Mortalité Hospitaliere

load dada deces

```
source("connection_db.R")
```

inspect data deces

```
nrow(data)
```

```
## [1] 13946
```

```
ncol(data)
```

```
## [1] 14
```

```
dim(data)
```

```
## [1] 13946    14
```

```
names(data)
```

```
##  [1] "DINS"          "WILAYAR"       "COMMUNER"      "LD"
##  [5] "STRUCTURED"    "SERVICEHOSPIT" "DUREEHOSPIT"   "SEX"
##  [9] "Years"         "Days"          "Profession"    "CD"
## [13] "CODECIM0"      "CODECIM"
```

view data deces

```
str(data)
```

```
## 'data.frame':    13946 obs. of  14 variables:
##  $ DINS         : Date, format: "2020-01-27" "2019-08-10" ...
##  $ WILAYAR      : int  17000 17000 17000 17000 17000 17000 17000 17000 17000 17000 ...
##  $ COMMUNER     : int  935 917 935 947 920 935 935 935 917 917 ...
##  $ LD           : Factor w/ 5 levels "AAP","DOM","SSP",..: 3 3 2 3 3 3 3 3 2 3 ...
##  $ STRUCTURED   : Factor w/ 10 levels "1","2","3","4",..: 3 6 3 3 3 3 3 3 6 6 ...
##  $ SERVICEHOSPIT: Factor w/ 23 levels "0","1","2","3",..: 8 20 20 20 11 20 20 11 21 20 ...
##  $ DUREEHOSPIT  : int  0 1 0 4 5 1 0 1 0 1 ...
##  $ SEX          : Factor w/ 2 levels "F","M": 2 2 2 2 1 1 2 1 1 1 ...
##  $ Years        : int  71 56 85 77 0 84 80 0 88 36 ...
##  $ Days         : int  26100 20517 31101 28125 5 31044 29454 1 32391 13366 ...
##  $ Profession   : Factor w/ 16 levels "0","1","3","4",..: 1 1 1 1 10 11 1 10 1 1 ...
##  $ CD           : Factor w/ 3 levels "CI","CN","CV": 2 2 1 2 2 2 2 2 1 2 ...
##  $ CODECIM0     : int  10 9 18 10 16 9 1 17 0 10 ...
##  $ CODECIM      : int  751 690 1370 751 1145 675 10 1271 0 751 ...
```

modalité des variable deces

```
unique(data$LD)
```

```
## [1] SSP  DOM  VP   AAP  SSPV
## Levels: AAP DOM SSP SSPV VP
```

```
unique(data$SEX)
```

```
## [1] M F
## Levels: F M
```

```
unique(data$Profession)
```

```
##  [1] 0  12 13 1  19 11 20 15 16 3  7  6  5  4  14 9
## Levels: 0 1 3 4 5 6 7 9 11 12 13 14 15 16 19 20
```

```
unique(data$CD)
```

```
## [1] CN CI CV
## Levels: CI CN CV
```

summary data deces

```
summary(data)
```

```
##       DINS                WILAYAR         COMMUNER         LD
##  Min.   :0009-11-20   Min.   : 1000   Min.   :   1.0   AAP :  537
##  1st Qu.:2018-01-30   1st Qu.:17000   1st Qu.: 916.0   DOM :  699
##  Median :2019-11-24   Median :17000   Median : 924.0   SSP :12548
##  Mean   :2019-04-17   Mean   :17166   Mean   : 935.6   SSPV:   11
##  3rd Qu.:2021-03-19   3rd Qu.:17000   3rd Qu.: 935.0   VP  :  151
##  Max.   :2022-05-31   Max.   :47000   Max.   :2297.0
##  NA's   :52
##    STRUCTURED   SERVICEHOSPIT   DUREEHOSPIT      SEX          Years
##  1      :4508   20     :5199   Min.   :-4886.0   F:5790   Min.   :  0.00
##  2      :2170   10     :3100   1st Qu.:    0.0   M:8156   1st Qu.:  0.00
##  4      :2065   18     :1427   Median :    1.0            Median : 52.00
##  5      :1824   7      :1126   Mean   :  591.1            Mean   : 43.39
##  3      :1714   21     : 913   3rd Qu.:    5.0            3rd Qu.: 76.00
##  6      : 814   15     : 709   Max.   :29220.0            Max.   :100.00
##  (Other): 851   (Other):1472
##       Days          Profession    CD           CODECIM0         CODECIM
##  Min.   : -279   0      :9423   CI:  592   Min.   : 0.000   Min.   :   0.0
##  1st Qu.:   52   1      :2415   CN:12992   1st Qu.: 4.000   1st Qu.: 350.0
##  Median :19226   20     : 726   CV:  362   Median :10.000   Median : 711.0
##  Mean   :15986   13     : 419              Mean   : 9.687   Mean   : 723.8
##  3rd Qu.:28106   12     : 408              3rd Qu.:16.000   3rd Qu.:1155.0
##  Max.   :36857   19     : 378              Max.   :23.000   Max.   :2039.0
##                  (Other): 177
```

lieux du deces

lieux:table

```
lieux <- table(data$LD)#,useNA = "always"
sort(lieux)
```

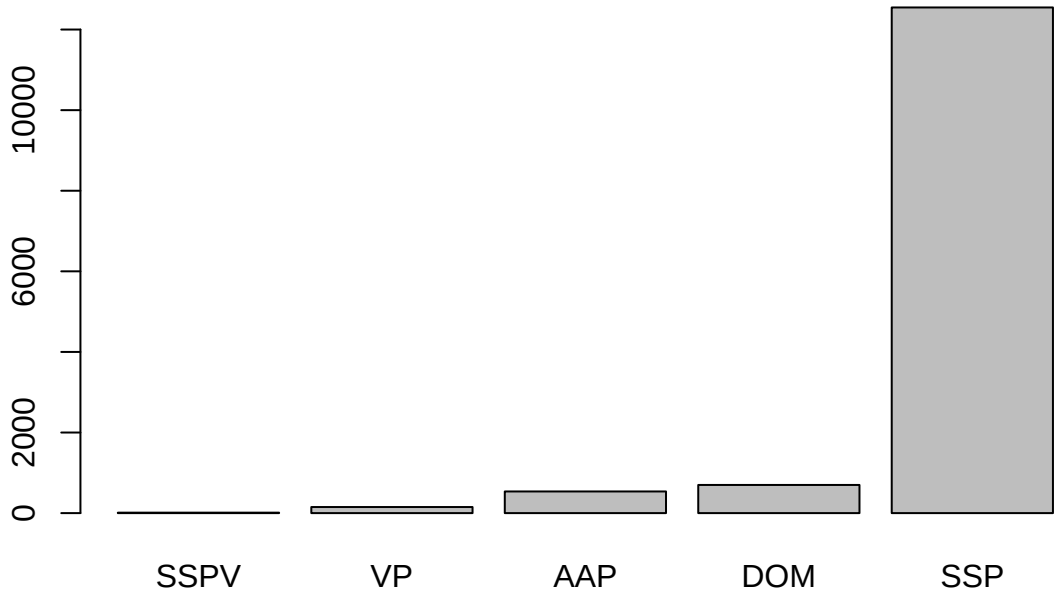
```
##
##  SSPV    VP   AAP   DOM   SSP
##    11   151   537   699 12548
```

```
summary(lieux)
```

```
## Number of cases in table: 13946
## Number of factors: 1
```

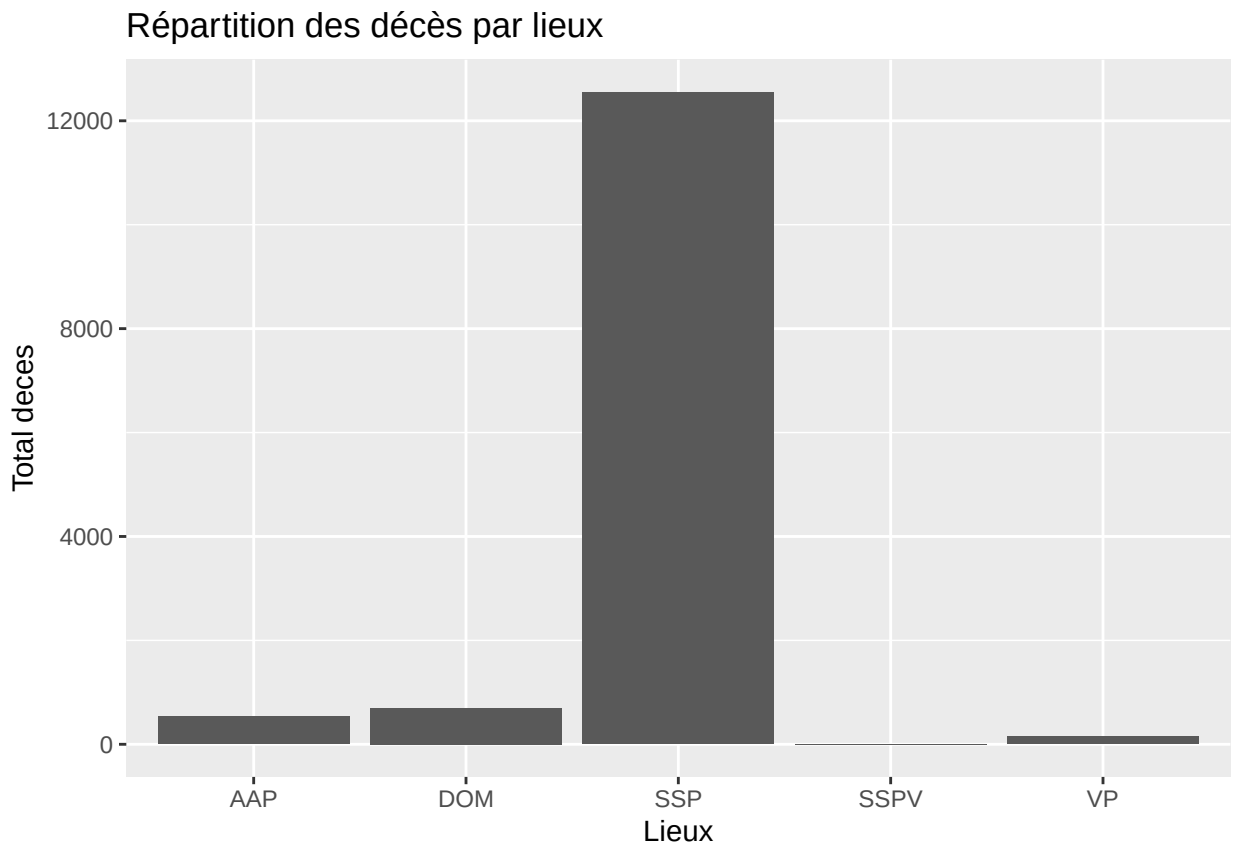
lieux:graphic:r-base

```
barplot(sort(lieux))
```



lieux:grafic:r-ggplot2

```
ggplot(data = data, mapping = aes(x = LD)) +
  geom_bar()+
  labs(title = "Répartition des décès par lieux",
       x = "Lieux",
       y = "Total deces")
```



structure sanitaire

structure:table

```
structure <- table(data$STRUCTURED)#,useNA = "always"
sort(structure)
```

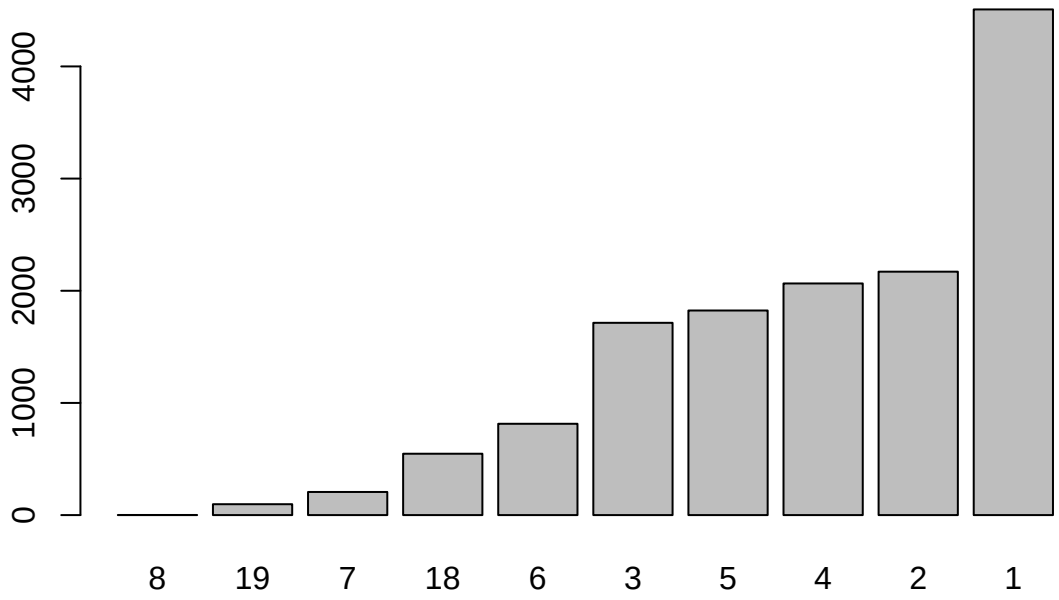
```
##
##    8   19    7   18    6    3    5    4    2    1
##    1   97  206  547  814 1714 1824 2065 2170 4508
```

```
summary(structure)
```

```
## Number of cases in table: 13946
## Number of factors: 1
```

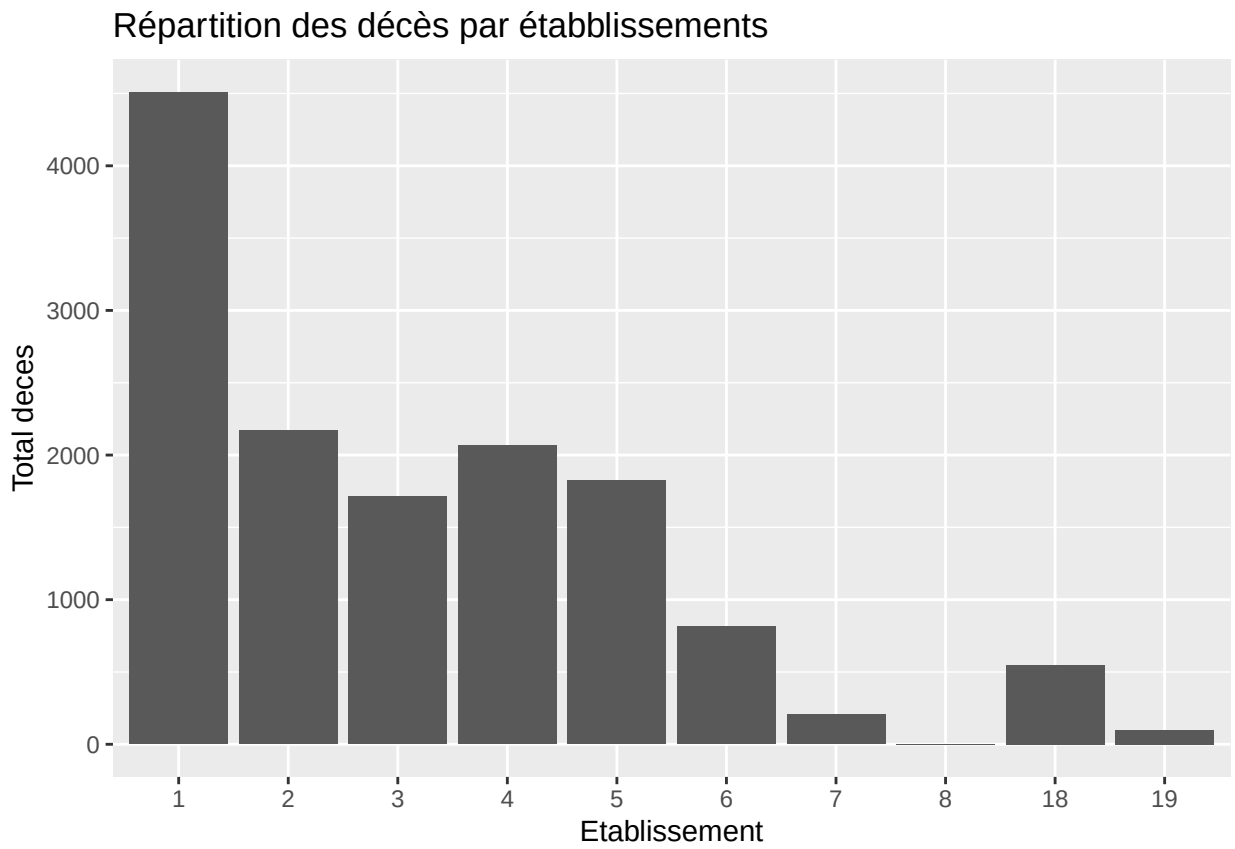
structure:graphic r-base

```
barplot(sort(structure))
```



structure:graphic:r-ggplot2

```
ggplot(data = data, mapping = aes(x = STRUCTURED)) +
  geom_bar()+
  labs(title = "Répartition des décès par étabblissements",
       x = "Etablissement",
       y = "Total deces")
```



service du deces

service:table

```
service <- table(data$SERVICEHOSPIT)#,useNA = "always"
sort(service)
```

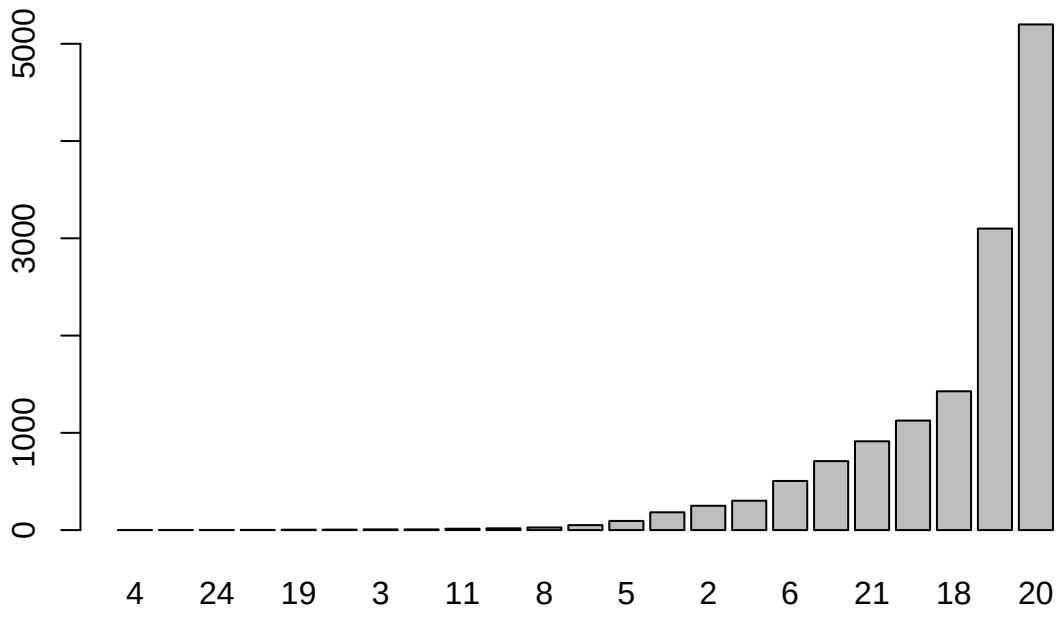
```
##
##    4   17   24   12   19   14    3    9   11   25    8    1    5    0    2   16
##    1    1    1    2    4    5    7    7   14   18   27   51   94  183  250  302
##    6   15   21    7   18   10   20
##  505  709  913 1126 1427 3100 5199
```

```
summary(service)
```

```
## Number of cases in table: 13946
## Number of factors: 1
```

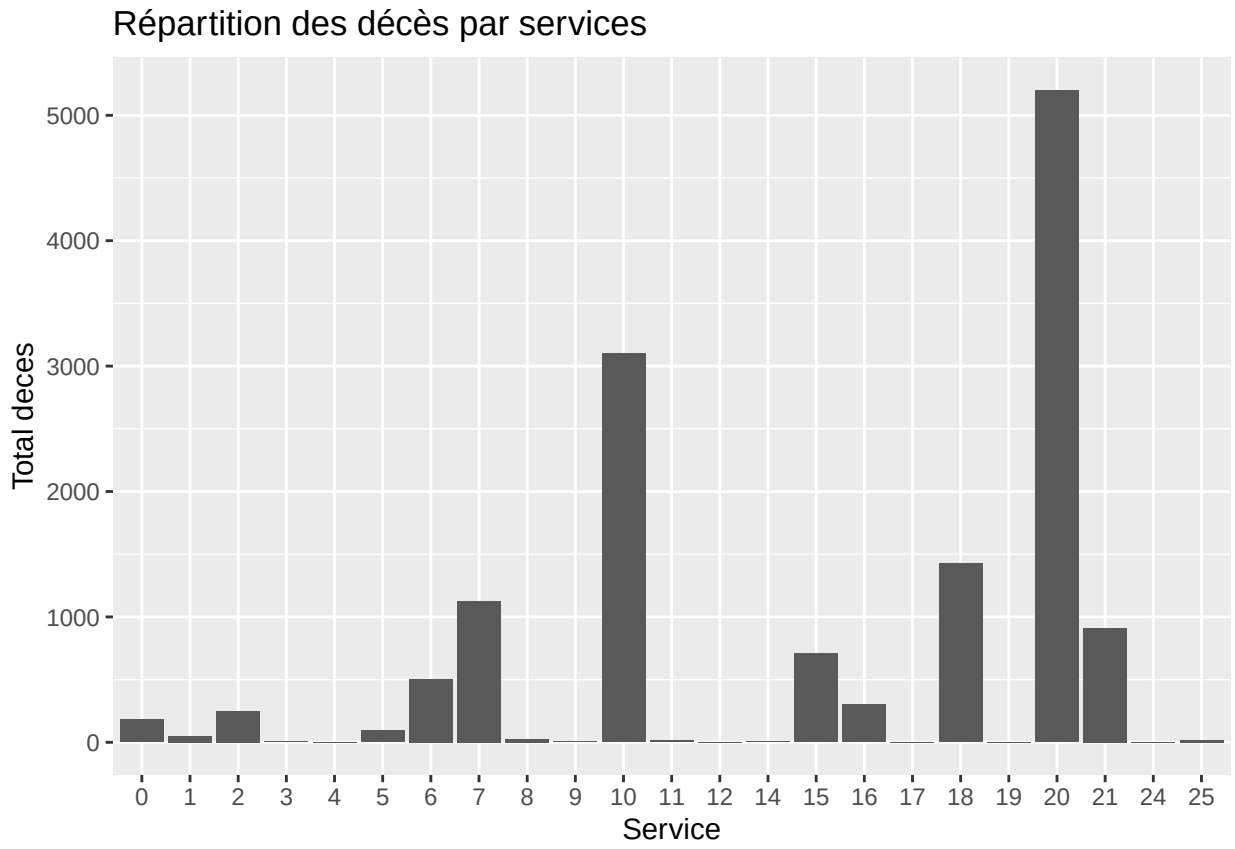
service:graphic:r-base

```
barplot(sort(service))
```



service:graphic:r-ggplot2

```
ggplot(data = data, mapping = aes(x = SERVICEHOSPIT)) +
  geom_bar()+
  labs(title = "Répartition des décès par services",
       x = "Service",
       y = "Total deces")
```



age

age:table

```
age <- table(data$Years)#,useNA = "always"
sort(age)
```

```
##
##   13   98   19   15   99  100    7    8    6   10    9   16   12   14   20   11
##   16   17   18   20   20   22   23   23   25   27   28   28   29   29   29   32
##   21   17   96   97   18    5   24    3   22   23   25   27    4   30   26   39
##   32   35   35   35   36   37   40   43   43   43   43   44   46   46   47   47
##   32   35   28   29   31   36   95   46   40   37   42   43   41   34   38   33
##   48   50   51   51   57   57   58   60   63   64   65   65   66   68   68   69
##   47   94   50   92   45    2   48   93   44   49   54   57   52   51   58   56
##   73   75   77   77   79   81   81   82   85   90   93   94   97   98  101  109
##   55   53   59   91   64   60   90   63   61   65   66   71   62   70   67   68
##  111  112  117  118  124  127  129  136  138  139  140  144  145  149  169  175
##   88   89   76   80   87   72   69   74   75    1   73   86   78   77   83   85
##  178  178  182  182  184  185  186  190  194  195  199  207  214  219  220  220
##   84   79   82   81    0
##  226  234  266  282 4242
```

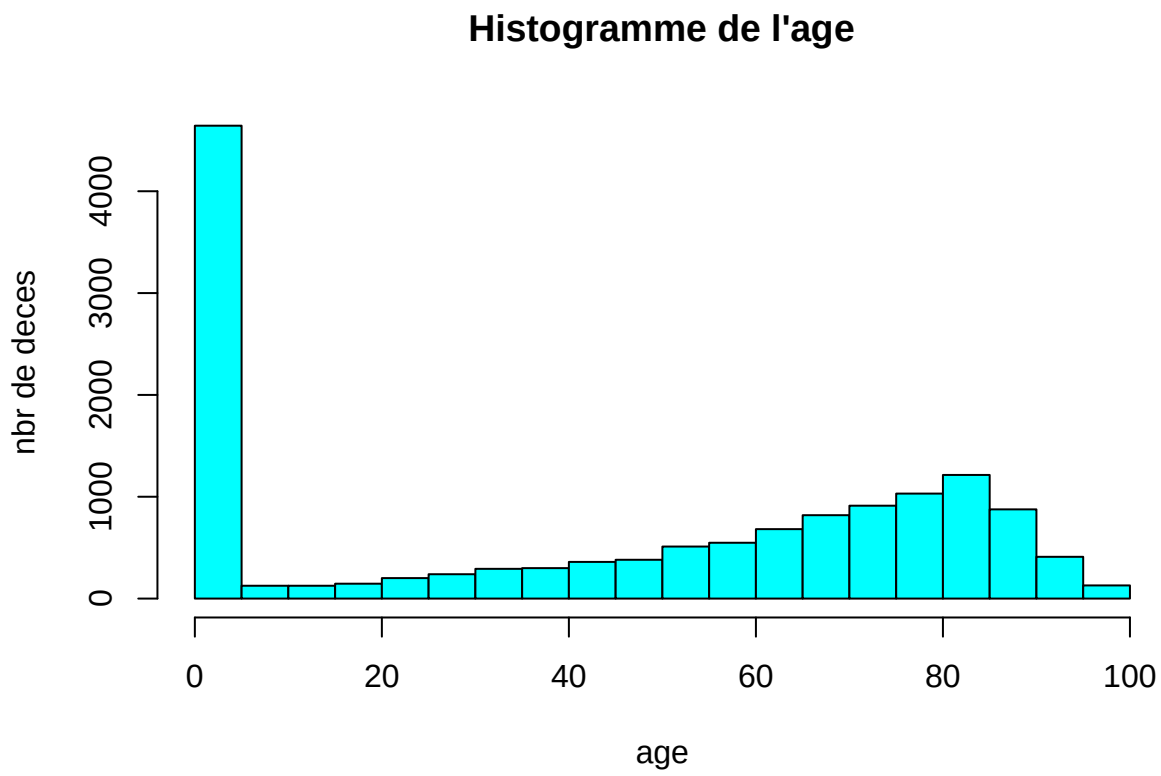
```
summary(age)
```

```
## Number of cases in table: 13946
```

```
## Number of factors: 1
```

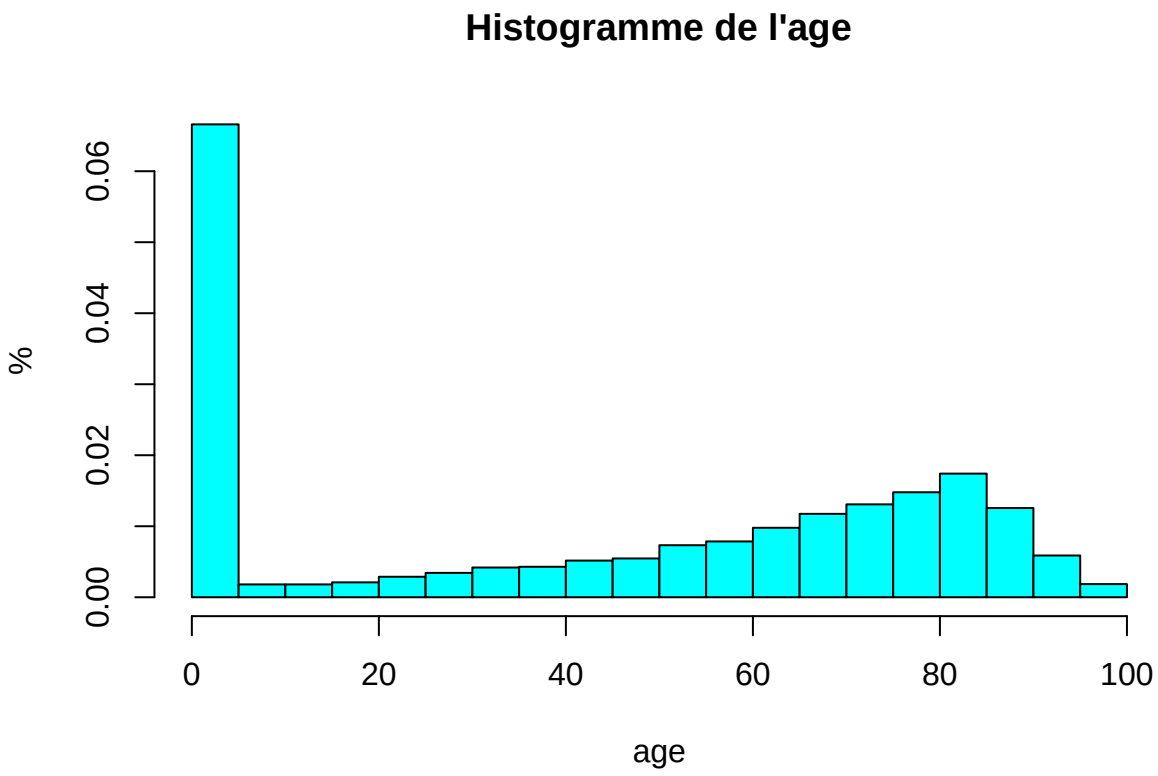
age:grafic:r-base

```
hist(data$Years,
     main = "Histogramme de l'age",
     xlab = "age",
     ylab= "nbr de deces",
     breaks = 15,
     col = "cyan")
```



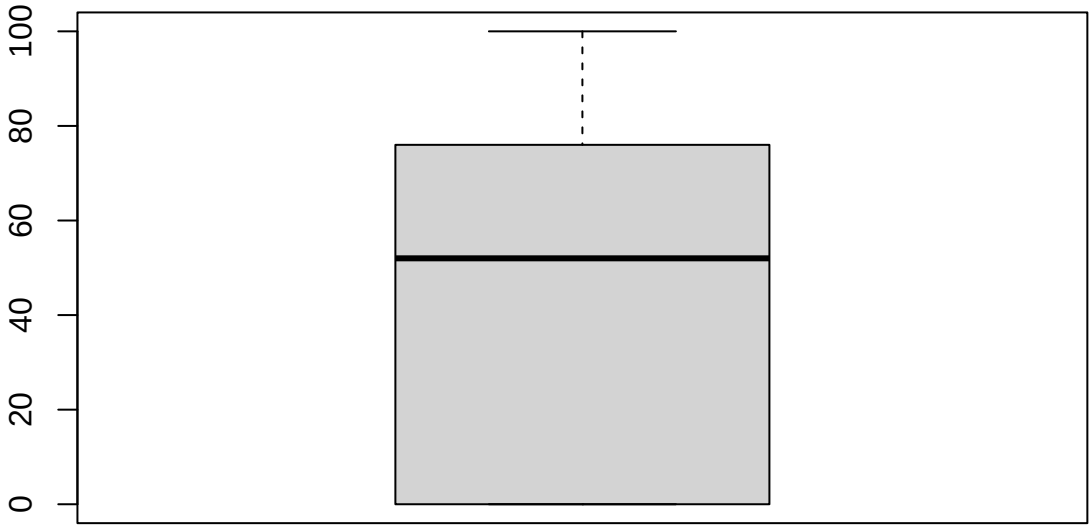
age:grafic:r-base

```
hist(data$Years,
     main = "Histogramme de l'age",
     xlab = "age",
     ylab= "%",
     breaks = 15,
     col = "cyan",
     probability = TRUE)
```



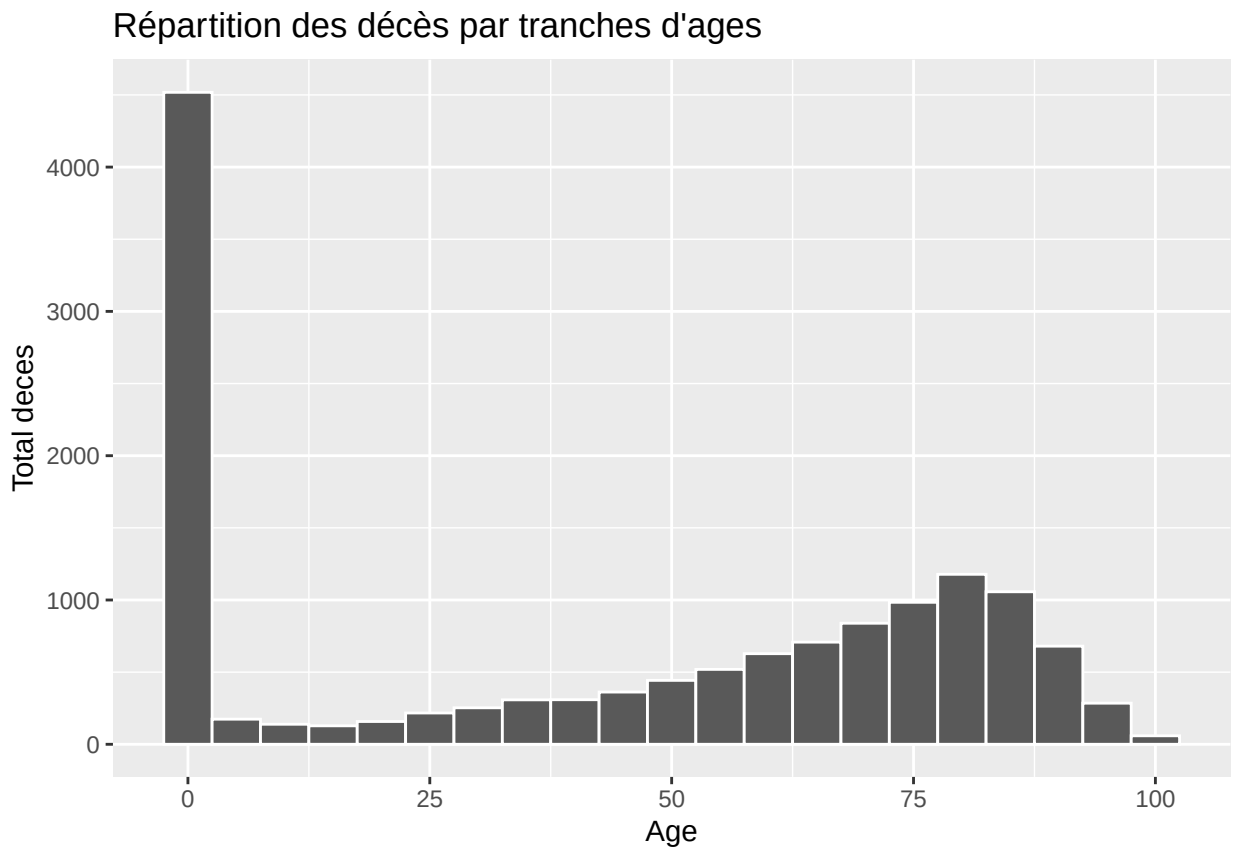
age:grafic:r-base

```
boxplot(data$Years)
```



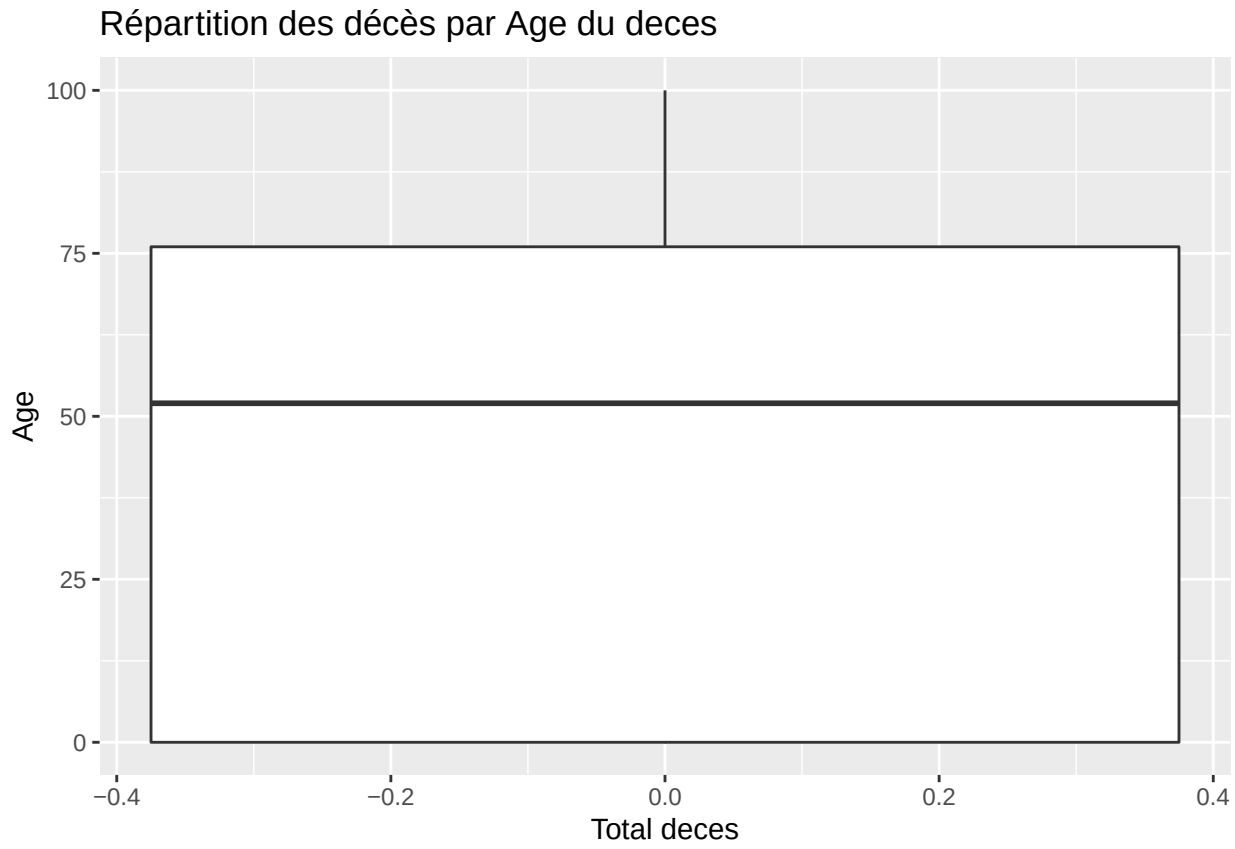
age:grafic:t-ggplot2

```
ggplot(data = data, mapping = aes(x = Years)) + #, fill = data$SEX
  geom_histogram(binwidth = 5, color = "white")+
  labs(title = "Répartition des décès par tranches d'ages",
       x = "Age",
       y = "Total deces")
```



age:grafic:t-ggplot2

```
ggplot(data = data, mapping = aes(y = Years)) + #, fill = data$SEX
  geom_boxplot()+
  labs(title = "Répartition des décès par Age du deces",
       x = "Total deces",
       y = "Age")
```



sexe

sexe:table

```
sexe <- table(data$SEX)#,useNA = "always"
sort(sexe)
```

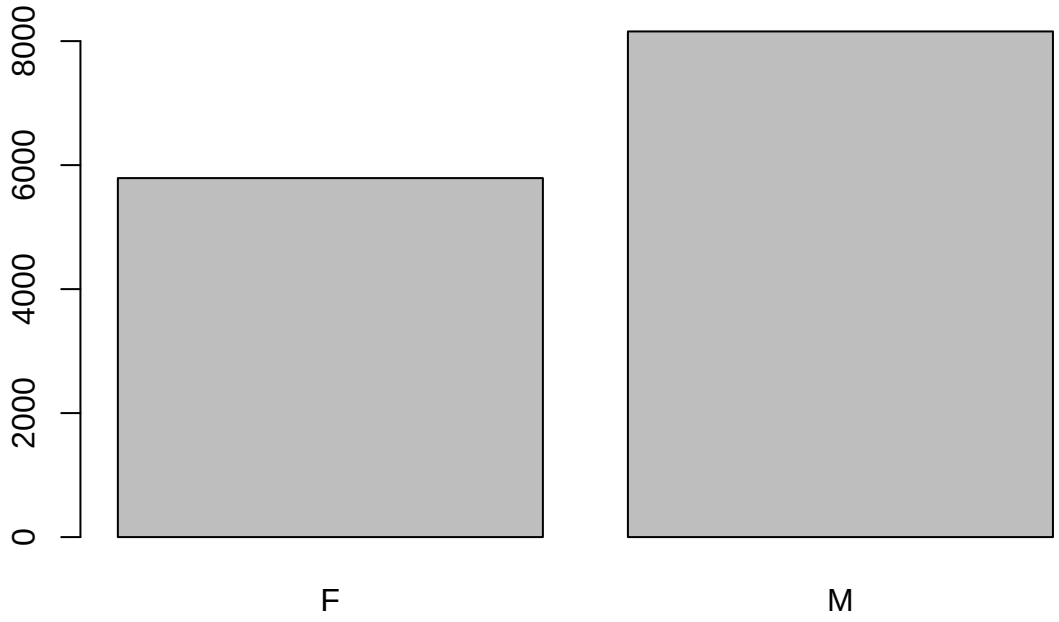
```
##
##    F    M
## 5790 8156
```

```
summary(sexe)
```

```
## Number of cases in table: 13946
## Number of factors: 1
```

sexe:graphic r-base

```
barplot(sort(sexe))
```



sexe:graphic r-ggplot2

```
ggplot(data = data, mapping = aes(x = SEX)) +
  geom_bar()+
  labs(title = "Répartition des décès par sexe",
       x = "sexe",
       y = "Total deces")
```

Profession

profession:table

```
Profession <- table(data$Profession)#,useNA = "always"
sort(Profession)
```

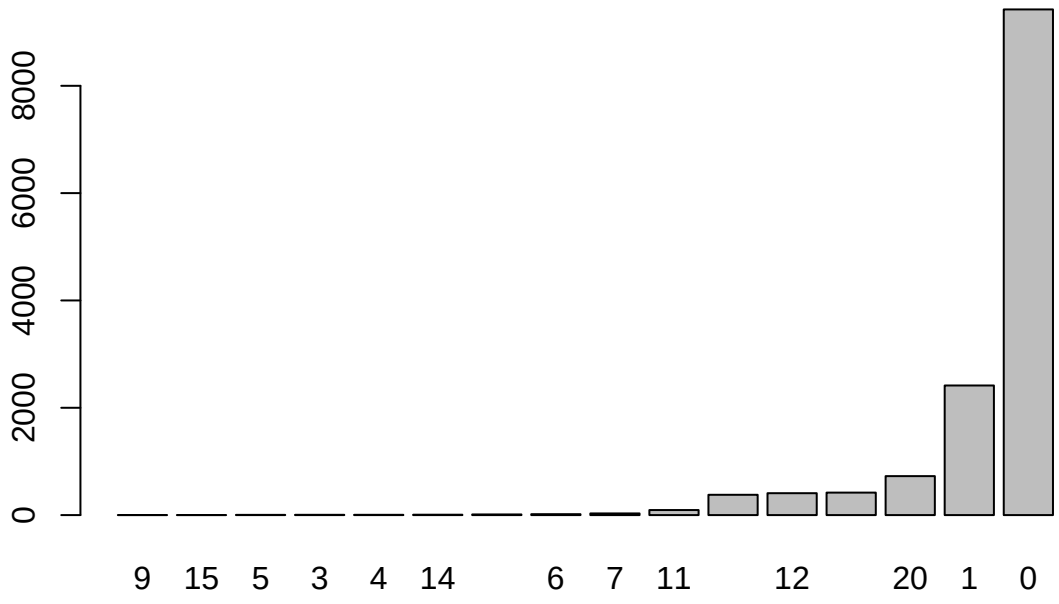
```
##
##    9   15    5    3    4   14   16    6    7   11   19   12   13   20    1    0
##    1    1    4    5    5    7   12   17   31   94  378  408  419  726 2415 9423
```

```
summary(Profession)
```

```
## Number of cases in table: 13946
## Number of factors: 1
```

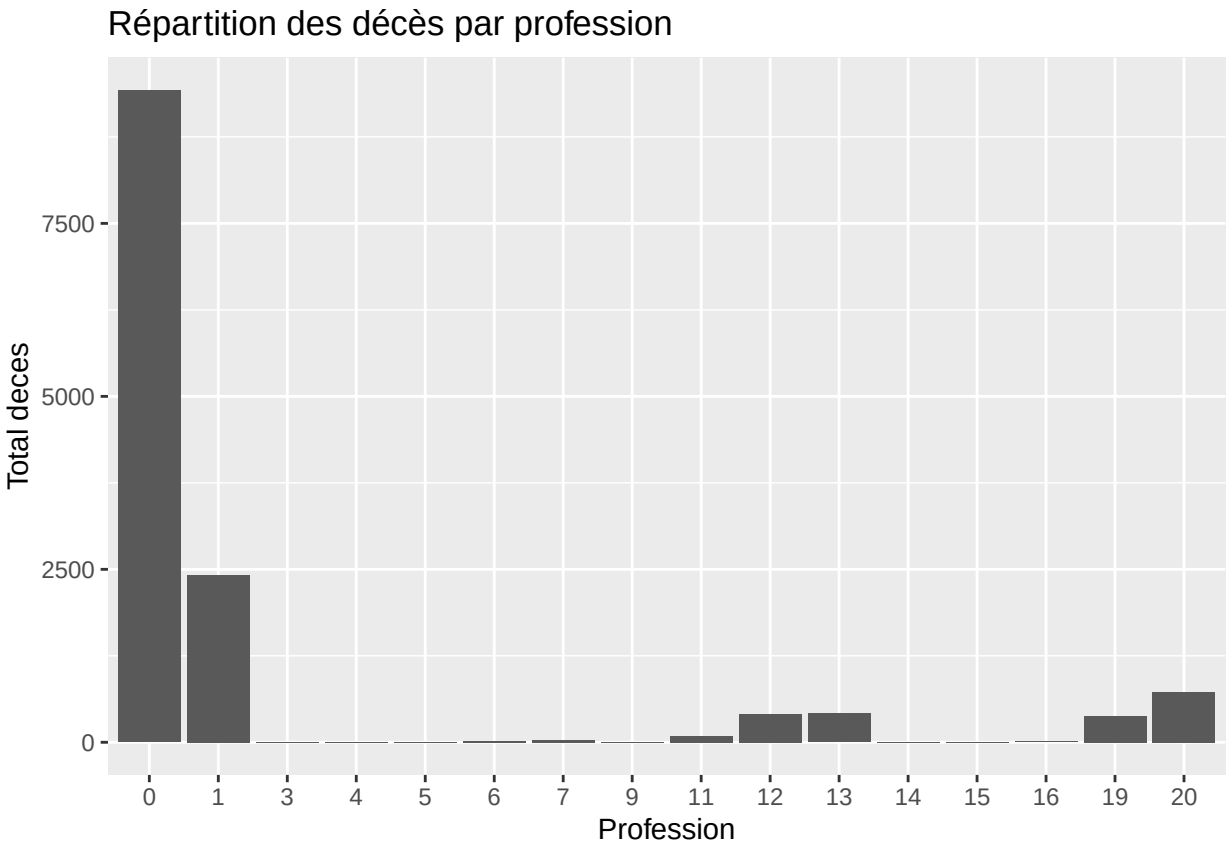
profession:graphic r-base

```
barplot(sort(Profession))
```



profession:graphic r-ggplot2

```
ggplot(data = data, mapping = aes(x = Profession)) +
  geom_bar()+
  labs(title = "Répartition des décès par profession",
       x = "Profession",
       y = "Total deces")
```


Répartition des décès par profession
Total deces
7500
5000
2500
0
0
1
3
4
5
6
7
9
11
12
13
14
15
16
19
20
Profession