Analyse des Correspondances

Jeu de données naissances

Ce jeu de données contient 9 lignes et 11 colonnes.

### 1. Observation d’individus extrêmes

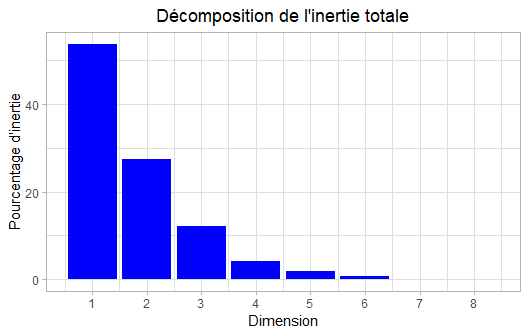
La détection des individus extrêmes ne s’applique pas à l’AFC.

### 2. Distribution de l’inertie

L’inertie des axes factoriels indique d’une part si les variables sont structurées et suggère d’autre part le nombre judicieux de composantes principales à étudier.

Les 2 premiers axes de l’ analyse expriment **81.05%** de l’inertie totale du jeu de données ; cela signifie que 81.05% de la variabilité totale du nuage des lignes (ou des colonnes) est représentée dans ce plan. C’est un pourcentage élevé, et le premier plan représente donc bien la variabilité contenue dans une très large part du jeu de données actif. Cette valeur est supérieure à la valeur référence de **67.44%**, la variabilité expliquée par ce plan est donc significative (cette intertie de référence est le quantile 0.95-quantile de la distribution des pourcentages d’inertie obtenue en simulant 258 jeux de données aléatoires de dimensions comparables sur la base d’une distribution uniforme).

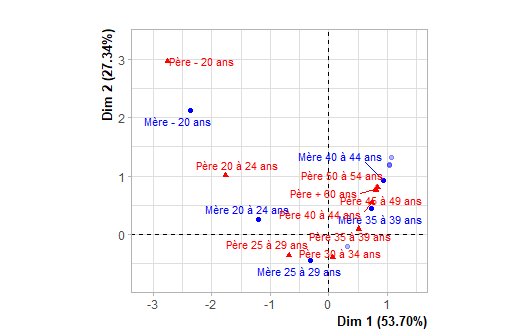
Du fait de ces observations, il n’est probablement pas nécessaire pour l’analyse d’interpréter les dimensions suivantes.



**Figure 2 - Décomposition de l’inertie totale**

Une estimation du nombre pertinent d’axes à interpréter suggère de restreindre l’analyse à la description des 2 premiers axes. Ces composantes révèlent un taux d’inertie supérieur à celle du quantile 0.95-quantile de distributions aléatoires (81.05% contre 67.44%). Cette observation suggère que seuls ces axes sont porteurs d’une véritable information. En conséquence, la description de l’analyse sera restreinte à ces seuls axes.

### 3. Description du plan 1:2



**Figure 3.1 - Graphe superposé (AFC)** *Les lignes libellées sont celles ayant la plus grande contribution à la construction du plan.* *Les colonnes libellées sont celles les mieux représentées sur le plan.*

La **dimension 1** oppose des facteurs tels que *Mère 35 à 39 ans* et *Mère 40 à 44 ans* (à droite du graphe, caractérisés par une coordonnée fortement positive sur l’axe) à des facteurs comme *Mère - 20 ans* (à gauche du graphe, caractérisés par une coordonnée fortement négative sur l’axe).

Le groupe auquel les facteurs *Mère 35 à 39 ans* et *Mère 40 à 44 ans* appartiennent (caractérisés par une coordonnée positive sur l’axe) partage :

* des modalité dont la fréquence ne diffèrent pas significativement de la moyenne.

Le groupe auquel le facteur *Mère - 20 ans* appartient (caractérisés par une coordonnées négative sur l’axe) partage :

* high frequency for the variable *Père.-.20.ans*.

La **dimension 2** oppose des facteurs tels que *Mère - 20 ans* (en haut du graphe, caractérisés par une coordonnées fortement positive sur l’axe) à des facteurs comme *Mère 20 à 24 ans* et *Mère 25 à 29 ans* (en bas du graphe, caractérisés par une coordonnées fortement négative sur l’axe).

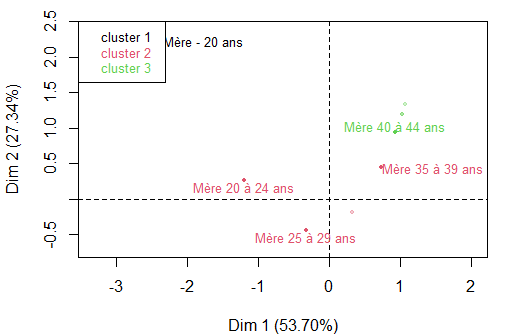
Le groupe auquel le facteur *Mère - 20 ans* appartient (caractérisés par une coordonnée positive sur l’axe) partage :

* high frequency for the variable *Père.-.20.ans*.

Le groupe auquel les facteurs *Mère 20 à 24 ans* et *Mère 25 à 29 ans* appartiennent (caractérisés par une coordonnées négative sur l’axe) partage :

* une forte fréquence des modalités *Père.25.à.29.ans* et *Père.30.à.34.ans* (du plus commun au plus rare).

### 4. Classification



**Figure 4 - Classification Ascendante Hiérachique des lignes.** *La classification réalisée sur les lignes fait apparaître 3 clusters.*

La classe 1 est composée de lignes telles que *Mère - 20 ans*. Ce groupe est caractérisé par :

* une forte fréquence des modalités *Père.20.à.24.ans* et *Père.-.20.ans* (du plus commun au plus rare).
* une faible fréquence des modalités *Père.30.à.34.ans*, *Père.35.à.39.ans*, *Père.40.à.44.ans*, *Père.25.à.29.ans*, *Père.45.à.49.ans*, *Père.50.à.54.ans*, *Père.+.55*, *Père.55.à.59.ans* et *Père.+.60.ans* (du plus rare au plus commun).

La classe 2 est composée de lignes telles que *Mère 20 à 24 ans*, *Mère 25 à 29 ans* et *Mère 35 à 39 ans*. Ce groupe est caractérisé par :

* une forte fréquence des modalités *Père.35.à.39.ans*, *Père.30.à.34.ans* et *Père.25.à.29.ans* (du plus commun au plus rare).
* une faible fréquence des modalités *Père.-.20.ans*, *Père.20.à.24.ans*, *Père.40.à.44.ans*, *Père.45.à.49.ans*, *Père.50.à.54.ans*, *Père.55.à.59.ans*, *Père.+.55* et *Père.+.60.ans* (du plus rare au plus commun).

La classe 3 est composée de lignes telles que *Mère 40 à 44 ans*. Ce groupe est caractérisé par :

* une forte fréquence des modalités *Père.+.55*, *Père.55.à.59.ans*, *Père.50.à.54.ans*, *Père.45.à.49.ans*, *Père.40.à.44.ans* et *Père.+.60.ans* (du plus commun au plus rare).
* une faible fréquence des modalités *Père.20.à.24.ans*, *Père.25.à.29.ans*, *Père.30.à.34.ans*, *Père.35.à.39.ans* et *Père.-.20.ans* (du plus rare au plus commun).

## Annexes

dimdesc(res, axes = 1:2)

$`Dim 1`  
$`Dim 1`$row  
 coord  
Mère - 20 ans -2.3671127  
Mère 20 à 24 ans -1.1949875  
Mère 25 à 29 ans -0.3189009  
Mère 30 à 34 ans 0.3242924  
Mère 35 à 39 ans 0.7327401  
Mère 40 à 44 ans 0.9337979  
Mère 45 à 49 ans 1.0292017  
Mère + 45 1.0314511  
Mère 50 à 60 ans 1.0682757  
  
$`Dim 1`$col  
 coord  
Père - 20 ans -2.7512005  
Père 20 à 24 ans -1.7608104  
Père 25 à 29 ans -0.6763055  
Père 30 à 34 ans 0.0666460  
Père 35 à 39 ans 0.5150954  
Père 40 à 44 ans 0.7446557  
Père + 60 ans 0.8124167  
Père 45 à 49 ans 0.8125884  
Père + 55 0.8155595  
Père 55 à 59 ans 0.8167967  
Père 50 à 54 ans 0.8379469

**Figure 5 - Liste des variables caractéristiques des dimensions de l’analyse.**

res.hcpc$desc.var

$`1`  
 Intern % glob % Intern freq Glob freq p.value  
Père.20.à.24.ans 66.921329754 6.0679458 9723 49704 0.000000e+00  
Père.-.20.ans 14.350609127 0.3838247 2085 3144 0.000000e+00  
Père.+.60.ans 0.006882786 0.1750651 1 1434 3.767746e-10  
Père.55.à.59.ans 0.041296717 0.4447434 6 3643 1.057845e-20  
Père.+.55 0.048179503 0.6198085 7 5077 6.327322e-30  
Père.50.à.54.ans 0.068827862 1.2759973 10 10452 6.499500e-66  
Père.45.à.49.ans 0.199600798 3.7636548 29 30829 1.558485e-195  
Père.25.à.29.ans 12.189414275 22.0933583 1771 180972 5.057568e-211  
Père.40.à.44.ans 0.481795031 10.6619999 70 87335 0.000000e+00  
Père.35.à.39.ans 1.232018721 22.2348509 179 182131 0.000000e+00  
Père.30.à.34.ans 4.460045426 32.2787514 648 264403 0.000000e+00  
 v.test  
Père.20.à.24.ans Inf  
Père.-.20.ans Inf  
Père.+.60.ans -6.263358  
Père.55.à.59.ans -9.330087  
Père.+.55 -11.363877  
Père.50.à.54.ans -17.148052  
Père.45.à.49.ans -29.830787  
Père.25.à.29.ans -31.002054  
Père.40.à.44.ans -Inf  
Père.35.à.39.ans -Inf  
Père.30.à.34.ans -Inf  
  
$`2`  
 Intern % glob % Intern freq Glob freq p.value  
Père.35.à.39.ans 22.9183997 22.2348509 174123 182131 0.000000e+00  
Père.30.à.34.ans 34.3238320 32.2787514 260776 264403 0.000000e+00  
Père.25.à.29.ans 23.4868747 22.0933583 178442 180972 0.000000e+00  
Père.+.60.ans 0.1295159 0.1750651 984 1434 4.501527e-160  
Père.+.55 0.4504101 0.6198085 3422 5077 0.000000e+00  
Père.55.à.59.ans 0.3208942 0.4447434 2438 3643 0.000000e+00  
Père.50.à.54.ans 0.9281976 1.2759973 7052 10452 0.000000e+00  
Père.45.à.49.ans 2.8286862 3.7636548 21491 30829 0.000000e+00  
Père.40.à.44.ans 9.2270899 10.6619999 70103 87335 0.000000e+00  
Père.20.à.24.ans 5.2468437 6.0679458 39863 49704 0.000000e+00  
Père.-.20.ans 0.1392560 0.3838247 1058 3144 0.000000e+00  
 v.test  
Père.35.à.39.ans Inf  
Père.30.à.34.ans Inf  
Père.25.à.29.ans Inf  
Père.+.60.ans -26.95877  
Père.+.55 -Inf  
Père.55.à.59.ans -Inf  
Père.50.à.54.ans -Inf  
Père.45.à.49.ans -Inf  
Père.40.à.44.ans -Inf  
Père.20.à.24.ans -Inf  
Père.-.20.ans -Inf  
  
$`3`  
 Intern % glob % Intern freq Glob freq p.value  
Père.+.55 3.675044043 0.6198085 1648 5077 0.000000e+00  
Père.55.à.59.ans 2.673772941 0.4447434 1199 3643 0.000000e+00  
Père.50.à.54.ans 7.559708316 1.2759973 3390 10452 0.000000e+00  
Père.45.à.49.ans 20.759092835 3.7636548 9309 30829 0.000000e+00  
Père.40.à.44.ans 38.271302098 10.6619999 17162 87335 0.000000e+00  
Père.+.60.ans 1.001271101 0.1750651 449 1434 3.484979e-206  
Père.-.20.ans 0.002230002 0.3838247 1 3144 3.461245e-75  
Père.35.à.39.ans 17.458689205 22.2348509 7829 182131 1.652536e-145  
Père.30.à.34.ans 6.643177307 32.2787514 2979 264403 0.000000e+00  
Père.25.à.29.ans 1.692571862 22.0933583 759 180972 0.000000e+00  
Père.20.à.24.ans 0.263140289 6.0679458 118 49704 0.000000e+00  
 v.test  
Père.+.55 Inf  
Père.55.à.59.ans Inf  
Père.50.à.54.ans Inf  
Père.45.à.49.ans Inf  
Père.40.à.44.ans Inf  
Père.+.60.ans 30.64098  
Père.-.20.ans -18.34743  
Père.35.à.39.ans -25.68654  
Père.30.à.34.ans -Inf  
Père.25.à.29.ans -Inf  
Père.20.à.24.ans -Inf  
  
attr(,"class")  
[1] "descfreq" "list"

**Figure 6 - Liste des variables caractéristiques des clusters de la classification.**