

# Étude satisfaction à l'hôpital

Josselin Araujo

18 septembre 2022

Cette étude évalue la qualité de relation avec le personnel soignant et la quantité d'informations reçue par des patients lors d'un séjour à l'hôpital. 534 patients ont été recrutés sur plusieurs hôpitaux de la région parisienne.

Nous allons travailler sur le jeu de données `satisfaction_hopital.csv` regroupant les 534 patients.

Contenu du fichier :

- service : *service ayant accueilli le patient allant de 1 à 8*
- sexe : *sexe du patient (0 homme, 1 fille)*
- age : *âge du patient en années*
- profession :
  - 1 : *agriculteur exploitant*
  - 2 : *artisan, commerçant, chef d'entreprise*
  - 3 : *cadre, profession intellectuelle ou artistique, profession libérale*
  - 4 : *profession de l'enseignement, de la santé, du travail social ou de la fonction publique*
  - 5 : *employé*
  - 6 : *ouvrier*
  - 7 : *étudiant, militaire, chômeur sans avoir jamais travaillé*
  - 8 : *autre*
- amelioration.sante : *impression d'amélioration de la santé du fait du séjour à l'hôpital (codé de 0 : aggravé, à 3 : nettement amélioré)*
- amelioration.moral : *impression d'amélioration du moral du fait du séjour à l'hôpital (codé de 0 : aggravé, à 3 : nettement amélioré)*
- recommander : *recommander le service à son entourage (codé 0: non, 1 : oui, probablement, 2 : oui sûrement)*
- score.information : *score relatif à la qualité de l'information reçue pendant le séjour (score variant de 10 à 40)*
- score.relatif : *score relatif à la qualité des relations avec le personnel soignant pendant le séjour (score variant de 10 à 40)*

Importation des données :

```
smp <- read.csv2("/Users/josselin/Desktop/ING4/Mooc Intro statistique avec R/Semaine 1/satisfaction_hopital.csv")
```

Affichage de l'ensemble des variables :

```
names(smp)
```

```
## [1] "service"          "sexe"             "age"
## [4] "profession"       "amelioration.sante" "amelioration.moral"
## [7] "recommander"      "score.relation"   "score.information"
```

Affichage des 10 premières observations de chaque variable :

```
str(smp)

## 'data.frame':    534 obs. of  9 variables:
## $ service      : int  3 3 3 3 3 3 3 3 3 3 ...
## $ sexe         : int  0 1 1 0 1 0 0 0 0 1 ...
## $ age          : int  41 29 83 66 84 84 60 85 28 35 ...
## $ profession   : int  4 8 2 3 NA 6 3 3 3 3 ...
## $ amelioration.sante: int  1 2 2 2 NA 3 2 3 3 3 ...
## $ amelioration.moral: int  0 3 1 0 NA 3 1 1 1 2 ...
## $ recommander  : int  1 2 2 2 NA 2 1 2 1 NA ...
## $ score.relation : int  36 33 40 32 NA 39 31 NA 36 NA ...
## $ score.information : int  22 36 37 35 NA 28 30 NA 29 NA ...
```

## Présentation des pourcentages des sujets pour des variables catégorielles du fichier

Étudions trois variables catégorielles du fichier : *sexe*, *profession*, *service*

```
smp$profession.c <-factor(smp$profession,
                          labels = c("agriculteur","artisan","cadre","intermédiaire","employé",
                                      "ouvrier", "sans emploi","autre"))

smp$sexe.c <- factor(smp$sexe, labels = c("homme","femme"))

smp$service.c <-factor(smp$service)
```

### Pour les professions :

Affichons le nombre de personnes pour chaque profession parmi tous les patients :

```
table(smp$profession.c)

##
##  agriculteur      artisan      cadre intermédiaire      employé
##           1          38          124           88          69
##   ouvrier  sans emploi      autre
##           44          22          41
```

Affichons les effectifs pour chaque profession :

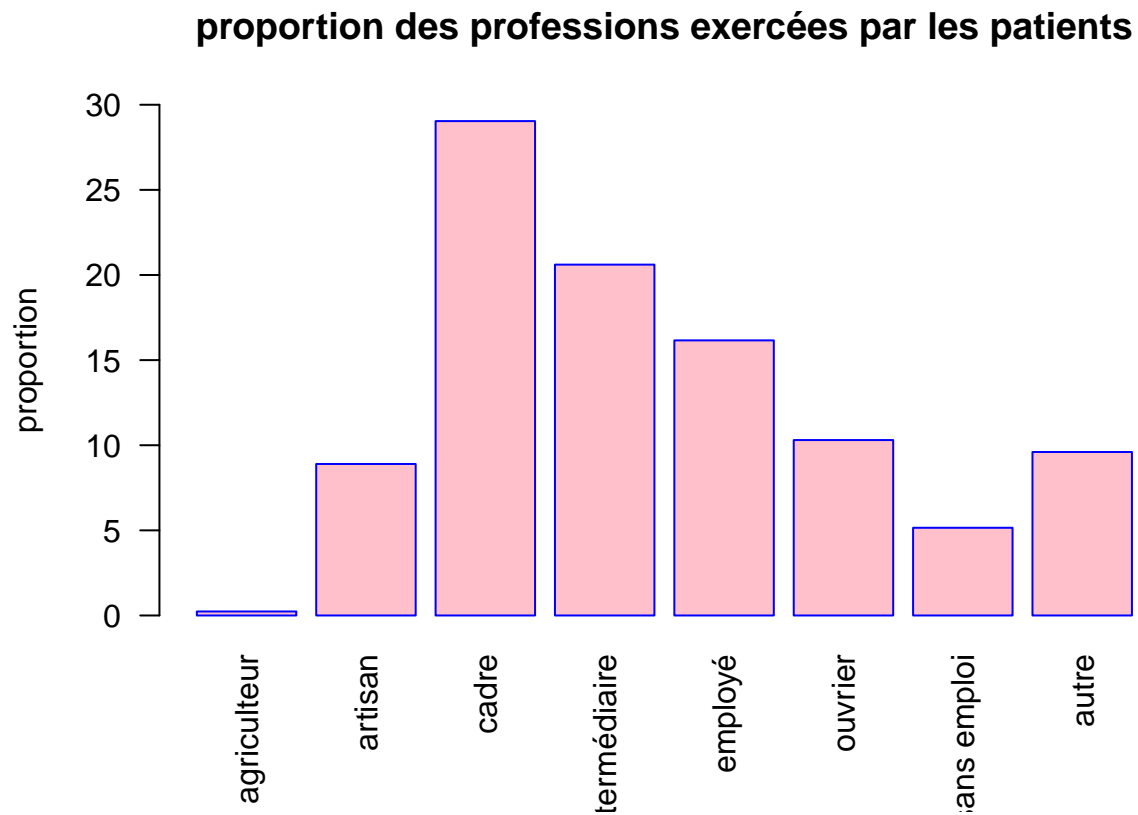
```
tabProf<-prop.table(table(smp$profession.c))
round(tabProf*100,2)

##
##  agriculteur      artisan      cadre intermédiaire      employé
##           0.23          8.90          29.04          20.61          16.16
##   ouvrier  sans emploi      autre
##           10.30          5.15          9.60
```

Visualisations les effectifs pour chaque profession :

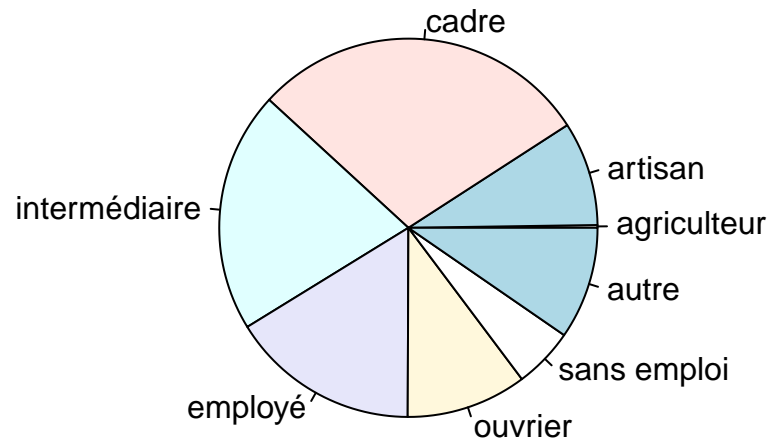
Nous apercevons que les cadres sont les plus représentés parmi les patients.

```
barplot(tabProf*100,  
        ylab = "proportion",  
        ylim = c(0,30),las=2,  
        border = "blue",  
        col="pink",  
        main="proportion des professions exercées par les patients"  
)
```



```
pie(table(smp$profession.c),  
     main="proportion des professions exercées par les patients"  
)
```

## proportion des professions exercées par les patients



Pour les services :

```
table(smp$service)
```

```
##  
##  1  2  3  4  5  6  7  8  
## 65 59 70 69 71 64 67 69
```

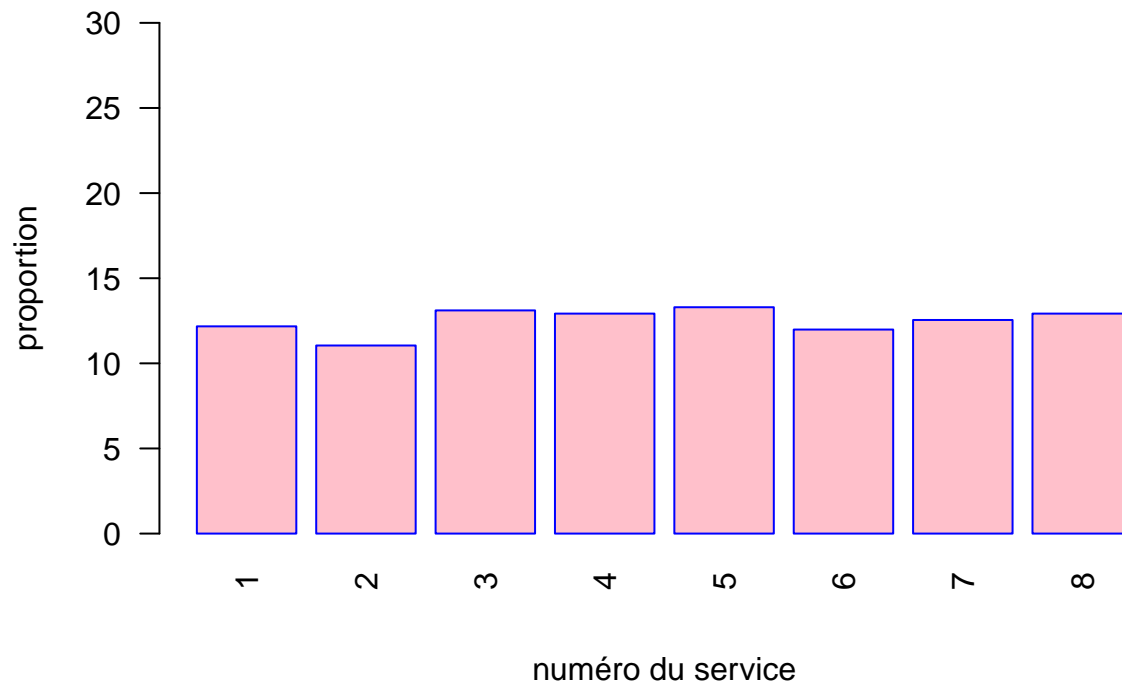
```
tabService<-prop.table(table(smp$service.c))  
round(tabService*100,2)
```

```
##  
##    1    2    3    4    5    6    7    8  
## 12.17 11.05 13.11 12.92 13.30 11.99 12.55 12.92
```

```
barplot(tabService*100,  
        ylab = "proportion",  
        ylim = c(0,30),  
        xlab = "numéro du service",  
        las=2,  
        main="proportion du nombre de patients accueilli par chaque service ",  
        border ="blue",
```

```
)
col="pink"
```

### proportion du nombre de patients accueilli par chaque service



Nous apercevons que les patients sont plutôt bien repartis équitablement au sein des services de l'hôpital.

Pour le sexe :

```
table(smp$sexe.c)
```

```
##
## homme femme
## 268 266
```

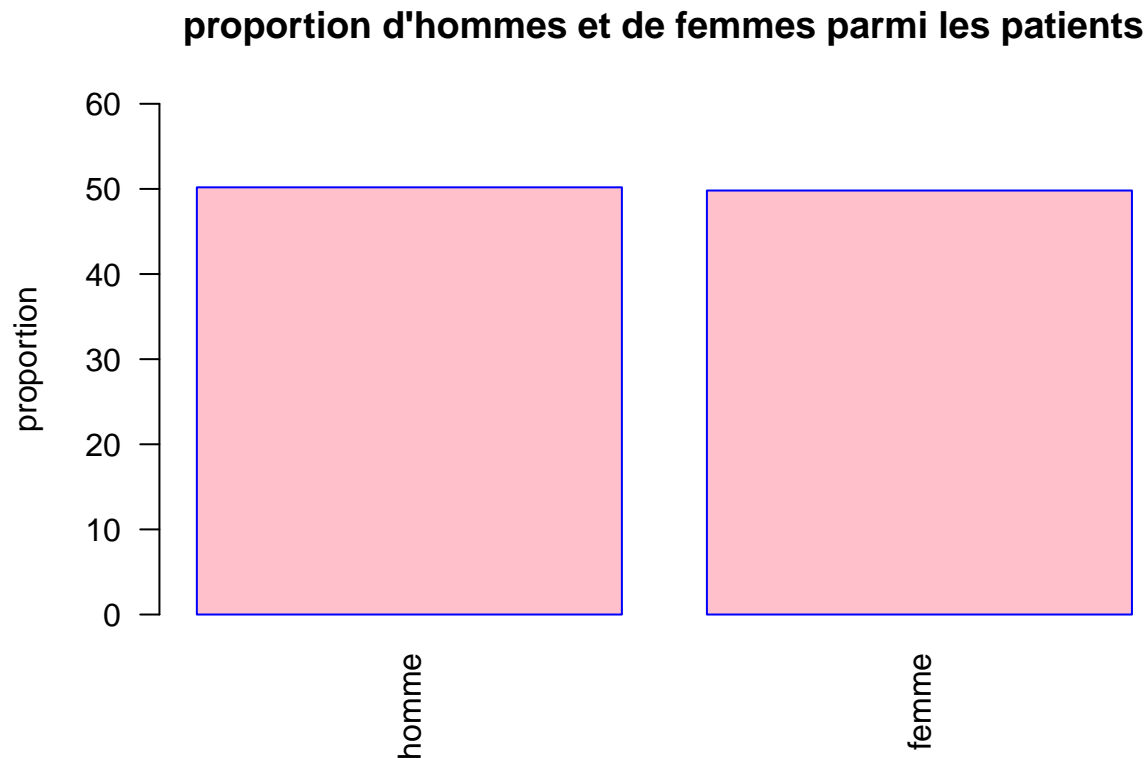
```
tabSexe<-prop.table(table(smp$sexe.c))
round(tabSexe*100,2)
```

```
##
## homme femme
## 50.19 49.81
```

```

barplot(tabSexe*100,ylab = "proportion",
        ylim = c(0,60),
        las=2,
        border ="blue",
        main="proportion d'hommes et de femmes parmi les patients",
        col="pink"
)

```



## Présentation de la moyenne, médiane, écart-type, minimum, maximum des autres variables

Pour l'âge, `summary()` nous offre un résumé de la distribution de cette variable :

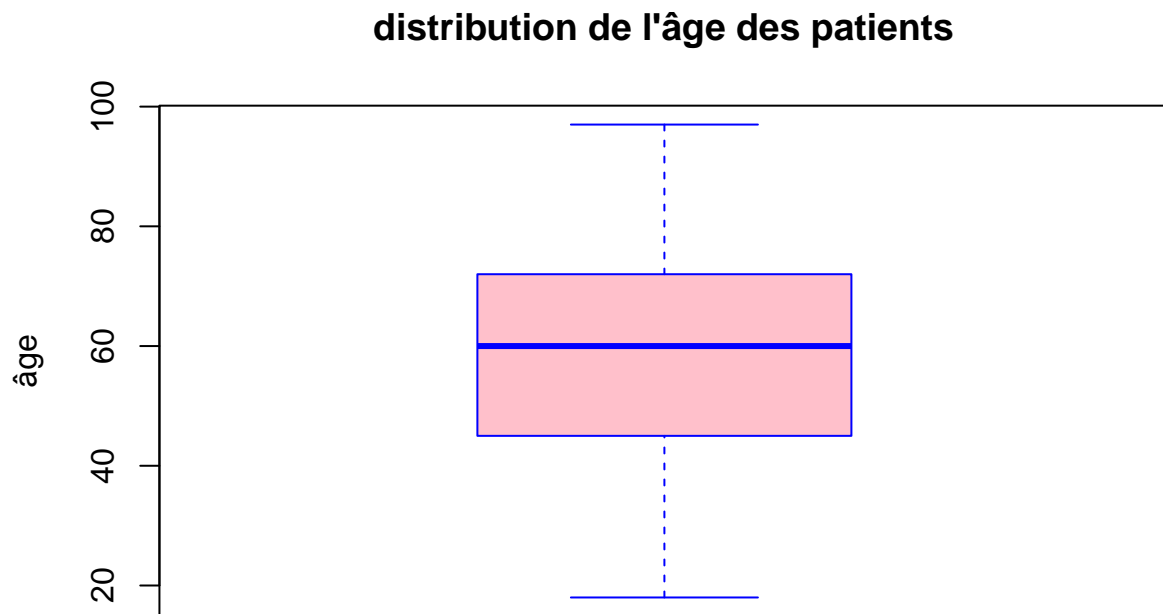
```
summary(smp$age)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	18.00	45.00	60.00	58.21	72.00	97.00	6

Visualisation de la distribution de la variable âge :

On remarque que la moitié des patients ont entre 45 et 72 ans, les patients ont donc majoritairement un âge avancé.

```
boxplot(smp$age,
        ylab="âge",
        main="distribution de l'âge des patients",
        border="blue",
        col="pink"
)
```



Autre méthode :

```
require(prettyR)
```

## Le chargement a nécessité le package : prettyR

```
describe(smp[,c(3,5,6,7,8,9)],num.desc = c("mean","median","sd","min","max","valid.n"))
```

## Description of smp[, c(3, 5, 6, 7, 8, 9)]

##

## Numeric

	mean	median	sd	min	max	valid.n
## age	58.21	60	17.81	18	97	528
## amelioration.sante	2.23	2	0.77	0	3	376
## amelioration.moral	1.68	1	0.95	0	3	383
## recommander	1.62	2	0.56	0	2	405
## score.relation	35.22	36	4.62	13	40	349
## score.information	31.91	33	6.79	13	40	358

## Relations avec les patients

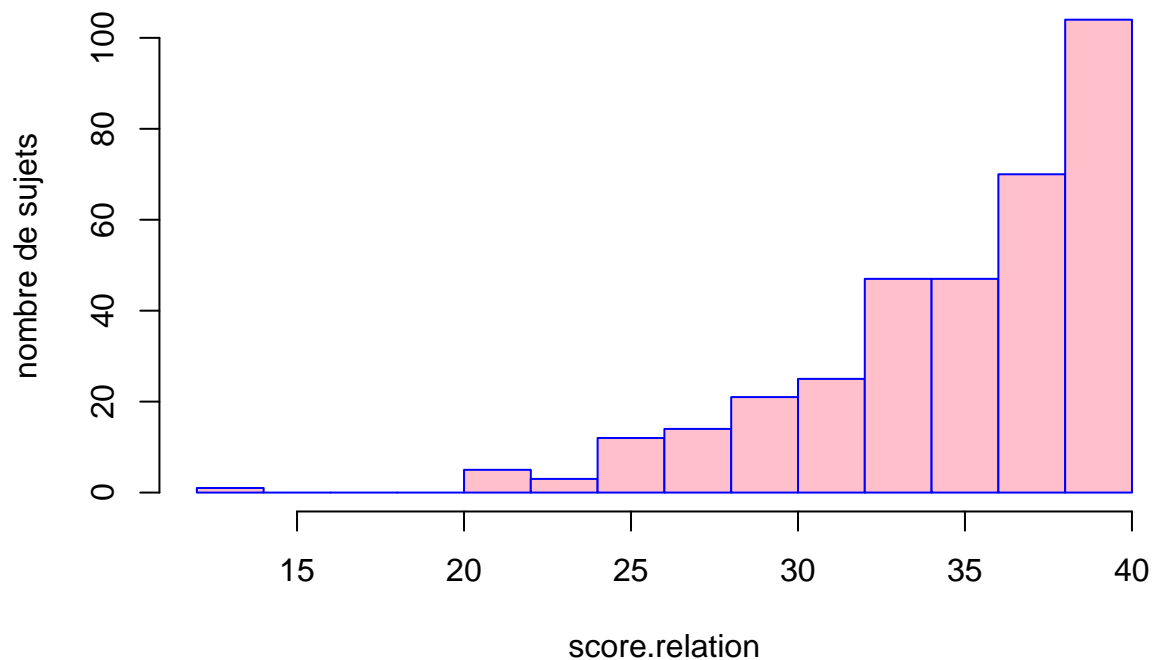
Distribution du score relatif à la qualité de la relation avec le personnel soignant pendant le séjour du patient

```
table(smp$score.relation)
```

```
##  
## 13 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40  
##  1  4  1  1  2  7  5  3 11  7 14 10 15 28 19 23 24 25 45 40 64
```

```
hist(smp$score.relation,  
     main = "Historigramme de la variable score.relation",  
     xlab = "score.relation",  
     ylab = "nombre de sujets",  
     border = "blue",  
     col = "pink"  
)
```

### Historigramme de la variable score.relation



```
summary(smp$score.relation)
```

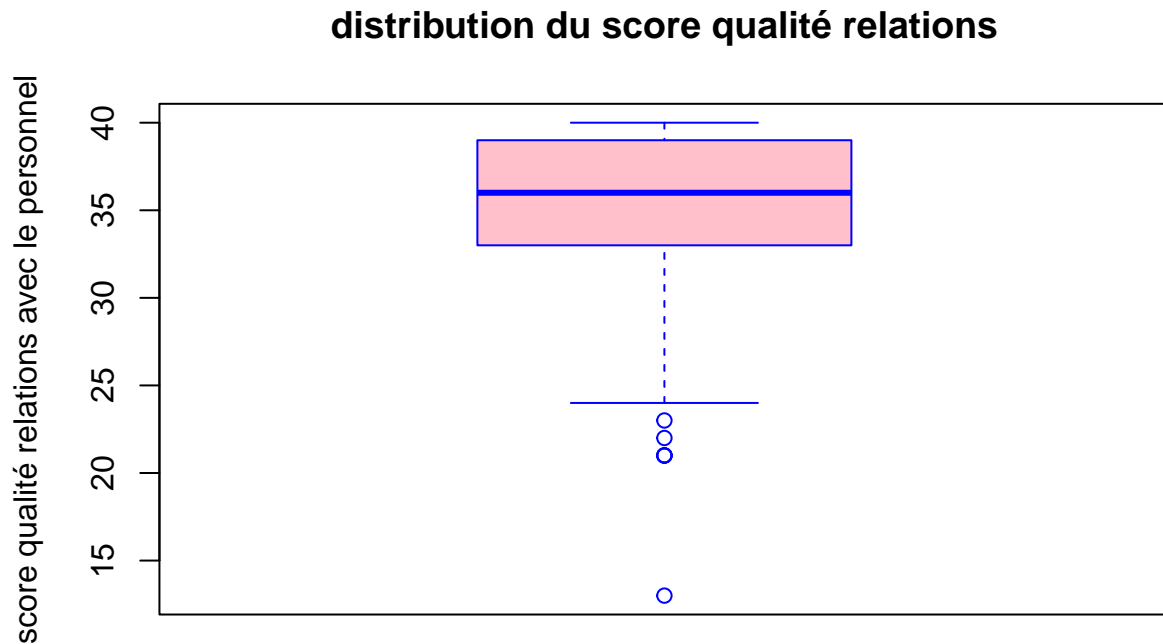
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
## 13.00   33.00   36.00   35.22  39.00   40.00   185
```



```

boxplot(smp$score.relation,
        ylab = "score qualité relations avec le personnel",
        main="distribution du score qualité relations",
        border ="blue",
        col="pink"
)

```



Nous pouvons donc constater que les patients sont en très grande majorité satisfaits de la qualité de la relation avec le personnel soignant pendant leur séjour à l'hôpital.

**Distribution du score relatif à la qualité de la relation avec le personnel soignant pendant le séjour à l'hôpital du patient en fonction du sexe.**

```

table(smp$score.relation[which(smp$sexe.c=="homme")])

```

```

##
## 13 21 22 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
##   1  2  1  2  3  1  1  4  2  6  6  5 16  9 15 13 15 27 24 30

```

```

table(smp$score.relation[which(smp$sexe.c=="femme")])

```

```
##
## 21 23 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
##  2  1  4  4  2  7  5  8  4 10 12 10  8 11 10 18 16 34
```

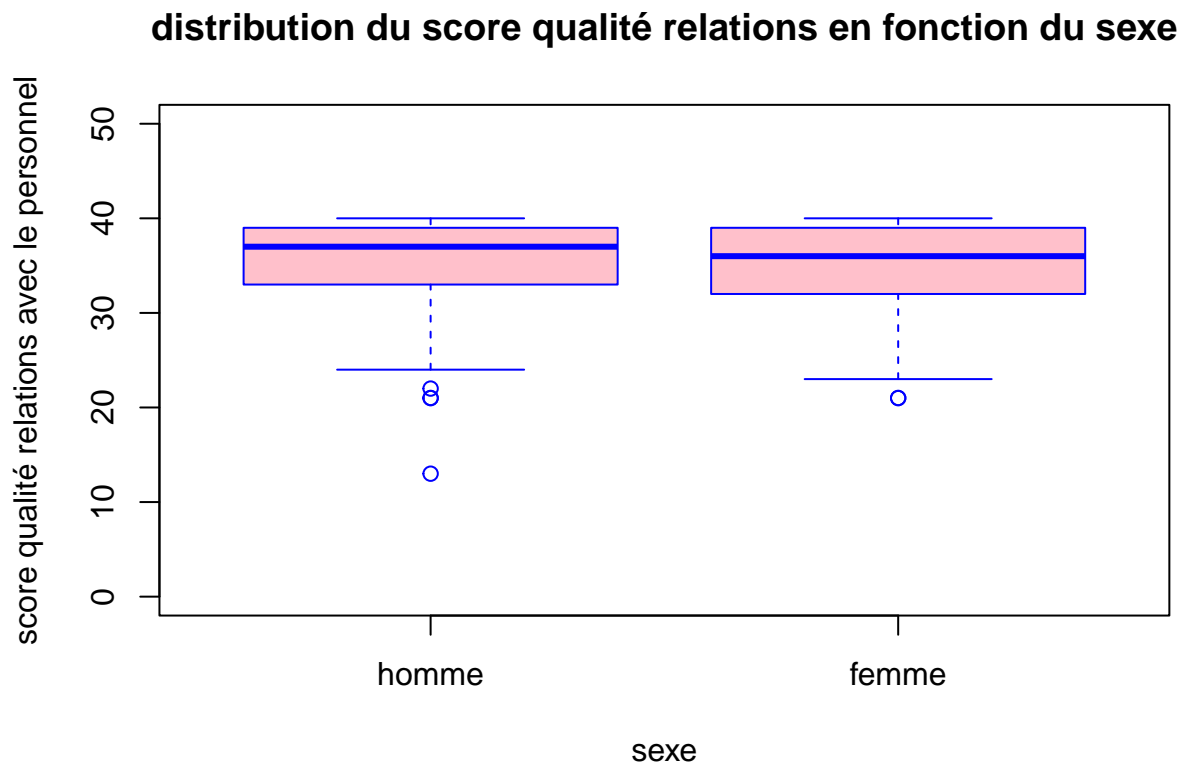
```
summary(smp$score.relation[which(smp$sexe.c=="homme")])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's
##      13.00   33.00   37.00   35.48   39.00   40.00     85
```

```
summary(smp$score.relation[which(smp$sexe.c=="femme")])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's
##      21.00   32.00   36.00   34.93   39.00   40.00    100
```

```
boxplot(smp$score.relation~smp$sexe.c,
        ylim=c(0,50),
        xlab = "sexe",
        ylab="score qualité relations avec le personnel",
        main="distribution du score qualité relations en fonction du sexe",
        border ="blue",
        col="pink" )
```



Nous pouvons donc constater que les hommes sont globalement autant satisfaits que les femmes sur la qualité de la relation personnel soignant pendant leur séjour à l'hôpital.

## Distribution du score relatif à la qualité de la relation avec le personnel soignant pendant le séjour à l'hôpital du patient en fonction de l'âge.

Découpons en 5 intervalles d'âges : les moins de 20 ans, les patients entre 20 et moins de 40 ans, ceux entre 40 et moins de 60 ans, puis ceux entre 60 et moins de 80 ans et enfin les personnes de plus de 80 ans.

```
smp$age.cat<-factor(smp$age)
table(smp$age.cat)
```

```
##
## 18 19 20 21 22 23 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
##  3  2  2  3  3  3  2  5  4  3  6  7  3  8  2  4  4  6  3  5  8 12  5  7 11  7
## 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70
## 12  7  7  7 11 11  6  8  7  8 14  7  9 11 10 15  8  8 11 14 11 11 14  7 14 12
## 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 95 96 97
##  4 10  4  6 18  7  8  9  6  7  7  7  8 12  5  1  3  5  2  3  3  2  1  1  1
```

```
table(smp$age.cat)[0:2]
```

```
##
## 18 19
##  3  2
```

```
levels(smp$age.cat)[0:2]<-"-20"
table(smp$age.cat)
```

```
##
## -20  20  21  22  23  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39
##  5  2  3  3  3  2  5  4  3  6  7  3  8  2  4  4  6  3  5  8
## 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
## 12  5  7 11  7 12  7  7  7 11 11  6  8  7  8 14  7  9 11 10
## 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79
## 15  8  8 11 14 11 11 14  7 14 12  4 10  4  6 18  7  8  9  6
## 80 81 82 83 84 85 86 87 88 89 90 91 92 95 96 97
##  7  7  7  8 12  5  1  3  5  2  3  3  2  1  1  1
```

```
table(smp$age.cat)[2:20]
```

```
##
## 20 21 22 23 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
##  2  3  3  3  2  5  4  3  6  7  3  8  2  4  4  6  3  5  8
```

```
levels(smp$age.cat)[2:20]<-"20 et 40-"
table(smp$age.cat)
```

```
##
##      -20 20 et 40-      40      41      42      43      44      45
##      5      81      12      5      7      11      7      12
##     46      47      48      49      50      51      52      53
##      7      7      7      11      11      6      8      7
```

```
##      54      55      56      57      58      59      60      61
##      8      14      7      9      11      10      15      8
##     62     63     64     65     66     67     68     69
##      8      11      14      11      11      14      7      14
##     70     71     72     73     74     75     76     77
##     12      4     10      4      6     18      7      8
##     78     79     80     81     82     83     84     85
##      9      6      7      7      7      8     12      5
##     86     87     88     89     90     91     92     95
##      1      3      5      2      3      3      2      1
##     96     97
##      1      1
```

```
table(smp$age.cat)[3:22]
```

```
##
## 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
## 12 5 7 11 7 12 7 7 7 11 11 6 8 7 8 14 7 9 11 10
```

```
levels(smp$age.cat)[3:22]<-"40 et 60-"
table(smp$age.cat)
```

```
##
##    -20 20 et 40- 40 et 60-      60      61      62      63      64
##      5      81      177      15      8      8      11      14
##     65     66     67     68     69     70     71     72
##     11     11     14      7     14     12      4     10
##     73     74     75     76     77     78     79     80
##      4      6     18      7      8      9      6      7
##     81     82     83     84     85     86     87     88
##      7      7      8     12      5      1      3      5
##     89     90     91     92     95     96     97
##      2      3      3      2      1      1      1
```

```
table(smp$age.cat)[4:23]
```

```
##
## 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79
## 15 8 8 11 14 11 11 14 7 14 12 4 10 4 6 18 7 8 9 6
```

```
levels(smp$age.cat)[4:23]<-"60 et 80-"
table(smp$age.cat)
```

```
##
##    -20 20 et 40- 40 et 60- 60 et 80-      80      81      82      83
##      5      81      177      197      7      7      7      8
##     84     85     86     87     88     89     90     91
##     12      5      1      3      5      2      3      3
##     92     95     96     97
##      2      1      1      1
```

```
table(smp$age.cat)[5:20]
```

```
##
## 80 81 82 83 84 85 86 87 88 89 90 91 92 95 96 97
## 7 7 7 8 12 5 1 3 5 2 3 3 2 1 1 1
```

```
levels(smp$age.cat)[5:20]<-"80+"
table(smp$age.cat)
```

```
##
##      -20 20 et 40- 40 et 60- 60 et 80-      80+
##          5      81      177      197      68
```

Nous venons donc de découper en 5 intervalles afin de visualiser plus facilement la variation du score relatif à la qualité de la relation avec le personnel soignant pendant le séjour à l'hôpital du patient en fonction de l'âge.

```
summary(smp$score.relation[which(smp$age.cat=="-20")])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
## 25.00   28.75   30.50   30.50   32.25   36.00         1
```

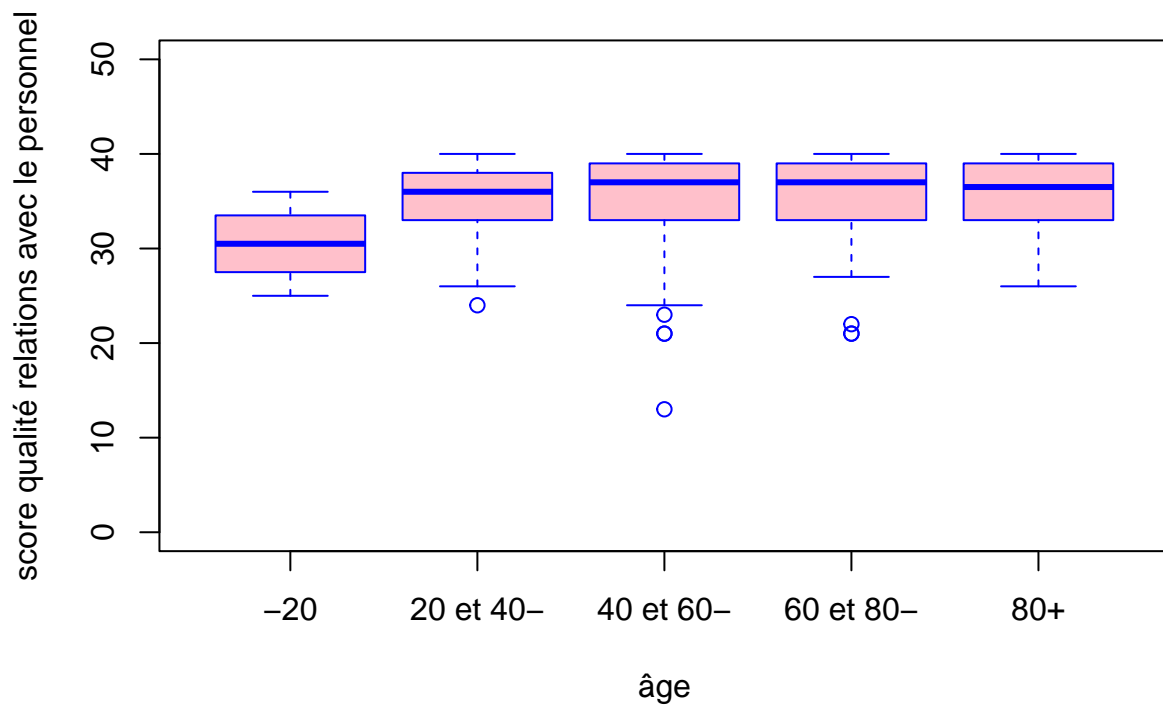
```
summary(smp$score.relation[which(smp$age.cat=="20 et 40-")])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
## 24.00   33.00   36.00   34.86   38.00   40.00        16
```

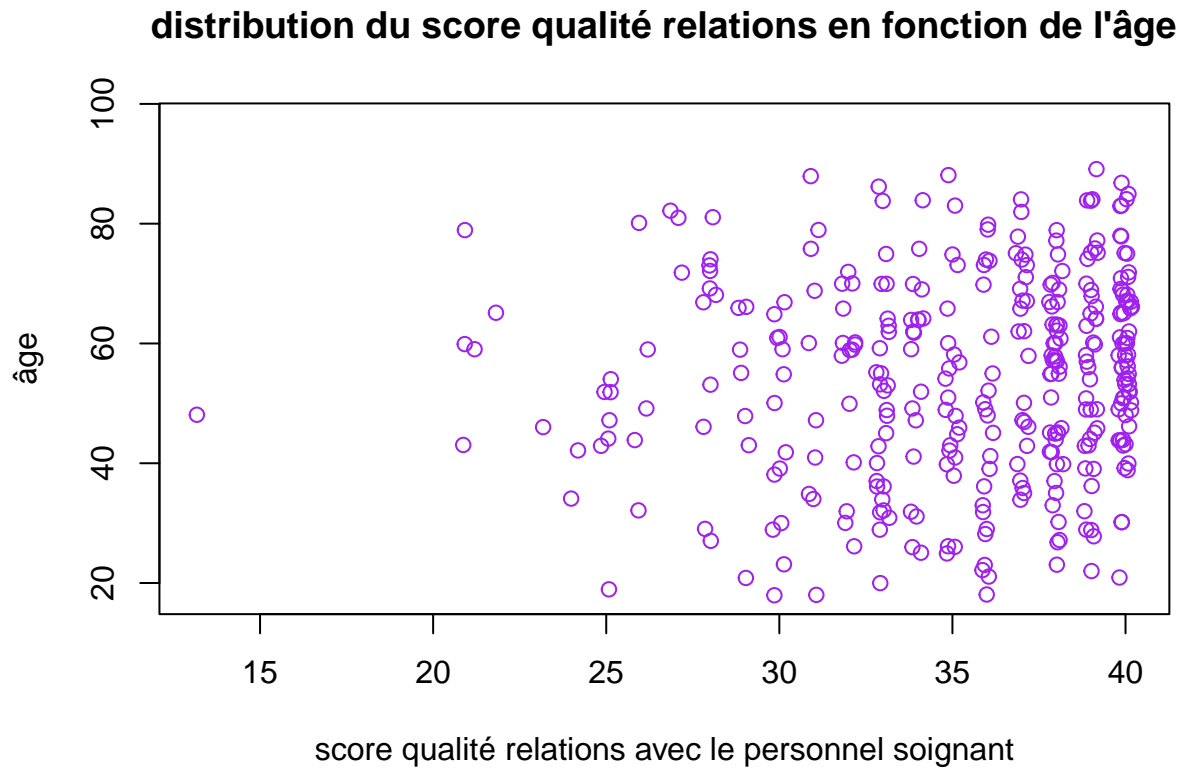
Visualisons ces informations à l'aide du graphique suivant :

```
boxplot(smp$score.relation~smp$age.cat,
        ylim=c(0,50),
        xlab = "âge",
        ylab="score qualité relations avec le personnel",
        border ="blue",
        col="pink",
        main="distribution du score qualité relations en fonction de l'âge"
        )
```

## distribution du score qualité relations en fonction de l'âge



```
plot(jitter(smp$score.relation),
     jitter(smp$age),
     xlab = "score qualité relations avec le personnel soignant",
     main="distribution du score qualité relations en fonction de l'âge",
     ylab = "âge",
     col="purple",
     )
```

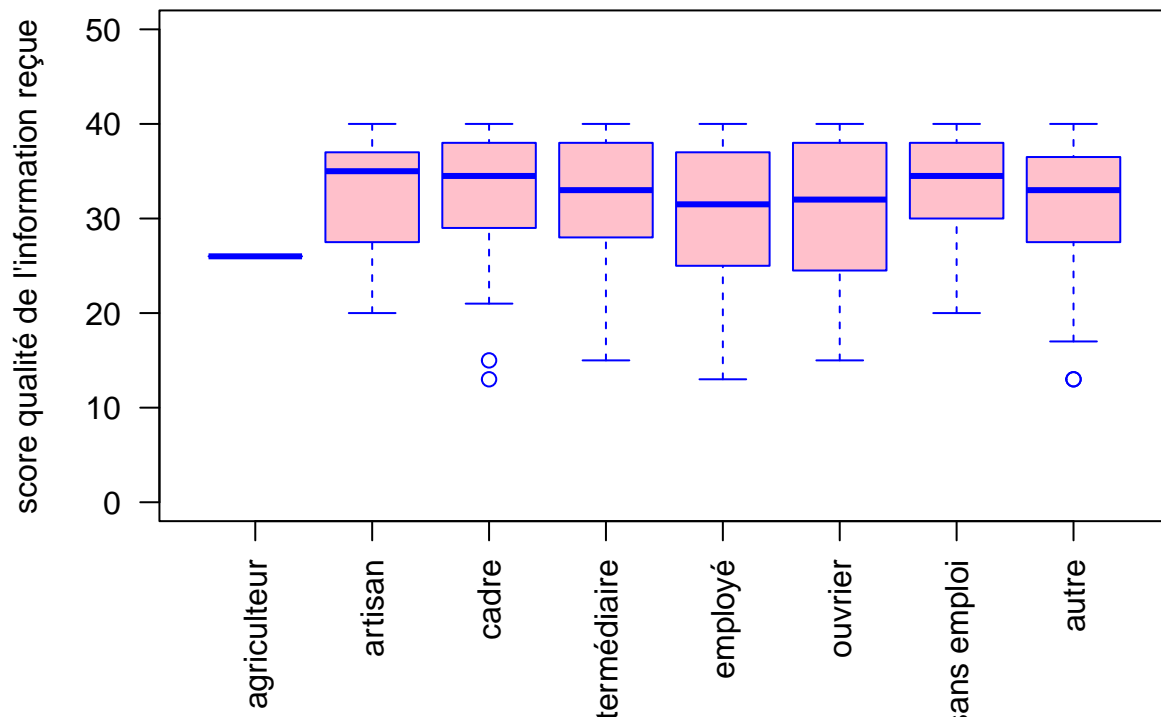


Nous pouvons donc dire que les patients de plus 20 ans sont en général plus satisfaits de la qualité de la relation avec le personnel soignant que les patients de moins de 20 ans.

### Distribution du score relatif à la qualité de l'information reçue pendant le séjour en fonction de la profession du patient

```
boxplot(smp$score.information~smp$profession.c,
        ylim=c(0,50),
        xlab = " ",
        ylab="score qualité de l'information reçue",
        border ="blue",
        col="pink",
        main="score qualité de l'information reçue en fonction de la profession",
        las=2,
        )
```

## score qualité de l'information reçue en fonction de la profession



Nous pouvons donc constater que globalement les patients exerçant des professions différentes sont globalement plutôt tous autant satisfaits sur la qualité l'information reçue pendant le séjour à l'hôpital.

## Conclusion

Lors de cette étude avons pu analyser différentes variables issus du jeu de données. Nous avons pu extraire des informations en étudiant plusieurs variables comme la profession, le sexe, l'âge des patients. Nous avons également pu voir les liens entre différentes variables, comme la variation du score de la qualité des relations en fonction du sexe ou de l'âge ou encore la variation du score de la qualité de l'information reçue par les patients en fonction de la profession.