c2'

Exercice II : Importation/exportation de fichier .txt

1. Importer dans un objet nommé A le jeu de donné nommé auto2004.txt
2. Vérifier la classe de l'objet A
2. Importer dans un objet nommé B le jeu de donnée auto2004_noColname.txt
3. Importer dans un objet nommé C le jeu de donnée auto2004_virg.txt
4. Importer dans un objet nommé D le jeu de donnée auto2004_NA.txt
5. Importer dans un objet nommé E le jeu de donnée auto2004_NA9.txt
6. Importer dans un objet nommé F le jeu de donnée auto2004_f.txt
7. Importer dans un objet nommé G le jeu de donnée auto2004_s.txt
9. Sauvegarder la data.frame A au format .txt sous le nom auto.txt
10. Sauver les objets présents en mémoire à l'adresse '/Outputs/Donnees.Rdata'
11. Sauver sous le nom « don.Rdata » les objets disponibles en mémoire à l'adresse "/Outputs" en utilisant la commande setwd()
12. Effacer tous les objets en mémoire en utilisant la fonction rm()
13. Charger les fichiers sauvegardés en question 10.
14. Restaurer le répertoire de travaille original.

Exercice III. Gestion des fenêtres graphiques

1. Créer 3 fenêtres graphiques vierges
2. Activer la deuxième fenêtre créée
3. Créer la partition suivante :



4. Activer la troisième fenêtre créée

Exercice IV : Les vins de bordeaux

Nous allons étudier à l'aide des outils graphiques disponibles sous R le jeu de données bordeaux_R.txt. Il s'agit d'un jeu de données qui fournit la qualité de vin de bordeaux (QUALITE) en fonction de 4 conditions climatiques (TEMPERAT, SOLEIL, CHALEUR, PLUIE)

TEMPERAT = somme des températures moyennes journalières (°C)

SOLEIL = Durée d'insolation (h)

CHALEUR = Nombre de jours de grande chaleur

PLUIE = hauteur de pluie en (mm)

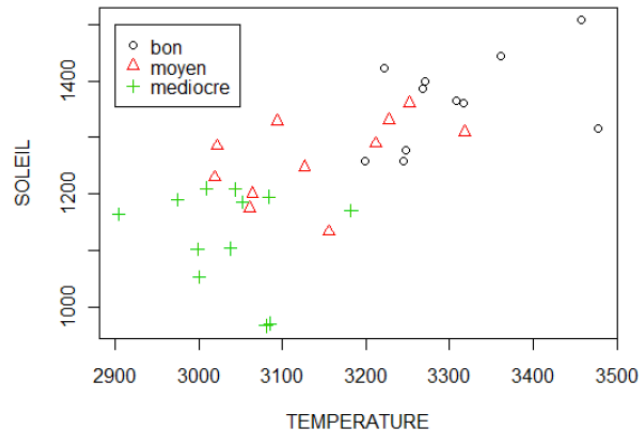
QUALITÉ : 1 = BON, 2 = MOYEN et 3 = MÉDIOCRE

On souhaite évaluer et visualiser l'impact des différents variables sur la qualité des vins de bordeaux.

1. Importer le jeu de données Bordeaux_R.txt.
2. Tracer les histogrammes de chacune des 4 conditions climatiques.
3. Tracer les boxplots de chacune des 4 conditions climatiques.
4. Tracer les boxplot de chacune des variables conditionnellement à la qualité. Qu'en concluez-vous?

Exercice V : Les légendes

A partir du jeu de données Bordeaux de l'exercice précédent, reconstruire le graphique suivant :



Exercice VI : Le package Lattice

1. Charger le package MASS
2. Dessiner un joli histogramme pour la variable Pulse du dataset Survey
3. Installer et charger le package Lattice
4. Dessiner un graphique avec les densités estimées de la variable Pulse conditionnées aux modalités de la variable 'Smoke'

Correction de l'exercice I : Création de data.frame

```
1. name = scan(what="character")
1: Giorgio Laura Marie Luc John
# Rappel: taper deux fois Enter pour passer à la variable suivante.
age = scan()
1: 21 18 18 24 20
taille = scan()
1: 190 167 164 166 172
poids = scan()
1: 90 56 68 78 82
nat = scan(what="character")
1: Ita Ita Fra Fra USA
niveau = scan(what="character")
1: Lic2 Lic3 Mr1 Mr2 Mr1
ave = scan()
1: 15 16.4 17.1 14.5 16

1. #Alternative :
name <- c("Giorgio","Laura","Marie","Luc","John")
age <- c(21, 18, 18, 24, 20)
taille = c(190, 167, 164, 166, 172)
poids = c(90, 56, 68, 78, 82)
nat = c("Ita","Ita","Fra","Fra","USA")
niveau = c("Lic2","Lic3","Mr1","Mr2","Mr1")
ave <- c(15, 16.4, 17.1, 14.5, 16)
etudiants = data.frame(name, age, taille, poids, nat, niveau, ave)

2. #Table de contingence
tab_nn <- table(name, niveau)

3. #Afficher les nommes des dimensions
dimnames(tab_nn)

4. #Renommer les dimensions (ici je le renomme toutes..)
dimnames(tab_nn) = list(nom = c("George","Jean" , "Laura", "Luc", "Marie"), niveau =
c("Lic2", "Lic3", "Mr1" , "Mr2" ))
tab_nn
# mais je peux renommer que les nommes qu'il faut
tab_nn <- table(name, niveau)
tab_nn
dimnames(tab_nn)[[1]][1] = "George"
dimnames(tab_nn)[[1]][2] = "Jean"
names(dimnames(tab_nn))[1] = "nom"
tab_nn

5. #Construction de l'array
array.gio <- array(data = 1:24, dim = c(4,3,2), dimnames =
list(paste("a",1:4,sep=' '),paste("b",1:3,sep=""),c("c1","c2")))
```

Correction de l'exercice II : Importation/exportation de fichier .txt

```
A = read.table('Datasets/auto2004.txt',row.names=1, header = TRUE, sep = '\t')

class(A)

B = read.table('Datasets/auto2004_noColname.txt', row.names=1, sep = '\t')

C = read.table('Datasets/auto2004_virg.txt', row.names=1,header = TRUE, sep = '\t',
dec = ",")

D = read.table('Datasets/auto2004_NA.txt', row.names=1, header = TRUE, sep = '\t')
```

```

E = read.table('Datasets/auto2004_NA9.txt', row.names=1, header = TRUE, sep = '\t',
na.strings = 99999)

F = read.table('Datasets/auto2004_f.txt', row.names=1, header = TRUE, sep = '\t', fill
= TRUE)

G = read.table('Datasets/auto2004_s.txt', row.names=1, header = TRUE, sep = '\t', skip
= 4)

write.table(A,"Outputs/auto.txt")

save(list = ls(), file = "Outputs/donnees.Rdata")

# La fonction getwd() renvoie le chemin de sauvegarde lorsque rien n'est spécifié. On
peut modifier le chemin par défaut grâce à la fonction setwd() :
setwd("Outputs")
save.image(file="Outputs/don.Rdata")
rm(list = ls())
load("Outputs/donnees.Rdata")
setwd("/Users/Giorgio/Desktop/lezioni università/CNAM/STA115/docs for R/corsoGio")

```

Correction de l'exercice III : Gestion des fenêtres graphiques

```

1. #
quartz() ;quartz() ;quartz() ; # Windows x11() ; x11() ; x11()

2. #
dev.set(2)

3. #
mat = matrix(c(1, 2, 2, 3, 3, 4), 3, 2)
layout(mat)
layout.show(4)

4. #
dev.set(3)

```

Correction de l'exercice IV : Les vins de bordeaux

```

1. #Importation des données
A = read.table('chemin/bordeaux_R.txt', header = TRUE, sep = '\t')

2. #Histogramme :
layout(matrix(1:4, 2, 2))
layout.show(4)
hist(A[, 2], 20, main = "température")
hist(A[, 3], 20, main = "soleil")
hist(A[, 4], 20, main = "chaleur")
hist(A[, 5], 20, main = "pluie")

3. #boxplot :
layout(matrix(1:4, 2, 2))
layout.show(4)
boxplot(A[, 2], main = "température")
boxplot(A[, 3], main = "soleil")
boxplot(A[, 4], main = "chaleur")
boxplot(A[, 5], main = "pluie")

4. #Boxplot conditionné :
layout(matrix(1:4, 2, 2))
layout.show(4)
boxplot(A[, 2] ~ c("bon", "moyen", "mediocre")[A[, 6]], main = "température")
boxplot(A[, 3] ~ c("bon", "moyen", "mediocre")[A[, 6]], main = "soleil")
boxplot(A[, 4] ~ c("bon", "moyen", "mediocre")[A[, 6]], main = "chaleur")

```

```
boxplot(A[, 5] ~ c("bon", "moyen", "mediocre")[A[, 6]], main = "pluie")
```

On peut interpréter l'influence de chaque variable sur la qualité du produit. Par exemple à l'aide des boxplots conditionnés, on peut remarquer que les meilleurs vins résultent d'année très chaleureux avec beaucoup de soleil.

Correction de l'exercice V : Les légendes

```
A = read.table("chemin/bordeaux.txt", header = TRUE, sep = "\t")
#windows
x11()
#mac
quartz()
levels(factor(A$QUALITE))
Y = factor(A$QUALITE, labels = c("bon", "moyen", "mediocre"))
plot(X[,2 :3], col = Y, pch = c(1:3)[Y], xlab="TEMPERATURE")
legend(x = 2900, y = 1500, legend = levels(Y), pch = c(1:3), col = 1:3)
```

Correction de l'exercice VI : Le package Lattice

1. #charger le paquet MASS
library(MASS)
2. #Histogramme
data(survey)
summary(survey\$Pulse)
summary(survey\$Pulse)
hist(survey\$Pulse,
 breaks = seq(from = 35, to = 104, length = 20),
 col = "pink", border = "darkblue",
 main = paste("Pulse de", nrow(survey), " étudiants"),
 xlab = "Battements", ylab = "densité",
 proba = TRUE, labels = TRUE, ylim = c(0, 0.06))
dst = density(survey\$Pulse, na.rm = TRUE, adjust = 1)
lines(dst\$x, dst\$y, lwd = 1.5, col='darkblue')
mtext("l'histogramme le plus beau\nc'est le mien!!", line = -2)
3. #Installer et charger le paquet LATTICE
install.packages("lattice")
library(lattice)
4. #graphique des densites
densityplot(~survey\$Pulse | survey\$Smoke, xlab = "nombre de Battements", col = "red")