

Compte Rendu TP1-2

Masson Kevin

2022-09-27

Question 1: Réalisation de l'ACP normée.

Préparation du jeu de donnée

Importation des librairies et packages nécessaires.

```
library("readxl")
library("Hmisc")

## Warning: The packages `ellipsis` (>= 0.3.2) and `vctrs` (>= 0.3.8) are required
## as of rlang 1.0.0.

## Warning: remplacement de l'importation précédente 'ellipsis::check_dots_unnamed'
## par 'rlang::check_dots_unnamed' lors du chargement de 'tibble'

## Warning: remplacement de l'importation précédente 'ellipsis::check_dots_used'
## par 'rlang::check_dots_used' lors du chargement de 'tibble'

## Warning: remplacement de l'importation précédente 'ellipsis::check_dots_empty'
## par 'rlang::check_dots_empty' lors du chargement de 'tibble'

## Warning: remplacement de l'importation précédente 'ellipsis::check_dots_unnamed'
## par 'rlang::check_dots_unnamed' lors du chargement de 'pillar'

## Warning: remplacement de l'importation précédente 'ellipsis::check_dots_used'
## par 'rlang::check_dots_used' lors du chargement de 'pillar'

## Warning: remplacement de l'importation précédente 'ellipsis::check_dots_empty'
## par 'rlang::check_dots_empty' lors du chargement de 'pillar'

library("ggplot2")
library("FactoMineR")
library("factoextra")
library("psych")
```

Lecture du tableau de données.

```
mydata = read_excel("/home/k/m/kmasson/4A/SE/TP1/Data_eleves.xls")
```

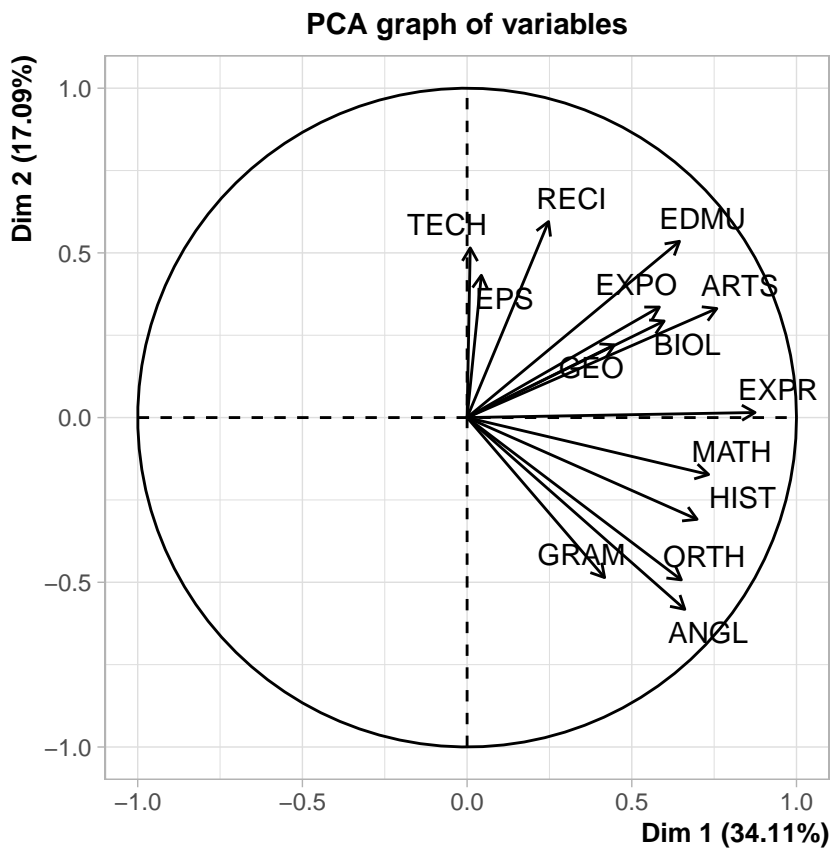
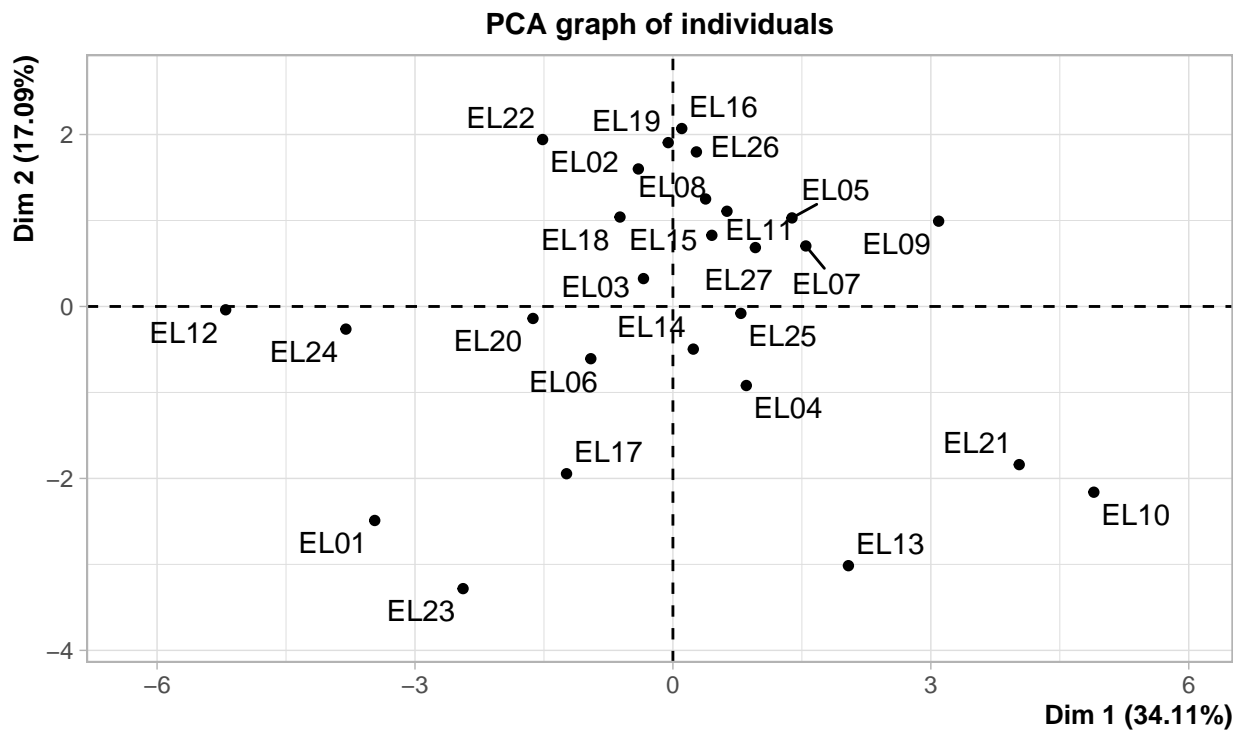
Changement des libellés et mise sous forme de Dataframe.

```
mydata = mydata[1:27,1:15]
mydata=data.frame(mydata)
rownames(mydata) = mydata$eleves
mydata = mydata[,-1]
```

Création de l'ACP normée

Pour ce faire nous faisons appel à la fonction *PCA* qui a pour paramètres le jeu de données.

```
mydata.pca = PCA(mydata, scale.unit=TRUE, graph=TRUE)
```



Question 2: Statistiques sommaires.

Statistiques sommaires sur les matières

```
summary(mydata)
```

```
##          ORTH          GRAM          EXPR          RECI
## Min.   : 0.00   Min.   : 2.000   Min.   : 2.000   Min.   : 4.00
## 1st Qu.: 5.75   1st Qu.: 6.500   1st Qu.: 7.500   1st Qu.: 6.50
## Median :11.00   Median : 7.500   Median : 9.000   Median :10.00
## Mean   :10.02   Mean    : 7.556   Mean    : 8.889   Mean    :10.07
## 3rd Qu.:14.00   3rd Qu.: 8.500   3rd Qu.:10.500   3rd Qu.:12.50
## Max.   :20.00   Max.    :14.500   Max.    :16.500   Max.    :16.00
##          MATH          ANGL          HIST          BIOL
## Min.   : 8.00   Min.   : 3.00   Min.   : 5.000   Min.   : 2.000
## 1st Qu.:10.00   1st Qu.: 8.75   1st Qu.: 7.750   1st Qu.: 7.000
## Median :12.00   Median :11.00   Median : 9.000   Median :10.000
## Mean   :12.57   Mean    :10.93   Mean    : 9.204   Mean    : 9.593
## 3rd Qu.:14.50   3rd Qu.:12.75   3rd Qu.:11.000   3rd Qu.:12.000
## Max.   :18.00   Max.    :17.00   Max.    :15.000   Max.    :17.000
##          EDMU          ARTS          TECH          EPS
## Min.   : 7.00   Min.   : 1.500   Min.   : 0.00   Min.   : 5.00
## 1st Qu.:13.75   1st Qu.: 6.750   1st Qu.:11.50   1st Qu.:11.00
## Median :16.00   Median : 9.000   Median :14.00   Median :13.50
## Mean   :15.20   Mean    : 9.019   Mean    :12.63   Mean    :13.07
## 3rd Qu.:17.50   3rd Qu.:12.500   3rd Qu.:16.00   3rd Qu.:15.25
## Max.   :19.00   Max.    :14.500   Max.    :18.00   Max.    :18.50
##          GEO          EXPO
## Min.   : 0.00   Min.   : 0.00
## 1st Qu.:13.00   1st Qu.:13.50
## Median :14.00   Median :15.00
## Mean   :13.80   Mean    :14.48
## 3rd Qu.:15.25   3rd Qu.:17.00
## Max.   :18.00   Max.    :18.00
```

Les matières GRAM,EXPR,HIST,BIOL et ARTS sont les matières les plus difficiles pour ces élèves d'après les résultats obtenues (moyennes et médianes faibles). Les matières ORTH et RECI n'ayant pas de très bons résultats, on peut se demander si les matières en rapport avec le Français sont en corrélations. EDMU et EXPO sont les matières où les élèves reussissent le mieux.

Statistiques sommaires sur les élèves

```
mydata_trp = data.frame(t(mydata))
summary(mydata_trp)
```

```
##          EL01          EL02          EL03          EL04
## Min.   : 1.500   Min.   : 4.000   Min.   : 5.00   Min.   : 0.00
## 1st Qu.: 7.125   1st Qu.: 8.125   1st Qu.: 8.00   1st Qu.: 9.25
## Median : 9.000   Median :12.000   Median :10.25   Median :11.00
## Mean   : 8.464   Mean    :11.679   Mean    :11.11   Mean    :10.93
## 3rd Qu.:10.375   3rd Qu.:15.750   3rd Qu.:14.00   3rd Qu.:14.50
## Max.   :14.000   Max.    :18.000   Max.    :18.50   Max.    :18.00
##          EL05          EL06          EL07          EL08
## Min.   : 7.50   Min.   : 5.000   Min.   : 6.50   Min.   : 2.500
## 1st Qu.:11.00   1st Qu.: 6.625   1st Qu.:10.25   1st Qu.: 9.125
## Median :13.25   Median :10.500   Median :13.50   Median :12.000
## Mean   :12.93   Mean    :10.464   Mean    :12.96   Mean    :11.464
```

##	3rd Qu.:15.38	3rd Qu.:13.375	3rd Qu.:15.75	3rd Qu.:13.375
##	Max. :17.00	Max. :17.500	Max. :18.00	Max. :17.000
##	EL09	EL10	EL11	EL12
##	Min. : 7.50	Min. :10.00	Min. : 4.000	Min. : 0.000
##	1st Qu.:13.12	1st Qu.:13.12	1st Qu.: 9.875	1st Qu.: 2.250
##	Median :15.00	Median :14.75	Median :12.000	Median : 6.000
##	Mean :14.21	Mean :14.79	Mean :11.786	Mean : 6.679
##	3rd Qu.:15.88	3rd Qu.:16.50	3rd Qu.:13.375	3rd Qu.:11.250
##	Max. :17.00	Max. :20.00	Max. :18.000	Max. :15.000
##	EL13	EL14	EL15	EL16
##	Min. : 4.50	Min. : 6.50	Min. : 7.50	Min. : 6.50
##	1st Qu.: 9.50	1st Qu.:10.25	1st Qu.: 9.50	1st Qu.: 9.25
##	Median :12.25	Median :12.00	Median :11.25	Median :11.25
##	Mean :12.07	Mean :11.46	Mean :12.21	Mean :11.75
##	3rd Qu.:15.88	3rd Qu.:13.75	3rd Qu.:15.00	3rd Qu.:14.25
##	Max. :17.00	Max. :15.00	Max. :18.00	Max. :18.00
##	EL17	EL18	EL19	EL20
##	Min. : 4.500	Min. : 7.000	Min. : 4.50	Min. : 5.00
##	1st Qu.: 6.625	1st Qu.: 8.125	1st Qu.: 9.00	1st Qu.: 8.00
##	Median : 8.750	Median :10.500	Median :12.00	Median :10.25
##	Mean : 9.607	Mean :11.179	Mean :11.50	Mean :10.21
##	3rd Qu.:12.875	3rd Qu.:13.500	3rd Qu.:12.88	3rd Qu.:12.00
##	Max. :16.000	Max. :17.500	Max. :18.00	Max. :18.00
##	EL21	EL22	EL23	EL24
##	Min. : 6.00	Min. : 2.000	Min. : 0.000	Min. : 0.000
##	1st Qu.:11.75	1st Qu.: 7.625	1st Qu.: 5.375	1st Qu.: 5.125
##	Median :14.00	Median :10.500	Median : 8.500	Median : 7.750
##	Mean :13.71	Mean :10.464	Mean : 8.429	Mean : 7.714
##	3rd Qu.:16.00	3rd Qu.:14.250	3rd Qu.:12.250	3rd Qu.: 8.750
##	Max. :19.00	Max. :16.000	Max. :18.500	Max. :15.000
##	EL25	EL26	EL27	
##	Min. : 4.000	Min. : 4.50	Min. : 7.50	
##	1st Qu.: 9.375	1st Qu.: 7.75	1st Qu.: 9.00	
##	Median :11.000	Median :12.25	Median :12.25	
##	Mean :11.286	Mean :11.57	Mean :12.21	
##	3rd Qu.:13.625	3rd Qu.:14.00	3rd Qu.:14.75	
##	Max. :17.500	Max. :18.00	Max. :17.50	

Certains élèves se démarquent des autres: EL10,EL09 avec les moyennes générales les plus élevées. Cependant, ils semblent avoir des profils différents, EL09 possède un profil homogène alors qu'EL10 possède un profil avec de grosses capacités dans certaines matières.

Les élèves 1,12,17,23,24 sont en difficultés (moyennes basses) notamment l'élève numéro 12 avec pour moyenne 6,679. Le profil de l'élève numéro 1 est aussi inquiétant, sa note maximum est 14, ainsi il possède un profil fragile dans l'ensemble des matières, vers quel filière le projeté dans l'avenir ?

Question 3: Corrélations.

```
rcorr(as.matrix(mydata))$r
```

##	ORTH	GRAM	EXPR	RECI	MATH	ANGL
## ORTH	1.00000000	0.53193641	0.512491039	-0.06584006	0.37741318	0.654419471
## GRAM	0.53193641	1.00000000	0.354117977	-0.10403022	0.34116941	0.403571855
## EXPR	0.51249104	0.35411798	1.000000000	0.23620258	0.58934714	0.538117448

```
## RECI -0.06584006 -0.10403022 0.236202578 1.00000000 0.18054329 -0.041416815
## MATH 0.37741318 0.34116941 0.589347137 0.18054329 1.00000000 0.696924068
## ANGL 0.65441947 0.40357186 0.538117448 -0.04141681 0.69692407 1.000000000
## HIST 0.54466302 0.29507903 0.591412881 -0.05760725 0.63763616 0.592184818
## BIOL 0.35035637 -0.02040237 0.401657962 0.13463036 0.11846745 0.171548732
## EDMU 0.18337932 0.20517943 0.634169061 0.41163083 0.27548988 0.004124413
## ARTS 0.30740988 0.14521896 0.724991566 0.32487576 0.37126688 0.286880270
## TECH -0.29362248 -0.16625700 0.003608444 0.34084641 0.20992505 -0.226097749
## EPS -0.03960858 -0.09472332 0.026370611 0.23648194 -0.06024121 -0.273408387
## GEO 0.16244818 -0.07176006 0.394955377 -0.11088774 0.32617140 0.134354152
## EXPO 0.14269742 0.05278196 0.314635869 0.32985689 0.41091247 0.253589128
## HIST BIOL EDMU ARTS TECH EPS
## ORTH 0.54466302 0.35035637 0.183379320 0.30740988 -0.293622483 -0.03960858
## GRAM 0.29507903 -0.02040237 0.205179429 0.14521896 -0.166257002 -0.09472332
## EXPR 0.59141288 0.40165796 0.634169061 0.72499157 0.003608444 0.02637061
## RECI -0.05760725 0.13463036 0.411630831 0.32487576 0.340846405 0.23648194
## MATH 0.63763616 0.11846745 0.275489880 0.37126688 0.209925049 -0.06024121
## ANGL 0.59218482 0.17154873 0.004124413 0.28688027 -0.226097749 -0.27340839
## HIST 1.00000000 0.25930812 0.221296794 0.26157530 -0.014742553 0.06731441
## BIOL 0.25930812 1.00000000 0.455598901 0.58337642 -0.165041767 0.09803248
## EDMU 0.22129679 0.45559890 1.000000000 0.70027355 0.176027437 0.29234997
## ARTS 0.26157530 0.58337642 0.700273555 1.00000000 -0.053943960 0.06551116
## TECH -0.01474255 -0.16504177 0.176027437 -0.05394396 1.000000000 0.06968198
## EPS 0.06731441 0.09803248 0.292349967 0.06551116 0.069681976 1.00000000
## GEO 0.32889942 0.46281961 0.322071990 0.23962044 0.154582835 0.34711802
## EXPO 0.25840339 0.56177062 0.432835103 0.57057209 0.288724270 -0.21845283
## GEO EXPO
## ORTH 0.16244818 0.14269742
## GRAM -0.07176006 0.05278196
## EXPR 0.39495538 0.31463587
## RECI -0.11088774 0.32985689
## MATH 0.32617140 0.41091247
## ANGL 0.13435415 0.25358913
## HIST 0.32889942 0.25840339
## BIOL 0.46281961 0.56177062
## EDMU 0.32207199 0.43283510
## ARTS 0.23962044 0.57057209
## TECH 0.15458283 0.28872427
## EPS 0.34711802 -0.21845283
## GEO 1.00000000 0.09439988
## EXPO 0.09439988 1.00000000
```

Le coefficient de corrélation linéaire le plus important est celui entre les variables ARTS et EXPR avec 0.7249916.

On réalise le test d'hypothèse suivant: H_0 : 'X1 et X2 ne sont pas corrélés' H_1 : 'X1 et X2 sont corrélés'

```
rcorr(as.matrix(mydata))$P
```

```
## ORTH GRAM EXPR RECI MATH ANGL
## ORTH NA 0.004293198 0.0062725660 0.74421188 5.228715e-02 2.129458e-04
## GRAM 0.0042931980 NA 0.0699523107 0.60558682 8.158363e-02 3.684120e-02
## EXPR 0.0062725660 0.069952311 NA 0.23556562 1.217572e-03 3.787899e-03
## RECI 0.7442118797 0.605586819 0.2355656178 NA 3.674945e-01 8.374843e-01
## MATH 0.0522871467 0.081583630 0.0012175723 0.36749447 NA 5.369169e-05
## ANGL 0.0002129458 0.036841195 0.0037878986 0.83748431 5.369169e-05 NA
```

```

## HIST 0.0033089340 0.135116908 0.0011585091 0.77533094 3.467399e-04 1.137084e-03
## BIOL 0.0731908336 0.919544417 0.0378302241 0.50316642 5.561797e-01 3.922288e-01
## EDMU 0.3598891937 0.304575783 0.0003821281 0.03290090 1.642749e-01 9.837107e-01
## ARTS 0.1188049750 0.469852587 0.0000188982 0.09824448 5.656475e-02 1.468252e-01
## TECH 0.1371458181 0.407212150 0.9857483051 0.08189150 2.932794e-01 2.567912e-01
## EPS 0.8444927033 0.638379422 0.8961200453 0.23499589 7.653354e-01 1.676155e-01
## GEO 0.4181919445 0.722071794 0.0414620006 0.58188877 9.683315e-02 5.040508e-01
## EXPO 0.4776804220 0.793732075 0.1099462947 0.09290177 3.323776e-02 2.018419e-01
##          HIST          BIOL          EDMU          ARTS          TECH          EPS
## ORTH 0.0033089340 0.073190834 3.598892e-01 1.188050e-01 0.1371458 0.84449270
## GRAM 0.1351169081 0.919544417 3.045758e-01 4.698526e-01 0.4072122 0.63837942
## EXPR 0.0011585091 0.037830224 3.821281e-04 1.889820e-05 0.9857483 0.89612005
## RECI 0.7753309438 0.503166422 3.290090e-02 9.824448e-02 0.0818915 0.23499589
## MATH 0.0003467399 0.556179660 1.642749e-01 5.656475e-02 0.2932794 0.76533545
## ANGL 0.0011370841 0.392228809 9.837107e-01 1.468252e-01 0.2567912 0.16761552
## HIST          NA 0.191511038 2.672989e-01 1.875185e-01 0.9418193 0.73867865
## BIOL 0.1915110378          NA 1.693208e-02 1.403175e-03 0.4106977 0.62663925
## EDMU 0.2672988677 0.016932078          NA 4.769473e-05 0.3797965 0.13893635
## ARTS 0.1875184895 0.001403175 4.769473e-05          NA 0.7892905 0.74544797
## TECH 0.9418192597 0.410697678 3.797965e-01 7.892905e-01          NA 0.72981974
## EPS 0.7386786537 0.626639254 1.389364e-01 7.454480e-01 0.7298197          NA
## GEO 0.0939113637 0.015059398 1.013513e-01 2.286582e-01 0.4413752 0.07606997
## EXPO 0.1931205532 0.002293884 2.412845e-02 1.885228e-03 0.1441305 0.27365228
##          GEO          EXPO
## ORTH 0.41819194 0.477680422
## GRAM 0.72207179 0.793732075
## EXPR 0.04146200 0.109946295
## RECI 0.58188877 0.092901774
## MATH 0.09683315 0.033237760
## ANGL 0.50405079 0.201841879
## HIST 0.09391136 0.193120553
## BIOL 0.01505940 0.002293884
## EDMU 0.10135130 0.024128448
## ARTS 0.22865818 0.001885228
## TECH 0.44137525 0.144130537
## EPS 0.07606997 0.273652278
## GEO          NA 0.639531560
## EXPO 0.63953156          NA

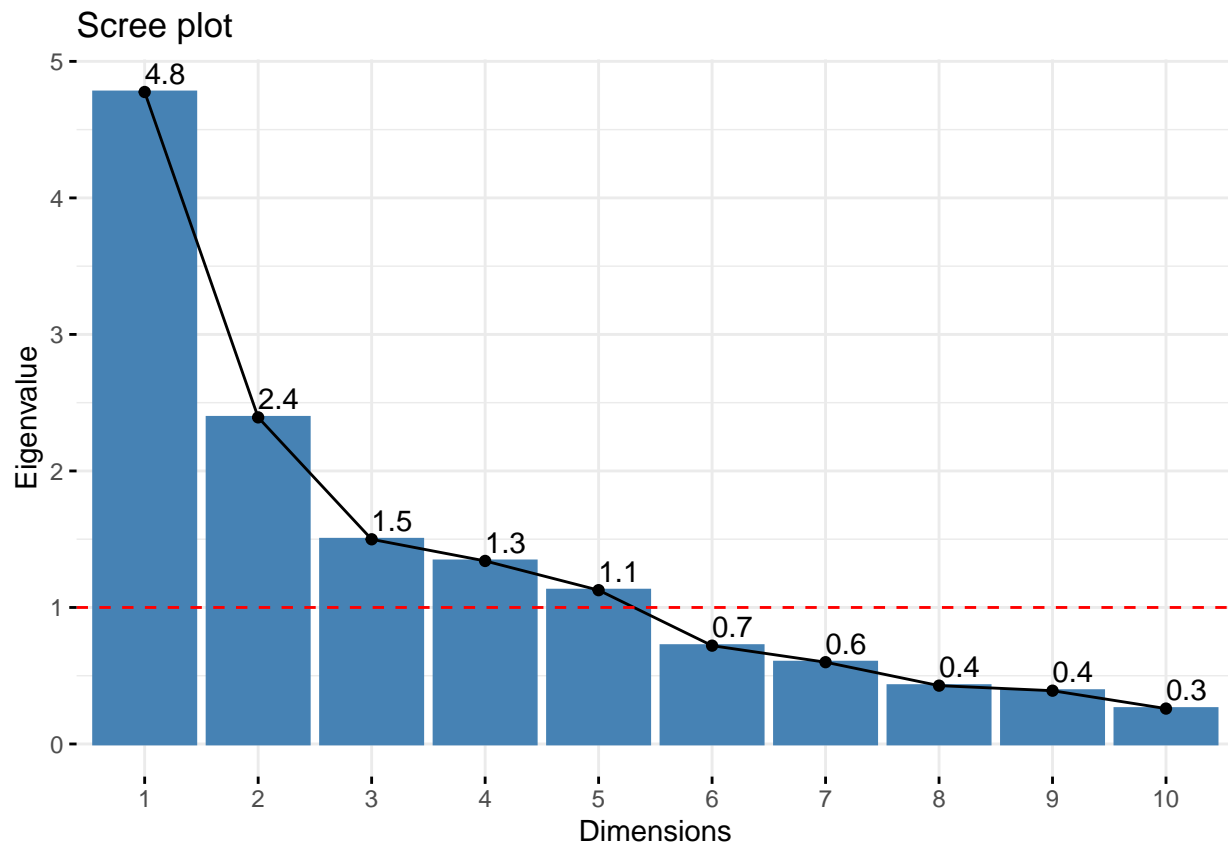
```

La P-valeur entre ARTS et EXPR est inférieure au seuil de signification donc on rejette H_0 et on favorise H_1 et donc la corrélation entre ARTS et EXPR.

Question 4: Composantes Principales.

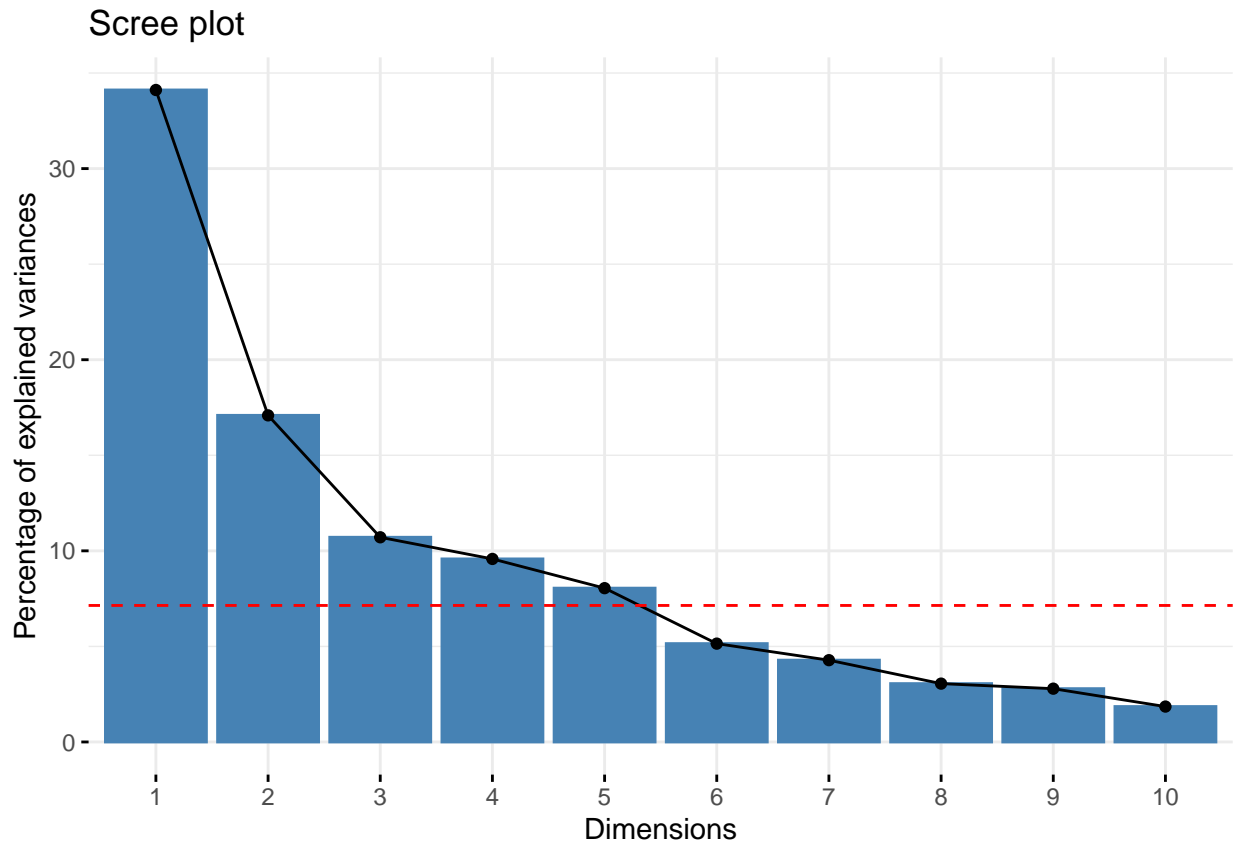
On va étudier le tableau obtenu lors de l'ACP (`mydata.pca`) afin de déterminer le nombre de composantes principales à retenir.

```
fviz_eig(mydata.pca, choice = "eigenvalue", addlabels = TRUE) + geom_hline (yintercept = 1, linetype = 2)
```



Les composantes 1 à 5 ont une inertie projetée supérieure à 1.

```
fviz_screepLOT (mydata.pca) +geom_hline (yintercept = 100/nrow(mydata.pca$eig), linetype = 2, color = "red")
```



Les composantes 1 à 5 ont une inertie projetée cumulée supérieure à 80%. De ce fait, j'ai fait le choix de garder les 3 premières composantes représentant 61,91% des valeurs ce qui est acceptable par rapport à la dimension du tableau.

Question 5: Interprétation.

Création d'une fonction donnant la valeur de la contribution totale en ne gardant que les matières respectant le critère de Kaiser (Inertie supérieur à l'inertie moyenne).

```
Ctb_tot = function(n){
  som = 0
  for (i in 1:14){
    if (mydata.pca$var$contrib[i,n]>mean(mydata.pca$var$contrib[,n]))
      som=som+mydata.pca$var$contrib[i,n]
  }
  return (som)
}
```

Interprétation Axe 1:

```
Ctb_tot(1)
```

```
## [1] 83.7054
```

```
mydata.pca$var$contrib[,1]>mean(mydata.pca$var$contrib[,1])
```

```
## ORTH GRAM EXPR RECI MATH ANGL HIST BIOL EDMU ARTS TECH EPS GEO
## TRUE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE FALSE FALSE FALSE
```

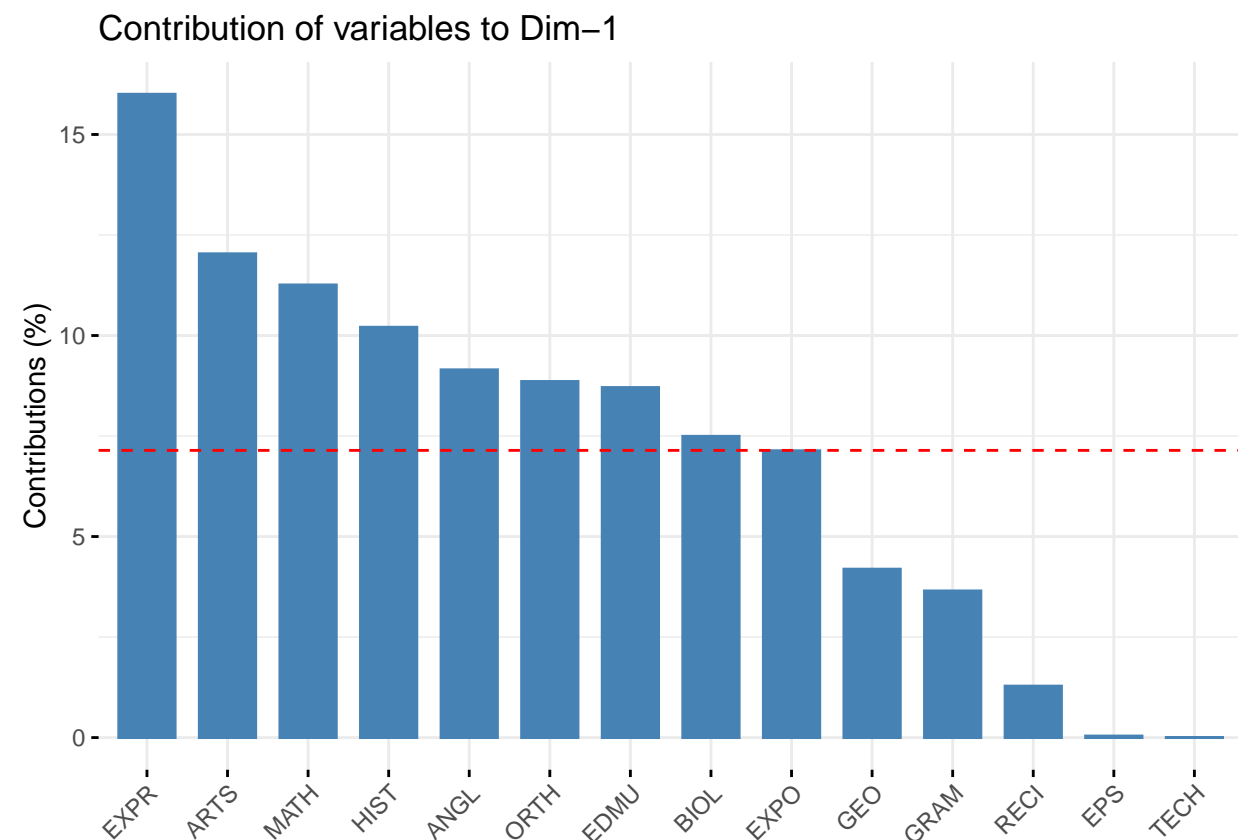


```
## EXPO
## FALSE
```

```
mydata.pca$var$contrib
```

##		Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
##	ORTH	8.857178434	10.147700841	2.011115774	0.819262759	1.9692093
##	GRAM	3.647913177	9.862949216	0.343291721	0.008243911	22.7076131
##	EXPR	16.001163011	0.009919583	0.146062754	0.142529916	1.9664615
##	RECI	1.280583878	14.800939040	9.201063052	0.030321946	13.6369489
##	MATH	11.258878062	1.249038115	6.282449446	13.644601594	0.6409718
##	ANGL	9.147755175	14.164029878	2.131100334	0.304427329	0.8831389
##	HIST	10.205943479	3.985966555	0.352172894	8.812289759	1.4405132
##	BIOL	7.493655757	3.591472838	7.035864125	14.725203802	10.7481175
##	EDMU	8.707419780	11.968896845	0.632205474	0.643621144	7.9649991
##	ARTS	12.033404324	4.583971428	0.002441541	9.735102286	0.5733860
##	TECH	0.001890561	11.091355234	14.496417584	23.845091334	1.3943743
##	EPS	0.039310859	7.801449997	24.155785266	10.697922436	11.1753449
##	GEO	4.189905868	2.037057095	19.492996119	10.971312784	15.2269386
##	EXPO	7.134997636	4.705253336	13.717033917	5.620069002	9.6719828

```
fviz_contrib(mydata.pca,choice="var", axes = 1 )
```



La composante numéro 1 est représentée à hauteur de 83,7% par les 8 matières suivantes : ORTH, EXPR, MATH, ANGL, HIST, BIOL, EDMU, ARTS.

```
sign(mydata.pca$var$coord)
```

```
## Dim.1 Dim.2 Dim.3 Dim.4 Dim.5
```

```
## ORTH      1      -1      1      -1      1
## GRAM      1      -1     -1     -1      1
## EXPR      1       1      1       1      1
## RECI      1       1     -1       1      1
## MATH      1      -1     -1       1     -1
## ANGL      1      -1     -1       1     -1
## HIST      1      -1      1       1     -1
## BIOL      1       1      1      -1     -1
## EDMU      1       1      1      -1      1
## ARTS      1       1      1      -1      1
## TECH      1       1     -1       1     -1
## EPS       1       1      1       1      1
## GEO       1       1      1       1     -1
## EXPO      1       1     -1      -1     -1
```

Du fait des signes de projection , on n'observe pas d'opposition entre ces différentes matières.

```
mydata.pca$var$cos2
```

```
##          Dim.1          Dim.2          Dim.3          Dim.4          Dim.5
## ORTH 4.229923e-01 0.2427720217 3.015351e-02 0.0109850169 0.022187119
## GRAM 1.742134e-01 0.2359596680 5.147118e-03 0.0001105378 0.255847113
## EXPR 7.641676e-01 0.0002373146 2.189980e-03 0.0019111006 0.022156159
## RECI 6.115685e-02 0.3540953710 1.379554e-01 0.0004065693 0.153647766
## MATH 5.376903e-01 0.0298817942 9.419541e-02 0.1829525118 0.007221842
## ANGL 4.368694e-01 0.3388580549 3.195249e-02 0.0040818887 0.009950343
## HIST 4.874053e-01 0.0953596459 5.280277e-03 0.1181588583 0.016230290
## BIOL 3.578746e-01 0.0859218394 1.054917e-01 0.1974416771 0.121099246
## EDMU 4.158403e-01 0.2863420326 9.478923e-03 0.0086299409 0.089741798
## ARTS 5.746793e-01 0.1096662218 3.660706e-05 0.1305323137 0.006460351
## TECH 9.028755e-05 0.2653478631 2.173509e-01 0.3197249347 0.015710442
## EPS  1.877369e-03 0.1866406803 3.621779e-01 0.1434422081 0.125912825
## GEO  2.000974e-01 0.0487342381 2.922667e-01 0.1471079399 0.171562210
## EXPO 3.407461e-01 0.1125677514 2.056653e-01 0.0753562303 0.108974416
```

Les modalités actives les plus importantes sont EXPR(16%), ARTS(12.03%), MATH(11.26%) car ce sont les 3 qui sont bien représentées (du fait du \cos^2).

Interprétation Axe 2:

```
Ctb_tot(2)
```

```
## [1] 79.83732
```

```
mydata.pca$var$contrib[,2]>mean(mydata.pca$var$contrib[,2])
```

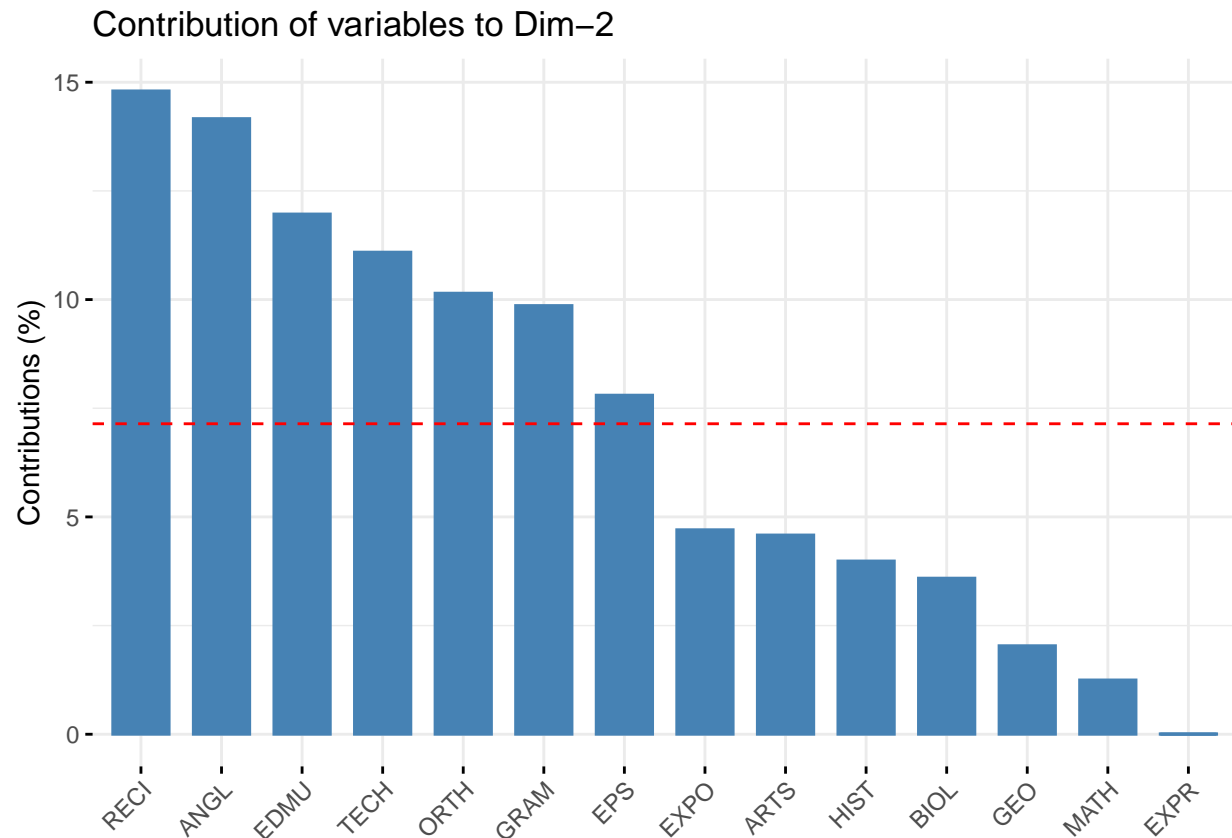
```
## ORTH GRAM EXPR RECI MATH ANGL HIST BIOL EDMU ARTS TECH EPS GEO
## TRUE TRUE FALSE TRUE FALSE TRUE FALSE FALSE TRUE FALSE TRUE TRUE FALSE
## EXPO
## FALSE
```

```
mydata.pca$var$contrib
```

```
##          Dim.1          Dim.2          Dim.3          Dim.4          Dim.5
## ORTH 8.857178434 10.147700841 2.011115774 0.819262759 1.9692093
## GRAM 3.647913177 9.862949216 0.343291721 0.008243911 22.7076131
## EXPR 16.001163011 0.009919583 0.146062754 0.142529916 1.9664615
## RECI 1.280583878 14.800939040 9.201063052 0.030321946 13.6369489
```

```
## MATH 11.258878062 1.249038115 6.282449446 13.644601594 0.6409718
## ANGL 9.147755175 14.164029878 2.131100334 0.304427329 0.8831389
## HIST 10.205943479 3.985966555 0.352172894 8.812289759 1.4405132
## BIOL 7.493655757 3.591472838 7.035864125 14.725203802 10.7481175
## EDMU 8.707419780 11.968896845 0.632205474 0.643621144 7.9649991
## ARTS 12.033404324 4.583971428 0.002441541 9.735102286 0.5733860
## TECH 0.001890561 11.091355234 14.496417584 23.845091334 1.3943743
## EPS 0.039310859 7.801449997 24.155785266 10.697922436 11.1753449
## GEO 4.189905868 2.037057095 19.492996119 10.971312784 15.2269386
## EXPO 7.134997636 4.705253336 13.717033917 5.620069002 9.6719828
```

```
fviz_contrib(mydata.pca,choice="var", axes = 2 )
```



La composante numéro 2 est représentée à hauteur de 79.83% par les 7 matières suivantes : RECI, ANGL, EDMU, TECH, ORTH, GRAM, EPS.

```
sign(mydata.pca$var$coord)
```

```
##      Dim.1 Dim.2 Dim.3 Dim.4 Dim.5
## ORTH      1    -1      1    -1      1
## GRAM      1    -1     -1    -1      1
## EXPR      1     1      1     1      1
## RECI      1     1     -1     1      1
## MATH      1    -1     -1     1     -1
## ANGL      1    -1     -1     1     -1
## HIST      1    -1      1     1     -1
## BIOL      1     1      1    -1     -1
## EDMU      1     1      1    -1      1
## ARTS      1     1      1    -1      1
```

```
## TECH      1      1     -1      1     -1
## EPS       1      1      1      1      1
## GEO       1      1      1      1     -1
## EXPO      1      1     -1     -1     -1
```

Du fait des signes de projection , on observe des oppositions entre les couples (ORTH,GRAM,ANGL) et (RECI,EDMU,TECH,EPS).

```
mydata.pca$var$cos2
```

```
##          Dim.1      Dim.2      Dim.3      Dim.4      Dim.5
## ORTH 4.229923e-01 0.2427720217 3.015351e-02 0.0109850169 0.022187119
## GRAM 1.742134e-01 0.2359596680 5.147118e-03 0.0001105378 0.255847113
## EXPR 7.641676e-01 0.0002373146 2.189980e-03 0.0019111006 0.022156159
## RECI 6.115685e-02 0.3540953710 1.379554e-01 0.0004065693 0.153647766
## MATH 5.376903e-01 0.0298817942 9.419541e-02 0.1829525118 0.007221842
## ANGL 4.368694e-01 0.3388580549 3.195249e-02 0.0040818887 0.009950343
## HIST 4.874053e-01 0.0953596459 5.280277e-03 0.1181588583 0.016230290
## BIOL 3.578746e-01 0.0859218394 1.054917e-01 0.1974416771 0.121099246
## EDMU 4.158403e-01 0.2863420326 9.478923e-03 0.0086299409 0.089741798
## ARTS 5.746793e-01 0.1096662218 3.660706e-05 0.1305323137 0.006460351
## TECH 9.028755e-05 0.2653478631 2.173509e-01 0.3197249347 0.015710442
## EPS  1.877369e-03 0.1866406803 3.621779e-01 0.1434422081 0.125912825
## GEO  2.000974e-01 0.0487342381 2.922667e-01 0.1471079399 0.171562210
## EXPO 3.407461e-01 0.1125677514 2.056653e-01 0.0753562303 0.108974416
```

Les modalités actives les plus importantes sont RECI(14,80%), ANGL(14.16%) car ce sont les 2 qui sont bien représentées (du fait du \cos^2).

Interprétation Axe 3:

```
Ctb_tot(3)
```

```
## [1] 81.0633
```

```
mydata.pca$var$contrib[,3]>mean(mydata.pca$var$contrib[,3])
```

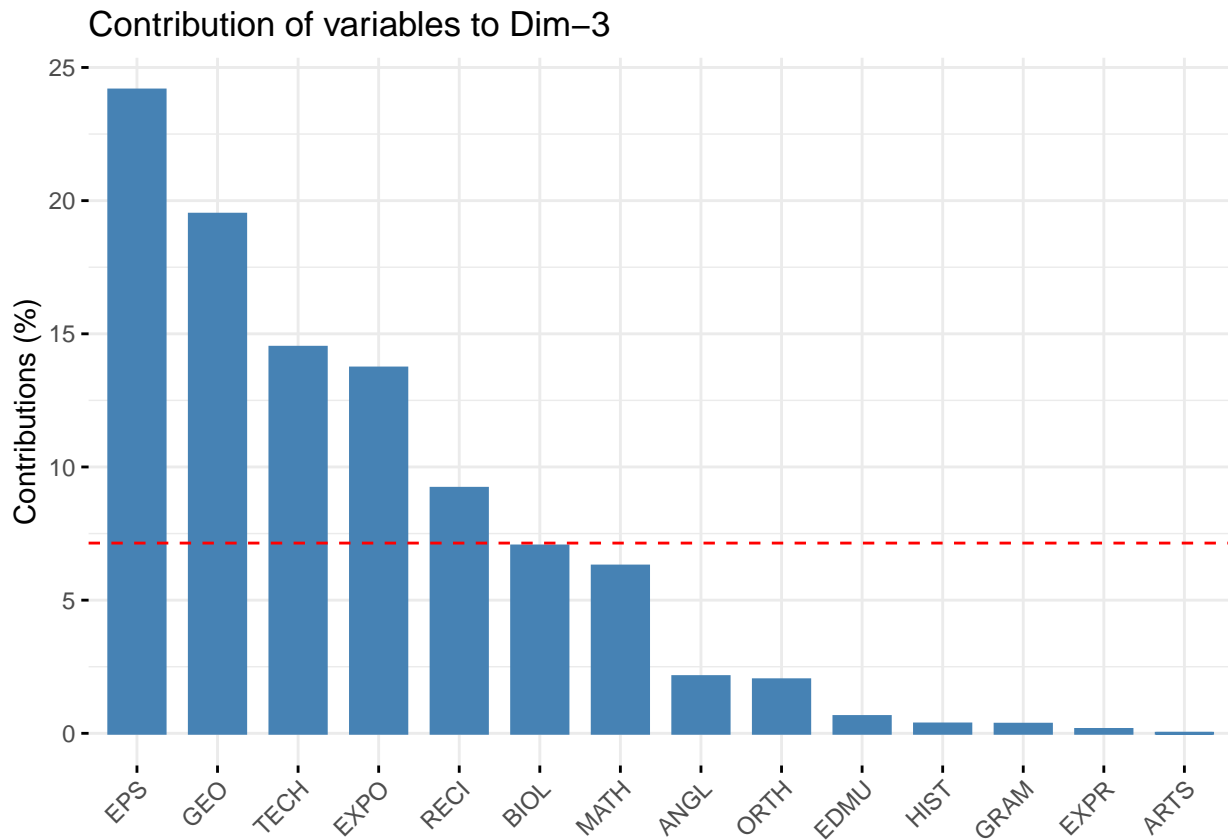
```
## ORTH GRAM EXPR RECI MATH ANGL HIST BIOL EDMU ARTS TECH EPS GEO
## FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE TRUE TRUE TRUE
## EXPO
## TRUE
```

```
mydata.pca$var$contrib
```

```
##          Dim.1      Dim.2      Dim.3      Dim.4      Dim.5
## ORTH 8.857178434 10.147700841 2.011115774 0.819262759 1.9692093
## GRAM 3.647913177 9.862949216 0.343291721 0.008243911 22.7076131
## EXPR 16.001163011 0.009919583 0.146062754 0.142529916 1.9664615
## RECI 1.280583878 14.800939040 9.201063052 0.030321946 13.6369489
## MATH 11.258878062 1.249038115 6.282449446 13.644601594 0.6409718
## ANGL 9.147755175 14.164029878 2.131100334 0.304427329 0.8831389
## HIST 10.205943479 3.985966555 0.352172894 8.812289759 1.4405132
## BIOL 7.493655757 3.591472838 7.035864125 14.725203802 10.7481175
## EDMU 8.707419780 11.968896845 0.632205474 0.643621144 7.9649991
## ARTS 12.033404324 4.583971428 0.002441541 9.735102286 0.5733860
## TECH 0.001890561 11.091355234 14.496417584 23.845091334 1.3943743
## EPS  0.039310859 7.801449997 24.155785266 10.697922436 11.1753449
## GEO  4.189905868 2.037057095 19.492996119 10.971312784 15.2269386
```

```
## EXPO 7.134997636 4.705253336 13.717033917 5.620069002 9.6719828
```

```
fviz_contrib(mydata.pca,choice="var", axes = 3 )
```



La composante numéro 3 est représentée à hauteur de 81.06% par les 5 matières suivantes : EPS, GEO, TECH, EXPO, RECI.

```
sign(mydata.pca$var$coord)
```

```
##      Dim.1 Dim.2 Dim.3 Dim.4 Dim.5
## ORTH      1    -1      1    -1      1
## GRAM      1    -1     -1    -1      1
## EXPR      1      1      1      1      1
## RECI      1      1     -1      1      1
## MATH      1     -1     -1      1     -1
## ANGL      1     -1     -1      1     -1
## HIST      1     -1      1      1     -1
## BIOL      1      1      1     -1     -1
## EDMU      1      1      1     -1      1
## ARTS      1      1      1     -1      1
## TECH      1      1     -1      1     -1
## EPS       1      1      1      1      1
## GEO       1      1      1      1     -1
## EXPO      1      1     -1     -1     -1
```

Du fait des signes de projection , on observe des oppositions entre les couples (TECH,EXPO,RECI) et (EPS,GEO).

```
mydata.pca$var$cos2
```

		Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
##	ORTH	4.229923e-01	0.2427720217	3.015351e-02	0.0109850169	0.022187119
##	GRAM	1.742134e-01	0.2359596680	5.147118e-03	0.0001105378	0.255847113
##	EXPR	7.641676e-01	0.0002373146	2.189980e-03	0.0019111006	0.022156159
##	RECI	6.115685e-02	0.3540953710	1.379554e-01	0.0004065693	0.153647766
##	MATH	5.376903e-01	0.0298817942	9.419541e-02	0.1829525118	0.007221842
##	ANGL	4.368694e-01	0.3388580549	3.195249e-02	0.0040818887	0.009950343
##	HIST	4.874053e-01	0.0953596459	5.280277e-03	0.1181588583	0.016230290
##	BIOL	3.578746e-01	0.0859218394	1.054917e-01	0.1974416771	0.121099246
##	EDMU	4.158403e-01	0.2863420326	9.478923e-03	0.0086299409	0.089741798
##	ARTS	5.746793e-01	0.1096662218	3.660706e-05	0.1305323137	0.006460351
##	TECH	9.028755e-05	0.2653478631	2.173509e-01	0.3197249347	0.015710442
##	EPS	1.877369e-03	0.1866406803	3.621779e-01	0.1434422081	0.125912825
##	GEO	2.000974e-01	0.0487342381	2.922667e-01	0.1471079399	0.171562210
##	EXPO	3.407461e-01	0.1125677514	2.056653e-01	0.0753562303	0.108974416

Les modalités actives les plus importantes sont EPS(24.16%), GEO(19.49%) car ce sont les 2 qui sont bien représentées (du fait du \cos^2).

Question 6: Corrélations.

Cercle des corrélations

```
fviz_pca_var(mydata.pca, col.var = "cos2",  
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),  
              repel = TRUE # Évite le chevauchement de texte  
)
```

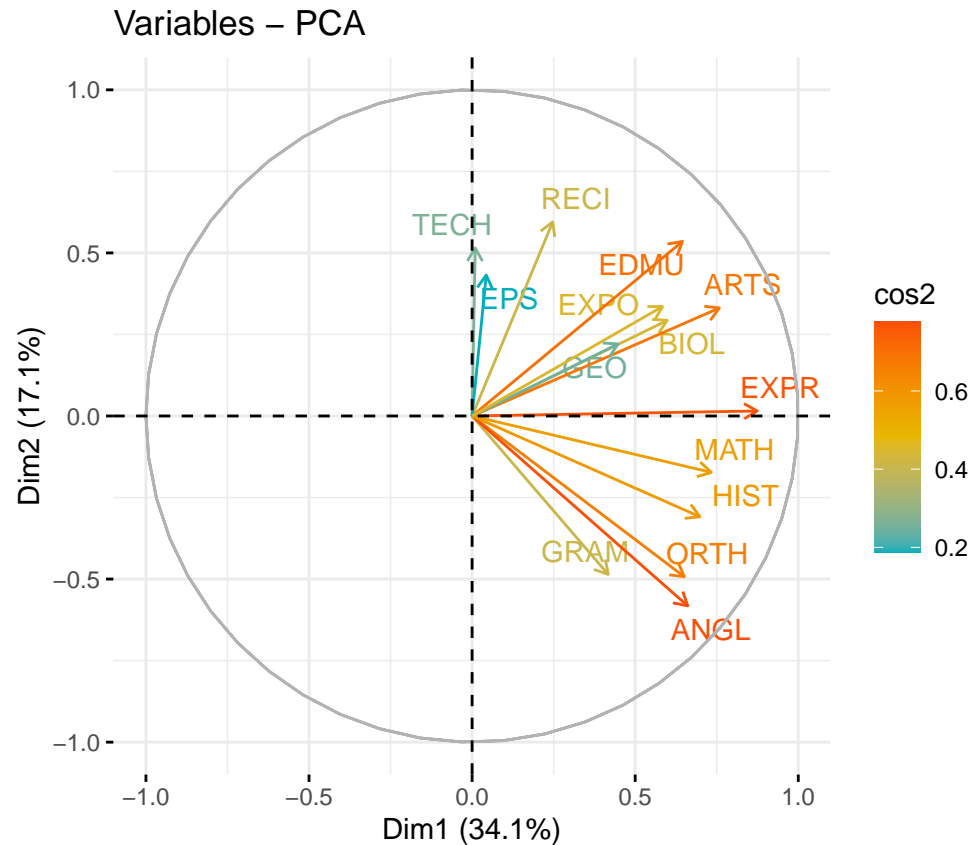


Tableau des corrélations

```
mydata.pca$var$cor
```

##	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
## ORTH	0.650378590	-0.49271901	0.173647652	-0.10480943	0.14895341
## GRAM	0.417388795	-0.48575680	-0.071743415	-0.01051370	0.50581332
## EXPR	0.874166811	0.01540502	0.046797227	0.04371614	0.14884945
## RECI	0.247299110	0.59505913	-0.371423503	0.02016356	0.39197929
## MATH	0.733273680	-0.17286351	-0.306912714	0.42772948	-0.08498142
## ANGL	0.660960956	-0.58211516	-0.178752583	0.06388966	-0.09975141
## HIST	0.698144175	-0.30880357	0.072665513	0.34374243	-0.12739816
## BIOL	0.598226171	0.29312427	0.324794825	-0.44434410	-0.34799317
## EDMU	0.644856795	0.53510937	0.097359762	-0.09289748	0.29956935
## ARTS	0.758076081	0.33115891	0.006050377	-0.36129256	0.08037631
## TECH	0.009501976	0.51511927	-0.466209076	0.56544225	-0.12534130
## EPS	0.043328615	0.43201931	0.601812160	0.37873765	0.35484197
## GEO	0.447322427	0.22075833	0.540616973	0.38354653	-0.41420069
## EXPO	0.583734627	0.33551118	-0.453503335	-0.27451089	-0.33011273

D'après le cercle des corrélations, on en déduit des résultats similaires à la question précédente. De plus on remarque le phénomène suivant, par des corrélations négatives de GRAM, ORTH, ANGL, HIST par rapport à F2, que la réussite dans ces matières ne sont pas en corrélation avec une réussite dans les matières TECH, EPS, RECI qui sont corrélées positivement par rapport à F2. Les matières EDMU, ARTS, EXPR, MATH, HIST sont bien corrélées par rapport à F1, résultat en adéquation avec l'interprétation de l'axe 1. Pour F1, les disciplines les mieux représentées sont EXPR, ARTS, MATH, HIST. Pour F2, les disciplines les mieux représentées sont RECI, ANGL, EDMU.

Question 7: Interprétation classe.

On en déduit pour les 27 élèves de cette classe que des élèves ayant des bonnes notes en littérature n'implique pas spécialement un bon niveau en EPS, TECH, matière sans lien direct avec la littérature (et inversement). Résultat confirmé par le tableau de valeur initial (ex : élève numéro 2 qui a un niveau faible en langue mais bon niveau en EPS et TECH).

Question 8: Interprétation élèves.

```
mydata.pca$ind$contrib[c(10,12),1:2]
```

```
##          Dim.1      Dim.2
## EL10 18.58838 7.22537004
## EL12 20.98589 0.00230761
```

```
sign(mydata.pca$ind$coord[c(10,12),1:2])
```

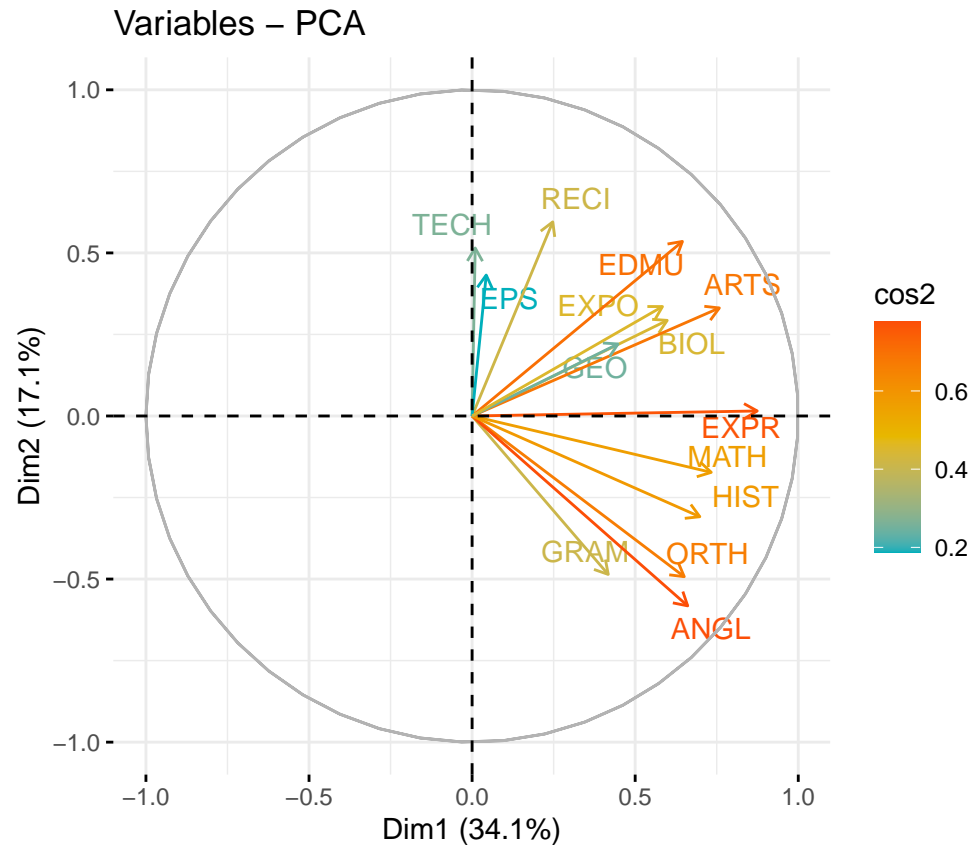
```
##          Dim.1 Dim.2
## EL10         1    -1
## EL12        -1    -1
```

D'après les interprétations que l'on a des axes 1 et 2, on peut dire que l'élève numéro 10 possède une forte contribution à la dimension 1 et 2, pour la dimension 2 il a un signe de coord négatif, ainsi il a un profil bon en général (surtout littéraire) avec certainement des difficultés en technologie, sport et recitation (corrélation négative par rapport à la dimension 2).

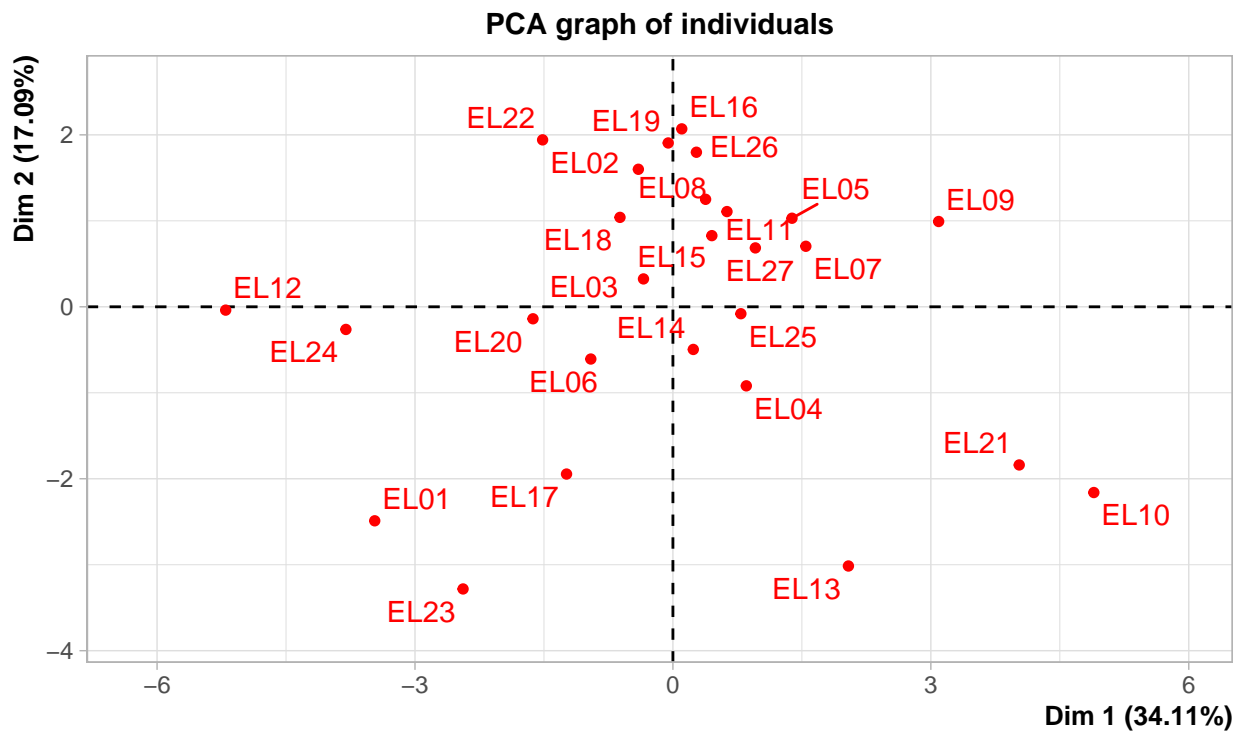
D'après les interprétations que l'on a des axes 1 et 2, on peut dire que l'élève numéro 12 possède une forte contribution à la dimension 1 et faible à la dimension 2, pour les 2 dimensions, on a un signe de coord négatif, ainsi il a un profil inquiétant en général (il contribue fortement à la dimension 1) avec certainement d'importantes difficultés dans l'ensemble des matières (MATH, ANGL, EXPR, EMU, ORTH..), les matières décrivant la dimension 2 peuvent quant à elles être raisonnables (contribution faible à la dimension 2).

Question 9: Elèves expliquants les axes.

```
fviz_pca_var(mydata.pca, col.var = "cos2",
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
              repel = TRUE # Évite le chevauchement de texte
)
```

```
plot.PCA(mydata.pca, axes=c(1,2), choix="ind", col.ind="red")
```



D'après la distance de certains élèves par rapport au centre, les élèves 12,24,10,21 représentent fortement la dimension 1. Alors que les élèves 23,13,1 représentent fortement la dimension 2.

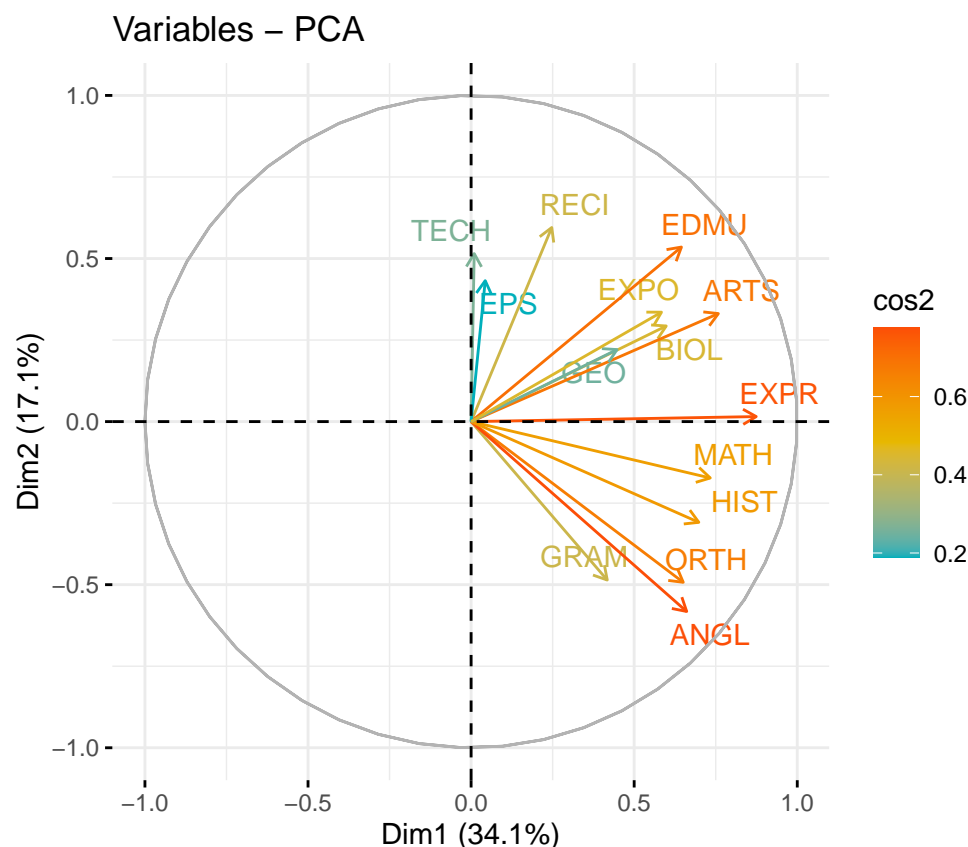
En mettant en relation avec le cercle des corrélations, on remarque par exemple que les 16,19 et 26 ont des profils scolaires similaires avec une bonne représentation de la dimension 2, c'est à dire de bonnes notes en technologie,eps,recitation mais des difficultés dans certaines matières littéraires comme l'orthographe,grammaire,histoire,anglais... De plus l'élève le plus intelligent semble etre l'élève numéro 10 qui représente le mieux la dimension 1, et par corrélation négative l'élève le plus en difficulté est le 12.

Cette interprétation est en cohérence avec le tableau de données initial et avec le tableau de corrélation.

Question 10: Résumé.

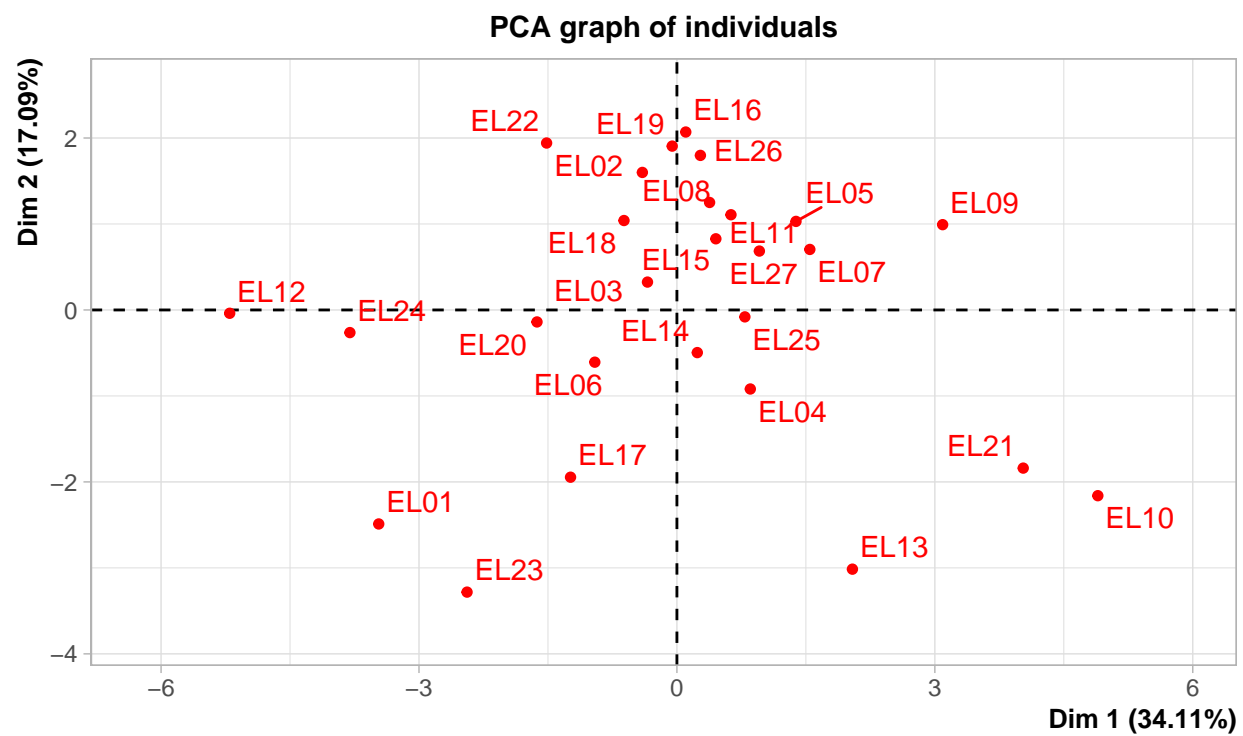
En résumé de cette ACP, on peut clairement voir 2 tendances différentes à l'aide du cercle des corrélations et des autres tableaux:

```
fviz_pca_var(mydata.pca, col.var = "cos2",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE # Évite le chevauchement de texte
)
```



Ce graphique nous montre qu'un élève se situant à droite de l'axe verticale aura une tendance à être un bon élève alors que du côté gauche, ce sera un élève en difficulté. On remarque aussi une tendance par rapport à la dimension 2 avec pour le côté négatif une tendance à être bon dans les matières littéraires.

```
plot.PCA(mydata.pca, axes=c(1,2), choix="ind", col.ind="red")
```



On remarque à l'aide de ce graphique une forte densité d'élèves autour de l'axe 2. Cela peut traduire une forte moyenne de classe pour les matières telles que TECH, EPS et RECI par exemple.