

Compte rendu TP1 Initiation à R

Laroussi Labid Bachri

01/09/2025

Contents

Premiers pas avec R et RMarkdown	1
Utilisation et calcul avec du code R	1
Utilisation de variables	2
Vecteurs	2
Fonctions	2
Lecture et écriture de fichiers contenant des données	3
Fonctions graphiques	3

Premiers pas avec R et RMarkdown

Ceci est un document de type R Markdown. Markdown est une mise en forme syntaxique simplifiée pour rédiger des documents HTML, PDF ou encore MsWord. Pour plus de détails sur l'utilisation de R Markdown cf. <http://rmarkdown.rstudio.com>.

Le fichier doit avoir une extension “.Rmd”. Quand vous cliquez sur le bouton **Knit**, le document sera généré et inclura le contenu ainsi que la sortie produite par le code R intégré dans le document. Le code R peut être intégré comme ceci :

```
-10 * log10(0.00666)
```

```
## [1] 21.76526
```

Il existe beaucoup de documents et tutoriels en ligne que vous pouvez explorer par vous-même. Exemples :

- Mise en forme du texte : <https://rmarkdown.rstudio.com/lesson-8.html>
- Bouts de code R (chunks) : <https://rmarkdown.rstudio.com/lesson-3.html>

Utilisation et calcul avec du code R

Cette partie est à compléter avec le sujet de TP en ligne : <https://moodle.utoulouse.fr/mod/page/view.php?id=21136>

Utilisation de variables

Les variables (ou objets) permettent de stocker des données qui peuvent être :

- une valeur simple de type numérique (**numeric**), logique (**logical**), chaîne de caractères (**character**) ou qualitative (**factor**).
- une liste (appelée **vector**).
- un tableau à 2 dimensions, les colonnes pouvant avoir des types différents (**data.frame**). Ce sont les plus utilisés en statistiques.
- un tableau à 2 dimensions, toutes les cases ayant le même type (**matrix**).
- un tableau à plus de 2 dimension (**array**).
- une combinaison des précédents (**list**).

Une variable a un nom (défini par l'utilisateur) qui permet d'accéder à son contenu.

Créer une variable appelée x qui stocke le résultat du calcul précédent $1 + 2 + 3$. L'affectation d'une valeur à une variable se fait avec `<-` ou le signe `=` comme suit dans la console :

A compléter

```
x = 1+2+3
```

Vecteurs

A compléter

```
x = c(2,3,5,8,4,6);longueur = length(x)
```

```
deuxieme_element = x[2]
```

```
acces_plusieurs_elements = x[c(2,4,1)]
```

```
valeurs_qui_se_suivent = x[2:4]
```

```
valeurs_superieures_a_4 = x[x>4]
```

```
inverse_de_x = 1/x
```

```
double_de_x = 2*x
```

```
arrondir_x_divise_par4 = round(x/4)
```

```
arrondir_x_divise_par3 = round(x/4 , 2)
```

```
# on ajoute ici l'argument qui permet de choisir combien de chiffres apres la virgule
```

Fonctions

A compléter

```
summary(x)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      2.000   3.250   4.500   4.667   5.750   8.000
```

```
valeur_min = min(x)
valeur_max = max(x)
mediane = median(x)
moyenne = mean(x)
variance = var(x)
ecart_type = sd(x)
```

Lecture et écriture de fichiers contenant des données

```
croissance = read.table(file = "/Users/bachri/Documents/M1_BBS/S7/traitemement_donnees_bio/TP1_TDB_Intro_1",
variables = names(croissance)
croissance$poids
```

```
##      [1] 383.5 395.8 406.8 387.7 407.6 387.6 393.3 402.4 392.9 396.5 383.2 407.7
##      [13] 402.6 422.1 402.9 403.1 390.2 417.3 396.9 398.4 403.3 410.4 389.6 398.3
##      [25] 389.6 400.3 402.9 417.5 392.8 415.4 412.7 420.8 412.0 403.1 421.8 423.2
##      [37] 421.8 418.5 436.8 420.7 414.2 415.5 436.1 412.5 423.3 431.9 431.6 403.4
##      [49] 421.1 420.2 413.6 433.5 413.5 404.8 431.0 422.3 424.1 418.4 437.7 442.0
##      [61] 422.1 421.3 410.2 433.8 431.5 426.2 429.9 427.8 425.6 430.0 415.4 427.1
##      [73] 422.5 422.5 416.1 429.1 437.4 453.7 434.8 437.1 441.5 416.4 432.5 440.7
##      [85] 424.3 435.1 440.8 446.9 420.7 415.6 446.5 436.6 443.8 415.4 446.9 442.7
##      [97] 449.6 457.3 459.1 443.7
```

```
attach(croissance)
is_it_from_pyr = origine_geo == 'pyr'

plantes_pyr = which(origine_geo == 'pyr') # identifier quels arbres ont pour originegeo et donc qui ont

pyr = croissance[origine_geo == 'pyr' , 1:2]

# on va sauvegarder ces données dans un autre fichier .txt

write.table(pyr , "croissance_plantes_pyr.txt" , quote = F , col.names = T , row.names = F , sep="\t")
```

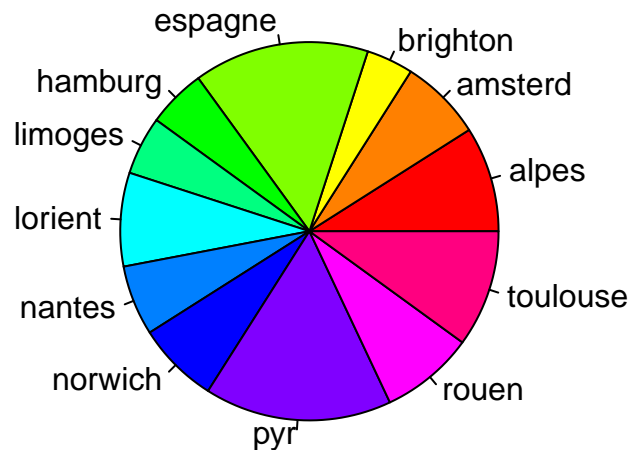
Fonctions graphiques

```
summary(origine_geo) # on synthetise les origines des plantes et compter combien de plantes ont tel ori.
```

```
##      alpes  amsterd brighton  espagne  hamburg  limoges  lorient  nantes
##          9          7          4          15          5          5          8          6
## norwich      pyr      rouen toulouse
##          7          16          8          10
```

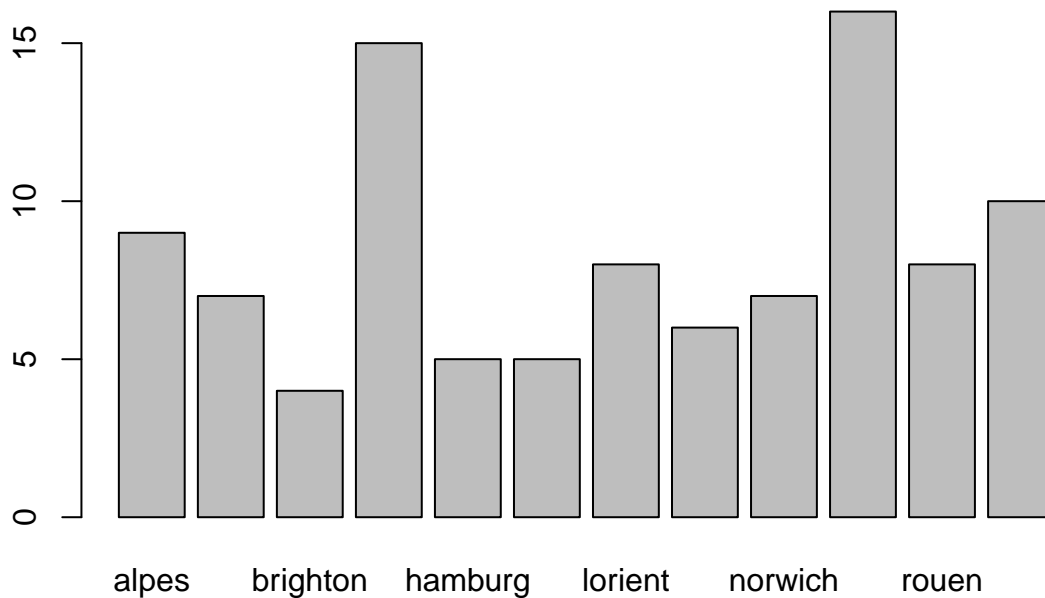
```
camembert_origine = pie(summary(origine_geo), main = "Origine géographique des plantes", col=rainbow(12))
```

Origine géographique des plantes



```
barplot(summary(origine_geo), main = "Origine géographique des plantes")
```

Origine géographique des plantes

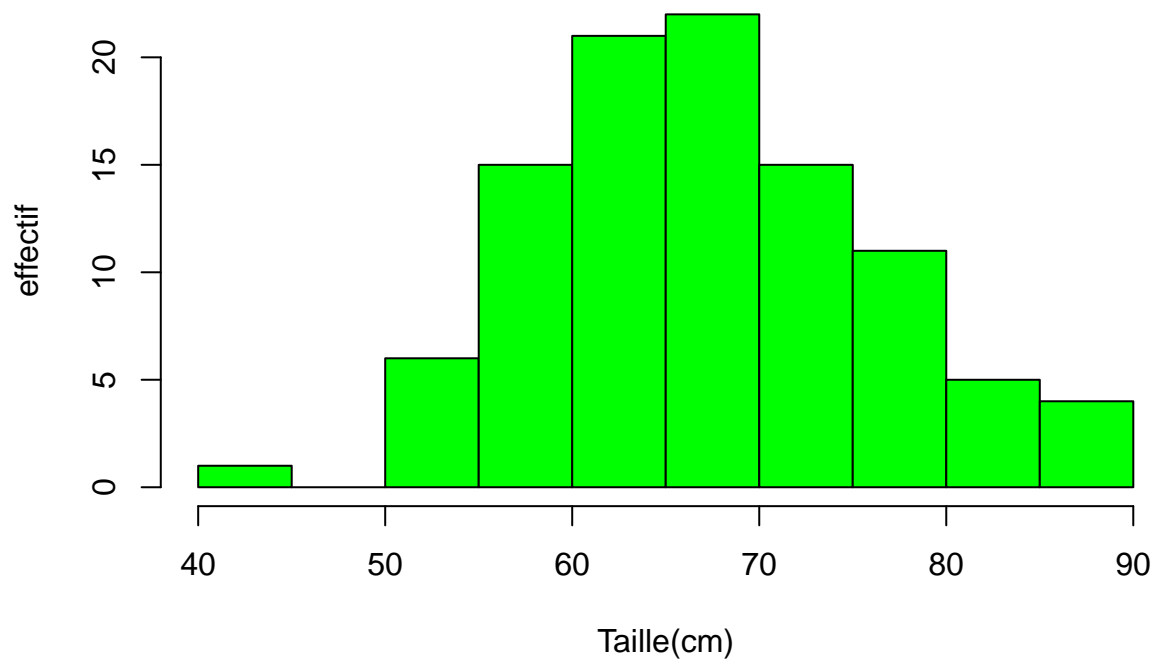


on est sur une variable qualitative, on ne va parler de histogramme mais de diagramme en barres

Histogrammes (variable quantitative)

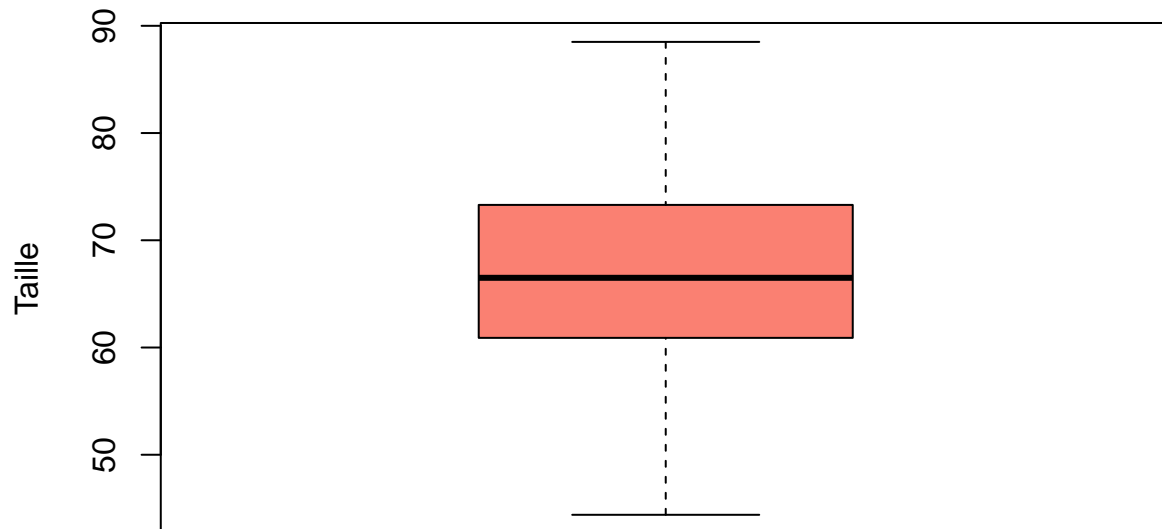
```
hist(taille, xlim=c(40,90), xlab="Taille(cm)" , ylab = "effectif" , freq=T , main = "Histogramme de la
```

Histogramme de la taille des plantes



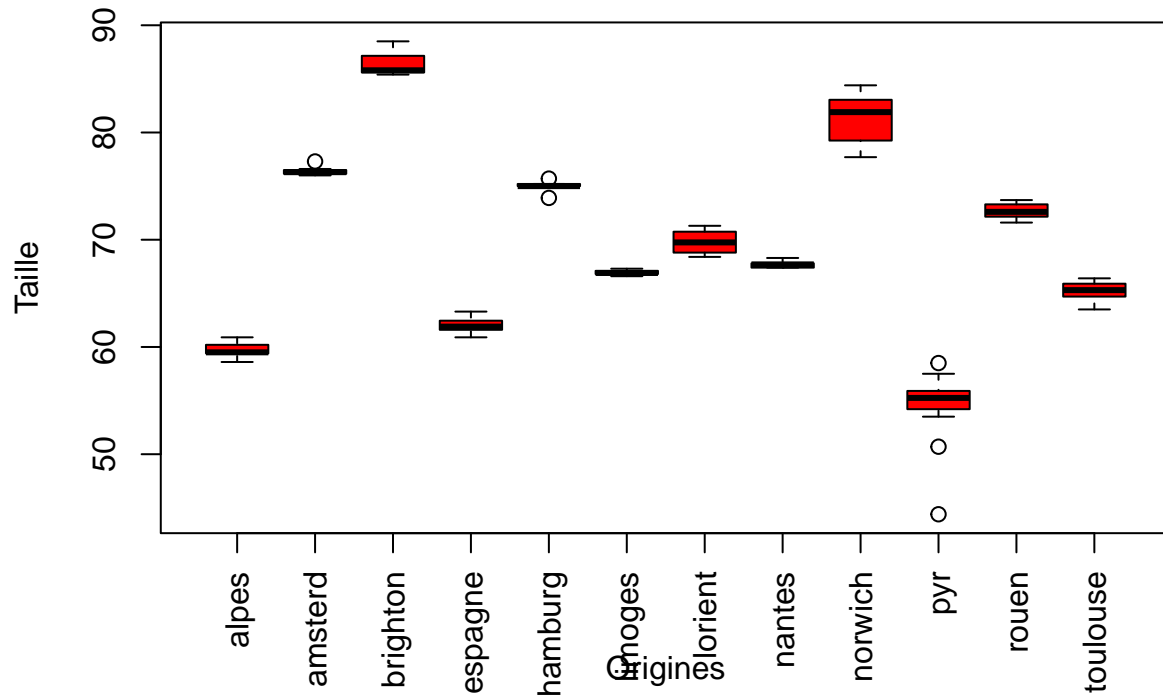
```
# Boite à moustaches (variable quantitative)
boxplot(taille, main="Boxplot de la taille des plantes" , ylab="Taille" , col='salmon')
```

Boxplot de la taille des plantes

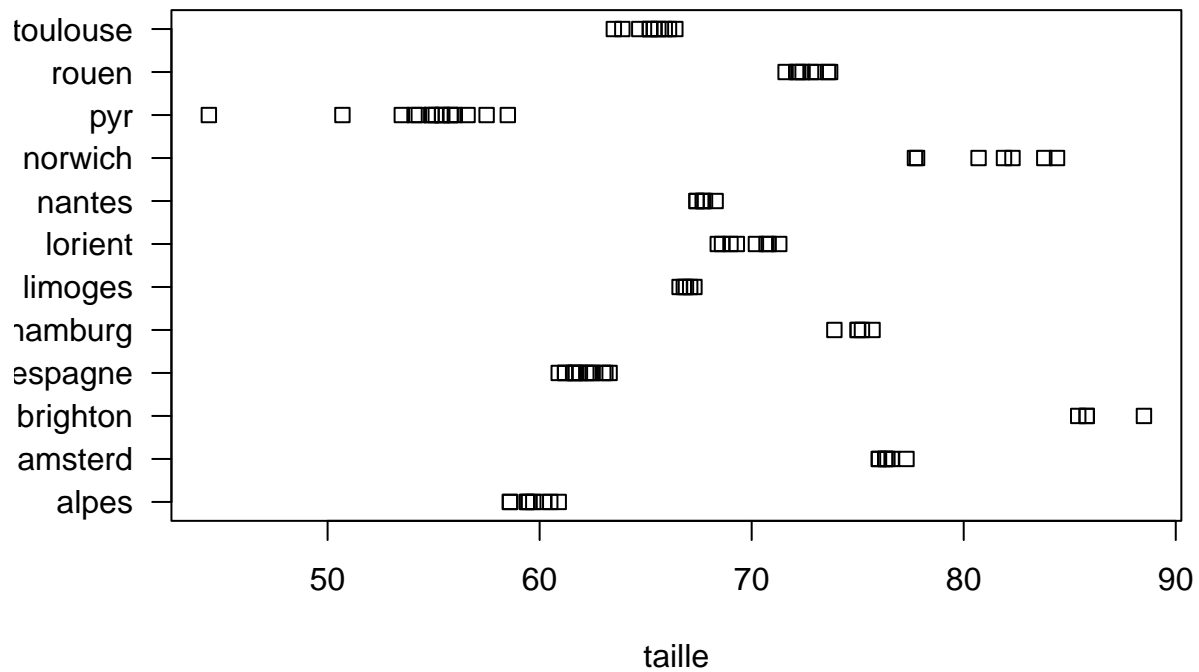


```
# Plusieurs boite à moustaches ( une variable quantitative en fonction des modalités d'une variable qualitative)
plot(origine_geo , taille, las=3 , main="Boxplot de la taille des plantes en fonction de l'origine géographique")
```

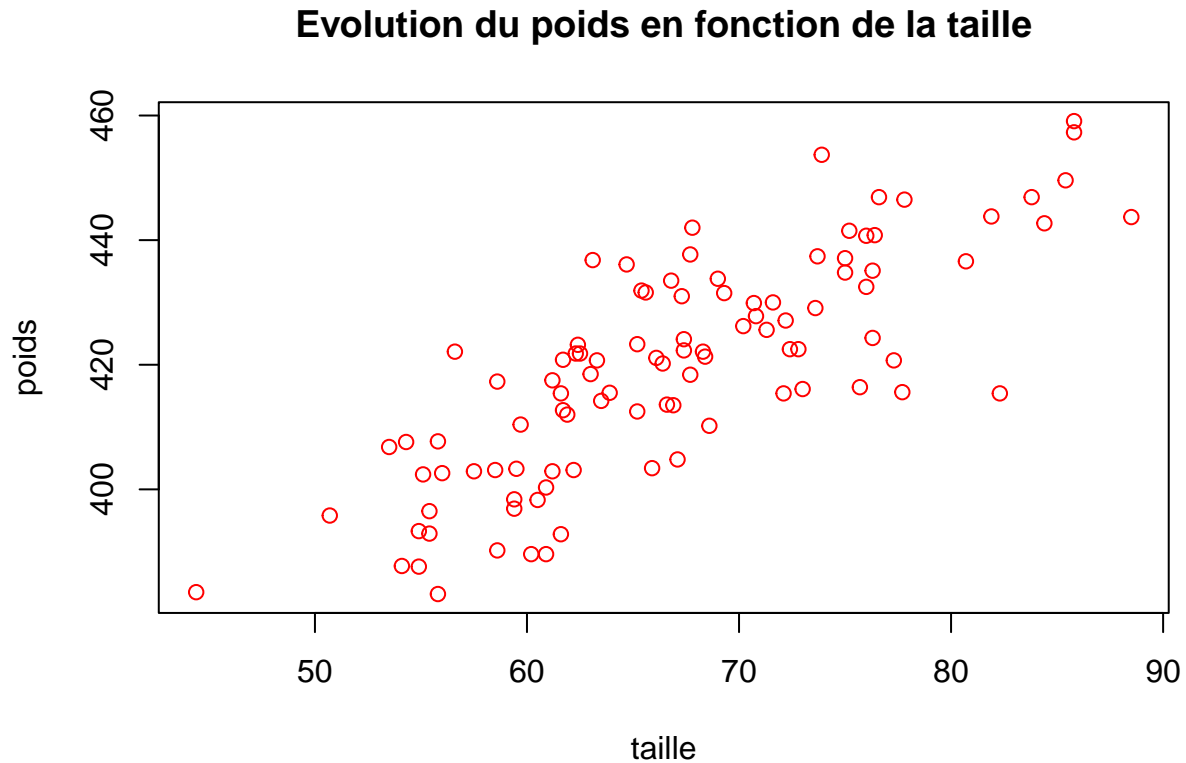
Boxplot de la taille des plantes en fonction de l'origine géographique



```
stripchart(taille~origine_geo , las=1)
```

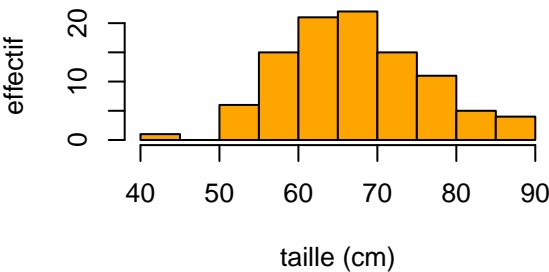


```
plot(taille,poids, col="red" , main = "Evolution du poids en fonction de la taille")
```



```
# Affichage de plusieurs graphiques dans la meme fenetre
par(mfrow=c(2,2)) # 2 en lignes et 2 en colonnes par exemple
hist(taille,xlim=c(40,90),xlab="taille (cm)",ylab="effectif",freq=T,main="histogramme de la taille des p")
# xlim=c(x,y) sert a limiter les valeurs affichées sur l'histogramme.
boxplot(taille,main="Boxplot de la taille des plantes",ylab="taille")
plot(taille~origine_geo,las=3)
stripchart(taille~origine_geo,las=1)
```

histogramme de la taille des plantes



Boxplot de la taille des plantes

