

Evaluation 2 - Rapport

Technologie de l'e-commerce et mobiles - Big Data

Khalladi Mohamed - B32-DA

2024-01-06

Contents

1	ANOVA	2	3
1.1	Bière et petits maux		3
1.2	Les médicaments contre la GCE		10
2	ACP et ACM		16
2.1	Eaux minérales		16

1 ANOVA 2

1.1 Bière et petits maux

L'Administration de la Santé Publique de Bidendumie a recensé le nombre de patients atteints de l'une des 4 maladies bénignes les plus fréquentes et ayant consommé l'une des 3 bières locales les plus répandues. Elle a mesuré un coefficient biochimique représentatif sur 6 patients (si possible) choisis aléatoirement.

Est-il possible d'interpréter de tels résultats ?

Nous allons former notre data-set et vérifier qu'il soit bien formé.

```
setwd("C:\\Users\\amine\\OneDrive\\Bureau\\EcomStat\\Labo\\Evaluation02\\datasets")

dataBiere <- read.csv("bieres_petits_maux.csv", h=TRUE, sep=";", fileEncoding="latin1")
dataBiere
```

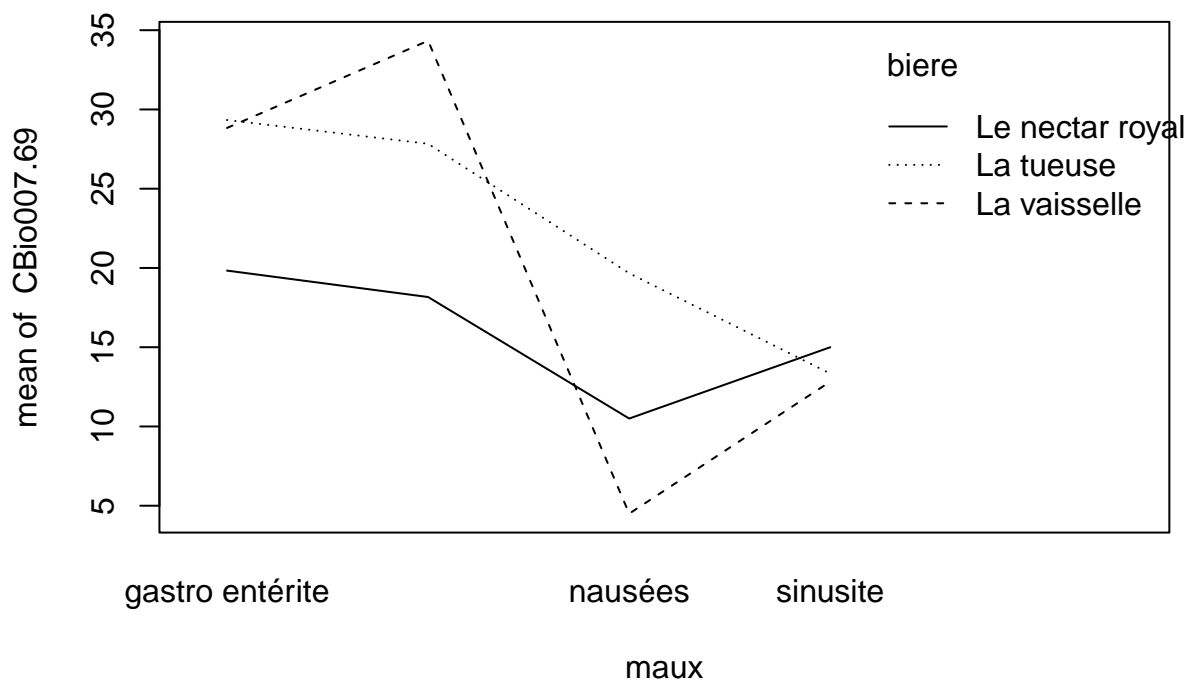
##	CBio007.69	maux	biere
## 1	42	gastro entérite	La tueuse
## 2	28	mal de tête	La tueuse
## 3	1	nausées	La tueuse
## 4	24	sinusite	La tueuse
## 5	44	gastro entérite	La tueuse
## 6	23	mal de tête	La tueuse
## 7	29	nausées	La tueuse
## 8	9	sinusite	La tueuse
## 9	36	gastro entérite	La tueuse
## 10	34	mal de tête	La tueuse
## 11	19	nausées	La tueuse
## 12	22	sinusite	La tueuse
## 13	13	gastro entérite	La tueuse
## 14	42	mal de tête	La tueuse
## 15	29	nausées	La tueuse
## 16	-2	sinusite	La tueuse
## 17	19	gastro entérite	La tueuse
## 18	13	mal de tête	La tueuse
## 19	18	nausées	La tueuse
## 20	15	sinusite	La tueuse
## 21	22	gastro entérite	La tueuse
## 22	27	mal de tête	La tueuse
## 23	22	nausées	La tueuse
## 24	12	sinusite	La tueuse
## 25	33	gastro entérite	La vaisselle
## 26	34	mal de tête	La vaisselle
## 27	11	nausées	La vaisselle
## 28	27	sinusite	La vaisselle
## 29	26	gastro entérite	La vaisselle
## 30	33	mal de tête	La vaisselle
## 31	9	nausées	La vaisselle
## 32	12	sinusite	La vaisselle
## 33	33	gastro entérite	La vaisselle
## 34	31	mal de tête	La vaisselle

```
## 35      7      nausées    La vaisselle
## 36     12      sinusite   La vaisselle
## 37     21 gastro entérite La vaisselle
## 38     36      mal de tête La vaisselle
## 39      1      nausées    La vaisselle
## 40     -5      sinusite   La vaisselle
## 41     29 gastro entérite La vaisselle
## 42     34      mal de tête La vaisselle
## 43     -6      nausées    La vaisselle
## 44     16      sinusite   La vaisselle
## 45     31 gastro entérite La vaisselle
## 46     38      mal de tête La vaisselle
## 47      5      nausées    La vaisselle
## 48     15      sinusite   La vaisselle
## 49     31 gastro entérite Le nectar royal
## 50      3      mal de tête Le nectar royal
## 51     21      nausées    Le nectar royal
## 52     22      sinusite   Le nectar royal
## 53     -3 gastro entérite Le nectar royal
## 54     26      mal de tête Le nectar royal
## 55      1      nausées    Le nectar royal
## 56      7      sinusite   Le nectar royal
## 57     25 gastro entérite Le nectar royal
## 58     28      mal de tête Le nectar royal
## 59      9      nausées    Le nectar royal
## 60     25      sinusite   Le nectar royal
## 61     25 gastro entérite Le nectar royal
## 62     32      mal de tête Le nectar royal
## 63      3      nausées    Le nectar royal
## 64      5      sinusite   Le nectar royal
## 65     24 gastro entérite Le nectar royal
## 66      4      mal de tête Le nectar royal
## 67     12      nausées    Le nectar royal
## 68     12      sinusite   Le nectar royal
## 69     17 gastro entérite Le nectar royal
## 70     16      mal de tête Le nectar royal
## 71     17      nausées    Le nectar royal
## 72     19      sinusite   Le nectar royal
```

```
dataBiere$maux <- as.factor(dataBiere$maux)
dataBiere$biere <- as.factor(dataBiere$biere)
summary(dataBiere)
```

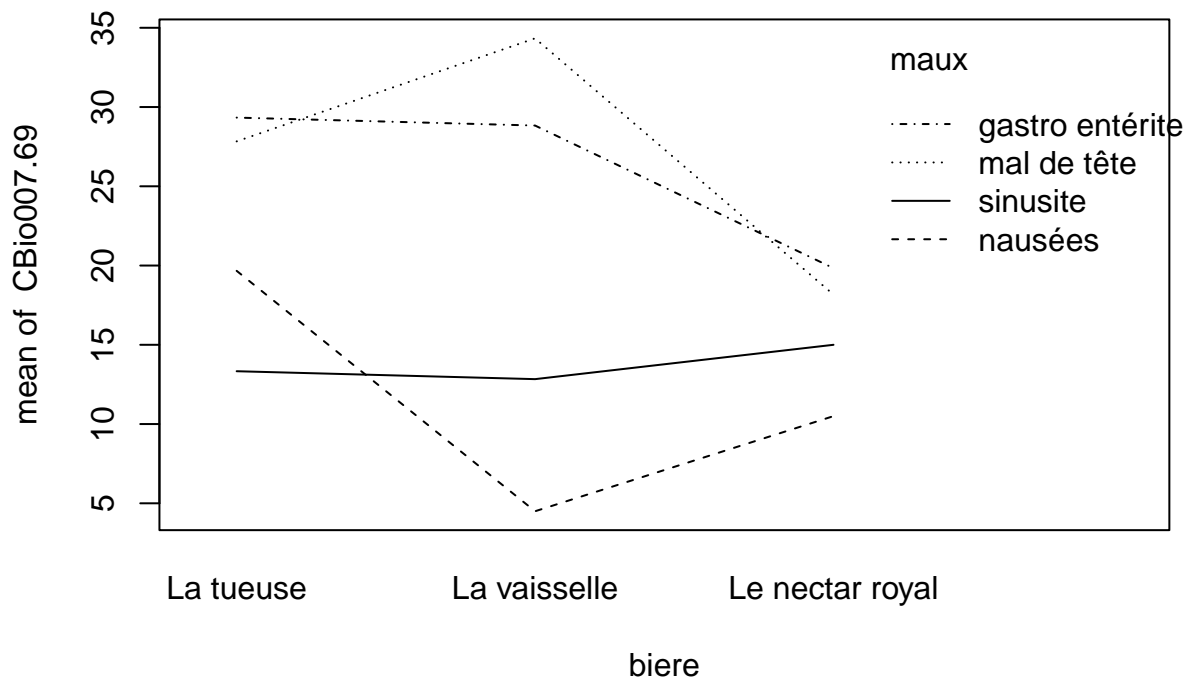
```
##      CBio007.69     iaux      biere
## Min.      :-6.00 gastro entérite:18 La tueuse      :24
## 1st Qu.:11.75 mal de tête      :18 La vaisselle    :24
## Median :21.00 nausées          :18 Le nectar royal:24
## Mean    :19.51 sinusite         :18
## 3rd Qu.:29.00
## Max.     :44.00
```

```
#Ce graphique montre comment le coefficient biochimique "CBio007-69" varie avec
#le type de maladie ("iaux") pour chaque type de bière.
with(dataBiere, interaction.plot(maux, biere, CBio007.69))
```



Chaque ligne du graphique représente un niveau du facteur “biere”, et les points sur les lignes représentent les moyennes du coefficient biochimique pour chaque combinaison de “biere” et “maux”. La non-parallélité des lignes suggère une interaction entre les types de bières et les maux en termes de coefficient biochimique.

```
#Ce graphique montre comment le coefficient biochimique varie avec
#le type de bière pour chaque type de maladie.
with(dataBiere, interaction.plot(biere, maux, CBio007.69))
```



Chaque ligne du graphique représente un niveau du facteur “maux”, et les points sur les lignes représentent les moyennes du coefficient biochimique pour chaque combinaison de “biere” et “maux”. Les lignes qui ne sont pas parallèles indiquent également une interaction entre le type de bière et le type de mal.

On va créer le modèle croisé pour pouvoir appliquer l’anova.

```
#y(ijk) = mu + alpha + beta(i) + gamma(j) + epsilon(ijk)

##Modèle avec interaction
modele_croise = lm(CBio007.69 ~ maux * biere, data = dataBiere)
modele_croise
```

```
##
## Call:
## lm(formula = CBio007.69 ~ maux * biere, data = dataBiere)
##
## Coefficients:
##              (Intercept)                  mauxmal de tête
##              2.933e+01                  -1.500e+00
##              mauxnausées                  mauxsinusite
##              -9.667e+00                  -1.600e+01
##              biereLa vaisselle              biereLe nectar royal
##              -5.000e-01                  -9.500e+00
##      mauxmal de tête:biereLa vaisselle      mauxnausées:biereLa vaisselle
##              7.000e+00                  -1.467e+01
##      mauxsinusite:biereLa vaisselle      mauxmal de tête:biereLe nectar royal
```

```
##                -1.305e-14                -1.667e-01
##    mauxnausées:biereLe nectar royal    mauxsinusite:biereLe nectar royal
##                3.333e-01                1.117e+01
```

(Mu): Le terme (Intercept) qui est de 29.33, représente la moyenne estimée du coefficient biochimique pour la catégorie de référence des maux et des bières ($\alpha_1 = 0$ et $\beta_1 = 0$).

(Alpha): Les coefficients liés à “maux” (par exemple, mauxsinusite de -16.00) représentent l’effet de chaque maladie sur le coefficient biochimique par rapport à la maladie de référence.

(Beta): Les coefficients liés à “biere” (par exemple, biereLe nectar royal de -9.50) indiquent l’effet de chaque type de bière sur le coefficient biochimique par rapport à la bière de référence.

(Gamma): Les coefficients d’interaction (par exemple, mauxmal de tête:biereLa vaisselle de 7.00) montrent l’effet combiné d’un certain mal avec une certaine bière sur le coefficient biochimique.

En d’autres termes, l’Intercept est notre point de départ, les coefficients alpha et beta nous disent comment chaque facteur change ce point de départ individuellement, et les coefficients gamma nous montrent ce qui se passe quand ces facteurs interagissent et se combinent de manière unique.

Effets principaux du facteur “maux”:

H0: Il n’y a pas de différence dans les moyennes du coefficient biochimique entre les différents types de “maux”.

H1: Il existe au moins une différence dans les moyennes du coefficient biochimique entre les différents types de “maux”.

Effets principaux du facteur “biere”: H0: Il n’y a pas de différence dans les moyennes du coefficient biochimique entre les différentes bières.

H1: Il existe au moins une différence dans les moyennes du coefficient biochimique entre les différentes bières.

Interaction entre “maux” et “biere”:

H0: Il n’y a pas d’interaction entre les “maux” et les “bières”, c’est-à-dire que l’effet d’un “mal” sur le coefficient biochimique est le même pour toutes les “bières”.

H1: Il existe une interaction entre les “maux” et les “bières”, c’est-à-dire que l’effet d’un “mal” sur le coefficient biochimique change selon la “bière” consommée.

```
anova(modele_croise)
```

```
## Analysis of Variance Table
##
## Response: CBio007.69
##          Df Sum Sq Mean Sq F value    Pr(>F)
## maux      3 3450.8  1150.27  12.9620 1.243e-06 ***
## biere     2   546.8   273.39   3.0807  0.05326 .
## maux:biere 6 1305.9   217.65   2.4526  0.03461 *
## Residuals 60 5324.5    88.74
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pour les effets principaux du facteur “maux”: Pour un seuil de 1% ou 5% on rejette H0, donc :

H1: Il existe au moins une différence dans les moyennes du coefficient biochimique entre les différents types de “maux”.

Pour les effets principaux du facteur “biere”:

Pour un seuil de 1% ou 5% on garde H0, donc :

H0: Il n’y a pas de différence dans les moyennes du coefficient biochimique entre les différentes bières.

Pour l'interaction entre "maux" et "biere":

Pour un seuil de 1% on garde H0 mais pour 5% on rejette H0, donc :

H1: Il existe une interaction entre les "maux" et les "bières", c'est-à-dire que l'effet d'un "mal" sur le coefficient biochimique change selon la "bière" consommée.

```
summary(modele_croise)
```

```
##
## Call:
## lm(formula = CBio007.69 ~iaux * biere, data = dataBiere)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -22.8333  -3.7083   0.3333   6.2500  14.6667
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.933e+01  3.846e+00   7.627 2.12e-10 ***
## mauxmal de tête    -1.500e+00  5.439e+00  -0.276  0.78365
## mauxnausées       -9.667e+00  5.439e+00  -1.777  0.08058 .
## mauxsinusite      -1.600e+01  5.439e+00  -2.942  0.00463 **
## biereLa vaisselle  -5.000e-01  5.439e+00  -0.092  0.92706
## biereLe nectar royal -9.500e+00  5.439e+00  -1.747  0.08581 .
## mauxmal de tête:biereLa vaisselle  7.000e+00  7.692e+00   0.910  0.36642
## mauxnausées:biereLa vaisselle  -1.467e+01  7.692e+00  -1.907  0.06133 .
## mauxsinusite:biereLa vaisselle  -1.305e-14  7.692e+00   0.000  1.00000
## mauxmal de tête:biereLe nectar royal -1.667e-01  7.692e+00  -0.022  0.98278
## mauxnausées:biereLe nectar royal   3.333e-01  7.692e+00   0.043  0.96558
## mauxsinusite:biereLe nectar royal   1.117e+01  7.692e+00   1.452  0.15177
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.42 on 60 degrees of freedom
## Multiple R-squared:  0.499, Adjusted R-squared:  0.4072
## F-statistic: 5.433 on 11 and 60 DF, p-value: 6.31e-06
```

rapport de corrélation = pourcentage de la variance expliquée par le modèle (donc toutes les contributions sauf la variance résiduelle)

R squared : proportion de la SCEf par rapport à la SCEt

La p-value nous indique ici qu'il pourrait y avoir une influence avec la sinusite et le coefficient biochimique vue que la p-value est très faible. Les autres en prenant un seuil de 5% n'auraient pas d'influence.

Pourtant, la p-value du modele est très faible également donc ça veut dire qu'il y a de l'interaction mais il faut en trouver plus. C'est pour cela que l'on va chercher à utiliser le modèle hiérarchisé pour rechercher d'autres interactions.

```
#y(ijk) = mu + alpha + beta(i) + gamma(j) + epsilon(ijk)
#Modèle sans interaction
modele_hierarchise = lm(CBio007.69 ~iaux + biere, data = dataBiere)
modele_hierarchise
```

```
##
## Call:
```



```
## lm(formula = CBio007.69 ~ maux + biere, data = dataBiere)
##
## Coefficients:
##      (Intercept)      mauxmal de tête      mauxnausées
##      29.0278      0.7778      -14.4444
##      mauxsinusite      biereLa vaisselle      biereLe nectar royal
##      -12.2778      -2.4167      -6.6667
```

```
anova(modele_hierarchise)
```

```
## Analysis of Variance Table
##
## Response: CBio007.69
##      Df Sum Sq Mean Sq F value    Pr(>F)
## maux      3 3450.8 1150.27 11.4500 3.895e-06 ***
## biere      2  546.8  273.39  2.7214  0.07317 .
## Residuals 66 6630.4  100.46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ici, on voit que la bière a une p-value supérieure à 7% , ce qui vaut dire qu'on pourrait accepter le H0 avec un seuil de 5% , donc que la bière n'aurait pas d'influence sur le coefficient biochimique.

```
summary(modele_hierarchise)
```

```
##
## Call:
## lm(formula = CBio007.69 ~ maux + biere, data = dataBiere)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.361  -5.424   1.792   6.701  14.972
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    29.0278     2.8934  10.032 6.63e-15 ***
## mauxmal de tête     0.7778     3.3410   0.233 0.816639
## mauxnausées     -14.4444     3.3410  -4.323 5.30e-05 ***
## mauxsinusite    -12.2778     3.3410  -3.675 0.000479 ***
## biereLa vaisselle  -2.4167     2.8934  -0.835 0.406598
## biereLe nectar royal -6.6667     2.8934  -2.304 0.024373 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.02 on 66 degrees of freedom
## Multiple R-squared:  0.3761, Adjusted R-squared:  0.3289
## F-statistic: 7.959 on 5 and 66 DF,  p-value: 6.466e-06
```

Grace à ce modele hierarchise, on voit que les sinusites ont toujours de l'influence mais que les nausées ont également une influence sur le coefficient biochimique ainsi que la bière « Le nectar Royal » (pour un seuil de 5%).

1.2 Les médicaments contre la GCE

Une entreprise pharmaceutique s'intéresse à une maladie tropicale (la Gengivite Cephalopodique Endiablée - GCE) et a mis au point trois molécules susceptibles de soigner cette maladie : AlphaVictoire, BetaTriomphe et GammaSucces. Les tests cliniques ont été pratiqués pour mesurer un coefficient relatif d'amélioration de l'état de patients gravement atteints (plus ce coefficient d'immunité est élevé et plus l'action sera considérée comme efficace). Mais, de plus, on souhaite également tenir compte du mode d'administration des différentes molécules (par voie orale ou par injection intraveineuse).

Observe-t-on une différence significative d'efficacité soit selon la molécule, soit selon le mode d'administration ou encore selon une combinaison des deux facteurs ?

Nous allons former notre data-set et vérifier qu'il soit bien formé.

```
#Les médicaments contre la GCE
setwd("C:\\Users\\amine\\OneDrive\\Bureau\\EcomStat\\Labo\\Evaluation02\\datasets")

dataMedicament <- data.frame(
  Amelioration = c(10, 12, 8, 10, 6, 13, 9, 10, 9, 8, 11, 18, 12, 15, 13, 8, 15, 16, 9, 13,
                  7, 14, 10, 11, 9, 10, 11, 7, 9, 9, 8, 9, 10, 9, 11, 13, 7, 14, 15, 12,
                  12, 9, 11, 27, 7, 8, 13, 14, 10, 11, 7, 6, 10, 7, 7, 5, 6, 7, 9, 6),
  Molecule = rep(c("AlphaVictoire", "BetaTriomphe", "GammaSucces"), each = 20),
  Administration = rep(c("Oral", "Injection"), each = 10, times = 3)
)
dataMedicament
```

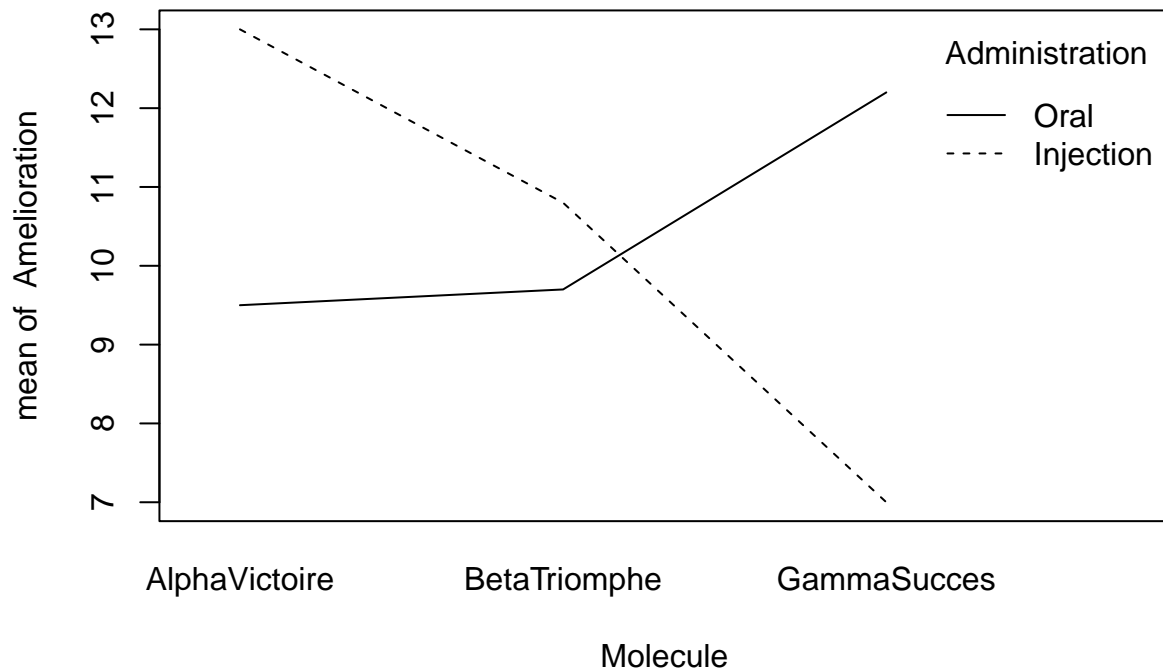
##	Amelioration	Molecule	Administration
## 1	10	AlphaVictoire	Oral
## 2	12	AlphaVictoire	Oral
## 3	8	AlphaVictoire	Oral
## 4	10	AlphaVictoire	Oral
## 5	6	AlphaVictoire	Oral
## 6	13	AlphaVictoire	Oral
## 7	9	AlphaVictoire	Oral
## 8	10	AlphaVictoire	Oral
## 9	9	AlphaVictoire	Oral
## 10	8	AlphaVictoire	Oral
## 11	11	AlphaVictoire	Injection
## 12	18	AlphaVictoire	Injection
## 13	12	AlphaVictoire	Injection
## 14	15	AlphaVictoire	Injection
## 15	13	AlphaVictoire	Injection
## 16	8	AlphaVictoire	Injection
## 17	15	AlphaVictoire	Injection
## 18	16	AlphaVictoire	Injection
## 19	9	AlphaVictoire	Injection
## 20	13	AlphaVictoire	Injection
## 21	7	BetaTriomphe	Oral
## 22	14	BetaTriomphe	Oral
## 23	10	BetaTriomphe	Oral
## 24	11	BetaTriomphe	Oral
## 25	9	BetaTriomphe	Oral
## 26	10	BetaTriomphe	Oral

```
## 27      11 BetaTriomphe      Oral
## 28       7 BetaTriomphe      Oral
## 29       9 BetaTriomphe      Oral
## 30       9 BetaTriomphe      Oral
## 31       8 BetaTriomphe      Injection
## 32       9 BetaTriomphe      Injection
## 33      10 BetaTriomphe      Injection
## 34       9 BetaTriomphe      Injection
## 35      11 BetaTriomphe      Injection
## 36      13 BetaTriomphe      Injection
## 37       7 BetaTriomphe      Injection
## 38      14 BetaTriomphe      Injection
## 39      15 BetaTriomphe      Injection
## 40      12 BetaTriomphe      Injection
## 41      12 GammaSucces      Oral
## 42       9 GammaSucces      Oral
## 43      11 GammaSucces      Oral
## 44      27 GammaSucces      Oral
## 45       7 GammaSucces      Oral
## 46       8 GammaSucces      Oral
## 47      13 GammaSucces      Oral
## 48      14 GammaSucces      Oral
## 49      10 GammaSucces      Oral
## 50      11 GammaSucces      Oral
## 51       7 GammaSucces      Injection
## 52       6 GammaSucces      Injection
## 53      10 GammaSucces      Injection
## 54       7 GammaSucces      Injection
## 55       7 GammaSucces      Injection
## 56       5 GammaSucces      Injection
## 57       6 GammaSucces      Injection
## 58       7 GammaSucces      Injection
## 59       9 GammaSucces      Injection
## 60       6 GammaSucces      Injection
```

```
dataMedicament$Molecule <- as.factor(dataMedicament$Molecule)
dataMedicament$Administration <- as.factor(dataMedicament$Administration)
summary(dataMedicament)
```

```
##   Amelioration      Molecule      Administration
##   Min.   : 5.00   AlphaVictoire:20   Injection:30
##   1st Qu.: 8.00   BetaTriomphe :20   Oral      :30
##   Median :10.00   GammaSucces  :20
##   Mean   :10.37
##   3rd Qu.:12.00
##   Max.    :27.00
```

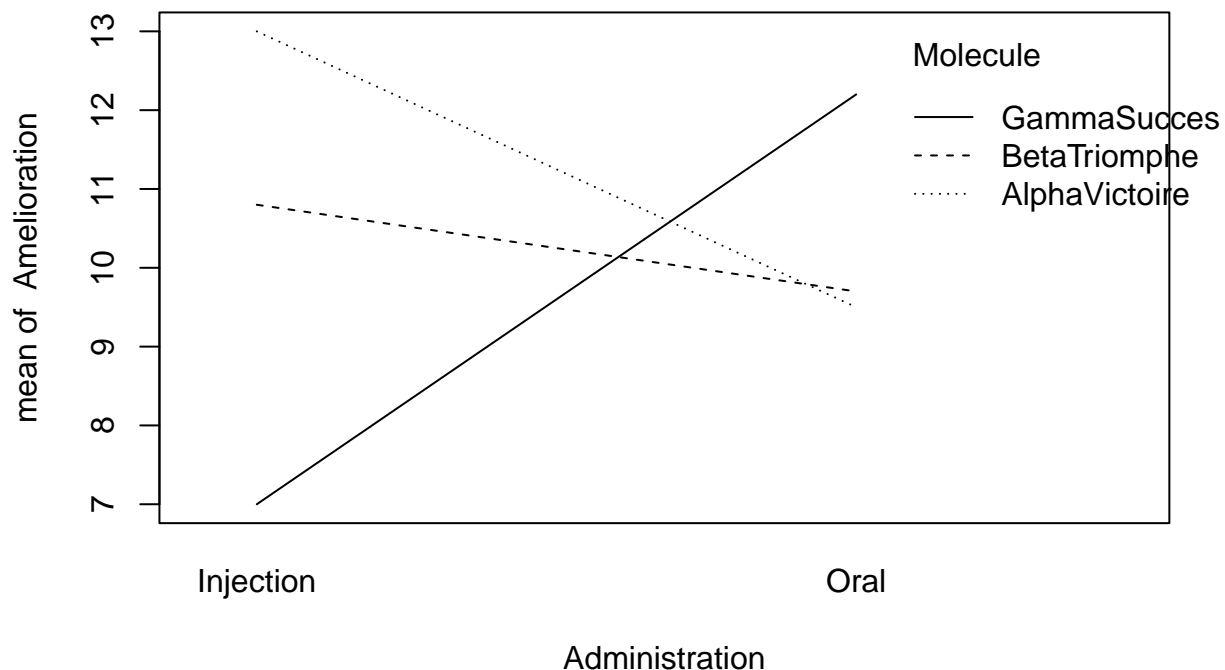
```
#Ce graphique montre comment le coefficient d'immunité varie avec
#le type de molecule pour chaque type d'administration.
with(dataMedicament, interaction.plot(Molecule, Administration, Amelioration))
```



Chaque ligne du graphique représente un niveau du facteur “Administration”, et les points sur les lignes représentent les moyennes du coefficient d’immunité pour chaque combinaison de “Administration” et “Molecule”.

La croisement des lignes suggère une interaction entre les types des administrations et les molecules en termes de coefficient d’immunité.

```
#Ce graphique montre comment le coefficient d'immunité varie avec
#le type d'administration pour chaque type de molecule.
with(dataMedicament, interaction.plot(Administration, Molecule, Amelioration))
```



Chaque ligne du graphique représente un niveau du facteur “Molecule”, et les points sur les lignes représentent les moyennes du coefficient d’immunité pour chaque combinaison de “Molecule” et “Administration”. Les lignes qui se croisent indiquent également une interaction entre le type de molecule et le type d’administration.

```
modele_croise <- lm(Amelioration ~ Molecule * Administration, data = dataMedicament)
modele_croise
```

```
##
## Call:
## lm(formula = Amelioration ~ Molecule * Administration, data = dataMedicament)
##
## Coefficients:
##              (Intercept)
##                   13.0
##      MoleculeBetaTriomphe
##                   -2.2
##      MoleculeGammaSucces
##                   -6.0
##      AdministrationOral
##                   -3.5
## MoleculeBetaTriomphe:AdministrationOral
##                   2.4
## MoleculeGammaSucces:AdministrationOral
##                   8.7
```

Effets principaux du facteur “Molecule”:

H0: Il n'y a pas de différence dans les moyennes du coefficient d'immunité entre les différents types de "Molecule".

H1: Il existe au moins une différence dans les moyennes du coefficient d'immunité entre les différents types de "Molecule".

Effets principaux du facteur "Administration": H0: Il n'y a pas de différence dans les moyennes du coefficient d'immunité entre les différentes types d'administrations.

H1: Il existe au moins une différence dans les moyennes du coefficient d'immunité entre les différentes types d'administrations.

Interaction entre "Molecule" et "Administration":

H0: Il n'y a pas d'interaction entre les "Molecule" et les "Administration", c'est-à-dire que l'effet d'un "Molecule" sur le coefficient d'immunité est le même pour toutes les "Administration".

H1: Il existe une interaction entre les "Molecule" et les "Administration", c'est-à-dire que l'effet d'un "Molecule" sur le coefficient d'immunité change selon l' "Administration".

```
anova(modele_croise)
```

```
## Analysis of Variance Table
##
## Response: Amelioration
##
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Molecule      2  27.63   13.817    1.4030 0.2546829
## Administration  1   0.60    0.600    0.0609 0.8059756
## Molecule:Administration  2 201.90  100.950  10.2507 0.0001683 ***
## Residuals     54 531.80    9.848
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pour les effets principaux du facteur "Molecule":

Pour un seuil de 1% ou 5% on garde H0, donc :

H0: Il n'y a pas de différence dans les moyennes du coefficient d'immunité entre les différents types de "Molecule".

Pour les effets principaux du facteur "Administration":

Pour un seuil de 1% ou 5% on garde H0, donc :

H0: Il n'y a pas de différence dans les moyennes du coefficient d'immunité entre les différentes types d'administrations.

Pour l'interaction entre "Molecule" et "Administration":

Pour un seuil de 1% on garde H0 mais pour 5% on rejette H0, donc :

H1: Il existe une interaction entre les "Molecule" et les "Administration", c'est-à-dire que l'effet d'un "Molecule" sur le coefficient d'immunité change selon l' "Administration".

```
summary(modele_croise)
```

```
##
## Call:
## lm(formula = Amelioration ~ Molecule * Administration, data = dataMedicament)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.200 -1.575 -0.100  1.300 14.800
##
## Coefficients:
```

```
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   13.0000     0.9924  13.100 < 2e-16
## MoleculeBetaTriomphe        -2.2000     1.4034  -1.568  0.1228
## MoleculeGammaSucces         -6.0000     1.4034  -4.275 7.82e-05
## AdministrationOral           -3.5000     1.4034  -2.494  0.0157
## MoleculeBetaTriomphe:AdministrationOral  2.4000     1.9848   1.209  0.2318
## MoleculeGammaSucces:AdministrationOral  8.7000     1.9848   4.383 5.43e-05
##
## (Intercept)                   ***
## MoleculeBetaTriomphe
## MoleculeGammaSucces         ***
## AdministrationOral           *
## MoleculeBetaTriomphe:AdministrationOral
## MoleculeGammaSucces:AdministrationOral ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.138 on 54 degrees of freedom
## Multiple R-squared:  0.302, Adjusted R-squared:  0.2374
## F-statistic: 4.674 on 5 and 54 DF, p-value: 0.001285
```

Les p-values nous indiquent ici que les molécules BetaTriomphe et GammaSucces ont de l'influence sur le coefficient d'immunité, et du côté des types des administrations "Oral" a bien une influence sur le coefficient d'immunité.

Cependant, on va quand même analyser l'anova pour chacune des variables qualitatives séparément pour essayer de voir si il n'y aurait pas encore plus d'interaction.

```
modele_hierarchise <- lm(Amelioration ~ Molecule + Administration, data = dataMedicament)
modele_hierarchise
```

```
##
## Call:
## lm(formula = Amelioration ~ Molecule + Administration, data = dataMedicament)
##
## Coefficients:
## (Intercept)  MoleculeBetaTriomphe  MoleculeGammaSucces
##          11.15                -1.00                -1.65
## AdministrationOral
##          0.20
```

```
anova(modele_hierarchise)
```

```
## Analysis of Variance Table
##
## Response: Amelioration
##          Df Sum Sq Mean Sq F value Pr(>F)
## Molecule    2  27.63  13.817  1.0546 0.3552
## Administration 1   0.60   0.600  0.0458 0.8313
## Residuals   56 733.70  13.102
```

Ici, on voit deux p-value sont à nouveau très élevées. Donc on peut déjà en conclure qu'il n'y aura aucune influence ... Mais nous allons le confirmer en analysant le summary.

```
summary(modele_hierarchise)
```

```
##
## Call:
## lm(formula = Amelioration ~ Molecule + Administration, data = dataMedicament)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.35  -2.50  -0.60   1.70  17.30
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      11.1500     0.9346  11.930 <2e-16 ***
## MoleculeBetaTriomphe -1.0000     1.1446  -0.874  0.386
## MoleculeGammaSucces -1.6500     1.1446  -1.442  0.155
## AdministrationOral    0.2000     0.9346   0.214  0.831
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.62 on 56 degrees of freedom
## Multiple R-squared:  0.03705,    Adjusted R-squared:  -0.01453
## F-statistic: 0.7183 on 3 and 56 DF,  p-value: 0.5452
```

Grace à ce summary, on ne constate aucune différence qu'avec l'anova au système croisé. Toutes les p-values sont plus élevées que le seuil « logique » qui est de 5%. Donc aucune des deux variables qualitatives influencent le coefficient d'immunité.

D'ailleurs cela se confirme en regardant la p-value qui est très élevée et qui montre qu'il n'y aucune influence !

2 ACP et ACM

2.1 Eaux minérales

Le fichier Eaux1.txt contient des données sur la teneur en divers éléments chimiques pour quelques eaux minérales commercialisées en France.

Quelles relations peut-on détecter ?

Peut-on donner une signification claire aux axes principaux ?

```
library(FactoMineR)
```

```
## Warning: le package 'FactoMineR' a été compilé avec la version R 4.2.3
```

```
library(factoextra)
```

```
## Warning: le package 'factoextra' a été compilé avec la version R 4.2.3
```

```
## Le chargement a nécessité le package : ggplot2
```



```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
replaceNAparMOY<-function(x)
{
  return ( ifelse(is.na(x), mean(x,na.rm = TRUE), x) )
}
```

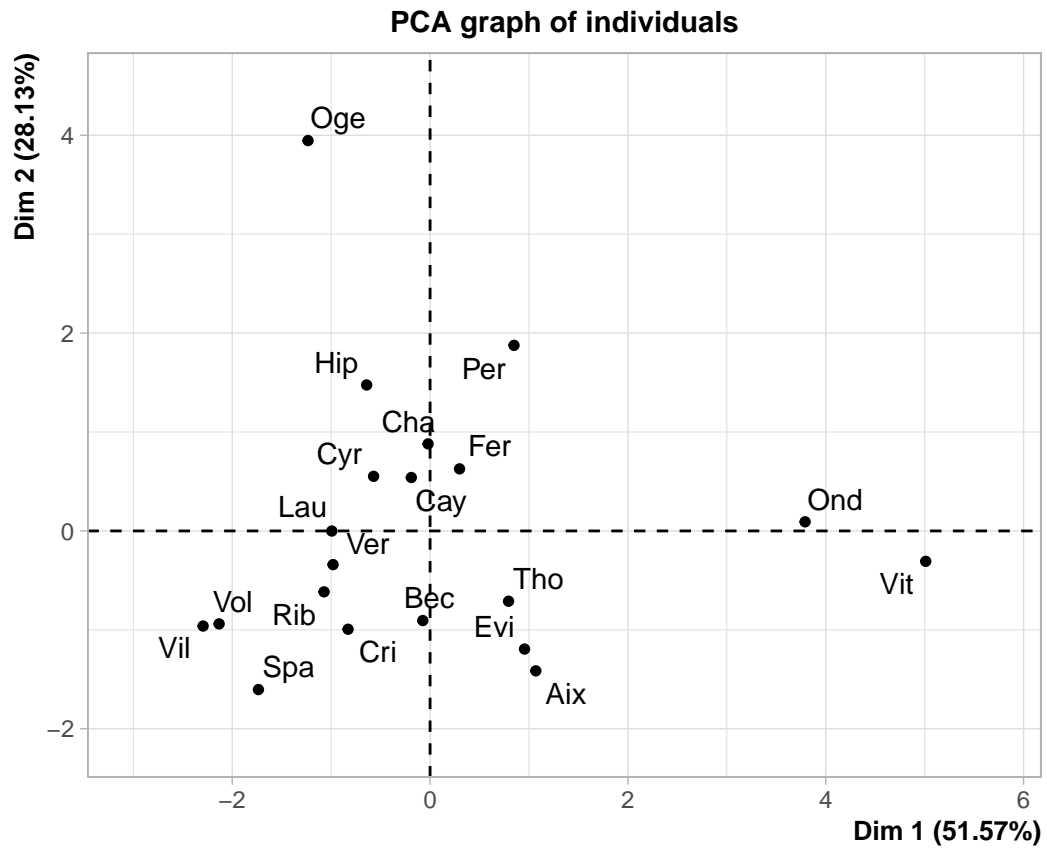
```
setwd("C:\\Users\\amine\\OneDrive\\Bureau\\EcomStat\\Labo\\Evaluation02\\datasets")
```

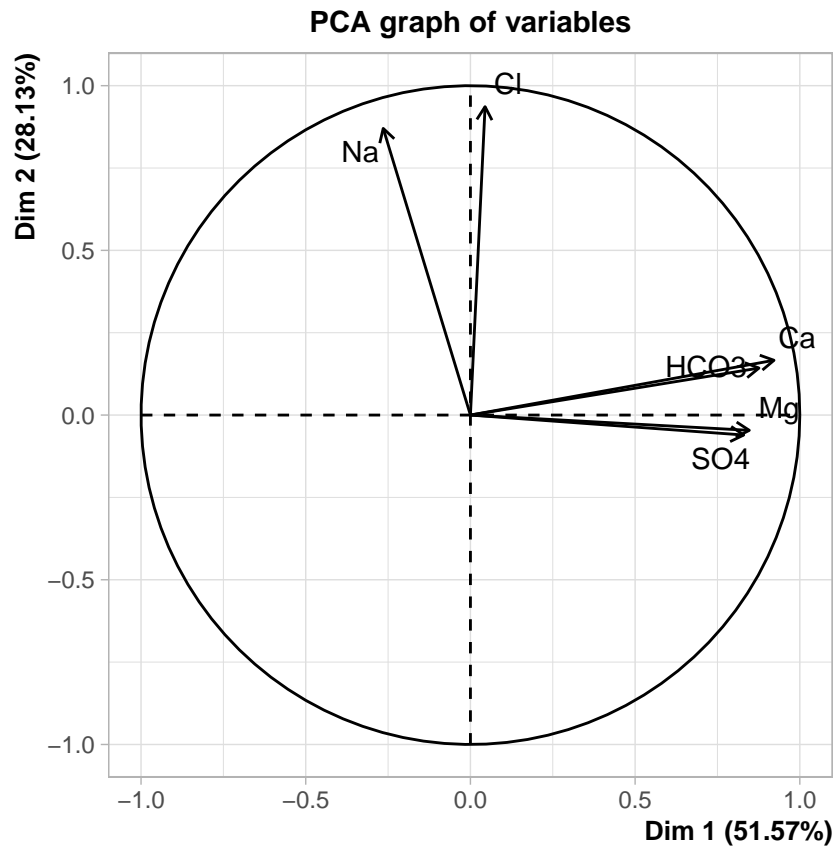
```
donnees <- read.table("Eaux1.txt", sep="\t", header=TRUE, row.names=7)
```

```
summary(donnees)
```

```
##      HCO3      SO4      Cl      Ca
## Min.   : 59.0   Min.   :  3.00   Min.   :  2.00   Min.   :  4.00
## 1st Qu.:185.2   1st Qu.:  8.50   1st Qu.:  6.00   1st Qu.: 53.25
## Median :259.5   Median : 14.50   Median :  8.50   Median : 72.00
## Mean   :250.4   Mean   : 42.40   Mean   :13.65   Mean   : 77.50
## 3rd Qu.:334.2   3rd Qu.: 24.75   3rd Qu.:18.50   3rd Qu.: 92.25
## Max.   :402.0   Max.   :306.00   Max.   :44.00   Max.   :202.00
##      Mg      Na
## Min.   : 1.00   Min.   :  2.00
## 1st Qu.: 4.00   1st Qu.:  4.75
## Median : 6.00   Median :  9.00
## Mean   :11.85   Mean   :10.10
## 3rd Qu.:19.25   3rd Qu.:13.00
## Max.   :36.00   Max.   :31.00
```

```
resultat_acp <- PCA(donnees)
```

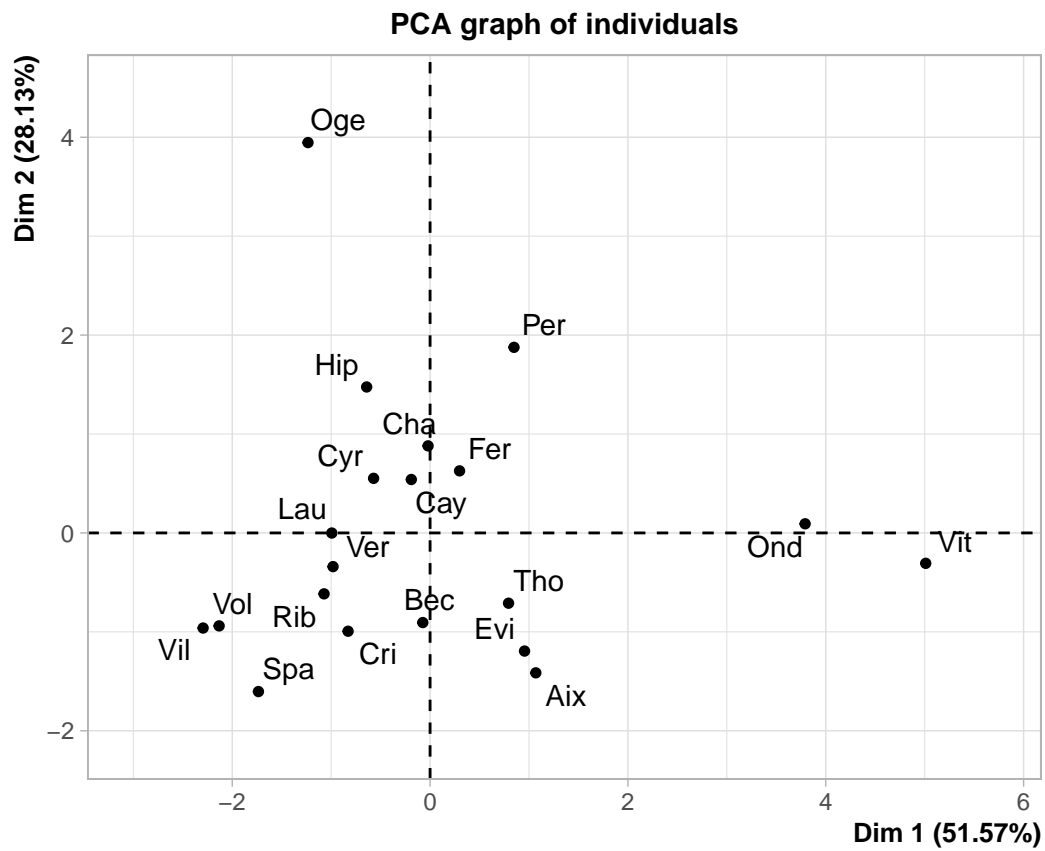




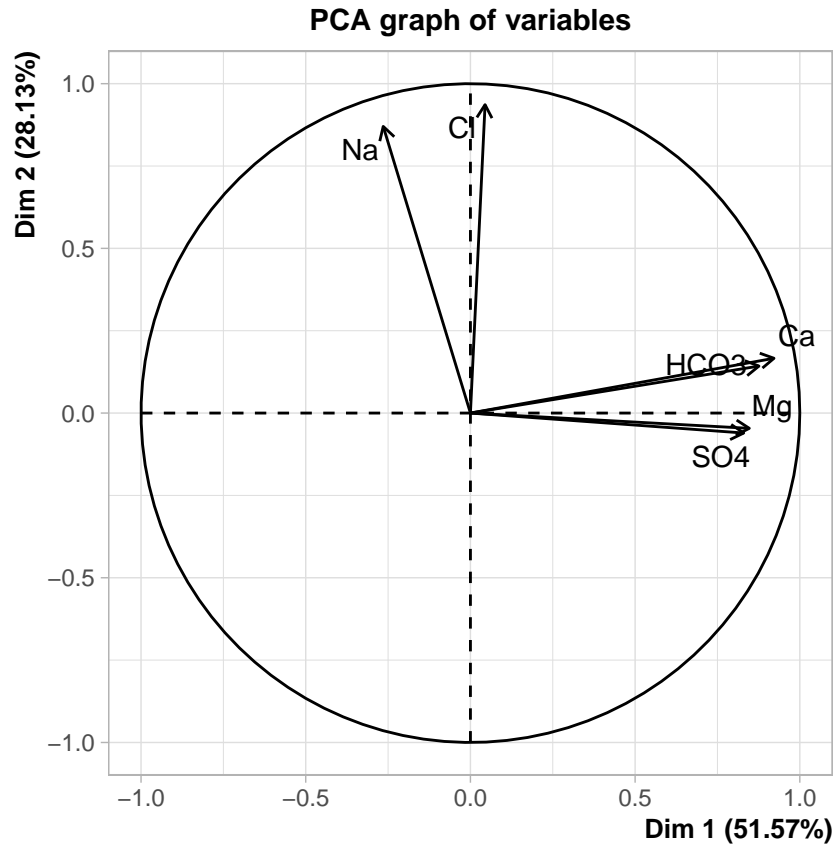
```
resultat_acp$eig
```

```
##      eigenvalue percentage of variance cumulative percentage of variance
## comp 1 3.09408747          51.5681245          51.56812
## comp 2 1.68756032          28.1260054          79.69413
## comp 3 0.59651319           9.9418865          89.63602
## comp 4 0.50284416           8.3807361          98.01675
## comp 5 0.09323922           1.5539871          99.57074
## comp 6 0.02575563           0.4292605          100.00000
```

```
plot(resultat_acp, choix="ind")
```



```
plot(resultat_acp, choix="var")
```



```
summary(resultat_acp)
```

```
##
## Call:
## PCA(X = donnees)
##
##
## Eigenvalues
##           Dim.1  Dim.2  Dim.3  Dim.4  Dim.5  Dim.6
## Variance      3.094   1.688   0.597   0.503   0.093   0.026
## % of var.     51.568  28.126   9.942   8.381   1.554   0.429
## Cumulative % of var. 51.568 79.694 89.636 98.017 99.571 100.000
##
## Individuals (the 10 first)
##           Dist  Dim.1  ctr  cos2  Dim.2  ctr  cos2  Dim.3  ctr
## Aix | 2.073 | 1.069 1.847 0.266 | -1.414 5.927 0.466 | 0.758 4.820
## Bec | 1.326 | -0.074 0.009 0.003 | -0.906 2.433 0.467 | -0.626 3.286
## Cay | 2.480 | -0.190 0.058 0.006 | 0.540 0.864 0.047 | 2.315 44.927
## Cha | 1.305 | -0.020 0.001 0.000 | 0.880 2.294 0.454 | -0.481 1.940
## Cri | 1.507 | -0.829 1.110 0.302 | -0.994 2.927 0.435 | -0.664 3.698
## Cyr | 0.974 | -0.571 0.527 0.344 | 0.552 0.902 0.321 | -0.294 0.724
## Evi | 2.089 | 0.955 1.474 0.209 | -1.194 4.226 0.327 | 1.199 12.044
## Fer | 1.087 | 0.298 0.144 0.075 | 0.627 1.166 0.333 | 0.668 3.745
## Hip | 1.727 | -0.641 0.664 0.138 | 1.476 6.451 0.730 | -0.226 0.430
## Lau | 1.133 | -0.993 1.595 0.769 | -0.001 0.000 0.000 | -0.510 2.183
```

```

##          cos2
## Aix  0.134 |
## Bec  0.223 |
## Cay  0.872 |
## Cha  0.136 |
## Cri  0.194 |
## Cyr  0.091 |
## Evi  0.329 |
## Fer  0.378 |
## Hip  0.017 |
## Lau  0.203 |
##
## Variables
##          Dim.1    ctr    cos2    Dim.2    ctr    cos2    Dim.3    ctr    cos2
## HC03 |  0.876 24.800 0.767 |  0.143  1.207 0.020 |  0.188  5.917 0.035 |
## S04  |  0.830 22.250 0.688 | -0.061  0.217 0.004 | -0.268 12.083 0.072 |
## Cl   |  0.044  0.064 0.002 |  0.936 51.937 0.876 | -0.301 15.223 0.091 |
## Ca   |  0.922 27.452 0.849 |  0.166  1.638 0.028 | -0.249 10.417 0.062 |
## Mg   |  0.847 23.169 0.717 | -0.046  0.127 0.002 |  0.464 36.161 0.216 |
## Na   | -0.265  2.265 0.070 |  0.870 44.873 0.757 |  0.347 20.198 0.120 |

```