

Analyse_Ciqual

Tchouenkou David Nathanäel

2023-10-15

```
## importation des données et creation d'un sous data frame
library(readxl)
url <- "Table Ciqual 2020_FR_2020 07 07.xls"
t<- read_excel(url, guess_max = 3000)
T<- as.data.frame(t, guess_max=3000)
```

```
## exploration du sous data frame
U<- T[ , c(5, 6, 8, 13, 14, 15, 17, 18, 19, 27)]
colnames(U)<- c("Groupes", "SousGroupe", "Aliment", "Energie",
               "Eau", "Proteines", "Glucides",
               "Lipides", "Sucres", "Fibres")
```

```
## Cette fonction permet de gérer Les données mal renseignés et Les données manquante dans Le
jeu de données
recod<- function(d){
  d<- gsub("[ -]", NA, d)
  d <- gsub(" ", NA, d)
  d <- gsub(",", ".", d)
  d<- gsub("^<.*", "0", d)
  d<- gsub("[a-z].*", "0", d)
  d<-as.numeric(d)
}

## utilisation de la fonction pour transformation des variables qui en ont besoin
U$Energie <- recod(U$Energie)
U$Eau<- recod(U$Eau)
U$Proteines <- recod(U$Proteines)
U$Glucides<- recod(U$Glucides)
U$Lipides<- recod(U$Lipides)
U$Sucres<- recod(U$Sucres)
U$Fibres<- recod(U$Sucres)
```

```
# Créer une Liste des groupes de desserts

dessert <- c( "fruits", "produits laitiers frais et assimilés",
             "chocolats et produits à base de chocolat",
             "confitures et assimilés",
             "gâteaux et pâtisseries", "glaces", "sorbets",
             "desserts glacés")

# Extraire Les aliments correspondant aux groupes de desserts
V <- U[U$Groupes %in% dessert, ]

V$Groupes <- factor(V$Groupes, levels = c("fruits",
                                          "produits laitiers frais et assimilés",
                                          "chocolats et produits à base de chocolat",
                                          "confitures et assimilés",
                                          "gâteaux et pâtisseries",
                                          "glaces", "sorbets",
                                          "desserts glacés"),
                   labels = c("fruits", "laitages", "chocolats",
                              "confitures", "pâtisseries",
                              "glaces", "sorbets",
                              "desserts glacés"))
```

Combien y a-t-il de dessert dans le jeu de données? combien avec une energie > 500 kcal? combien y a-t-il de dessert pour les quels l'energie n'est pas connue?

```
## nombre de desserts
summary(V$Groupes)
```

```
##      fruits      laitages      chocolats      confitures      pâtisseries
##      170        137         31          14             85
##      glaces      sorbets desserts glacés
##      11          5          12
```

```
##
subset(V, Energie > 500, select = c(Energie, Groupes))
```

```
##      Energie  Groupes
## 2541      523 chocolats
## 2544      537 chocolats
## 2548      551 chocolats
## 2552      525 chocolats
## 2554      505 chocolats
## 2559      567 chocolats
## 2563      591 chocolats
## 2569      532 chocolats
## 2571      557 chocolats
```

```
##(on a 272 qui ont l'energie inconnue )
summary(V$Energie)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's
##      27.6   64.8   97.5   161.4  226.0   591.0   272
```

```
## ou
energie_inconnue <- subset(V, is.na(Energie), select = c(Energie, Groupes))
summary(energie_inconnue)
```

```
##      Energie      Groupes
## Min.   : NA  fruits    :109
## 1st Qu.: NA  pâtisseries: 76
## Median : NA  laitages   : 50
## Mean   :NaN  chocolats  : 17
## 3rd Qu.: NA  glaces     : 9
## Max.   : NA  confitures : 6
## NA's   :272  (Other)   : 5
```

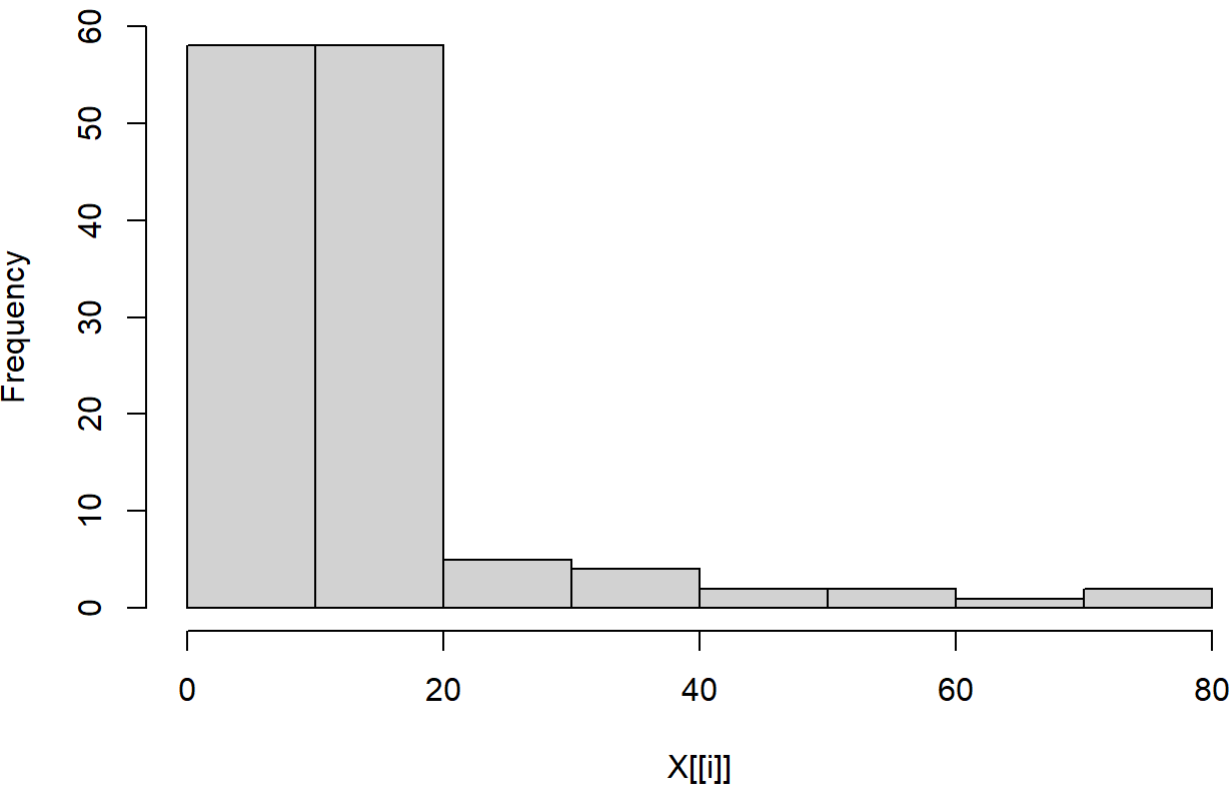
Quel est le contenu moyen en sucres des desserts de chaque groupe ?

```
tapply(V$Sucres,V$Groupes, mean)
```

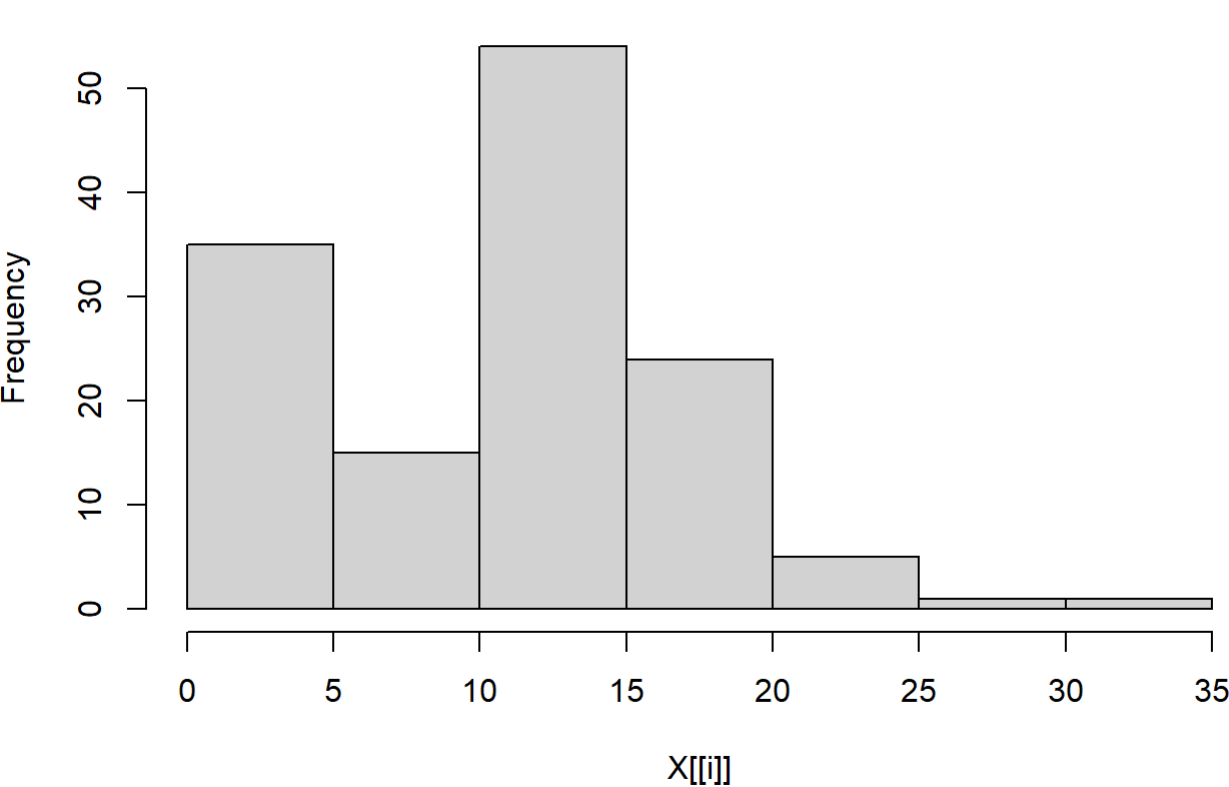
```
##      fruits      laitages      chocolats      confitures      pâtisseries
##      NA          NA          43.95806      55.46429      NA
##      glaces      sorbets desserts glacés
##      23.76364      23.96000      24.90833
```

```
tapply(V$Sucres,V$Groupes, hist)
```

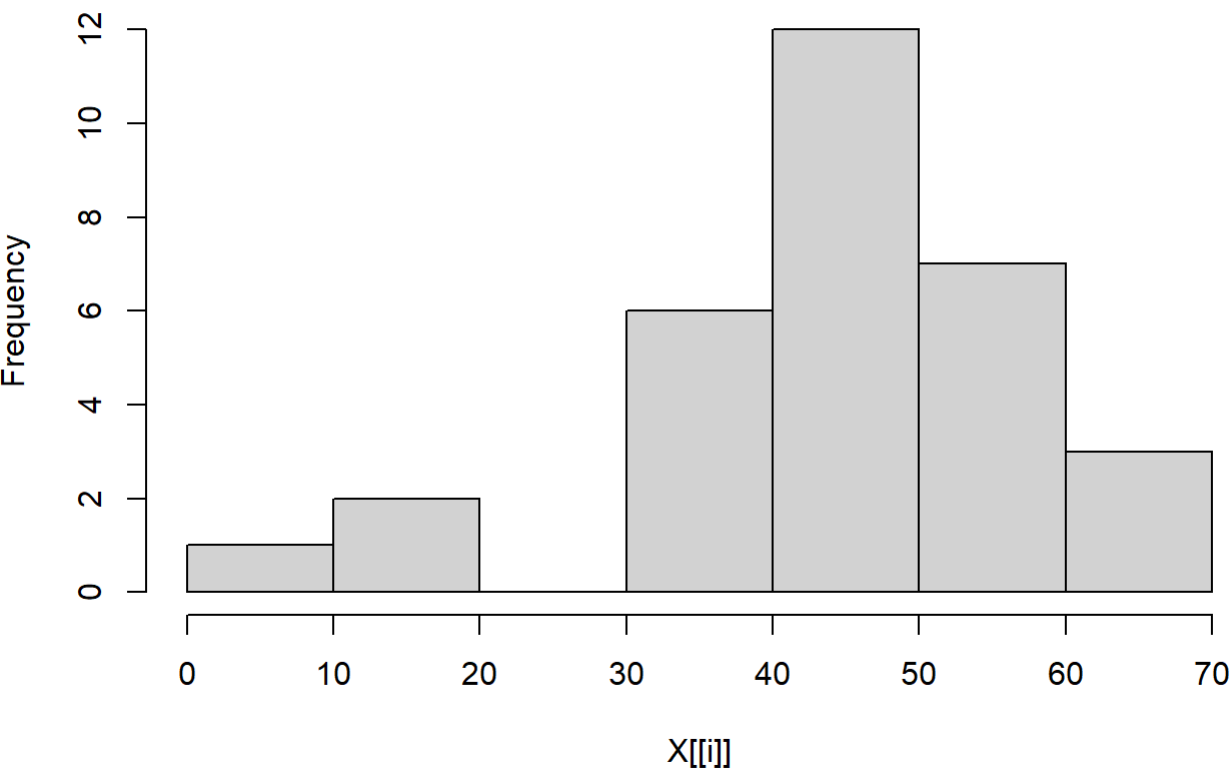
Histogram of X[[i]]



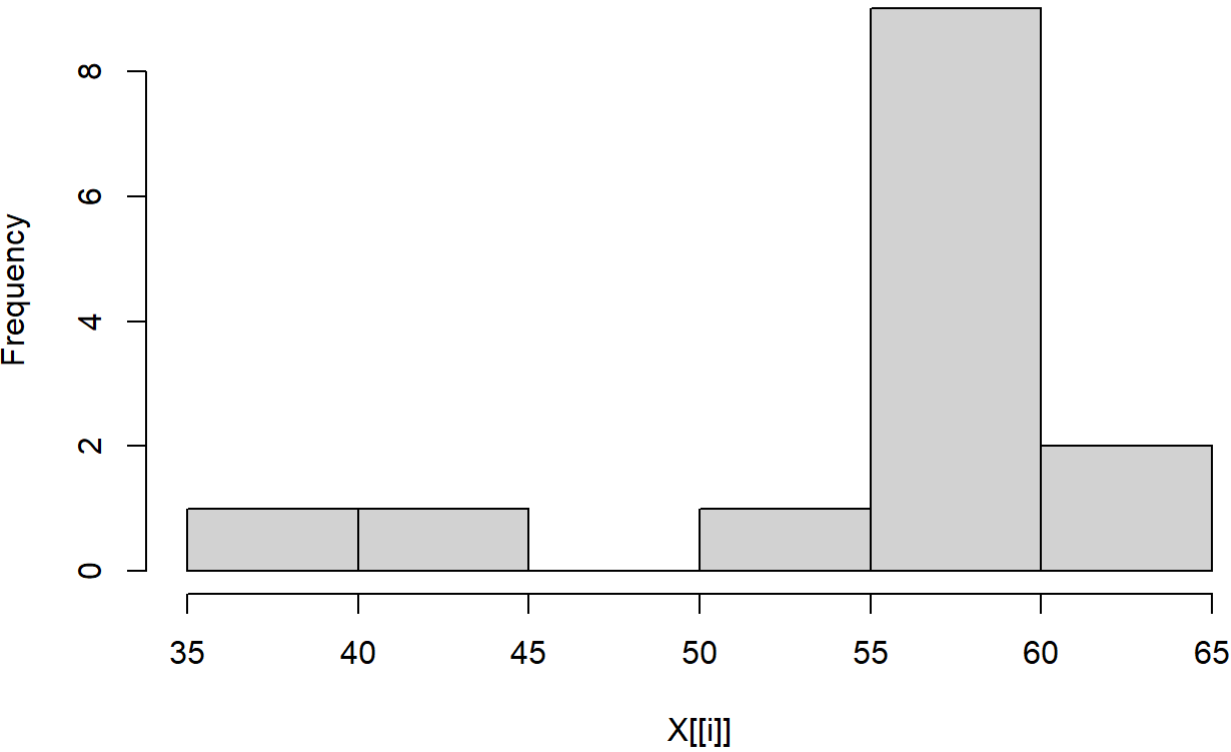
Histogram of X[[i]]



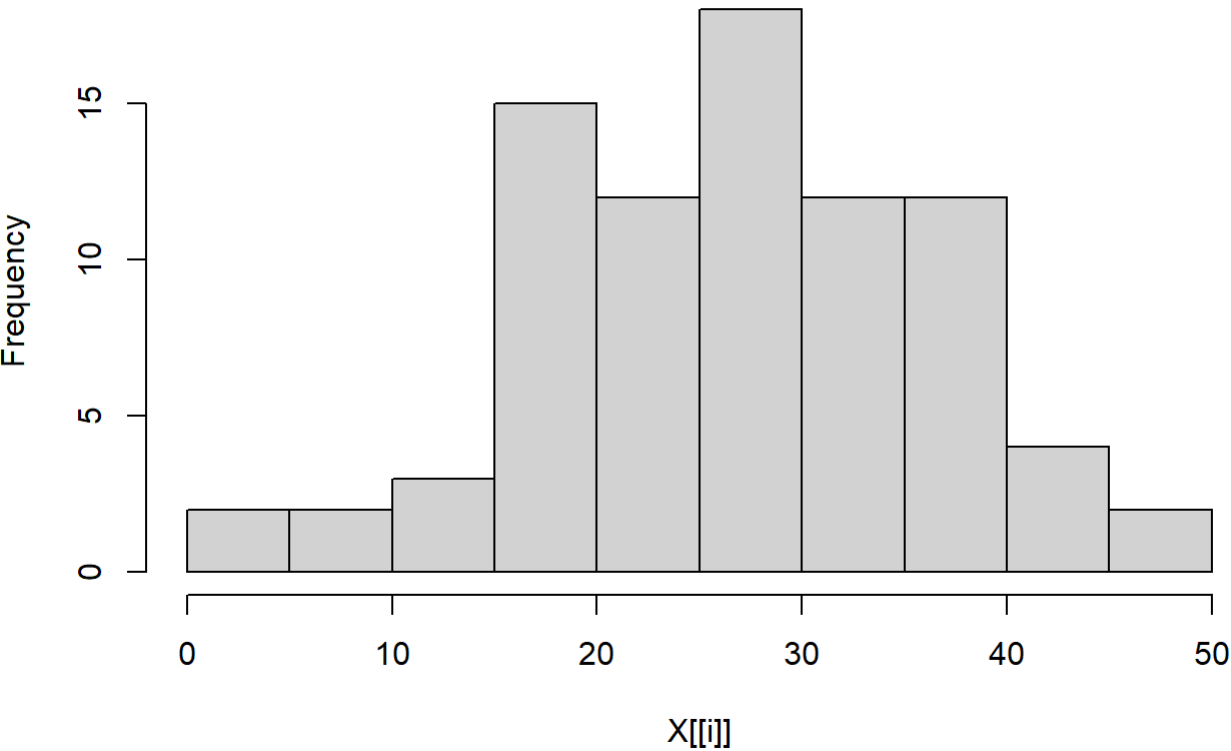
Histogram of X[[i]]



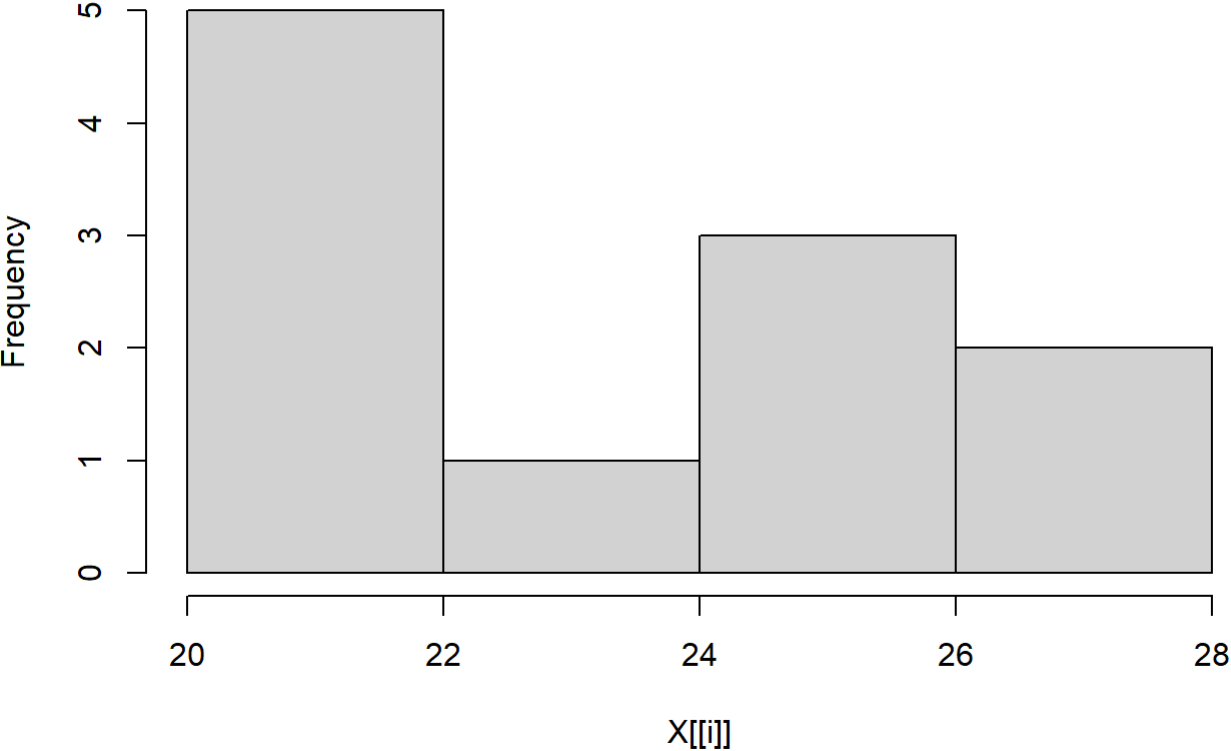
Histogram of X[[i]]



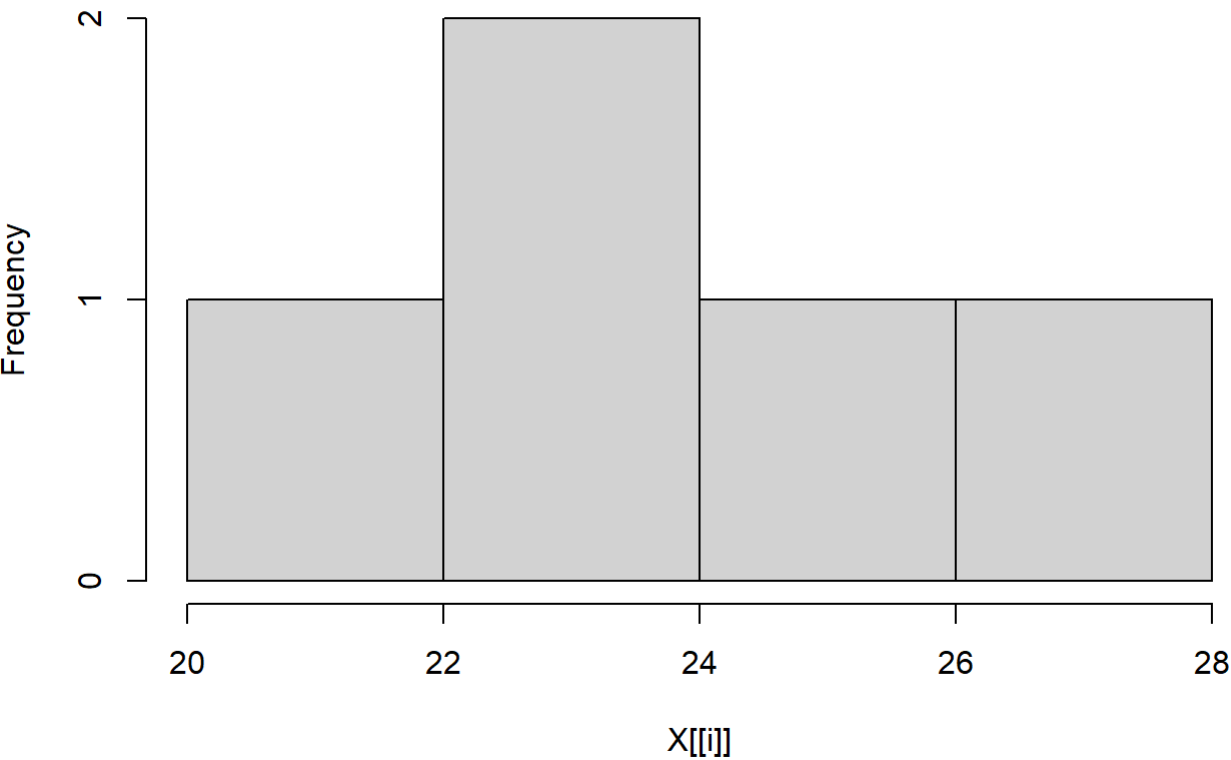
Histogram of X[[i]]



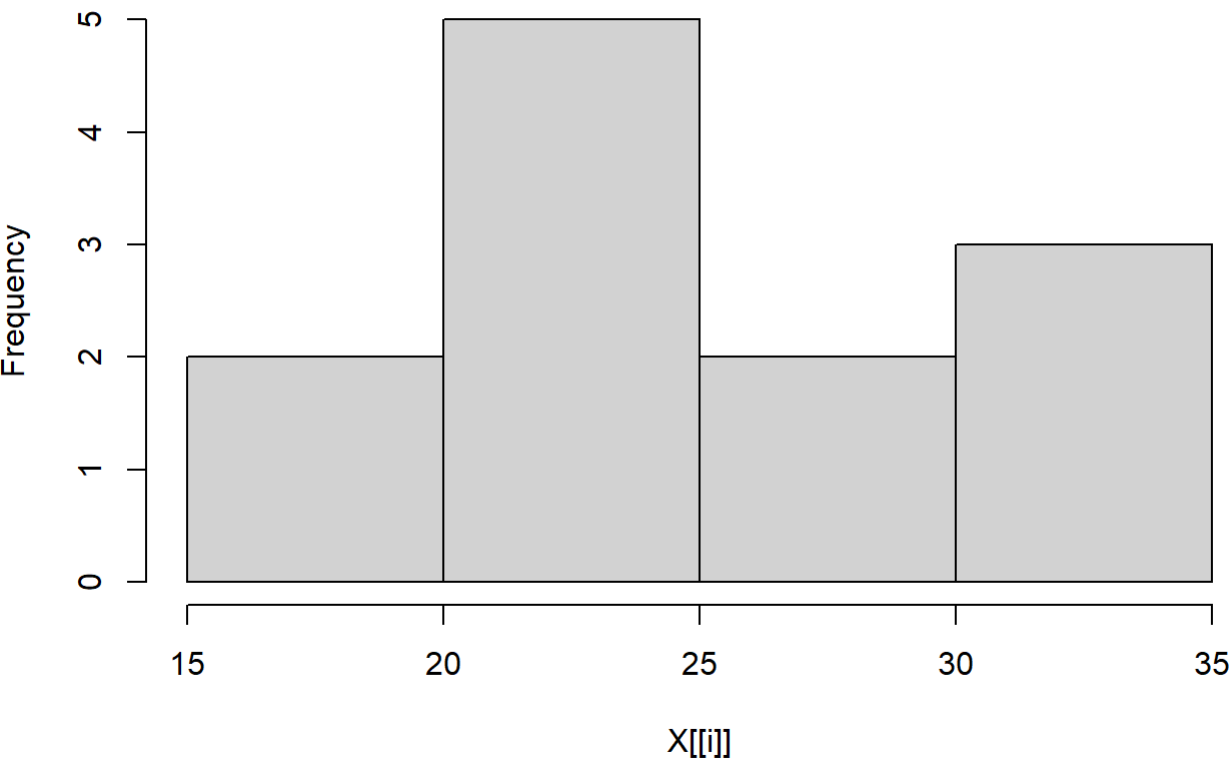
Histogram of X[[i]]



Histogram of X[[i]]



Histogram of X[[i]]



```

## $fruits
## $breaks
## [1]  0 10 20 30 40 50 60 70 80
##
## $counts
## [1] 58 58  5  4  2  2  1  2
##
## $density
## [1] 0.0439393939 0.0439393939 0.0037878788 0.0030303030 0.0015151515
## [6] 0.0015151515 0.0007575758 0.0015151515
##
## $mids
## [1]  5 15 25 35 45 55 65 75
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##
## attr("class")
## [1] "histogram"
##
## $laitages
## $breaks
## [1]  0  5 10 15 20 25 30 35
##
## $counts
## [1] 35 15 54 24  5  1  1
##
## $density
## [1] 0.051851852 0.022222222 0.080000000 0.035555556 0.007407407 0.001481481
## [7] 0.001481481
##
## $mids
## [1]  2.5  7.5 12.5 17.5 22.5 27.5 32.5
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##
## attr("class")
## [1] "histogram"
##
## $chocolats
## $breaks
## [1]  0 10 20 30 40 50 60 70
##
## $counts
## [1]  1  2  0  6 12  7  3
##
## $density
## [1] 0.003225806 0.006451613 0.000000000 0.019354839 0.038709677 0.022580645

```



```

## [7] 0.009677419
##
## $mids
## [1] 5 15 25 35 45 55 65
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##
## attr("class")
## [1] "histogram"
##
## $confitures
## $breaks
## [1] 35 40 45 50 55 60 65
##
## $counts
## [1] 1 1 0 1 9 2
##
## $density
## [1] 0.01428571 0.01428571 0.00000000 0.01428571 0.12857143 0.02857143
##
## $mids
## [1] 37.5 42.5 47.5 52.5 57.5 62.5
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##
## attr("class")
## [1] "histogram"
##
## $pâtisseries
## $breaks
## [1] 0 5 10 15 20 25 30 35 40 45 50
##
## $counts
## [1] 2 2 3 15 12 18 12 12 4 2
##
## $density
## [1] 0.004878049 0.004878049 0.007317073 0.036585366 0.029268293 0.043902439
## [7] 0.029268293 0.029268293 0.009756098 0.004878049
##
## $mids
## [1] 2.5 7.5 12.5 17.5 22.5 27.5 32.5 37.5 42.5 47.5
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##

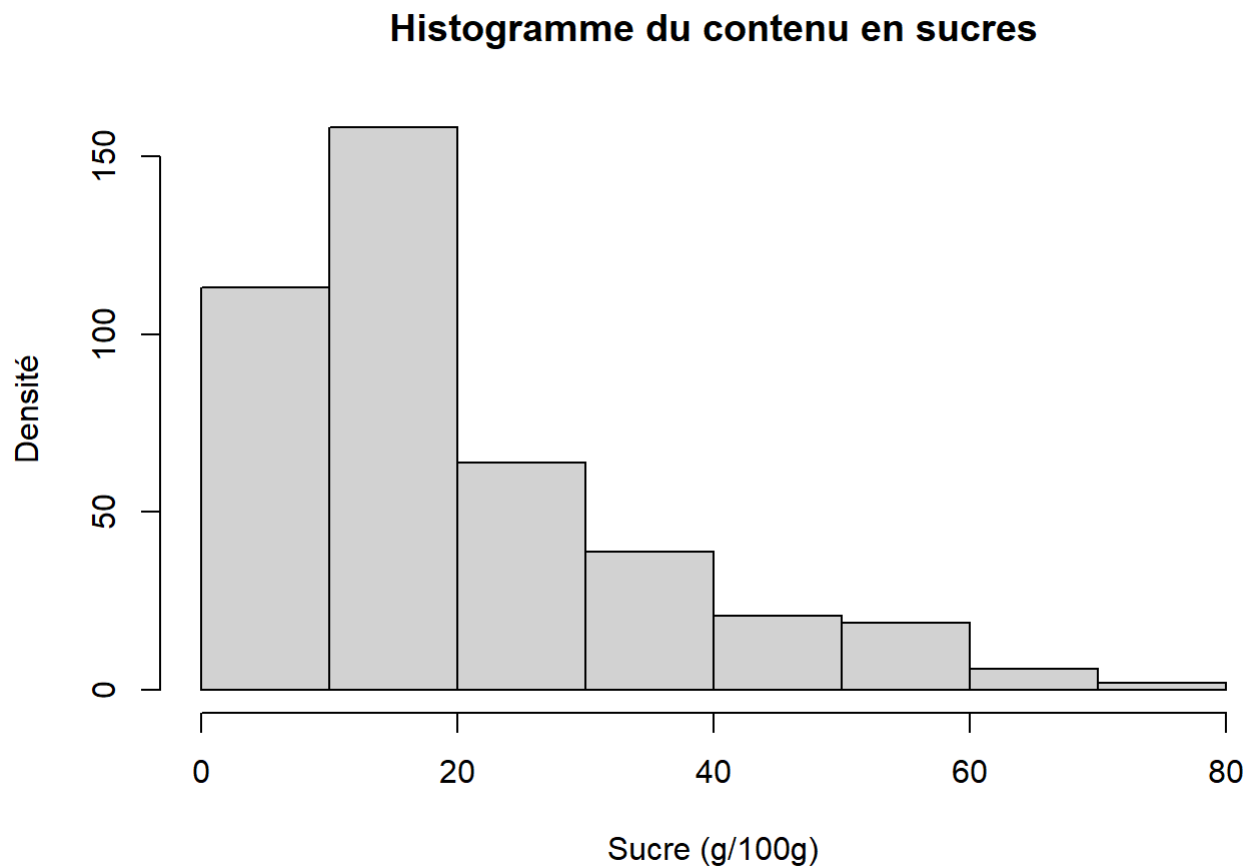
```

```
## attr(,"class")
## [1] "histogram"
##
## $glaces
## $breaks
## [1] 20 22 24 26 28
##
## $counts
## [1] 5 1 3 2
##
## $density
## [1] 0.22727273 0.04545455 0.13636364 0.09090909
##
## $mids
## [1] 21 23 25 27
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## $sorbets
## $breaks
## [1] 20 22 24 26 28
##
## $counts
## [1] 1 2 1 1
##
## $density
## [1] 0.1 0.2 0.1 0.1
##
## $mids
## [1] 21 23 25 27
##
## $xname
## [1] "X[[i]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## `$desserts glacés`
## $breaks
## [1] 15 20 25 30 35
##
## $counts
## [1] 2 5 2 3
##
## $density
## [1] 0.03333333 0.08333333 0.03333333 0.05000000
```

```
##  
## $mids  
## [1] 17.5 22.5 27.5 32.5  
##  
## $xname  
## [1] "X[[i]]"  
##  
## $equidist  
## [1] TRUE  
##  
## attr("class")  
## [1] "histogram"
```

Histogramme du contenu en sucres des desserts

```
hist(V$Sucres, xlab= "Sucre (g/100g)", ylab= "Densité", main="Histogramme du contenu en sucres")
```

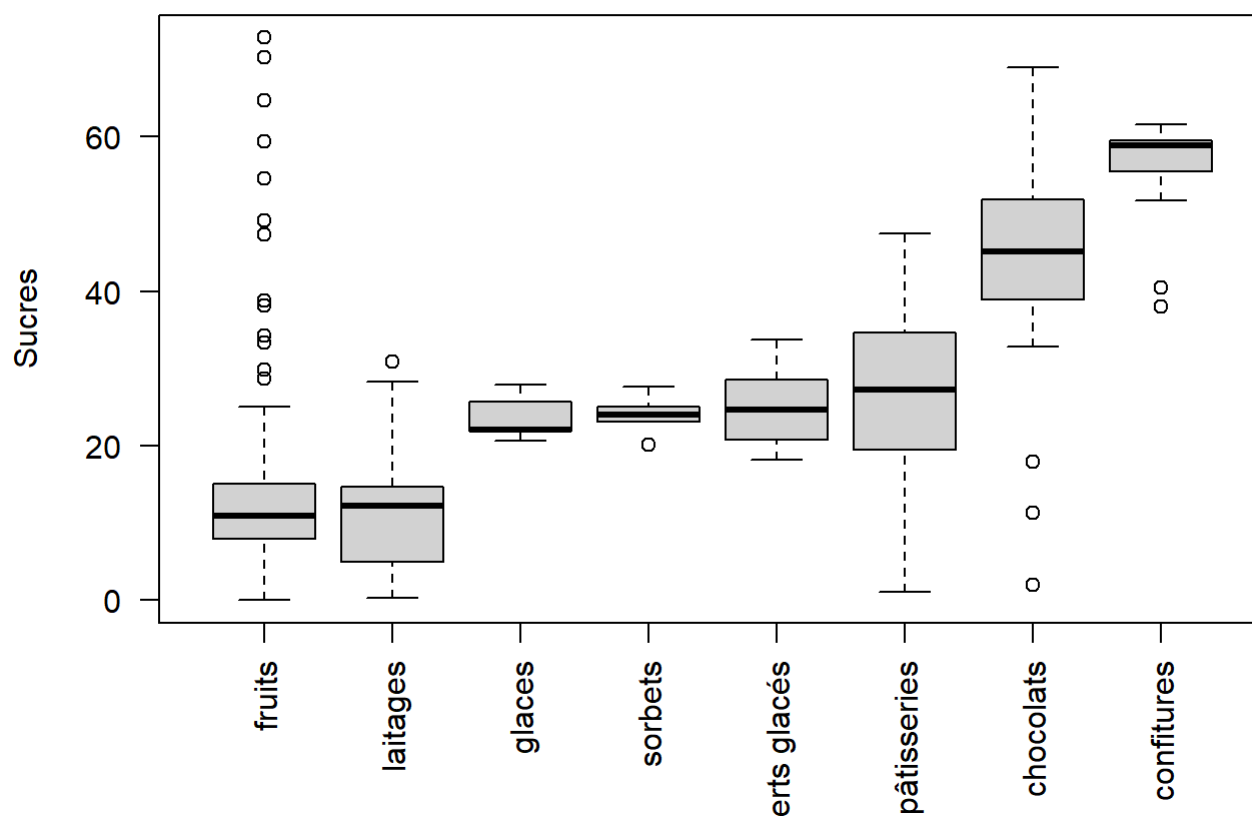


boxplot du contenu en sucres selon le groupe du dessert

```
pos<- as.numeric(factor(V$Groupes))
```

```
Groupeslev <- factor(V$Groupes, levels = c("fruits", "laitages","glaces", "sorbets",  
      "desserts glacés","pâtisseries", "chocolats",  
      "confitures"))
```

```
boxplot(V$Sucres~Groupeslev, ylab = "Sucres", xlab = "", las = 2)
```



```

par(mfrow=c(1, 2))
plot(V$Sucres, V$Energie,xlab = "Sucres", ylab = "Energie", type= "n")
points(V$Sucres[V$Groupes == "fruits"], V$Energie[V$Groupes == "fruits"], col= "black", pch=20)
points(V$Sucres[V$Groupes == "laitages"], V$Energie[V$Groupes == "laitages"], col= "deeppink", pch=20)
points(V$Sucres[V$Groupes == "glaces"], V$Energie[V$Groupes == "glaces"], col= "forestgreen", pch=20)
points(V$Sucres[V$Groupes == "sorbets"], V$Energie[V$Groupes == "sorbets"], col= "blue2", pch=20)
points(V$Sucres[V$Groupes == "desserts glacés"], V$Energie[V$Groupes == "desserts glacés"], col= "aquamarine1", pch=20)
points(V$Sucres[V$Groupes == "pâtisseries"], V$Energie[V$Groupes == "pâtisseries"], col= "maroon3", pch=20)
points(V$Sucres[V$Groupes == "chocolats"], V$Energie[V$Groupes == "chocolats"], col= "black", pch=18)
points(V$Sucres[V$Groupes == "confitures"], V$Energie[V$Groupes == "confitures"], col= "maroon", pch=18)
legend("bottomright", legend = c("fruits", "laitages", "glaces", "sorbets", "desserts glacés", "pâtisseries", "chocolats", "confitures"),
      pch= c(20, 20, 20, 20, 20, 20, 18, 18),
      col = c("black", "deeppink", "forestgreen", "blue2", "aquamarine1", "maroon3", "black", "maroon"), cex = 0.6)

plot(V$Lipides, V$Energie,xlab = "Graisse", ylab = "Energie", type= "n")
points(V$Lipides[V$Groupes == "fruits"], V$Energie[V$Groupes == "fruits"], col= "black", pch=20)
points(V$Lipides[V$Groupes == "laitages"], V$Energie[V$Groupes == "laitages"], col= "deeppink", pch=20)
points(V$Lipides[V$Groupes == "glaces"], V$Energie[V$Groupes == "glaces"], col= "forestgreen", pch=20)
points(V$Lipides[V$Groupes == "sorbets"], V$Energie[V$Groupes == "sorbets"], col= "blue2", pch=20)
points(V$Lipides[V$Groupes == "desserts glacés"], V$Energie[V$Groupes == "desserts glacés"], col= "aquamarine1", pch=20)
points(V$Lipides[V$Groupes == "pâtisseries"], V$Energie[V$Groupes == "pâtisseries"], col= "maroon3", pch=20)
points(V$Lipides[V$Groupes == "chocolats"], V$Energie[V$Groupes == "chocolats"], col= "black", pch=18)
points(V$Lipides[V$Groupes == "confitures"], V$Energie[V$Groupes == "confitures"], col= "maroon", pch=18)
legend("bottomright", legend = c("fruits", "laitages", "glaces", "sorbets", "desserts glacés", "pâtisseries", "chocolats", "confitures"),
      pch= c(20, 20, 20, 20, 20, 20, 18, 18),
      col = c("black", "deeppink", "forestgreen", "blue2", "aquamarine1", "maroon3", "black", "maroon"), cex = 0.6)

```

