TP2-1

Slim Kammoun

```
library("readxl")
mydata=read_excel("data-look-virgule.xls")
head(mydata)
```

```
## # A tibble: 6 x 13
##   libellé  ORLY  ALEZ  CORS  DIRE  DUCA  FONT  ICAR  ZODI  PAVI  COCK  ESCA
##   <chr>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 viel        1     2    14    38    18    10     9     5     9     4     0
## 2 nouv       20     9     1    11    10     9     1     1    20     9     7
## 3 sobr        9    23     1    15     7    11     6     2     7    12     3
## 4 coca        1     3    15    15     6     5    12    18     4    25     2
## 5 racé        4    33     7     8     3     6     6     4     5    15     5
## 6 miev        3     9     1     7     7     5    12     9     6     9     6
## # ... with 1 more variable: HOTE <dbl>
```

```
mydata=data.frame(mydata) ## il faut imposer le format dataframe
rownames(mydata)<-mydata$libellé
mydata<-mydata[,-1]
head(mydata)
```

```
##      ORLY ALEZ CORS DIRE DUCA FONT ICAR ZODI PAVI COCK ESCA HOTE
## viel    1    2   14   38   18   10    9    5    9    4    0    1
## nouv   20    9    1   11   10    9    1    1   20    9    7   12
## sobr    9   23    1   15    7   11    6    2    7   12    3   17
## coca    1    3   15   15    6    5   12   18    4   25    2    2
## racé    4   33    7    8    3    6    6    4    5   15    5    3
## miev    3    9    1    7    7    5   12    9    6    9    6   13
```

```
library(xlsx)
library(ggplot2)
library(FactoMineR)
library(factoextra)
```

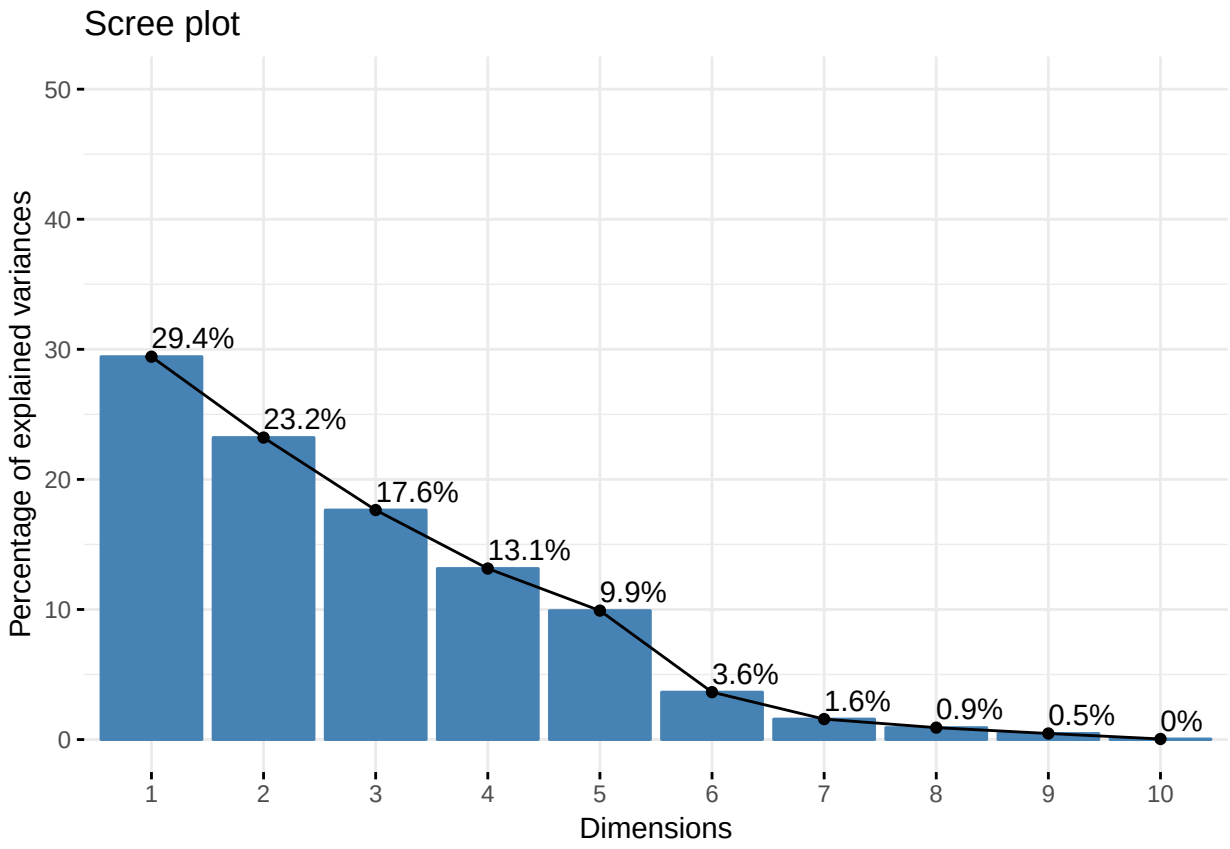
```
## Welcome! Related Books: `Practical Guide To Cluster Analysis in R` at https://goo.gl/13EFCZ
```

```
mydata.ca = CA(mydata, ncp=4, graph=FALSE)

####Q3#####
library(plyr)
#3.1
### RQ nombre de valeurs prpres = min(n,p)-1
mydata.ca$eig
```

```
##          eigenvalue percentage of variance
## dim 1  0.1698450535             29.43385786
## dim 2  0.1339617724             23.21534650
## dim 3  0.1018198712             17.64520989
## dim 4  0.0758603111             13.14646243
## dim 5  0.0572008706              9.91281325
## dim 6  0.0210353403              3.64538858
## dim 7  0.0090835971              1.57417188
```

```
## dim 8  0.0052824869             0.91544597
## dim 9  0.0026878540             0.46580051
## dim 10 0.0002625712             0.04550313
##        cumulative percentage of variance
## dim 1                           29.43386
## dim 2                           52.64920
## dim 3                           70.29441
## dim 4                           83.44088
## dim 5                           93.35369
## dim 6                           96.99908
## dim 7                           98.57325
## dim 8                           99.48870
## dim 9                           99.95450
## dim 10                         100.00000
```

```
fviz_eig(mydata.ca, addlabels = TRUE, ylim = c(0, 50))
```




```
eig=mydata.ca$eig
write.xlsx(as.data.frame(eig),file="TP2-1.xlsx",sheetName="eig")
```

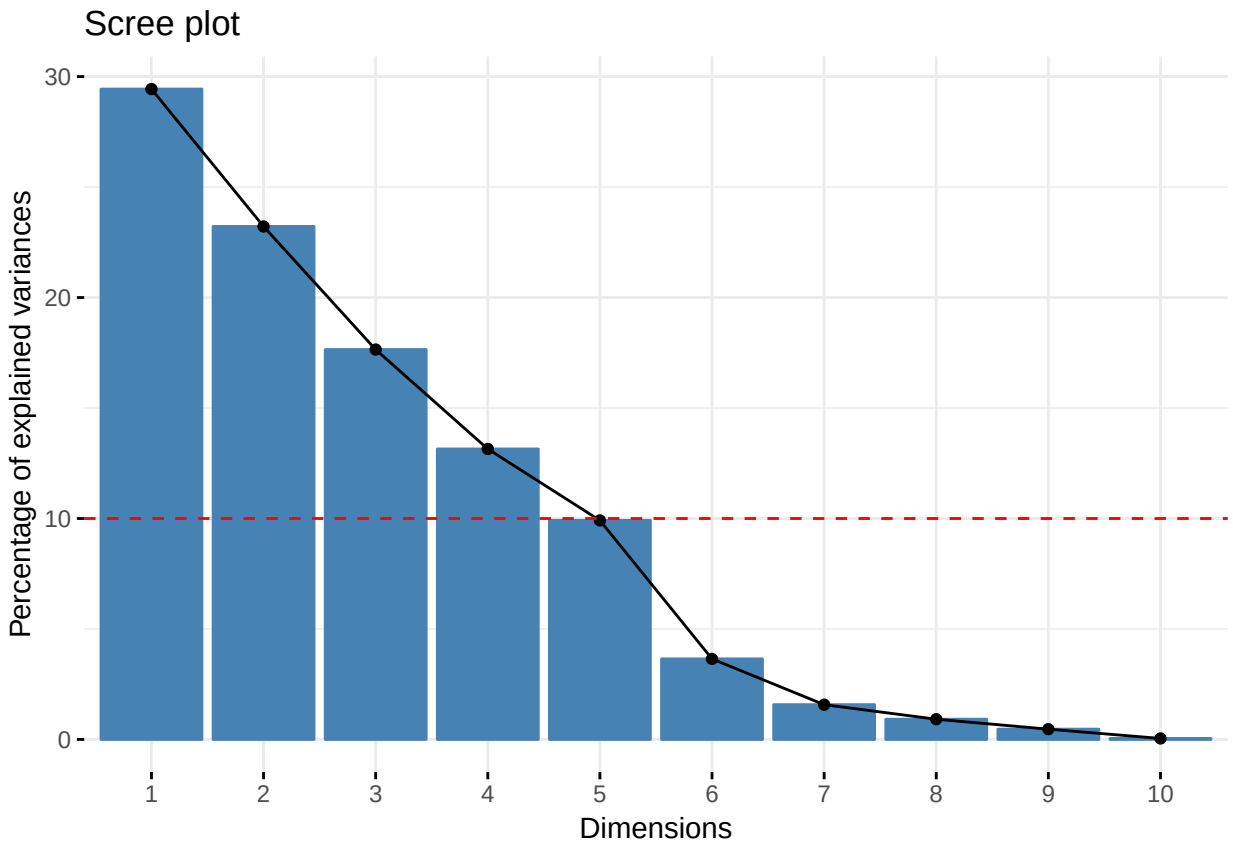
```
## inertie moyenne (Critère de Kaiser)

# code
```

```
## [1] 4
```

```
sum(mydata.ca$eig[,1]>(sum(mydata.ca$eig[,1])/nrow(mydata.ca$eig)), na.rm=TRUE)
```

```
## [1] 4
```

```
### graphiquement
fviz_screeplot (mydata.ca) +
  geom_hline (yintercept = 100/nrow(mydata.ca$eig), linetype = 2, color = "red")
```



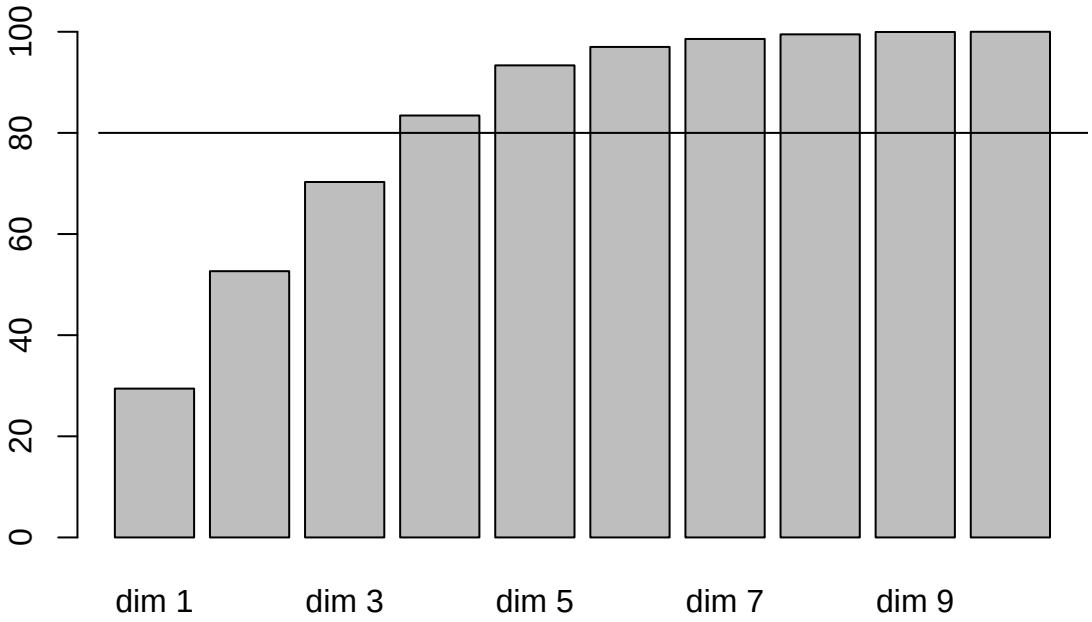

```
png("eig11.png", height=1000, width=1200, res=250, pointsize=8)
fviz_screeplot (mydata.ca) +
  geom_hline (yintercept = 100/nrow(mydata.ca$eig), linetype = 2, color = "red")

dev.off()
```

```
## pdf
##   2
```

```
## inertie totale,

# code
which(mydata.ca$eig[,3]>80)[1]
```

```
## dim 4
##     4
```

```
##graphiquement
barplot(mydata.ca$eig[,3])
lines(c(0,20),c(80,80))
```




```
png("eig12.png", height=1000, width=1200, res=250, pointsize=8)
barplot(mydata.ca$eig[,3])
lines(c(0,20),c(80,80))
dev.off()
```

```
## pdf
##   2
```

```
# 3.2
mydata.ca$col$contrib
```

```
##           Dim 1        Dim 2       Dim 3        Dim 4
## ORLY  9.0499380  2.463961589  0.03895623  0.015418874
## ALEZ  7.1285315  1.762530709 49.97165557 15.108259271
## CORS 37.5690403  4.908029390  1.76871878 37.240255834
## DIRE  0.3207198 25.772609744 19.52052778  0.001817997
## DUCA  0.4606663  5.482306035  4.27919305  0.266309673
## FONT  8.9873762  2.073748890  2.51543518  5.074887806
## ICAR  3.3416135  0.334516627  0.00330969  1.577375255
## ZODI 11.5008630  0.004730709  0.74399813 17.246031462
## PAVI  1.9041985  2.629894899  0.03676345  0.079254657
## COCK  5.0434099  1.874501025 16.32350725 19.126945023
## ESCA  0.3170196 27.965891963  0.01918867  0.457516990
## HOTE 14.3766234 24.727278420  4.77874623  3.805927158
```

```
mydata.ca$col$cos2
```

```
##            Dim 1        Dim 2        Dim 3        Dim 4
## ORLY 0.465518364 0.0999661876 0.0012012896 3.542465e-04
## ALEZ 0.148279059 0.0289164133 0.6231364098 1.403641e-01
## CORS 0.630072208 0.0649225186 0.0177827388 2.789553e-01
## DIRE 0.008730227 0.5533325132 0.3185451566 2.210318e-05
## DUCA 0.039945571 0.3749500492 0.2224455275 1.031409e-02
## FONT 0.422943985 0.0769722323 0.0709647545 1.066690e-01
## ICAR 0.357078365 0.0281937612 0.0002120188 7.528424e-02
## ZODI 0.543689067 0.0001763902 0.0210849052 3.641422e-01
```

```
## PAVI 0.113339259 0.1234623458 0.0013117890 2.106952e-03
## COCK 0.186235189 0.0545947822 0.3613521062 3.154604e-01
## ESCA 0.012230240 0.8509527688 0.0004437857 7.883476e-03
## HOTE 0.336140787 0.4560039611 0.0669819710 3.974537e-02
```

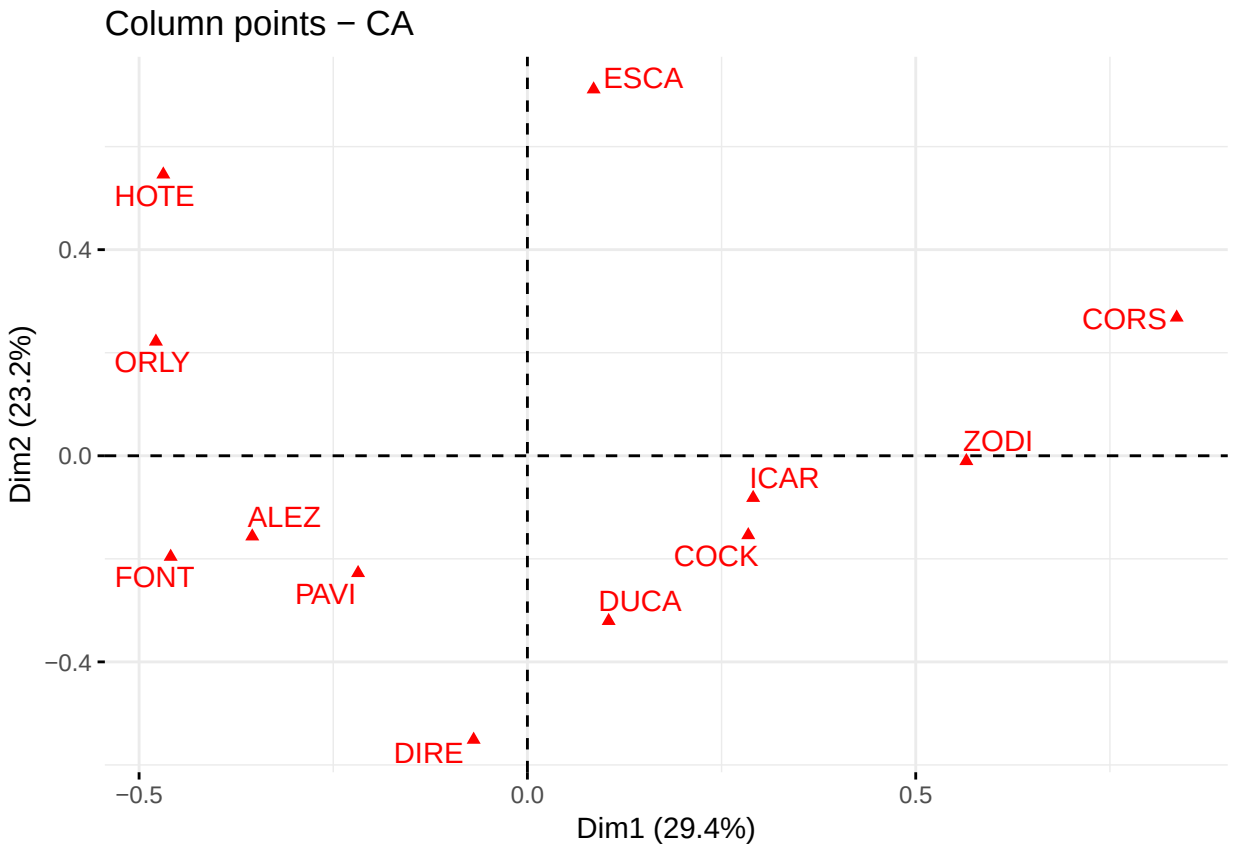
```
mydata.ca$col$coord
```

```
##            Dim 1       Dim 2        Dim 3        Dim 4
## ORLY -0.47831947  0.22165433  0.024298167 -0.013194789
## ALEZ -0.35426935 -0.15644651 -0.726248485 -0.344684396
## CORS  0.83599692  0.26835346  0.140445828 -0.556258698
## DIRE -0.06921791 -0.55105987  0.418110615  0.003482836
## DUCA  0.10461567 -0.32051558  0.246873387  0.053159128
## FONT -0.45932425 -0.19594988  0.188147869 -0.230673125
## ICAR  0.29065203 -0.08167104 -0.007082372  0.133457698
## ZODI  0.56516935 -0.01017983 -0.111298453  0.462528856
## PAVI -0.21801437 -0.22754233  0.023454541 -0.029725028
## COCK  0.28434929 -0.15395608 -0.396083330  0.370078341
## ESCA  0.08525825  0.71116711  0.016240731  0.068450686
## HOTE -0.46878764  0.54600907  0.209263994  0.161197685
```

```
sign(mydata.ca$col$coord)
```

```
##      Dim 1 Dim 2 Dim 3 Dim 4
## ORLY    -1     1     1    -1
## ALEZ    -1    -1    -1    -1
## CORS     1     1     1    -1
## DIRE    -1    -1     1     1
## DUCA     1    -1     1     1
## FONT    -1    -1     1    -1
## ICAR     1    -1    -1     1
## ZODI     1    -1    -1     1
## PAVI    -1    -1     1    -1
## COCK     1    -1    -1     1
## ESCA     1     1     1     1
## HOTE    -1     1     1     1
```

```
write.xlsx(mydata.ca$col,file="TP2-1.xlsx",sheetName="col",append=T)
fviz_ca_col(mydata.ca, repel = TRUE)
```




```
png("plot11.png", height=1000, width=1200, res=250, pointsize=8)
fviz_ca_col(mydata.ca, repel = TRUE)
dev.off()
```

```
## pdf
##   2
```

```
# 3.3
mydata.ca$row$contrib
```

```
##            Dim 1        Dim 2        Dim 3       Dim 4
## viel   2.3389610 3.367508e+01 32.905007073  3.13257559
## nouv  12.8311607 2.179490e-01  0.386543045  0.01369445
## sobr  10.4941917 1.813913e+00  4.160956318  0.43460838
## coca  20.7893180 5.182747e+00  0.602994678 12.10522487
## racé   0.6239266 2.894855e+00 45.458560027  8.82130156
## miev   0.0871466 9.164543e-06  0.559301958  8.35288271
## dist  20.9274239 2.873400e-01  7.527435823  0.15911577
## vulg  28.7696635 1.147597e+01  0.074909451 12.36513457
## hom    0.8043667 4.131328e+00  0.008212738 21.28333964
## fem    0.5772638 3.989257e+01  4.754556232  6.93848116
## petit  1.7565776 4.282360e-01  3.561522656 26.39364129
```

```
mydata.ca$row$cos2
```

```
##             Dim 1        Dim 2        Dim 3        Dim 4
## viel  0.044665918 5.072125e-01 0.3766993824 0.0267187758
## nouv  0.402265070 5.389262e-03 0.0072648037 0.0001917577
## sobr  0.636961579 8.683777e-02 0.1514038925 0.0117821367
```

```
## coca  0.598084107 1.176007e-01 0.0103995642 0.1555451941
## racé  0.017020861 6.228778e-02 0.7434356179 0.1074836571
## miev  0.009082806 7.533701e-07 0.0349458467 0.3888370922
## dist  0.639686995 6.927488e-03 0.1379362616 0.0021723345
## vulg  0.647541532 2.037278e-01 0.0010107625 0.1243063773
## hom   0.046456218 1.881947e-01 0.0002843527 0.5490239296
## fem   0.013615659 7.421375e-01 0.0672285884 0.0730954752
## petit 0.083094134 1.597770e-02 0.1009993060 0.5576530275
```

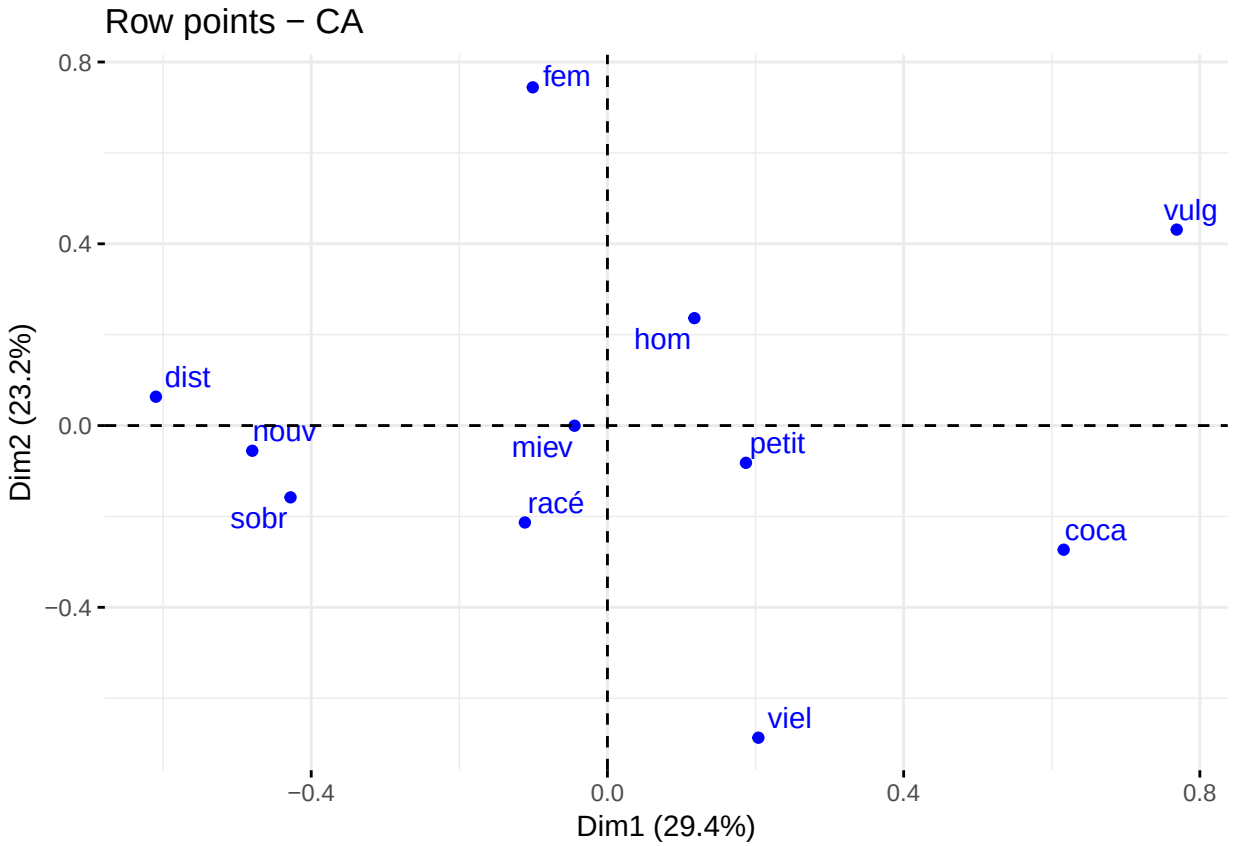
```
mydata.ca$row$coord
```

```
##             Dim 1         Dim 2        Dim 3       Dim 4
## viel   0.20384148 -0.6869093196  0.591972983 -0.15765680
## nouv  -0.47960002 -0.0555121024  0.064451788  0.01047127
## sobr  -0.42793503 -0.1580067211 -0.208636236 -0.05820138
## coca   0.61609943 -0.2731962130 -0.081241371  0.31419406
## racé  -0.11147865 -0.2132565406 -0.736754202 -0.28013809
## miev  -0.04444349 -0.0004047644 -0.087175797  0.29079169
## dist  -0.60973196  0.0634516763  0.283135564 -0.03553191
## vulg   0.76873090  0.4311869642  0.030371418 -0.33681174
## hom    0.11744125  0.2363755072  0.009188137 -0.40373313
## fem   -0.10081397  0.7442923337  0.224015652  0.23358590
## petit  0.18705039 -0.0820220311 -0.206220882  0.48456883
```

```
sign(mydata.ca$row$coord)
```

```
##       Dim 1 Dim 2 Dim 3 Dim 4
## viel      1    -1     1    -1
## nouv     -1    -1     1     1
## sobr     -1    -1    -1    -1
## coca      1    -1    -1     1
## racé     -1    -1    -1    -1
## miev     -1    -1    -1     1
## dist     -1     1     1    -1
## vulg      1     1     1    -1
## hom       1     1     1    -1
## fem      -1     1     1     1
## petit     1    -1    -1     1
```

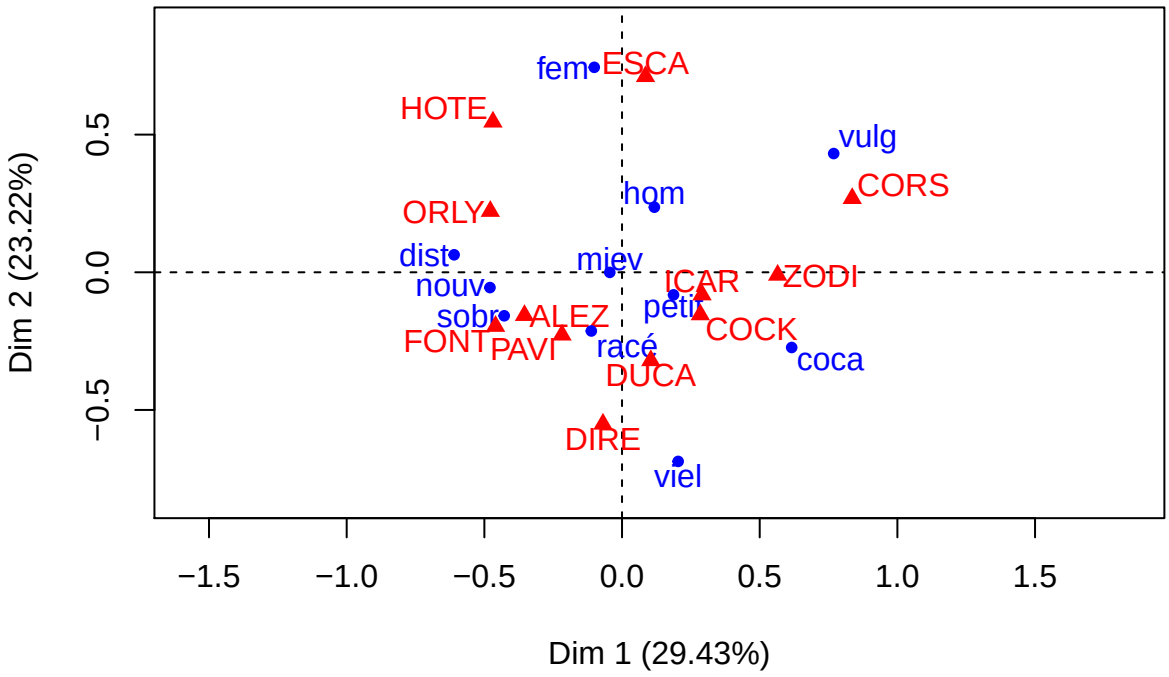
```
write.xlsx(mydata.ca$row,file="TP2-1.xlsx",sheetName="row",append=T)
fviz_ca_row(mydata.ca, repel = TRUE)
```




```
png("plot12.png", height=1000, width=1200, res=250, pointsize=8)
fviz_ca_row(mydata.ca, repel = TRUE)
dev.off()
```

```
## pdf
##   2
```

```
plot(mydata.ca)
```



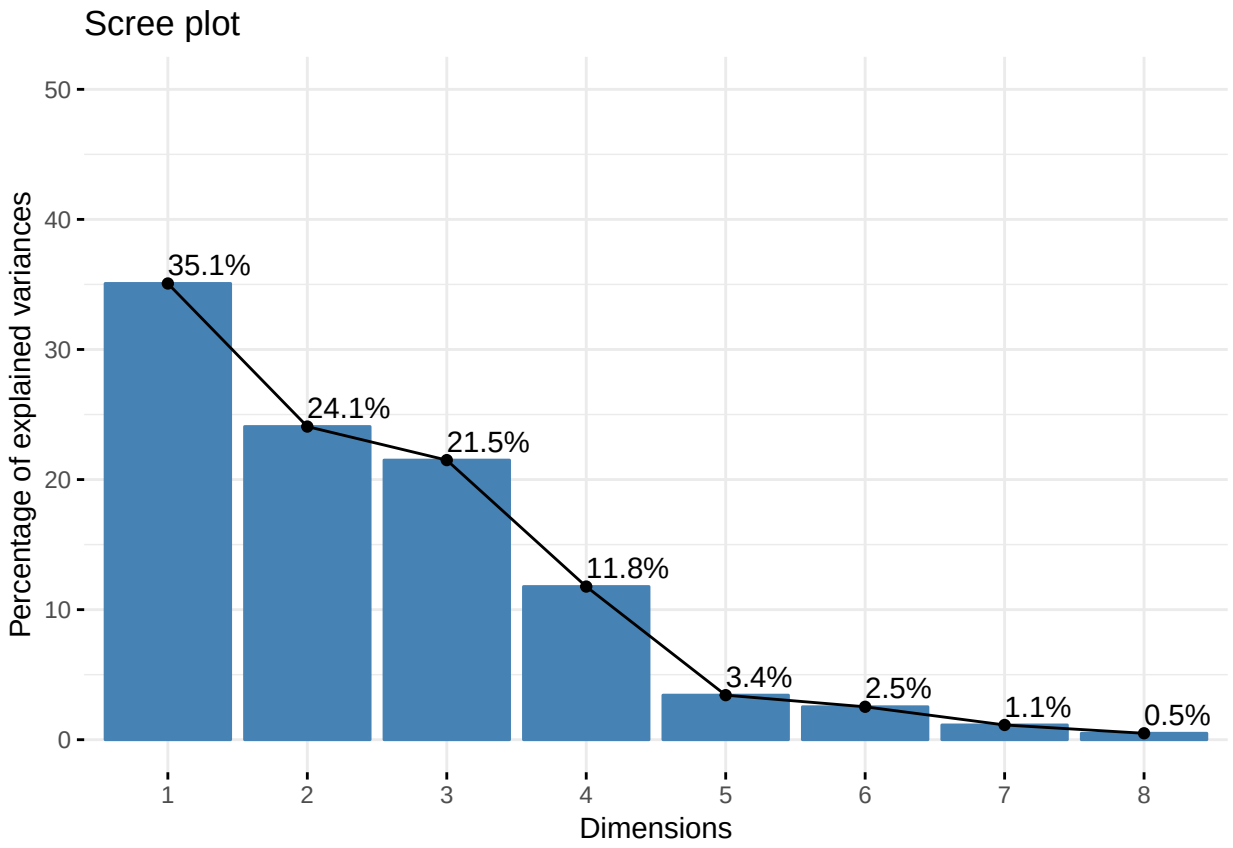

```
png("plot13.png", height=1000, width=1200, res=250, pointsize=8)
plot(mydata.ca)
dev.off()
```

```
## pdf
##   2
```

```
## 4
mydata.ca = CA(mydata, ncp=4, graph=F, col.sup=3, row.sup=c(7,8))



#4.1
mydata.ca$eig
```

```
##        eigenvalue percentage of variance cumulative percentage of variance
## dim 1 0.182323374             35.0725401                            35.07254
## dim 2 0.125149948             24.0744040                            59.14694
## dim 3 0.111743416             21.4954635                            80.64241
## dim 4 0.061202561             11.7731985                            92.41561
## dim 5 0.017806987              3.4254317                            95.84104
## dim 6 0.013157044              2.5309479                            98.37199
## dim 7 0.005898460              1.1346542                            99.50664
## dim 8 0.002564715              0.4933601                           100.00000
```

```
fviz_eig(mydata.ca, addlabels = TRUE, ylim = c(0, 50))
```



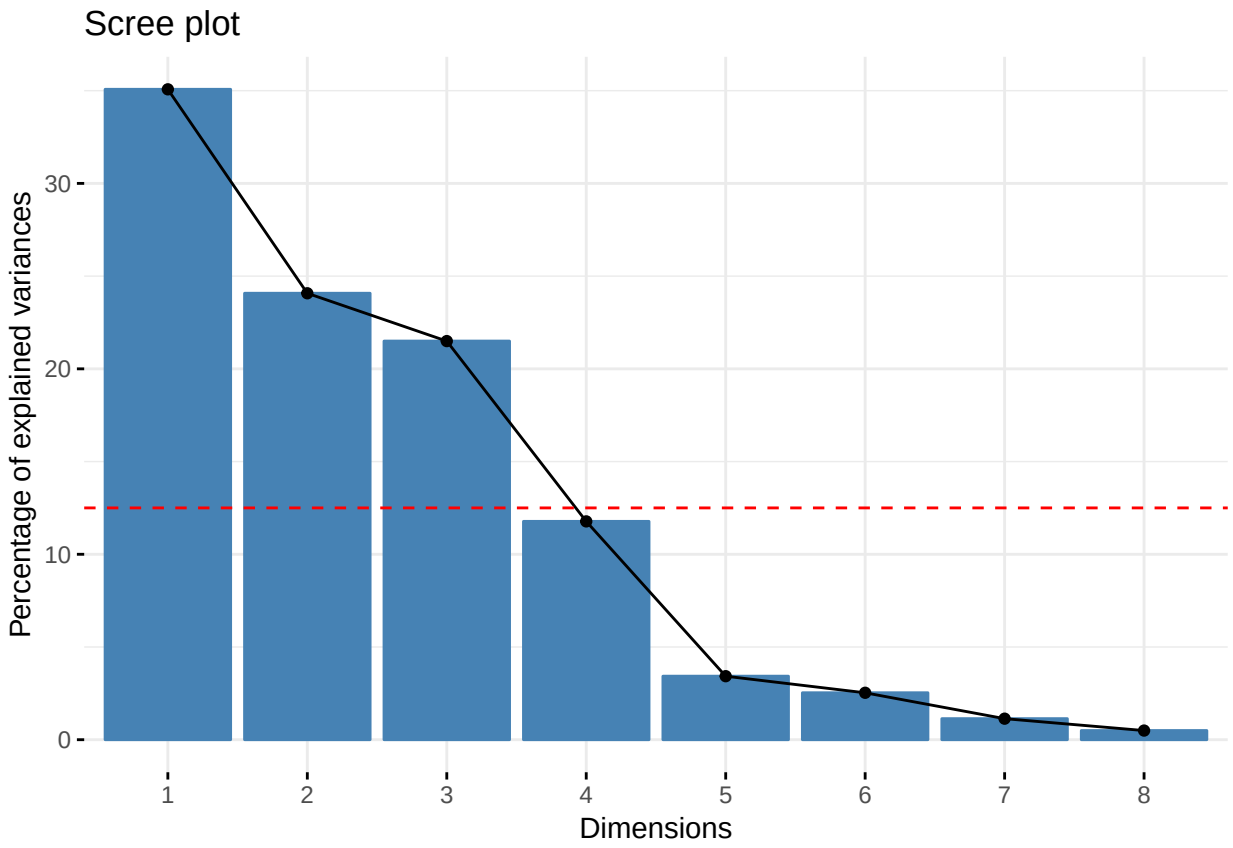

```
## Critère de Kaiser

# code
## les commandes sont équivalents
sum(mydata.ca$eig[,2]>(100/nrow(mydata.ca$eig)), na.rm=TRUE)
```

```
## [1] 3
```

```
sum(mydata.ca$eig[,1]>(sum(mydata.ca$eig[,1])/nrow(mydata.ca$eig)), na.rm=TRUE)
```

```
## [1] 3
```

```
### graphiquement
### graphiquement
fviz_screeplot (mydata.ca) +
  geom_hline (yintercept = 100/nrow(mydata.ca$eig), linetype = 2, color = "red")
```



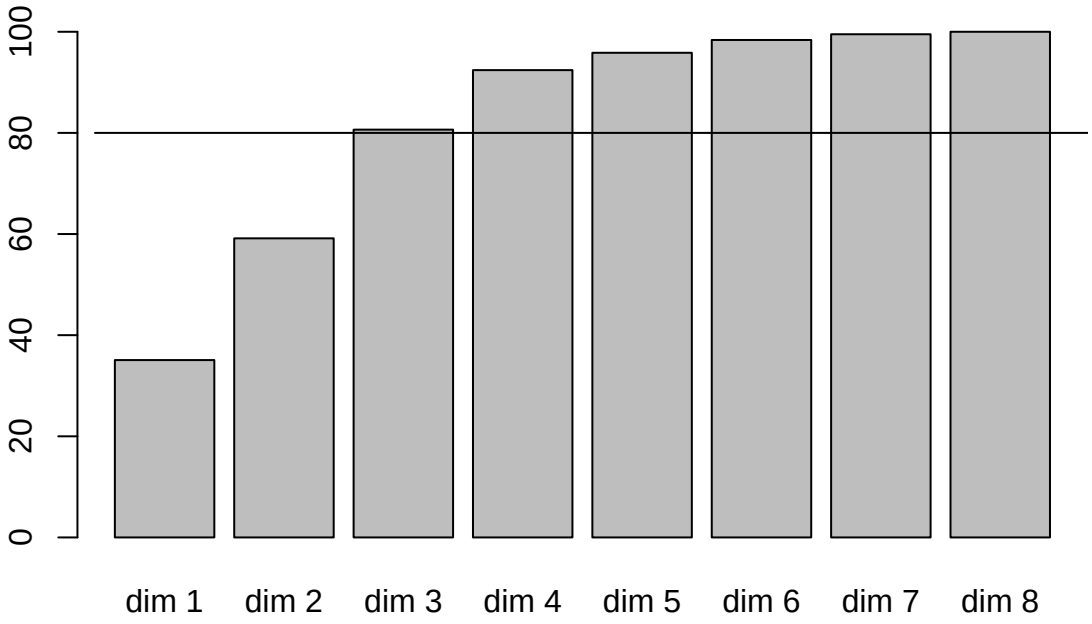

```
png("eig21.png", height=1000, width=1200, res=250, pointsize=8)
fviz_screeplot (mydata.ca) +
  geom_hline (yintercept = 100/nrow(mydata.ca$eig), linetype = 2, color = "red")

dev.off()
```

```
## pdf
##   2
```

```
## inertie totale,

# code
which(mydata.ca$eig[,3]>80)[1]
```

```
## dim 3
##     3
```

```
##graphiquement
barplot(mydata.ca$eig[,3])
lines(c(0,20),c(80,80))
```




```
png("eig22.png", height=1000, width=1200, res=250, pointsize=8)
barplot(mydata.ca$eig[,3])
lines(c(0,20),c(80,80))
dev.off()
```

```
## pdf
##   2
```

```
# 4.2
mydata.ca$col$contrib
```

```
##           Dim 1      Dim 2       Dim 3       Dim 4
## ORLY  7.7471185  2.0850958  6.19378744 20.04604507
## ALEZ  0.8676802 57.8384602 10.27411881 11.03864702
## DIRE 24.5048586  1.3994273 15.41081400 13.70653589
## DUCA  5.5837995  0.8114284  6.91260284  0.79324796
## FONT  0.5354740  4.2287182  4.07348630  0.09189403
## ICAR  2.6759548  2.9520411  3.66554830  4.93022573
## ZODI  2.0553371 14.0841747 15.15455766  0.65356283
## PAVI  0.2163343  2.9186755  7.40197916 24.16975033
## COCK  2.8568192  0.5732590 29.18024880 12.22992669
## ESCA 21.7363824  6.4484537  0.02275667  0.01002361
## HOTE 31.2202413  6.6602660  1.71010003 12.33014084
```

```
mydata.ca$col$cos2
```

```
##           Dim 1      Dim 2        Dim 3        Dim 4
## ORLY 0.37714771 0.06967638 0.1848022122 0.3275874265
## ALEZ 0.01708176 0.78158832 0.1239644596 0.0729484565
## DIRE 0.59416105 0.02329114 0.2290114259 0.1115595959
## DUCA 0.45415911 0.04530194 0.3445874115 0.0216578039
## FONT 0.05053348 0.27392894 0.2356061969 0.0029110883
## ICAR 0.24258830 0.18369689 0.2036618185 0.1500324246
## ZODI 0.09303137 0.43758867 0.4204062861 0.0099302672
## PAVI 0.01390160 0.12873997 0.2915188886 0.5213611416
## COCK 0.10562045 0.01454804 0.6612014818 0.1517806242
```

```
## ESCA 0.70787567 0.14414966 0.0004542114 0.0001095775
## HOTE 0.72248698 0.10579708 0.0242546425 0.0957831361
```

```
mydata.ca$col$coord
```

```
##            Dim 1       Dim 2       Dim 3        Dim 4
## ORLY  0.44443572 -0.19102756  0.31110486  0.414206422
## ALEZ  0.11865099 -0.80259057 -0.31963448 -0.245195829
## DIRE -0.59019565  0.11685283  0.36641437 -0.255739248
## DUCA -0.35051984  0.11070493  0.30532214  0.076544819
## FONT -0.11684444 -0.27204289  0.25229682 -0.028044398
## ICAR -0.26120305  0.22729727 -0.23933054 -0.205416762
## ZODI -0.22891823  0.49647652 -0.48663157  0.074790451
## PAVI -0.06995885 -0.21289577  0.32036398  0.428429788
## COCK -0.20518106  0.07614922 -0.51336963  0.245963933
## ESCA  0.71134161  0.32100111 -0.01801892 -0.008850349
## HOTE  0.72654991  0.27802719  0.13312140 -0.264542204
```

```
sign(mydata.ca$col$coord)
```

```
##      Dim 1 Dim 2 Dim 3 Dim 4
## ORLY     1    -1     1     1
## ALEZ     1    -1    -1    -1
## DIRE    -1     1     1    -1
## DUCA    -1     1     1     1
## FONT    -1    -1     1    -1
## ICAR    -1     1    -1    -1
## ZODI    -1     1    -1     1
## PAVI    -1    -1     1     1
## COCK    -1     1    -1     1
## ESCA     1     1    -1    -1
## HOTE     1     1     1    -1
```

```
# 4.3
mydata.ca$row$contrib
```

```
##             Dim 1     Dim 2      Dim 3     Dim 4
## viel  39.37049525  1.951714 30.3303412  9.784778
## nouv   2.46450274  4.652623 19.3497917 34.826529
## sobr   0.40516078 11.229758  0.2454375  7.250084
## coca  15.47276590 11.928469 22.0282016  1.812999
## racé   0.03385064 37.258799 16.9938034  6.257910
## miev   0.08744569  1.768014  1.9608691  2.879160
## hom    4.24981266  2.047774  1.4402837  1.740661
## fem   37.50672972 24.984707  0.5032481 12.304077
## petit  0.40923663  4.178144  7.1480236 23.143802
```

```
mydata.ca$row$cos2
```

```
##              Dim 1      Dim 2       Dim 3      Dim 4
## viel  0.6238251986 0.02122739 0.294543148 0.05204395
## nouv  0.0810546048 0.10503513 0.390035625 0.38449069
## sobr  0.0296274306 0.56367049 0.010999861 0.17796593
## coca  0.3940127789 0.20850447 0.343796134 0.01549770
## racé  0.0008594498 0.64933804 0.264437994 0.05333475
## miev  0.0115221497 0.15990759 0.158351915 0.12734685
## hom   0.3018354675 0.09983222 0.062694311 0.04149941
```

```
## fem   0.6263590976 0.28640248 0.005150815 0.06897480
## petit 0.0230876421 0.16179935 0.247155713 0.43829552
```

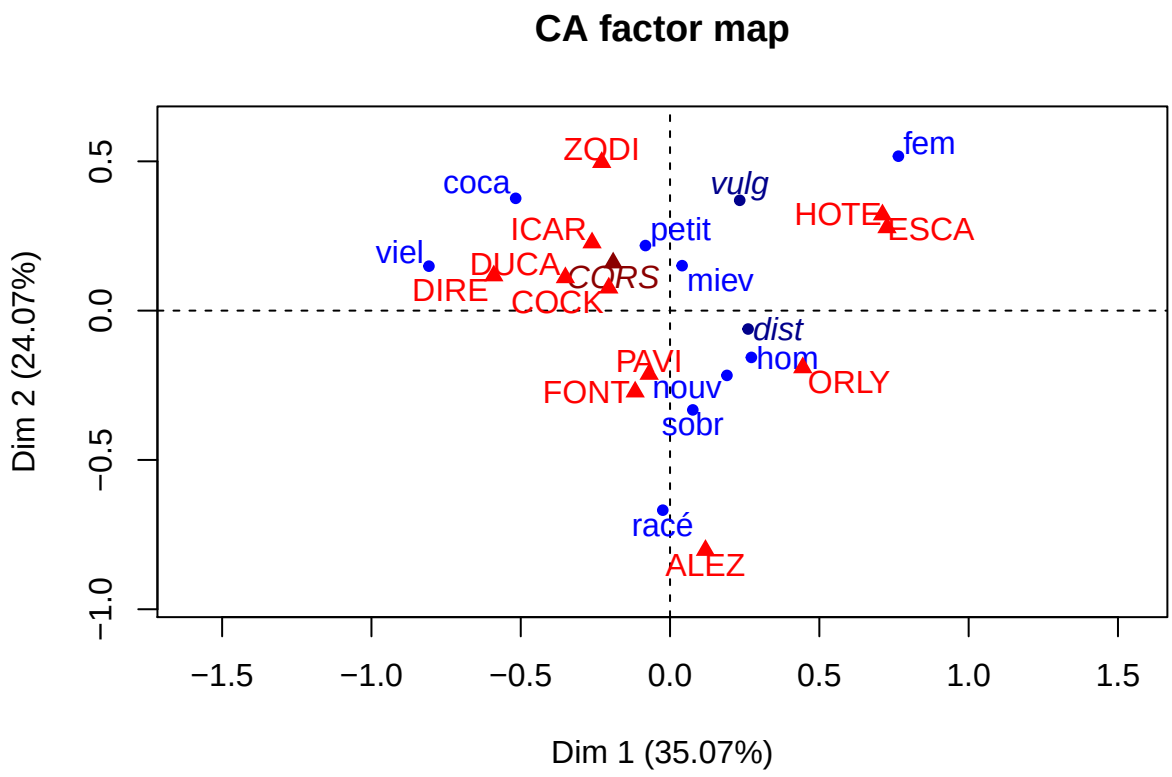
```
mydata.ca$row$coord
```

```
##             Dim 1      Dim 2       Dim 3      Dim 4
## viel  -0.80743710  0.1489449  0.55481949 -0.2332182
## nouv   0.19057263 -0.2169397  0.41804576  0.4150636
## sobr   0.07622788 -0.3324907  0.04644729 -0.1868251
## coca  -0.51695368  0.3760574 -0.48288862  0.1025251
## racé  -0.02431077 -0.6682267 -0.42643251 -0.1915109
## miev   0.04041374  0.1505556 -0.14982145 -0.1343558
## hom    0.27239561 -0.1566572  0.12414494  0.1010034
## fem    0.76479513  0.5171563  0.06935398 -0.2537924
## petit -0.08232098  0.2179260 -0.26934322  0.3586776
```

```
sign(mydata.ca$row$coord)
```

```
##       Dim 1 Dim 2 Dim 3 Dim 4
## viel     -1     1     1    -1
## nouv      1    -1     1     1
## sobr      1    -1     1    -1
## coca     -1     1    -1     1
## racé     -1    -1    -1    -1
## miev      1     1    -1    -1
## hom       1    -1     1     1
## fem       1     1     1    -1
## petit    -1     1    -1     1
```

```
plot(mydata.ca)
```



```r
png("plot23.png", height=1000, width=1200, res=250, pointsize=8)
plot(mydata.ca)
dev.off()
```

```
## pdf
##   2
```

Bonus

```r
mydata.ca$row.sup$coord
```

```
##          Dim 1       Dim 2      Dim 3       Dim 4
## dist 0.2613286 -0.06157351  0.3922651 -0.27558117
## vulg 0.2333117  0.36948404 -0.3032042  0.06856188
```

```r
mydata.ca$row.sup$cos2
```

```
##          Dim 1       Dim 2     Dim 3      Dim 4
## dist 0.1186788 0.006588508 0.2673982 0.13197705
## vulg 0.1262052 0.316516131 0.2131448 0.01089858
```

```r
mydata.ca$col.sup$coord
```

```
##           Dim 1     Dim 2      Dim 3      Dim 4
## CORS -0.1904932 0.1601855 0.03256312 -0.1227919
```

```r
mydata.ca$col.sup$cos2
```

```
##           Dim 1      Dim 2       Dim 3      Dim 4
## CORS 0.03968382 0.02806086 0.001159595 0.01648899
```

```r
write.xlsx(mydata.ca$eig,file="TP2-1.xlsx",sheetName="eig_2",append=T)
write.xlsx(mydata.ca$row,file="TP2-1.xlsx",sheetName="row_2",append=T)
write.xlsx(mydata.ca$col,file="TP2-1.xlsx",sheetName="col_2",append=T)
write.xlsx(mydata.ca$row.sup,file="TP2-1.xlsx",sheetName="row.sub_2",append=T)
write.xlsx(mydata.ca$col.sup,file="TP2-1.xlsx",sheetName="col.sub_2",append=T)
```