

DS - 1

Exercice 1 : Les données `poison` proviennent d'une enquête menée auprès d'enfants de l'école primaire qui ont subi des intoxications alimentaires. Ils ont été interrogés sur leurs symptômes et sur ce qu'ils ont mangé. Les données contiennent 55 lignes (individus) et 11 colonnes (variables).

Nous avons effectué une Analyse des Correspondances Multiples sur ces données. Les résultats de cette ACM ainsi que les statistiques descriptives élémentaires sur ces données sont présentés à l'Annexe.

1. Calculer le taux d'inertie cumulé des 2 premiers axes de cette ACM.
2. Discuter le nombre d'axes à retenir en vous basant sur 3 critères différents.
3. Donner une interprétation de la première carte des modalités.
4. Donner une interprétation de la première carte des individus. On commentera uniquement les individus dont le numéro apparaît sur la carte.

Exercice 2 : On considère la base de données `donnees_pays` datant de 1991. Pour chacun des 10 pays suivants : Af. du Sud, Algérie, Allemagne, Arabie S., Egypte, Ethiopie, Finlande, France, Koweït et Tunisie on dispose des valeurs des 6 variables suivantes : le PNB/h. mesuré en \$ US, les taux d'inflation (T. Inflat.) et de chômage (T. Chom.) en pourcentage, "Com." désigne la balance des échanges commerciaux mesurée en Milliards de \$ US, "Popu." désigne la population mesurée en millions d'habitants et "Sup." désigne la superficie mesurée en millions de km². On a effectué une ACP sur les 6 variables décrites ci-dessus en exécutant le script suivant :

```
> library(FactoMineR)
> summary(donnees_pays)
      PNB      T.INFLAT      T.CHOM      COM      POPU
Min.   : 110   Min.   : 3.20   Min.   : 0.0   Min.   : -10.100   Min.   : 1.30
1st Qu.: 1398   1st Qu.: 3.65   1st Qu.: 0.0   1st Qu.: -0.810   1st Qu.: 10.12
Median : 5069   Median : 6.30   Median : 6.4   Median : 3.190   Median : 31.46
Mean   : 9872   Mean   : 14.69   Mean   : 7.9   Mean   : 6.077   Mean   : 34.15
3rd Qu.: 19273  3rd Qu.: 18.52   3rd Qu.: 13.6   3rd Qu.: 16.567   3rd Qu.: 55.73
Max.   : 25800   Max.   : 50.00   Max.   : 24.3   Max.   : 23.500   Max.   : 81.00
SUP
Min.   : 0.020
1st Qu.: 0.335
Median : 0.775
Mean   : 0.938
3rd Qu.: 1.220
Max.   : 2.380
> res.pca <- PCA(donnees_pays, graph=TRUE)
> round(res.pca$eig[1:4,1],3)
comp 1 comp 2 comp 3 comp 4
2.559  1.296  1.000  0.598
> round(res.pca$var$coord[,1:3],2)
      Dim.1 Dim.2 Dim.3
PNB    -0.77 -0.29  0.12
```

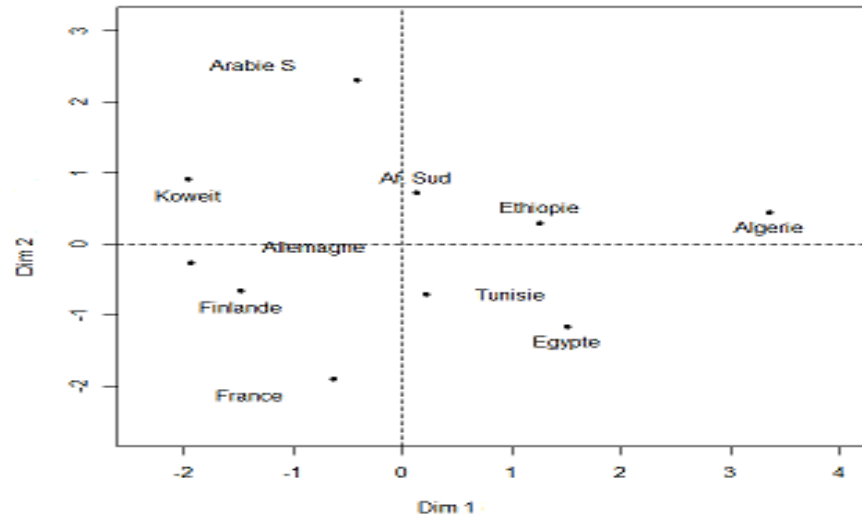


FIGURE 1 – Carte des individus (Axes1-2)

```
INFLAT  0.91  0.10  0.13
CHOM    0.64 -0.47 -0.26
COM     -0.48  0.74  0.19
POPU    0.09 -0.39  0.91
SUP     0.70  0.53  0.21
```

1. Combien de composantes principales devrait-on retenir ?

Dans la suite, on suppose que l'on retient 3 composantes principales.

2. Donner une interprétation des axes retenus.

3. Donner une interprétation de la première carte des individus.

Annexe

```
> library(FactoMineR)
> library(factoextra)
> summary(poison)
  Nausea      Vomiting  Abdominals    Fever      Diarrhae      Potato
Nausea_n:43  Vomit_n:33  Abdo_n:18  Fever_n:20  Diarrhea_n:20  Potato_n: 3
Nausea_y:12  Vomit_y:22  Abdo_y:37  Fever_y:35  Diarrhea_y:35  Potato_y:52
  Fish      Mayo    Courgette    Cheese      Icecream
Fish_n: 1   Mayo_n:10  Courg_n: 5  Cheese_n: 7  Icecream_n: 4
Fish_y:54   Mayo_y:45  Courg_y:50  Cheese_y:48  Icecream_y:51

> res.mca <- MCA (poison, graph = TRUE)
```

```
> round(res.mca$eig[1:5,1],3)
[1] 0.335 0.129 0.107 0.096 0.079
```

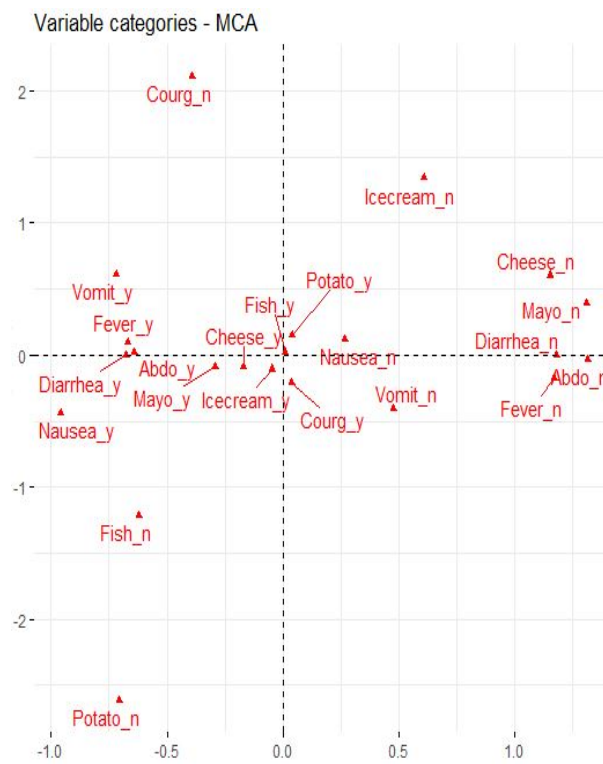


FIGURE 2 – Carte des modalités (Axes1-2)

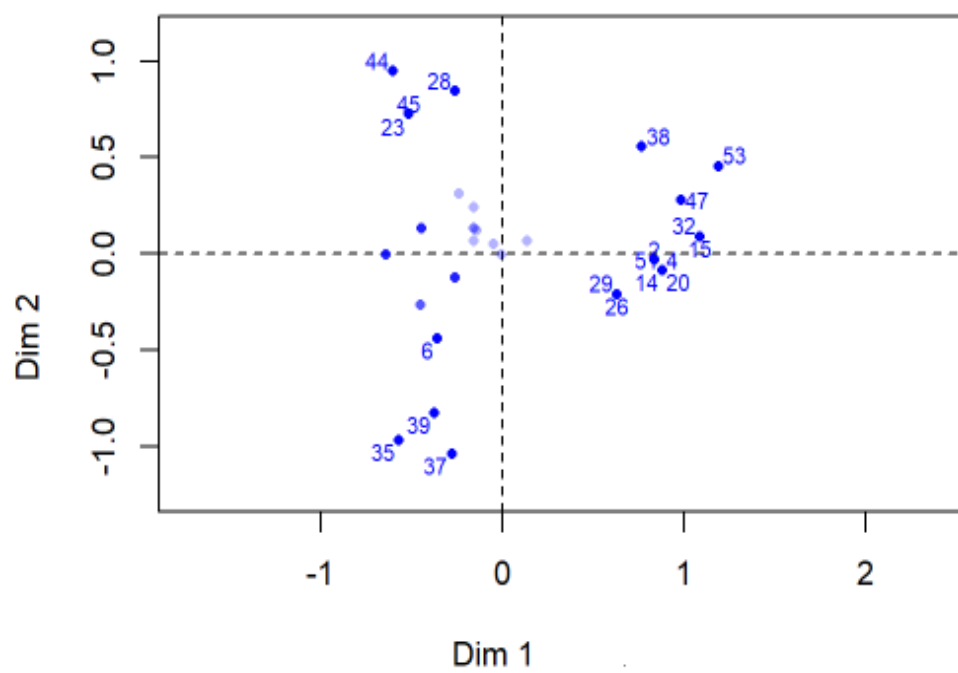


FIGURE 3 – Carte des individus (Axes1-2)