

TP1_MAP201

Groupe: INF2

Binôme: Shaghayagh HAJMOHAMMADKASHI

Kaiwen ZHENG

#exo6

#Q1

```
TI = read.table(file="data/titanic.csv", header=TRUE, sep=";")
```

#La fonction `read.table()` permet de lire et importer des fichiers .txt et .csv. Le résultat va se trouver dans une structure de type "data frame" dans l'environnement de RStudio et pour obtenir un dataframe séparé on utilise « `sep=";"` » et pour avoir la première ligne en tête on utilise « `header=TRUE` ».

#Q2

```
TI[1:6,]
```

```
head(TI)
```

console:

```
> TI[1:6,]
  pclass survived gender   age
1      1      yes     F 29.0000
2      1      yes     M  0.9167
3      1      no      F  2.0000
4      1      no      M 30.0000
5      1      no      F 25.0000
6      1      yes     M 48.0000
> head(TI)
  pclass survived gender   age
1      1      yes     F 29.0000
2      1      yes     M  0.9167
3      1      no      F  2.0000
4      1      no      M 30.0000
5      1      no      F 25.0000
6      1      yes     M 48.0000
```

#on trouve le même résultat avec la fonction « `head()` » parce qu'il affiche aussi les 6 premières lignes de la base

#Q3

```
P = TI[, 1]
```

```
P = TI[, "pclass"]
```

```
P = TI$pclass
```

```
S = TI[, 2]
```

```
G = TI[, "gender"]
```

```
A = TI$age
```

```
#P, S, G, A sont affichés dans l'environnement de RStudio
```

```
class(P)
```

```
class(S)
```

```
class(G)
```

```
class(A)
```

```
console:
```

```
> class(P)
[1] "integer"
> class(S)
[1] "character"
> class(G)
[1] "character"
> class(A)
[1] "numeric"
```

```
#le type de P,A est Numérique et le type de S,G est Caractère
```

```
#Q4
```

```
TI[c(3,45,73),]
```

```
TI[,c(1,3)]
```

```
TI[c(67,82,101),c(2,4)]
```

```
#Q5
```

```
TI[P==1, ]
```

```
TI[(P==2) | (P==3), ] # les données des passagers de deuxième et troisième classe
```

```
#Q6
```

```
TI[P==1 & G=="F", ]
```

```
TI[P==2 & G=="M", ] # les données des hommes de deuxième classe
```

```
#Q7
```

```
TI[A<1, ]
```

```
TI[A<18, ] # les données des enfants (moins de 18 ans)
```

```
TI[(A<18) & (A>12), ] # les données des adolescents (12-18 ans)
```

```
#Q8
```

```
table(P)
```

```
nrow(TI)
```

```
# l'effectif total de l'échantillon (c'est le nombre de valeurs dans P) est 1046
```

```
console:
```

```
> table(P)
P
  1    2    3
284 261 501
> nrow(TI)
[1] 1046
```

```
P_1=nrow(TI[P==1, ])/nrow(TI)
```

```
P_1# la proportion des passagers en première classe est 0.2715105
```

```
P_2=nrow(TI[P==2, ])/nrow(TI)
```

```
P_2# la proportion des passagers en deuxième classe est 0.249522
```

```
P_3=nrow(TI[P==3, ])/nrow(TI)
```

```
P_3# la proportion des passagers en troisième classe est 0.4789675
```

```
e_tot_p = c(P_1,P_2,P_3)
```

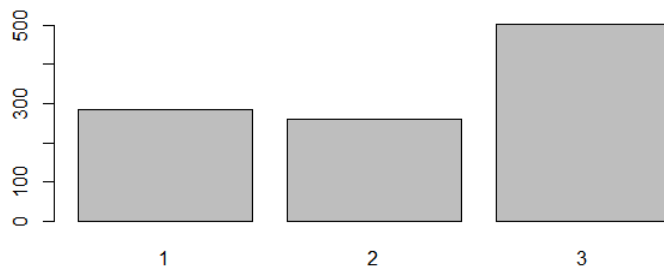
```
e_tot_p
```

```
prop.table(table(P))
```

```
console:
```

```
> table(P)
P
  1    2    3
284 261 501
> P_1=nrow(TI[P==1, ])/nrow(TI)
> P_1# la proportion des passagers en première class est 0.2715105
[1] 0.2715105
> P_2=nrow(TI[P==2, ])/nrow(TI)
> P_2
[1] 0.249522
> P_3=nrow(TI[P==3, ])/nrow(TI)
> P_3
[1] 0.4789675
> e_tot_p = c(P_1,P_2,P_3)
> e_tot_p
[1] 0.2715105 0.2495220 0.4789675
> prop.table(table(P))
P
      1      2      3
0.2715105 0.2495220 0.4789675
```

```
barplot(table(P))
```



```
barplot(prop.table(P))
```



#Q9

```
table(S)
```

console:

```
> table(S)
```

```
S
no yes
619 427
```

#les effectifs des survivants:427

#les effectifs des décédés:619

#donc 427 personnes ont survécus

```
prop.table(table(G))
```

console :

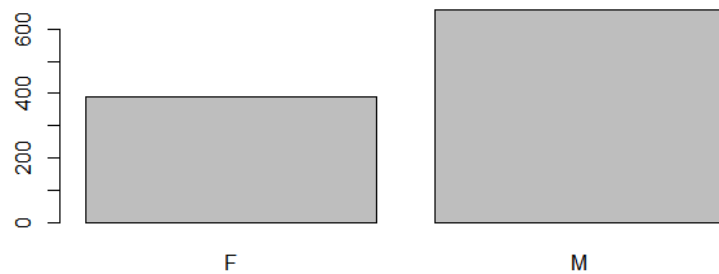
```
> prop.table(table(G))
```

```
G
      F      M
0.3709369 0.6290631
```

la proportion de hommes est 0.6290631

la proportion de femmes est 0.3709369

```
barplot(table(G))
```



#Q10

```
table(P, S)
```

```
prop.table(table(P,S))
```

console:

```
> table(P, S)
```

```
  S
P   no  yes
1  103  181
2  146  115
3  370  131
```

```
> prop.table(table(P,S))
```

```
  S
P   no      yes
1 0.09847036 0.17304015
2 0.13957935 0.10994264
3 0.35372849 0.12523901
```

#Q11

```
SP1 = (nrow(TI[S=="yes" & P==1,]))/(nrow(TI[S=="yes",]))
```

SP1 #Parmi les survivants, 0.4238876 est la proportion de passagers de première classe

```
P1S = (nrow(TI[S=="yes" & P==1,]))/(nrow(TI[P==1,]))
```

P1S #Parmi les passagers de première classe, 0.6373239 est la proportion de survivants

#Q12

```
prop.table(table(P,S),1)
```

console:

```
> prop.table(table(P,S),1)
```

```
  S
P   no      yes
1 0.3626761 0.6373239
2 0.5593870 0.4406130
3 0.7385230 0.2614770
```

#Parmi les passagers de première classe, 0.6373239 est la proportion de survivants

#Car les sommes des proportions de chaque linge(pclass) est égale à 1

```
prop.table(table(P,S),2)
```

console:

```
> prop.table(table(P,S),2)
```

```
  S
P      no      yes
1 0.1663974 0.4238876
2 0.2358643 0.2693208
3 0.5977383 0.3067916
```

#Parmi les survivants, 0.4238876 est la proportion de passagers de première classe

#Car les sommes des proportions de chaque colonne(survived) est égale à 1

#a:

```
table(P,S)
```

#181 passagers de première classe ont survécu

#b:

```
PS = (nrow(TI[S=="yes",]))/(nrow(TI))
```

PS

#Parmi tous les passagers, 0.4082218 est la proportion des survivants.

#c:

```
DP2 = (nrow(TI[S=="no" & P==2,]))/(nrow(TI[P==2,]))
```

DP2

#Parmi les passagers de deuxième classe, 0.559387 est la proportion des décédés

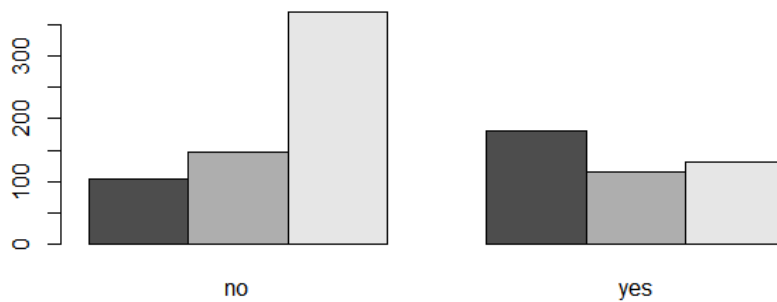
#d:

```
P2D = (nrow(TI[S=="no" & P==2,]))/(nrow(TI[S=="no",]))
```

P2D

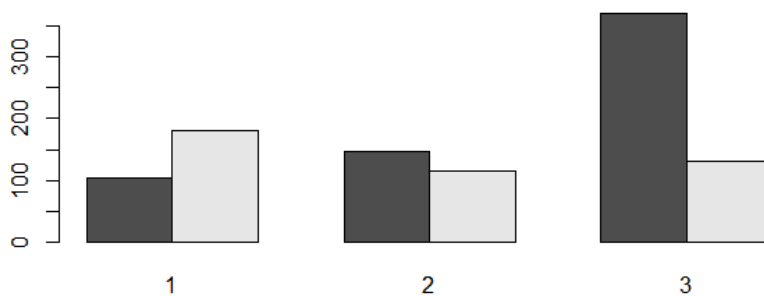
#Parmi les décédés, 0.2358643 est la proportion des passagers de deuxième classe

```
barplot(table(P, S), beside=TRUE)
```



ici on a dans l'ordre les passagers en première, deuxième et troisième classe qui sont décédés ou ils ont survécu.

`barplot(table(S, P), beside=TRUE)`



ici on a dans l'ordre les passagers en première classe qui sont décédés ou ils ont survécu, les passagers en deuxième classe qui sont décédés ou ils ont survécu et les passagers en troisième classe qui sont décédés ou ils ont survécu .

Commenter :

En ajoutant « `beside=TRUE` », il trace les barres les unes à côtés des autres plutôt que les unes au-dessus des autres (en cumulé).

#Q13

#a:

`table(P,G)`

console :

`> table(P,G)`

```

  G
P  F  M
1 133 151
2 103 158

```

```
3 152 349
```

```
#133 passagers de première classe sont des femmes
```

```
#b:
```

```
FP = (nrow(TI[G=="F",]))/(nrow(TI))
```

```
FP
```

```
#Parmi tous les passagers, 0.3709369 est la proportion de femmes
```

```
#c:
```

```
FP1 = (nrow(TI[G=="F" & P==1,]))/(nrow(TI[P==1,]))
```

```
FP1
```

```
#Parmi les passagers de première classe, 0.4683099 est la proportion de femmes
```

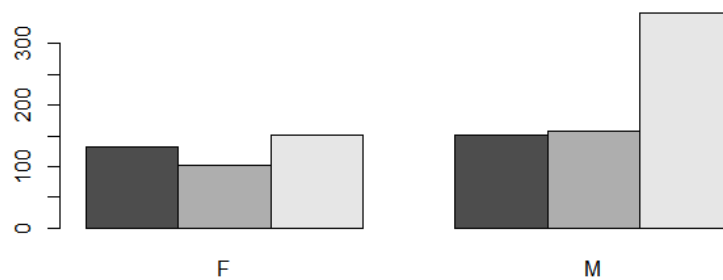
```
#d:
```

```
P1F = (nrow(TI[G=="F" & P==1,]))/(nrow(TI[G=="F",]))
```

```
P1F
```

```
#Parmi les femmes, 0.3427835 est la proportion de passagères de première classe
```

```
barplot(table(P, G), beside=TRUE)
```



```
barplot(table(G, P), beside=TRUE)
```