

# **L'Analyse Factorielle multiple avec les macros SAS**

## **%AFMULT et %PLOTAFM**

*Brigitte GELEIN (Ensa), Olivier SAUTORY (Insee)*

version août 2010  
SAS 9.1

# Sommaire

Introduction .....	3
1. Aspects théoriques de l'AFM .....	3
1.1 Principe de la méthode.....	3
1.2 Notations.....	4
1.3 Étape préliminaire : les ACP séparées .....	5
1.4 Les individus .....	5
1.5 Les variables.....	6
1.6 Les groupes de variables .....	8
1.7 Variables qualitatives et tableaux mixtes .....	8
1.8 Éléments supplémentaires .....	9
2. La macro AFMULT.....	10
2.1. Les paramètres de la macro AFMULT.....	10
2.1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse : .	10
2.1.2. Paramètres spécifiant les sorties éditées : .....	13
2.1.3. Exemple de programmation : AFM temporelle.....	14
2.2. La boîte de dialogue de la macro AFMULT .....	16
2.2.1. Navigation et fonctionnalités de la boîte de dialogue : .....	16
2.2.2. Le paramétrage : les quatre fenêtres de la boîte de dialogue.....	18
2.2.3. Exemple de saisie : AFM temporelle .....	22
2.3. Les sorties éditées par la macro AFMULT .....	25
2.3.1. Liste des sorties.....	25
2.3.2. Explication de ces sorties .....	26
2.3.3. Les sorties de l'exemple AFM temporelle .....	31
3. La macro PLOTAFM.....	43
3.1. Les paramètres de la macro PLOTAFM .....	43
3.1.1. Paramètres généraux .....	43
3.1.2. Paramètres spécifiant les graphiques.....	44
3.1.3. Exemple de programmation : AFM temporelle.....	48
3.2. La boîte de dialogue de la macro PLOTAFM.....	49
3.3. Exemple de graphiques obtenus avec PLOTAFM .....	51
Bibliographie.....	54

# Introduction

L'analyse factorielle multiple (AFM) permet de traiter conjointement plusieurs tableaux de données décrivant un même ensemble d'individus. Chaque tableau se rapporte à un thème particulier (analyse thématique) ou à une date (analyse des données évolutives).

Cette méthode d'analyse des données a été proposée par Brigitte Escofier et Jérôme Pagès dans les années 80 [1].

L'AFM peut être mise en œuvre avec le logiciel SAS à l'aide des macros :

- **%AFMULT** qui permet de réaliser des AFM sur des variables quantitatives et qualitatives ;
- **%PLOTAFM** qui permet d'obtenir des graphiques illustrant les résultats de l'analyse.

## 1. Aspects théoriques de l'AFM

### 1.1 Principe de la méthode

L'analyse factorielle multiple (AFM) est une méthode permettant d'analyser des tableaux dans lesquels des individus sont décrits par plusieurs groupes de variables, quantitatives ou qualitatives. Ces groupes (un groupe correspond à un tableau) peuvent être constitués :

- de variables identiques mesurées sur les mêmes individus à différentes dates (AFM "temporelle") ;
- de variables regroupées par "thème" : chaque groupe correspond alors à un thème (AFM "thématique").

De plus, l'étude est enrichie (par rapport à une analyse traditionnelle) par la possibilité de comparer les groupes de variables eux-mêmes.

La spécificité de l'AFM par rapport aux méthodes classiques (analyse en composantes principales ou analyse des correspondances multiples) est d'équilibrer l'influence des différents groupes, grâce à des pondérations attribuées à ces groupes. Ce point est important dans la mesure où un groupe de variables peut jouer un rôle prépondérant dans une analyse traditionnelle, pour deux raisons :

1. le nombre de variables du groupe : plus ce nombre est élevé, plus l'influence du groupe sur les résultats risque d'être forte ;
2. la structure du groupe : plus ses variables sont liées, plus il va influencer la détermination des premiers axes principaux de l'analyse globale.

L'AFM offre des représentations factorielles, construites à partir d'un système d'axes, où les individus sont décrits selon deux points de vue :

- globalement, en prenant en compte l'ensemble des variables – on parlera alors **d'individus moyens ou compromis**;
- au travers de chacun des thèmes pris séparément – on parlera alors **d'individus partiels**.

Chaque individu est donc représenté, dans un plan factoriel, par  $1 + J$  points, où  $J$  est le nombre de groupes. On peut donc savoir si un individu se distingue des autres globalement ou bien sur un thème particulier.

De plus, l'étude est enrichie (par rapport à une analyse traditionnelle) par la possibilité de comparer les groupes de variables eux-mêmes.

L'AFM peut traiter en même temps des groupes de variables qualitatives et des groupes de variables quantitatives. Par contre, au sein d'un même groupe, les variables doivent être de même nature.

L'AFM se comporte comme une ACP pour les variables quantitatives et comme une ACM pour les variables qualitatives.

L'AFM équilibre l'influence des groupes de variables en donnant à chaque variable un poids : ces poids sont identiques au sein de chaque groupe, afin de conserver la structure interne de chacun des groupes.

Le poids donné par l'AFM à chacune des variables d'un groupe est égal à l'inverse de l'inertie du premier axe principal de ce groupe (obtenue par des ACP ou des ACM réalisées sur les différents groupes de façon séparée). On a donc une normalisation des nuages représentant les individus pour chaque groupe.

Cette pondération rend possible la présence simultanée de variables continues et nominales parmi les éléments actifs.

Les éléments théoriques qui suivent portent, dans un premier temps, sur l'application de la méthode à des variables quantitatives. Sont ensuite traités les variables qualitatives et les éléments supplémentaires.

## **1.2 Notations**

Les données portent sur  $I$  individus décrits par  $K$  variables formant  $J$  groupes. À chaque groupe correspond un tableau  $X_j$ ,  $j = 1 \dots J$ . Mais l'ensemble des  $J$  tableaux initiaux peut être traité comme un unique tableau structuré en sous-tableaux. En effet, les  $J$  tableaux portent sur les mêmes individus et peuvent, par juxtaposition, former un tableau croisant les  $I$  individus et les  $K$  variables.

On note  $X$  le tableau complet,  $I$  l'ensemble des individus,  $K$  l'ensemble des variables,  $J$  l'ensemble des sous-tableaux,  $K_j$  l'ensemble des variables du groupe  $j$ . Les symboles  $I$ ,  $J$ ,  $K$  ou  $K_j$  désignent donc à la fois l'ensemble et son cardinal.

On considère les trois espaces suivants :

- $\mathbb{R}^K$  dans lequel est situé le nuage des individus
- $\mathbb{R}^I$  dans lequel est situé le nuage des variables
- $\mathbb{R}^{I^2}$  dans lequel est situé le nuage des groupes de variables.

On associe à chaque individu  $i$  un poids  $p_i$ , avec  $\sum_i p_i = 1$ , et on note  $D$  la matrice diagonale (I,I) contenant ces poids.

On associe à chaque variable  $v_k$  appartenant au groupe  $j$  un poids  $m_k^j$ , et on note  $M_j$  la matrice diagonale des poids des variables du groupe  $j$ . Les variables numériques ont souvent un poids égal à 1.

### 1.3 Étape préliminaire : les ACP séparées

On commence par effectuer  $J$  analyses en composantes principales (ACP) sur chacun de tableaux définis par les groupes de variables.

Ces ACP reviennent à déterminer les valeurs et vecteurs propres des matrices  $X_j M_j X_j' D = W_j D$ . On note  $\lambda_j^i$  la première valeur propre de l'ACP du groupe  $j$ . L'inverse de cette valeur propre est le poids donné au groupe  $j$  dans l'analyse : la variable  $v_k$  appartenant au groupe  $j$  est donc munie du poids  $m_k = m_k^j / \lambda_j^i$ .

**Cette pondération permet d'équilibrer le rôle des groupes.**

### 1.4 Les individus

Dans l'espace  $\mathbb{R}^K$ , on cherche :

- une représentation du nuage  $N_j$  des individus caractérisés par les  $K$  variables (= **individus compromis ou moyen**),
- une représentation superposée des  $J$  nuages  $N_j^i$  ( $j = 1 \dots J$ ) d'individus caractérisés chacun par un groupe de variables (= **individus partiels**).

$\mathbb{R}^K$  peut se décomposer en somme directe de  $J$  sous-espaces orthogonaux deux à deux isomorphes aux espaces  $\mathbb{R}^{K^j}$  engendrés chacun par un des  $J$  groupes de variables :  $\mathbb{R}^K = \oplus \mathbb{R}^{K^j}$

Les coordonnées des points de  $N_j^i$  sont contenues dans le tableau  $X_j$ . On définit le tableau  $\tilde{X}_j$ , de dimensions  $(I,K)$ , en complétant  $X_j$  par des zéros pour les variables n'appartenant pas au groupe  $j$  :

$$\tilde{X}_j = \begin{array}{|c|c|c|} \hline & & \\ \hline 0 & X_j & 0 \\ \hline & & \\ \hline \end{array}$$

Un individu  $i$ , décrit par les variables du groupe  $j$ , est alors représenté par un point de  $\mathbb{R}^K$ , noté  $i^j$ . Dans  $\mathbb{R}^K$ , on note  $N_j^*$  le nuage des centres de gravité,

notés  $i^*$ , des J points  $i^j$ , représentant le même individu i dans les  $N_i^j$  (J individus partiels).  $N_i^*$  se déduit de  $N_i$  par une homothétie de rapport  $1/J$  : il constitue un nuage moyen pour les  $N_i^j$ .

**On effectue une ACP du tableau X**, l'objectif étant de projeter le nuage  $N_i$  (et donc  $N_i^*$ ) sur un sous-espace de faible dimension tel que le nuage projeté ressemble le plus possible au nuage initial  $N_i$  (respectivement  $N_i^*$ ). Dans cette ACP, chaque variable  $v_k$  est munie du poids  $m_k$  défini précédemment.

Les J nuages  $N_i^j$  étant également tous situés dans  $\mathbb{R}^K$ , il est possible de les projeter sur un même sous-espace de  $\mathbb{R}^K$ . Ce sous-espace doit permettre de comparer les positions d'un même individu dans les différents nuages. La projection sur ce sous-espace doit permettre d'obtenir :

- une bonne représentation de chaque  $N_i^j$
- une ressemblance entre les nuages  $N_i^j$  : éviter que les symétries, rotations ou homothéties ne masquent les ressemblances. Il ne serait en effet pas possible de comparer les positions d'un même individu dans les différents nuages si ces représentations étaient très différentes.

On montre que la représentation optimale est celle obtenue par l'ACP pondérée du tableau X. La représentation superposée des J nuages définis par chaque groupe de variables s'obtient en traitant en éléments supplémentaires de cette ACP les individus partiels. On projette donc les nuages  $N_i^j$  sur les axes factoriels de  $N_i$ .

## 1.5 Les variables

Dans l'espace  $\mathbb{R}^I$ , on veut représenter :

- les K variables initiales,
- les composantes principales obtenues par des ACP séparées de chacun des groupes.

La représentation des variables s'obtient par l'ACP du tableau X. Duale de l'image de  $N_i$  obtenue dans  $\mathbb{R}^K$ , elle offre une aide à l'interprétation de la représentation du nuage des individus. Elle constitue également une représentation optimale des corrélations entre variables. Elle permet donc une comparaison fine des groupes de variables.

La pondération des groupes par  $1/\lambda_i^j$  est telle que :

- aucun groupe ne détermine à lui seul le premier axe (sauf situation de symétrie)
- un groupe contribue à la détermination de nombreux axes s'il est composé de nombreuses variables indépendantes (il est de grande "dimensionnalité").

On obtient la représentation des composantes principales de chaque groupe en les introduisant en éléments supplémentaires dans l'ACP de  $X$ . On peut ainsi repérer des facteurs communs à plusieurs groupes de variables.

## 1.6 Les groupes de variables

Dans  $\mathbb{R}^{I^2}$ , on cherche à comparer globalement les groupes. Pour les représenter, on va construire un nuage de J points, noté  $N_j$ . Un groupe j est représenté par la matrice  $(I, I) X_j M_j X_j' D = W_j D$ , ou plus précisément, en faisant apparaître explicitement la pondération de l'AFM,  $W_j D / \lambda_j^j$ . Cette matrice de  $I^2$  scalaires est considérée comme un élément de l'espace vectoriel  $\mathbb{R}^{I^2}$ , muni du produit scalaire de Hilbert-Schmidt  $\langle A, B \rangle = Tr(AB)$ .

On obtient ainsi une mesure de liaison entre les groupes j et  $\ell$  :

$$L(K_j, K_\ell) = \left\langle \frac{W_j D}{\lambda_j^j}, \frac{W_\ell D}{\lambda_\ell^\ell} \right\rangle$$

Pour comparer les groupes ainsi représentés dans  $\mathbb{R}^{I^2}$ , on décrit les proximités entre les  $W_j D$  en les projetant sur un espace de faible dimension de  $\mathbb{R}^{I^2}$ . On cherche, dans  $\mathbb{R}^{I^2}$ , un repère orthonormé qui ajuste au mieux les  $W_j D$  et dont les axes sont des éléments symétriques de rang 1. On montre que ces éléments sont de la forme  $z_s z_s' D$ , où les  $z_s$  sont les composantes principales issues de l'ACP de X. Ils maximisent la somme des projections des  $W_j D$  sur les axes. Cette somme est égale à l'inertie dans  $\mathbb{R}^I$  des K variables projetées sur  $z_s$  :

$$\sum_j \langle W_j D, z_s z_s' D \rangle.$$

La suite des éléments recherchés est associée à celle des composantes principales  $z_s$ . L'orthonormalité des  $z_s$  dans  $\mathbb{R}^I$  est équivalente à celle des  $z_s z_s' D$ , dans  $\mathbb{R}^{I^2}$ . La coordonnée de  $W_j D$  sur  $z_s z_s' D$  peut aussi s'interpréter comme la contribution du groupe j à l'inertie de  $z_s$ .

## 1.7 Variables qualitatives et tableaux mixtes

L'AFM peut s'appliquer aux tableaux mixtes, dans lesquels certains groupes sont constitués de variables qualitatives, grâce à l'équivalence suivante entre ACM et ACP : les résultats d'une ACM peuvent être obtenus en réalisant une ACP normée des indicatrices relatives aux modalités des variables, à condition d'affecter à ces indicatrices des poids adéquats.

Pour réaliser une AFM sur des tableaux mixtes, on commence donc par construire les indicatrices associées aux modalités des variables qualitatives, qui seront ensuite correctement pondérées dans l'analyse.

## **1.8 Éléments supplémentaires**

Les individus, affectés d'un poids nul, peuvent intervenir en tant qu'éléments supplémentaires. Sont alors calculées les projections de leurs points représentatifs dans  $N_i^*$  et dans les  $N_i^j$ .

Un groupe de variables  $j'$  peut également être traité en élément supplémentaire. On effectue alors :  
une normalisation du nuage associé semblable à celle des autres groupes, la projection des composantes principales du groupe sur les axes de l'analyse, et la projection de  $W_{j'} D$ .

## 2. La macro AFMULT

La macro SAS **AFMULT** réalise une **Analyse Factorielle Multiple** : analyse conjointe de plusieurs tableaux.

Les résultats de l'analyse sont stockés dans une table SAS, et peuvent être ainsi réutilisés par une autre procédure, ou par la macro **PLOTAFM**.

En ce qui concerne les variables qualitatives, la macro **AFMULT n'utilise que les tableaux disjonctifs complets**. Avant l'appel de la macro, l'utilisateur devra donc construire pour chaque variable qualitative les variables indicatrices associées aux modalités. La construction d'un bloc d'indicatrices à partir d'une colonne contenant les modalités d'une même variable peut se faire avec la syntaxe des variables logiques ou bien avec la proc transreg (voir 2.1.3 *Exemple*).

On dispose de deux possibilités pour utiliser la macro AFMULT :

- soit renseigner les paramètres de la macro dans la fenêtre éditeur de programmes, en utilisant les règles habituelles d'écriture des paramètres d'une macro (voir 2.1 *les paramètres de la macro*) :  
**%AFMULT(datact=...,IDENT=... , ....)**
- soit utiliser la boîte de dialogue, qui permet de spécifier les valeurs prises par les paramètres de la macro dans des fenêtres de saisie apparaissant à l'écran (voir 2.2 *la boîte de dialogue*) :  
**%AFMULT(fen=o)** pour lancer la macro et ses boîtes de dialogue

Quel que le soit mode d'utilisation retenu, il faut soumettre la ligne suivante pour indiquer à SAS la librairie (nommée par exemple « compil ») dans laquelle se situe votre catalogue de macros.

**options sasstore=compil mstored;**

### 2.1. Les paramètres de la macro AFMULT

#### 2.1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse :

**DATAACT = nom de table SAS**

nom de la table SAS contenant les individus actifs (**obligatoire**).

**DATASUP = nom de table SAS**

nom de la table SAS contenant les individus supplémentaires. On peut déclarer les individus supplémentaires d'une autre façon (voir paramètre poids) et dans ce cas ne pas utiliser le paramètre DATASUP.

**IDENT = variable**

variable servant à identifier les individus dans les éditions et récupérée dans la table en sortie (voir paramètre DATASOR)

**Ce paramètre est obligatoire, et la variable identifiant doit être de type "caractère"<sup>1</sup>.**

Cette variable doit figurer dans la (ou les deux) table(s) en entrée (paramètres DATAACT et DATASUP).

**POIDS = variable**

nom d'une variable numérique servant à pondérer les individus (les valeurs de cette variable peuvent ne pas être des entiers). Cette variable doit figurer dans la table en entrée et spécifiée dans le paramètre DATAACT. On peut mettre un poids négatif aux individus qu'on souhaite traiter en éléments supplémentaires sans créer une table à part : on laisse alors le paramètre DATASUP non renseigné et tous les individus actifs et supplémentaires figurent dans DATAACT.

**Par défaut : tous les individus ont le même poids.**

**ACTQUAL = liste des indicatrices des modalités des variables qualitatives actives**

Dans la liste, les variables indicatrices doivent être regroupées par variable. Elles doivent être séparées par des blancs ou par des tirets, et séparées par des / si on souhaite définir plusieurs groupes.

Exemple :

Supposons que l'on dispose de 3 variables qualitatives, A, B et C, ayant respectivement 2, 2 et 3 modalités. Les variables indicatrices associées, sont nommées par exemple A1 et A2 pour la variable A, B1 et B2 pour la variable B, C1, C2 et C3 pour la variable C.

**actqual= A1 A2 B1 B2 / C1 C2 C3;**

ou bien

**actqual =A1 -- B2 /C1-C3;** si on souhaite utiliser la syntaxe compactée (attention dans le cas du double tiret à vérifier l'ordre des indicatrices dans la table)

Ici on a déclaré deux groupes de variables qualitatives. Le premier groupe contient les indicatrices **A1 A2 B1 B2** correspondant aux variables A et B et le second **C1 C2 C3** correspondant à la seule variable C.

**MODACT = liste de nombres de modalités actives**

pour chacune des variables qualitatives actives, on indique le nombre de modalités de la variable ; l'ordre dans lequel sont donnés ces nombres doit être le même que l'ordre dans lequel ont été spécifiées les indicatrices des variables dans le paramètre ACTQUAL.

---

<sup>1</sup> Pour créer une variable caractère Y à partir d'une variable numérique X, (prenant par exemple des valeurs entières comprises entre 0 et 999), il suffit d'écrire une instruction de la forme : **Y=PUT(X,3.).**

Dans la liste, les nombres de modalités doivent être séparés par des blancs, et séparés par des / pour distinguer les groupes.

Exemple : **modact= 2 2 / 3**

### **VARQUALACT = liste de libellés des variables qualitatives actives**

liste des libellés des variables qualitatives actives, séparés par des blancs (les blancs ne sont pas autorisés dans les libellés), et séparés par des / pour distinguer les groupes. Ces libellés seront utilisés dans les tableaux édités par la macro.

Note : il s'agit bien de **libellés**, qui peuvent donc être distincts des noms SAS des variables qualitatives initiales. Néanmoins **leur longueur est limitée à 32 caractères**.

Exemple : **varqualact= varA varB / varC**

Les paramètres suivants, relatifs aux variables qualitatives supplémentaires, présentent des règles d'écriture similaires à celles des paramètres relatifs aux variables qualitatives actives.

### **SUPQUAL = liste des indicatrices des modalités des variables qualitatives supplémentaires**

### **MODSUP = liste de nombres de modalités supplémentaires**

### **VARQUALSUP = liste de libellés des variables qualitatives supplémentaires**

### **ACTNUM = liste de variables actives numériques**

Dans la liste, les variables doivent être séparées par des blancs ou bien par des tirets, et séparées par des / si on souhaite définir plusieurs groupes.

**actnum= XX YY ZZ / JJ1 JJ2 JJ3;**

ou bien

**actnum= XX --ZZ / JJ1-JJ3;** si on souhaite utiliser la syntaxe compactée (attention dans le cas du double tiret à vérifier l'ordre des variables dans la table)

### **SUPNUM = liste de variables supplémentaires numériques**

Dans la liste, les variables doivent être séparées par des blancs ou bien par des tirets, et séparées par des / si on souhaite définir plusieurs groupes.

**supnum= XX YY ZZ / JJ1 JJ2 JJ3;**

ou bien

**supnum= XX --ZZ / JJ1-JJ3;** si on souhaite utiliser la syntaxe compactée (attention dans le cas du double tiret à vérifier l'ordre des variables dans la table)

### **2.1.2. Paramètres spécifiant les sorties éditées :**

#### **NBAXE = xx**

Nombre d'axes qui seront édités dans les aides à l'interprétation pour les variables, les individus, les centres de gravité, les groupes.  
Par défaut, NBAXE=2.

#### **NBAXEPART = xx**

Nombre d'axes qui seront édités dans les aides à l'interprétation pour les axes partiels.

#### **AIDEINDACT = O ou N**

Si AIDEINDACT vaut O (=oui), il y a édition complète des aides à l'interprétation pour tous les individus actifs.  
Par défaut : AIDEINDACT = N.

#### **NINDCONTR = xx**

Si on souhaite n'éditer les aides à l'interprétation que pour les xx individus actifs ayant la plus forte contribution (pour chaque axe).

#### **AIDEINDSUP = O ou N**

Si AIDEINDSUP vaut O (=oui), il y a édition complète des aides à l'interprétation pour tous les individus supplémentaires.  
Par défaut : AIDEINDSUP = N.

#### **CORRVAR = O ou N**

Si CORRVAR vaut O (=oui), il y a édition de la matrice des corrélations entre les variables actives (si l'on souhaite obtenir les corrélations avec les variables supplémentaires, il faut utiliser une PROC CORR).  
Par défaut : CORRVAR = N.

#### **CORRFAC = O ou N**

Si CORRFAC vaut O (=oui), il y a édition de la matrice des corrélations entre les facteurs partiels.  
Par défaut : CORRFAC = N.

## **LINESIZE = xx**

Valeur du paramètre *linesize*<sup>2</sup> de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LINESIZE = 80

## **DATASOR = nom de table SAS**

nom de la table SAS en sortie contenant les résultats de l'analyse (coordonnées, contributions ...etc).

### **2.1.3. Exemple de programmation : AFM temporelle**

La stratégie de l'Union Européenne en faveur du développement durable prévoit qu'Eurostat mette à jour tous les deux ans un rapport de suivi.

Une liste d'indicateurs a donc été développée par la Commission européenne, en coopération avec les États membres, au travers du groupe de travail d'Eurostat sur les indicateurs de développement durable (IDD).

Ces indicateurs portent sur différents domaines :

- développement économique
- consommation et production durable
- inclusion sociale
- changements démographiques
- santé publique
- **changement climatique et énergie**
- transport durable
- ressources naturelles
- partenariat global
- bonne gouvernance

On retient ici comme individus statistiques 15 pays de l'UE et comme variables actives (XX=année) celles relatives au changement climatique et à l'énergie:

- Émissions de gaz à effet de serre (indice base 100 année de référence =1990) = *serreXX*
- Proportion des énergies renouvelables dans la consommation brute d'énergie intérieure = *renouXX*
- Dépendance énergétique = *depenXX*
- Consommation brute d'énergie intérieure = *consoXX*
- Électricité provenant des énergies renouvelables = *elecXX*

*Les variables serreXX, renouXX, depenXX, consoXX, elecXX* constituent les variables actives de l'année XX.

On a trois tableaux actifs: pour 1996, 2001 et 2006.

---

<sup>2</sup> i.e. la longueur des lignes qui est utilisée dans les fenêtres de l'*output* et de la *log*

pays	serre96	...	depen96	serre01	...	depen01	serre06	...	depen06
Allemagne	90.4	...	59.2	84.1	...	61.0	81.5	...	61.3
Autriche	105.9	...	69.6	107.9	...	65.0	115.2	...	72.9
Belgique	106.0	...	79.3	99.6	...	78.2	94.0	...	77.9
...	...	...	...	...	...	...	...	...	...
Suède	107.2	...	40.4	95.6	...	36.6	91.1	...	37.4

On utilise un tableau supplémentaire contenant une seule variable qualitative à 3 modalités : date d'entrée dans l'UE antérieure à 1960 (av1960), date d'entrée comprise entre 1960 et 1980, date d'entrée dans l'UE postérieure à 1980.

Les trois indicatrices de cette variable qualitative sont construites grâce à la syntaxe de type variable logique (si la condition entre parenthèse est vraie alors on a la valeur 1, sinon on a la valeur 0).

```
data climat;
set lib.climat;
av1960=(adhesion<1960);
av1980=(adhesion>=1960 and adhesion<1980);
ap1980=(adhesion>1980);
run;
```

Pour appeler la macro %AFMULT sans les boîtes de dialogue (sans écran de saisie) par exemple :

```
options sasstore=compil mstored;

%afmult(dataact=climat,ident=pays,
actnum=serre96--elec96/serre01 renou01 depen01 conso01 elec01/serre06--
elec06,
supqual=av1960--ap1980,
modsup=3,
varqualsup=ancienneté,
aideindact=o,
nbaxepart=2,
datasor=_datasor_,
linesize=100,
corrvar=o,
corrfac=o,
nbaxe=2);
```

On remarquera la présence d'un caractère accentué dans la valeur prise par le paramètre varqualsup : il s'agit bien d'un libellé et non pas d'un nom de variable SAS (qui n'autorise pas de caractère accentué).

Il est possible d'encadrer l'appel de la macro avec les instructions relatives à l'ODS (Output Delivery System) si on souhaite obtenir les sorties éditées dans des fichiers de type rtf ou html pour une plus belle présentation que celle fournie dans les listings classiques de la fenêtre output.

## 2.2. La boîte de dialogue de la macro AFMULT

La macro SAS AFMULT est dotée d'une boîte de dialogue qui facilite la saisie des différents paramètres de la macro.

L'ouverture de la boîte de dialogue s'effectue en donnant la valeur "o" (pour "oui") au paramètre fen (pour "fenêtre") de la macro %AFMULT. Il est alors inutile de renseigner les autres paramètres de la macro.

Si la librairie qui contient le catalogue de macros est appelée compil par exemple

```
options sasstore=compil mstored;
```

```
%afmult(fen=o)
```

Cette boîte de dialogue comporte un enchaînement de quatre fenêtres :

- 1 / 4 : appel des paramètres d'une analyse déjà réalisée et sauvegarde de paramètres
- 2 / 4 : saisie des paramètres concernant les variables actives
- 3 / 4 : saisie des paramètres concernant les variables supplémentaires
- 4 / 4 : paramétrage des éditions de la macro

La boîte de dialogue est une interface qui facilite la saisie des paramètres : chaque champ de saisie correspond donc à un des paramètres décrit précédemment. Le lecteur pourra aisément se reporter à la description des différents paramètres dans le paragraphe précédent s'il souhaite davantage de précisions sur celui-ci.

### 2.2.1. Navigation et fonctionnalités de la boîte de dialogue :

#### Navigation :

- **les déplacements au sein d'une fenêtre** peuvent s'effectuer de deux manières différentes :
  - soit "à la souris" : on clique sur le champ que l'on veut renseigner,
  - soit automatiquement : après avoir rempli un champ et tapé sur Entrée, on est directement renvoyé au champ suivant.
- **les déplacements entre deux fenêtres ou la sortie de la boîte de dialogue :**  
On place le curseur sur le petit rectangle bleu en bas de page et on suit les instructions relatives à ces déplacements (spécifiées en bas de chaque fenêtre).

### **IL NE FAUT PAS :**

- utiliser les flèches de "saut de page" : on descend  *systématiquement* d'une fenêtre pour chaque frappe (jusqu'à la dernière),
- fermer la boîte de dialogue "à la souris" en cliquant sur la croix en haut à droite de la fenêtre.

### Fonctionnalités :

- On peut renseigner les champs à partir de l'endroit où l'on a positionné le curseur, même si l'on n'est pas au début du champ.
- On ne peut sélectionner des caractères déjà écrits dans un champ et les supprimer en bloc : il faut corriger le texte déjà écrit caractère par caractère.
- Lorsqu'un champ occupe plusieurs lignes, il faut saisir les paramètres de telle sorte qu'il ne doit jamais y avoir de chevauchement d'écriture entre deux lignes : SAS considèrera cela comme une erreur.  
Par exemple lorsque l'on saisit la liste des variables qualitatives actives et que l'on arrive au bout de la première ligne, il faut écrire :

Variable1 Variable2 Variable3 Variable4 Variable5 Variable6 Variable7 Variable8 Variable9
---

ET IL NE FAUT PAS ÉCRIRE :

Variable1 Variable2 Variable3 Variable4 Variable5 Variable6 Variable7 Variable8 Var iable9 -
--

**MÊME SI C'EST CE QUE SAS PROPOSE PAR DÉFAUT.**

## 2.2.2. Le paramétrage : les quatre fenêtres de la boîte de dialogue

### La première fenêtre : appel des paramètres d'une analyse passée et sauvegarde de paramètres

*Il existe en fait deux versions de la première fenêtre.*

La version suivante apparaît lorsque l'on appelle la macro AFMULT pour la première fois dans la session SAS en cours :

```
Command ===>

***  ANALYSE FACTORIELLE MULTIPLE (fenêtre 1/4)  ***

Si vous voulez utiliser des paramètres stockés dans une table existant déjà,
indiquez son nom complet ci-dessous (sous la forme librairie.table) :

_____

Les valeurs des paramètres seront sauvegardées dans la table _PARAM_.
Si vous voulez aussi les sauvegarder dans une autre table,
indiquez son nom complet ci-dessous :

_____

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - valider : tapez sur Entrée
                  - mettre fin au dialogue : tapez f
```

L'autre version apparaît pour tout appel ultérieur de la macro au cours de la même session SAS :

```
Command ===>

***  ANALYSE FACTORIELLE MULTIPLE (fenêtre 1/4)  ***

Par défaut, les paramètres qui vont être affichés sont ceux qui ont été
stockés dans la table _PARAM_ déjà créée au cours de cette session SAS.

Vous pouvez faire appel à des paramètres stockés dans une autre table.
Dans ce cas, modifiez le nom _PARAM_ ci-dessous, en indiquant le nom complet
de la table (sous la forme librairie.table) :

PARAM_____

Vous pouvez aussi réinitialiser tous les paramètres.
Dans ce cas, mettez à blanc la ligne ci-dessus.

Les valeurs des paramètres seront sauvegardées dans la table _PARAM_.
Si vous voulez aussi les sauvegarder dans une autre table,
indiquez son nom complet ci-dessous :
fcda.param_____

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - valider : tapez sur Entrée
                  - mettre fin au dialogue : tapez f
```

## La deuxième fenêtre : saisie des paramètres concernant les variables actives

The screenshot shows a terminal window titled 'AFMULT\_SAISIE1' with the following content:

```
Command ==>          ***  ANALYSE FACTORIELLE MULTIPLE (fenêtre 2/4)  ***

Table contenant les individus actifs : _____
Table contenant les individus supplémentaires : _____
Variable identifiant les individus dans ces tables : _____
Variable de pondération des individus actifs : _____

- LISTES DES VARIABLES NUMÉRIQUES ET DES MODALITÉS DES VARIABLES QUALITATIVES ACTIVES -

Les variables ou les modalités doivent être classées par groupes,
et les groupes doivent être séparés par des /
Exemple : X Y Z / A B / I J K L

Variables numériques actives
_____

Modalités des variables qualitatives actives, classées par variable
Exemple : A1 A2 A3 B1 B2 / C1 C2
A1, A2 et A3 sont les 3 modalités d'une 1ère variable,
B1 et B2 les 2 modalités d'une 2ème variable, etc.
_____

Nombre de modalités des variables qualitatives
Exemple : 3 2 / 2
_____

Libellés (sans espaces) des variables qualitatives dans les éditions
Exemple : varA varB / varC
_____

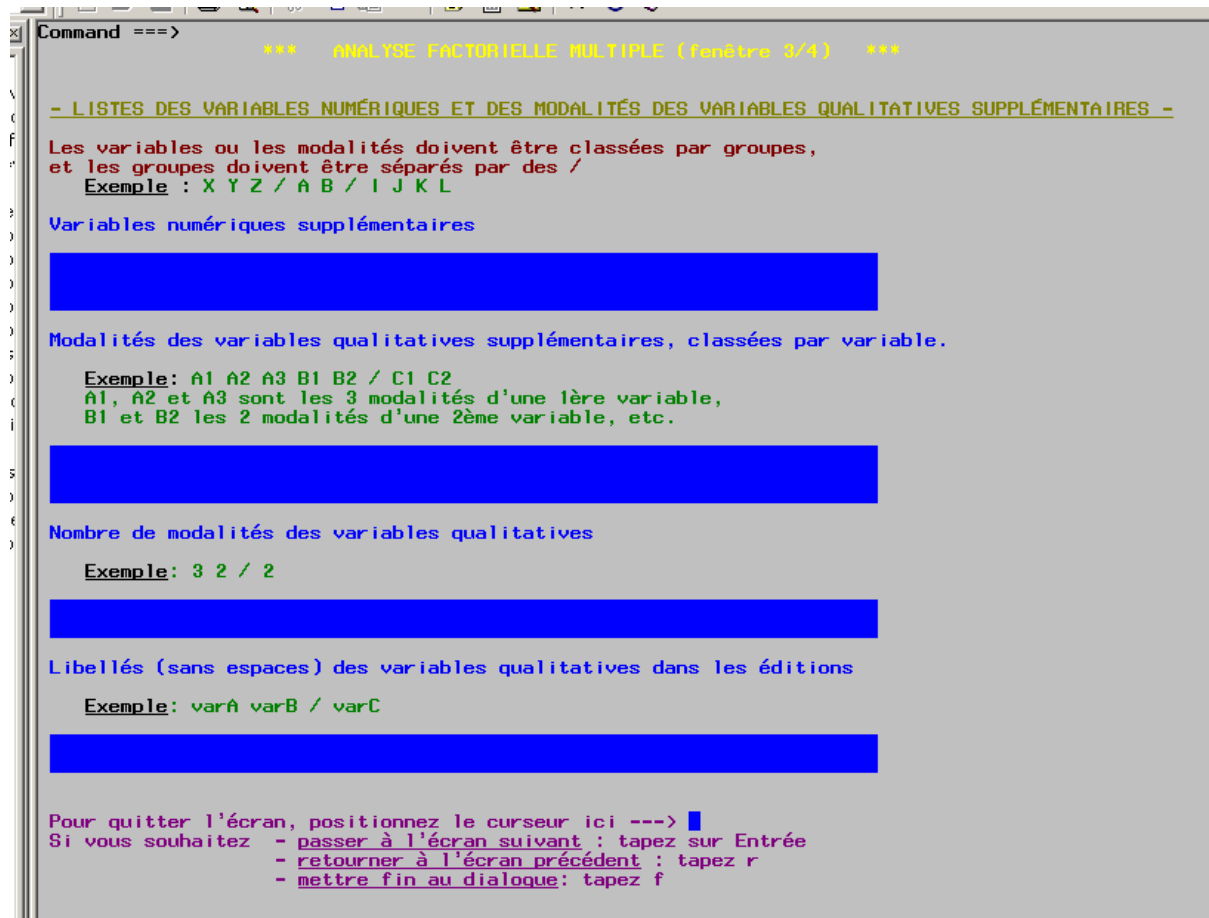
Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - passer à l'écran suivant : tapez sur Entrée
                  - retourner à l'écran précédent : tapez r
                  - mettre fin au dialogue: tapez f
```

Il existe des champs obligatoires, même si tous ne le sont pas :

- le nom de la **table SAS contenant les individus actifs** (...alors que le nom de la table contenant les individus supplémentaires ne l'est pas )
- L'identifiant **doit être une variable caractère** : il est obligatoire.

La variable de pondération doit être une variable numérique : lorsqu'elle n'est pas spécifiée, tous les individus sont affectés du même poids (= 1)

## La troisième fenêtre : saisie des paramètres concernant les variables supplémentaires



## Quatrième fenêtre : le paramétrage de l'édition des sorties

```
*** ANALYSE FACTORIELLE MULTIPLE (fenêtre 4/4) ***
- PARAMÈTRES D'ÉDITION -

N.B. : Dans tout ce qui suit : - tout ce qui est en rouge souligné est paramétrable
      - o = oui et n = non

Édition de la matrice des corrélations :
- entre les variables (et modalités) actives : o
- entre les facteurs partiels : o

*** AIDES À L'INTERPRETATION ***

* Nombre d'axes pour les aides à l'interprétation : 2

* Les aides à l'interprétation seront éditées pour les groupes, les variables,
  et les centres de gravité associés aux modalités des variables qualitatives.
  Les paramètres suivants sont relatifs aux individus et aux axes partiels.

* Édition complète des aides à l'interprétation :
  - pour les individus actifs : o
  - pour les individus supplémentaires : n

* Édition partielle des aides à l'interprétation des individus actifs :
  pour chaque axe sont édités les n individus ayant la plus forte contribution.
  Indiquez le nombre n d'individus édités :     

* Les aides à l'interprétation pour les axes partiels (i.e. les axes des analyses par groupe)
  seront éditées pour les p premiers axes de chaque analyse.
  Indiquez le nombre p d'axes partiels édités : 2

Nom de la table contenant les résultats de la macro : datasor
Indiquez la valeur du paramètre linesize désirée : 100

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - EXECUTER LA MACRO : tapez e
                  - retourner à l'écran précédent : tapez r
                  - mettre fin au dialogue: tapez f
Tout autre caractère vous renverra en haut de l'écran
```

### Remarques :

- Les valeurs affichées sont celles données par défaut aux paramètres
- Les aides à l'interprétation sont toujours éditées pour :
  - les groupes
  - les variables
  - les centres de gravité associés aux variables qualitatives
- Pour les individus, les éditions sont paramétrables et partielles ou complètes sur demande.
- Les matrices de corrélation ne sont éditées que sur demande.

### 2.2.3. Exemple de saisie : AFM temporelle

On souhaite sauvegarder les valeurs des paramètres dans la table fcda.param :

```
Command ==>

*** ANALYSE FACTORIELLE MULTIPLE (fenêtre 1/4) ***

Si vous voulez utiliser des paramètres stockés dans une table existant déjà,
indiquez son nom complet ci-dessous (sous la forme librairie.table) :

_____

Les valeurs des paramètres seront sauvegardées dans la table _PARAM_.
Si vous voulez aussi les sauvegarder dans une autre table,
indiquez son nom complet ci-dessous :
fcda.param

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - valider : tapez sur Entrée
- mettre fin au dialogue : tapez f
```

Pour obtenir les mêmes sorties qu'avec le programme donné au paragraphe 2.1.3 on saisit les valeurs suivantes.

```
Command ==>

*** ANALYSE FACTORIELLE MULTIPLE (fenêtre 2/4) ***

Table contenant les individus actifs : climat
Table contenant les individus supplémentaires : _____
Variable identifiant les individus dans ces tables : pays
Variable de pondération des individus actifs : _____

- LISTES DES VARIABLES NUMÉRIQUES ET DES MODALITÉS DES VARIABLES QUALITATIVES ACTIVES -

Les variables ou les modalités doivent être classées par groupes,
et les groupes doivent être séparés par des /
Exemple : X Y Z / A B / I J K L

Variables numériques actives

serre96--elec96/ serre01--elec01/
serre06--elec06

Modalités des variables qualitatives actives, classées par variable
Exemple : A1 A2 A3 B1 B2 / C1 C2
A1, A2 et A3 sont les 3 modalités d'une 1ère variable,
B1 et B2 les 2 modalités d'une 2ème variable, etc.

_____

Nombre de modalités des variables qualitatives
Exemple : 3 2 / 2

_____

Libellés (sans espaces) des variables qualitatives dans les éditions
Exemple : varA varB / varC

_____

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - passer à l'écran suivant : tapez sur Entrée
- retourner à l'écran précédent : tapez r
- mettre fin au dialogue : tapez f
```

```
Command ==> *** ANALYSE FACTORIELLE MULTIPLE (fenêtre 3/4) ***

- LISTES DES VARIABLES NUMÉRIQUES ET DES MODALITÉS DES VARIABLES QUALITATIVES SUPPLÉMENTAIRES -

Les variables ou les modalités doivent être classées par groupes,
et les groupes doivent être séparés par des /
Exemple : X Y Z / A B / I J K L

Variables numériques supplémentaires

[redacted]

Modalités des variables qualitatives supplémentaires, classées par variable.

Exemple: A1 A2 A3 B1 B2 / C1 C2
A1, A2 et A3 sont les 3 modalités d'une 1ère variable,
B1 et B2 les 2 modalités d'une 2ème variable, etc.

av1960--ap1980

Nombre de modalités des variables qualitatives

Exemple: 3 2 / 2

[redacted]
3

Libellés (sans espaces) des variables qualitatives dans les éditions

Exemple: varA varB / varC

[redacted]
ancienneté

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez - passer à l'écran suivant : tapez sur Entrée
- retourner à l'écran précédent : tapez r
- mettre fin au dialogue: tapez f
```

Command ==>

\*\*\* ANALYSE FACTORIELLE MULTIPLE (fenêtre 4/4) \*\*\*

- PARAMÈTRES D'ÉDITION -

N.B. : Dans tout ce qui suit : - tout ce qui est en rouge souligné est paramétrable  
- o = oui et n = non

Édition de la matrice des corrélations :

- entre les variables (et modalités) actives : n
- entre les facteurs partiels : n

\*\*\* AIDES À L'INTERPRÉTATION \*\*\*

\* Nombre d'axes pour les aides à l'interprétation : 2

\* Les aides à l'interprétation seront éditées pour les groupes, les variables, et les centres de gravité associés aux modalités des variables qualitatives. Les paramètres suivants sont relatifs aux individus et aux axes partiels.

\* Édition complète des aides à l'interprétation :

- pour les individus actifs : o
- pour les individus supplémentaires : n

\* Édition partielle des aides à l'interprétation des individus actifs : pour chaque axe sont édités les n individus ayant la plus forte contribution. Indiquez le nombre n d'individus édités :     

\* Les aides à l'interprétation pour les axes partiels (i.e. les axes des analyses par groupe) seront éditées pour les p premiers axes de chaque analyse. Indiquez le nombre p d'axes partiels édités : 2

Nom de la table contenant les résultats de la macro : datasor

Indiquez la valeur du paramètre linesize désirée : 100

Pour quitter l'écran, positionnez le curseur ici ---> █

Si vous souhaitez - EXECUTER LA MACRO : tapez e  
- retourner à l'écran précédent : tapez r  
- mettre fin au dialogue: tapez f

Tout autre caractère vous renverra en haut de l'écran

## **2.3. Les sorties éditées par la macro AFMULT**

### **2.3.1. Liste des sorties**

La macro édite :

- la matrice des corrélations entre variables actives (\*)
- la matrice des corrélations entre facteurs partiels (\*)
- le tableau résumé des valeurs propres des analyses partielles
- le diagramme des inerties (i.e. des valeurs propres) de l'analyse globale
- les coefficients de liaison entre les groupes :
  - coefficients Lg
  - coefficients RV
- les coordonnées et les aides à l'interprétation des variables :
  - actives
  - supplémentaires (s'il y en a)
- les coordonnées et les aides à l'interprétation des groupes :
  - actifs
  - supplémentaires (s'il y en a)
- les rapports inertie inter / inertie totale, axe par axe
- les corrélations entre les variables canoniques et les facteurs de l'analyse globale
- les coordonnées et les aides à l'interprétation des individus (\*) :
  - actifs
  - supplémentaires (s'il y en a)
- les individus ayant : (\*)
  - les plus fortes contributions
  - les plus fortes inerties intra
  - les plus faibles inerties intra
- les individus partiels ayant les plus fortes inerties intra (\*)
- les coordonnées et aides à l'interprétation des centres de gravité (s'il y a des variables qualitatives)
- les rapports de corrélation entre les variables qualitatives et les axes
- les coordonnées et aides à l'interprétation des axes partiels (\*)

(\*) signifie : édité seulement si explicitement demandé (dans la 4<sup>ème</sup> fenêtre de la boîte de dialogue)

### 2.3.2. Explicitation de ces sorties

- **La matrice des corrélations entre variables actives :**
  - seules les colonnes actives interviennent dans son calcul
  - seules les valeurs concernant les variables continues sont facilement interprétables
- **La matrice des corrélations entre facteurs partiels :**
  - les facteurs partiels sont les facteurs des analyses séparées (i.e. les analyses par groupe)
  - les facteurs d'une même analyse sont non corrélés
  - les facteurs étant définis au signe près, le signe des corrélations ne s'interprète pas
- **Le tableau résumé des valeurs propres et des analyses partielles :**
  - le numéro du groupe,
  - sa nature (quantitatif ou qualitatif, actif ou supplémentaire),
  - dim : sa dimension,
  - valp : les valeurs propres de l'analyse du groupe
  - pct : le pourcentage d'inertie expliqué par chaque axe
  - pctcum : les pourcentages cumulés de l'inertie

Compte tenu de la pondération utilisée en AFM, la première valeur propre ne peut excéder le nombre de groupes actifs

- **Le diagramme des inerties (i.e. des valeurs propres) de l'analyse globale, contient :**
  - les valeurs propres classées dans l'ordre décroissant
  - Diff . : les différences entre deux valeurs propres successives
  - Pct . : le pourcentage d'inertie expliqué par l'axe
  - Pct cumulé : les pourcentages cumulés de l'inertie, i.e. la part d'inertie expliquée par le sous-espace engendré par l'axe et les axes précédents
- **Les coefficients de liaison entre les groupes**
  - **Coefficient  $L_g$  :**
    - L'indice de liaison  $L_g(j,l)$  entre les groupes  $j$  et  $l$  est la somme, pondérée par les poids des variables de l'AFM, des carrés des covariances<sup>3</sup> entre chaque colonne du groupe  $j$  (variable continue ou indicatrice) et chaque colonne du groupe  $l$

---

<sup>3</sup> Ici des coefficients de corrélation car les variables continues sont réduites

- Ce coefficient est d'autant plus grand que les deux groupes ont des directions d'inertie importante en commun.
- $Lg(j,l)$  vaut 0 si chaque colonne du groupe  $j$  est non corrélée avec chaque colonne du groupe  $l$ . Il est d'autant plus élevé que les groupes possèdent des directions d'inertie importante en commun.
- Sur la diagonale,  $Lg$  est la somme des carrés des valeurs propres d'un seul groupe (en tenant compte de la pondération de l'AFM) : il s'interprète alors comme un indicateur de multidimensionalité dont la racine est notée  $Ng(j)=\sqrt{Lg(j,j)}$ .

- **Coefficient RV :**

- L'indice de liaison  $RV(j,l)$  se définit par :  

$$RV(j,l) = Lg(i,l) / [ Ng(j) Ng(l) ]$$
- Il est compris entre 0 et 1, et plus il est proche de 1, plus les groupes ont des structures proches
- Il prend la valeur 1 si les nuages d'individus induits par les deux groupes sont homothétiques.

- **Les coordonnées et les aides à l'interprétation des variables**

- **Actives**

Pour chaque axe :

- coord : coordonnée du point sur l'axe
- ctr : contribution relative à l'inertie expliquée par l'axe (en %).
- cos : qualité de représentation du point sur l'axe (en %).

- **Supplémentaires**

Pour chaque axe :

- coord : coordonnée du point sur l'axe
- cos : qualité de représentation du point sur l'axe (en %).

Pour les variables continues centrées et réduites, la distance à l'origine est constante et égale à 1. Lorsque les variables ne sont pas réduites, on retrouve la variance.

**La coordonnée sur un axe  $s$**  d'une variable continue centrée réduite est égale au coefficient de corrélation entre cette variable et l'axe  $s$ .

**La contribution** est égale à l'inertie projetée du point-variable divisée par l'inertie projetée de l'ensemble des points-variables actives.

- La somme des contributions des colonnes (variables continues ou indicatrices) d'un groupe est égale à la contribution du groupe.
- Pour les variables continues illustratives, ce rapport est calculé alors qu'il ne s'agit pas d'une contribution au plein sens du terme.

**Le cosinus carré** est égal au rapport : inertie projetée/inertie totale du point (égal au carré de la coordonnée dans le cas centré-réduit).  
 Pour un groupe : il mesure la qualité de représentation de l'ensemble des variables ce groupe par un facteur.

- **Les coordonnées et les aides à l'interprétation des groupes**

- disto : carré de la distance dans  $R^{I^2}$  entre le point représentant un groupe et l'origine. On retrouve la mesure  $[Ng]^2$

Pour chaque axe :

- coord : coordonnée du point sur l'axe, toujours comprise entre 0 et 1 (du fait de l'équilibrage du poids des groupes). La coordonnée d'un groupe  $j$  sur un axe  $s$  s'interprète, dans l'espace des variables, comme l'inertie projetée de l'ensemble des variables du groupe  $j$  pour l'axe  $s$ . Cette coordonnée est égale à 1 si l'axe  $s$  de l'analyse globale coïncide avec la première composante principale du groupe  $j$ .

La somme, pour tous les groupes actifs, des coordonnées sur l'axe  $s$  est égale à l'inertie de l'axe  $s$  de l'analyse globale.

- ctr : contribution relative à l'inertie expliquée par l'axe (coordonnée divisée par la somme des coordonnées des groupes actifs en %), seulement pour les groupes actifs
- cos : qualité de représentation du point sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %).
  - pour un groupe : il mesure la qualité de représentation de ce groupe (en tant que point unique dans l'espace des groupes) par un facteur.
  - pour l'ensemble : il mesure la qualité de représentation du nuage  $NJ$  de l'ensemble des groupes par un facteur.

- **Rapport entre l'inertie inter et l'inertie totale**

Le rapport (inertie inter)/(inertie totale) permet de mesurer l'intérêt d'une représentation simultanée de l'ensemble des nuages d'individus induits par chaque groupe de variables (nuages partiels).

En effet, l'ensemble des points de ces nuages peut être partitionné en regroupant dans une même classe les points représentant un même individu. Chaque classe contient autant de points que de groupes actifs. Il y a autant de classes que d'individus actifs. Si l'on applique le théorème de Huygens à cette partition (inertie totale = inertie inter + inertie intra), le rapport (inertie inter/inertie totale) montre dans quelle mesure chaque axe met en évidence une structure commune aux groupes. En cas de structure

commune, les différentes représentations d'un même individu sont proches : le rapport (inertie inter/inertie totale) est élevé.

Ce rapport ne diminue pas forcément avec le rang de l'axe (car ce n'est pas le critère utilisé pour leur ordonnancement). Il ne s'additionne pas d'un axe à l'autre.

- **Les corrélations entre les variables canoniques et les facteurs de l'analyse globale**

La s-ième variable canonique du groupe j est la projection du nuage des individus (décrits uniquement par le groupe j) sur le s-ième axe de l'analyse globale. La s-ième variable générale est le s-ième facteur sur les individus de l'analyse globale. Le coefficient de corrélation (toujours positif ou nul) entre ces deux facteurs montre dans quelle mesure le facteur de l'analyse globale peut être considéré comme une structure du groupe j. Les coefficients de corrélation s'interprètent conjointement aux coordonnées des groupes.

- **Les coordonnées et les aides à l'interprétation des individus**

Chaque individu est représenté par autant de points qu'il y a de groupes actifs : les individus partiels. A ceux-ci, s'ajoute un point représentant l'individu moyen (centre de gravité des individus partiels correspondants).

- **Actifs**

- poids : poids de l'individu (en %)
- disto : carré de la distance entre le point et l'origine.

Pour chaque axe :

- coord : coordonnée du point (individu moyen) sur l'axe.
- ctr : contribution relative à l'inertie expliquée par l'axe (en %).
- cos : qualité de représentation du point sur l'axe (en %).

La macro édite également pour chaque axe les coordonnées, contributions et qualités de représentation pour chacun des individus partiels, ces individus partiels étant identifiés par le numéro du groupe.

- **Supplémentaires**

- disto : carré de la distance entre le point et l'origine.

Pour chaque axe :

- coord : coordonnée du point (individu moyen) sur l'axe.
- cos : qualité de représentation du point sur l'axe (en %).

La macro édite également pour chaque axe les coordonnées et qualités de représentation pour chacun des individus partiels, ces individus partiels étant identifiés par le numéro du groupe.

Pour chaque axe :

- **les individus ayant les plus fortes contributions à l'inertie de l'axe**

Ce sont les individus prépondérants dans la détermination des axes.

- **les individus ayant les plus faibles inerties intra**

Ce sont les individus ayant des images partielles "cohérentes" entre elles, ce sont donc ceux qui illustrent le mieux la structure commune aux groupes actifs.

- **les individus ayant les plus fortes inerties intra**

Ce sont les individus qui s'inscrivent le moins bien dans la structure commune aux groupes actifs.

- **les individus partiels ayant les plus fortes inerties intra**

Ce sont les individus partiels les plus éloignés de l'individu moyen qui leur est associé, et qui donc illustrent le mieux les écarts à la structure commune aux groupes actifs. Ces inerties figurent déjà dans les aides à l'interprétation des individus sous le nom de contributions.

- **Les coordonnées et les aides à l'interprétation des centres de gravité**

Les centres de gravité des individus présentant une même modalité d'une variable qualitative sont considérés comme des individus supplémentaires. Les centres de gravité sont regroupés par variable.

La macro éditée en sortie :

- poids : poids de la modalité (en %), égal à la somme des poids des individus possédant la modalité.
- disto : carré de la distance entre le point et l'origine

Pour chaque axe :

- coord : coordonnée du point sur l'axe.
- ctr : c'est l'inertie de la modalité rapportée à l'inertie totale. Il ne s'agit donc pas d'une véritable contribution (même dans le cas où les variables qualitatives sont considérées comme actives : elle n'est que proportionnelle à la contribution au sens usuel)
- cos : qualité de représentation du point sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %).
- La macro éditée également pour chaque axe les coordonnées et qualités de représentation pour chacun des centres de gravité partiels, ces centres de gravité partiels étant identifiés par le numéro du groupe.

- Les **rapports de corrélation entre les variables qualitatives et les axes**

Le rapport de corrélation entre une variable qualitative et un axe mesure la liaison entre la variable et l'axe. Il est égal à la somme des contributions des centres de gravité moyens associés à la variable.

- Les **coordonnées et aides à l'interprétation des axes partiels**

Les axes partiels sont issus des analyses séparées des groupes. Ils sont considérés comme des variables quantitatives supplémentaires (centrées réduites) : ils donnent donc lieu aux mêmes aides à l'interprétation (coordonnées, contributions et qualités de représentation).

La macro édite aussi le poids de chaque axe, égal à la valeur propre qui lui est associée divisée par la première valeur propre de l'analyse du groupe dont l'axe est issu.

### 2.3.3. Les sorties de l'exemple AFM temporelle

Il est possible d'avoir des sorties dans l'output : sorties de type listings.

Coefficients Lg de liaison entre groupes 9

Groupe	1	2	3	4
n°1 - num act	1.51	1.38	1.30	0.78
n°2 - num act	1.38	1.59	1.55	0.83
n°3 - num act	1.30	1.55	1.76	0.78
n°4 - qual sup	0.78	0.83	0.78	2.00

On peut obtenir leur équivalent avec mise en forme HTML soit en utilisant les instructions utilisant l'ODS Output Delivery System , soit en cochant dans les préférences la production d'une sortie HTML (tools/options/preferences /onglet results) :

#### Coefficients Lg de liaison entre groupes

Groupe	1	2	3	4
n°1 - num act	1.51	1.38	1.30	0.78
n°2 - num act	1.38	1.59	1.55	0.83
n°3 - num act	1.30	1.55	1.76	0.78
n°4 - qual sup	0.78	0.83	0.78	2.00

Les sorties obtenues dans notre exemple sont donc si on retient la forme listing (moins belle mais qui prendra ici moins de place) :

Matrice des corrélations entre variables actives

Variable	Émissions de gaz à effet de serre 1996	Part des énergies renouv. dans conso éner. 1996	Dépendance énergétique 1996	Conso brute d'énergie intérieure 1996	Electricité d'énergies renouvelables 1996
serre96	1.00	0.32	-0.23	-0.43	0.24
renou96	0.32	1.00	0.05	-0.33	0.87
depen96	-0.23	0.05	1.00	-0.31	0.15
conso96	-0.43	-0.33	-0.31	1.00	-0.27
elec96	0.24	0.87	0.15	-0.27	1.00

Variable	Émissions de gaz à effet de serre 2001	Part des énergies renouv. dans conso éner. 2001	Dépendance énergétique 2001	Conso brute d'énergie intérieure 2001	Electricité d'énergies renouvelables 2001
serre96	0.55	0.40	-0.38	-0.42	0.28
renou96	0.17	0.96	-0.04	-0.35	0.87
depen96	0.32	-0.01	0.90	-0.29	0.09
conso96	-0.42	-0.38	-0.20	1.00	-0.25
elec96	0.27	0.87	0.08	-0.27	0.99

...etc

Tableau résumé des valeurs propres des analyses partielles

309

groupe	dim	valp1	valp2	valp3	valp4	valp5	pct1	pct2	pct3
n°1 - num act	5	2.29	1.22	1.02	0.35	0.12	45.9	24.4	20.5
n°2 - num act	5	2.22	1.44	0.71	0.55	0.08	44.4	28.8	14.2
n°3 - num act	5	2.11	1.66	0.65	0.43	0.14	42.3	33.1	13.1
n°4 - qual sup	2	1.00	1.00				50.0	50.0	

groupe	pct4	pct5	pctcum1	pctcum2	pctcum3	pctcum4	pctcum5
n°1 - num act	7.0	2.3	45.9	70.3	90.7	97.7	100.0
n°2 - num act	11.1	1.6	44.4	73.1	87.4	98.4	100.0
n°3 - num act	8.6	2.9	42.3	75.4	88.5	97.1	100.0
n°4 - qual sup			50.0	100.0			

## Matrice des corrélations entre facteurs partiels

310

Facteur	F1_1	F1_2	F1_3	F1_4	F1_5	F2_1	F2_2	F2_3	F2_4
F1_1	1.00	0.00	0.00	0.00	0.00	0.98	-0.07	-0.06	-0.01
F1_2	0.00	1.00	0.00	0.00	0.00	0.07	0.60	0.39	-0.63
F1_3	0.00	0.00	1.00	0.00	0.00	0.02	-0.56	0.76	-0.06
F1_4	0.00	0.00	0.00	1.00	0.00	-0.02	0.21	0.46	0.50
F1_5	0.00	0.00	0.00	0.00	1.00	-0.09	-0.08	-0.09	-0.22
F2_1	0.98	0.07	0.02	-0.02	-0.09	1.00	0.00	0.00	0.00
F2_2	-0.07	0.60	-0.56	0.21	-0.08	0.00	1.00	0.00	0.00
F2_3	-0.06	0.39	0.76	0.46	-0.09	0.00	0.00	1.00	0.00
F2_4	-0.01	-0.63	-0.06	0.50	-0.22	0.00	0.00	0.00	1.00
F2_5	0.01	-0.06	0.00	0.04	0.84	0.00	0.00	0.00	0.00
F3_1	0.90	-0.21	0.03	-0.17	-0.06	0.88	-0.37	-0.22	-0.05
F3_2	0.30	0.64	-0.43	0.07	-0.21	0.39	0.89	0.03	-0.03
F3_3	0.10	0.28	0.81	0.40	-0.21	0.17	-0.12	0.96	0.06
F3_4	-0.00	-0.52	-0.12	0.48	-0.25	0.02	-0.02	-0.06	0.91
Facteur	F2_5	F3_1	F3_2	F3_3	F3_4	F3_5	F4_1	F4_2	
F1_1	0.01	0.90	0.30	0.10	-0.00	-0.16	-0.20	-0.73	
F1_2	-0.06	-0.21	0.64	0.28	-0.52	-0.01	-0.54	0.12	
F1_3	0.00	0.03	-0.43	0.81	-0.12	-0.08	-0.33	-0.02	
F1_4	0.04	-0.17	0.07	0.40	0.48	-0.06	-0.15	0.11	
F1_5	0.84	-0.06	-0.21	-0.21	-0.25	0.23	-0.08	0.12	
F2_1	0.00	0.88	0.39	0.17	0.02	-0.09	-0.23	-0.78	
F2_2	0.00	-0.37	0.89	-0.12	-0.02	0.04	-0.19	-0.01	
F2_3	0.00	-0.22	0.03	0.96	-0.06	-0.01	-0.55	0.09	
F2_4	0.00	-0.05	-0.03	0.06	0.91	-0.02	0.34	-0.31	
F2_5	1.00	-0.02	-0.07	-0.07	-0.04	0.62	-0.17	-0.11	
F3_1	-0.02	1.00	0.00	0.00	0.00	0.00	-0.02	-0.71	
F3_2	-0.07	0.00	1.00	0.00	0.00	0.00	-0.27	-0.35	
F3_3	-0.07	0.00	0.00	1.00	0.00	0.00	-0.50	-0.11	
F3_4	-0.04	0.00	0.00	0.00	1.00	0.00	0.30	-0.31	

## Matrice des corrélations entre facteurs partiels

311

Facteur	F1_1	F1_2	F1_3	F1_4	F1_5	F2_1	F2_2	F2_3	F2_4
F3_5	-0.16	-0.01	-0.08	-0.06	0.23	-0.09	0.04	-0.01	-0.02
F4_1	-0.20	-0.54	-0.33	-0.15	-0.08	-0.23	-0.19	-0.55	0.34
F4_2	-0.73	0.12	-0.02	0.11	0.12	-0.78	-0.01	0.09	-0.31
Facteur	F2_5	F3_1	F3_2	F3_3	F3_4	F3_5	F4_1	F4_2	
F3_5	0.62	0.00	0.00	0.00	0.00	1.00	-0.09	-0.13	
F4_1	-0.17	-0.02	-0.27	-0.50	0.30	-0.09	1.00	0.00	
F4_2	-0.11	-0.71	-0.35	-0.11	-0.31	-0.13	0.00	1.00	

N°	Valeur propre	Diff.	Pct	Pct cumulé	
1	2.91		42.9	42.9	!*****
2	1.81	1.11	26.6	69.4	!*****
3	1.06	0.75	15.6	85.0	!*****
4	0.63	0.43	9.2	94.2	!*****
5	0.13	0.49	2.0	96.2	!**
6	0.10	0.03	1.5	97.7	!*
7	0.09	0.01	1.3	99.0	!*
8	0.05	0.04	0.8	99.8	!
9	0.01	0.04	0.1	99.9	!
10	0.00	0.00	0.1	99.9	!
11	0.00	0.00	0.0	100.0	!
12	0.00	0.00	0.0	100.0	!
13	0.00	0.00	0.0	100.0	!
14	0.00	0.00	0.0	100.0	!

Coefficients Lg de liaison entre groupes

Groupe		1	2	3	4
n°1 - num act		1.51	1.38	1.30	0.78
n°2 - num act		1.38	1.59	1.55	0.83
n°3 - num act		1.30	1.55	1.76	0.78
n°4 - qual sup		0.78	0.83	0.78	2.00

Coefficients RV de liaison entre groupes

Groupe		1	2	3	4
n°1 - num act		1.00	0.89	0.80	0.45
n°2 - num act		0.89	1.00	0.93	0.47
n°3 - num act		0.80	0.93	1.00	0.41
n°4 - qual sup		0.45	0.47	0.41	1.00

Coordonnées et aides à l'interprétation des variables actives

315

variable	coord1	coord2	ctr1	ctr2	cos1	cos2
serre96	0.54	-0.14	4.3	0.5	28.9	1.9
renou96	0.87	-0.24	11.3	1.4	75.3	5.7
depen96	0.16	0.81	0.4	15.7	2.7	65.3
conso96	-0.66	-0.44	6.6	4.7	43.9	19.6
elec96	0.86	-0.13	11.1	0.4	74.2	1.6
groupe 1			33.6	22.7	45.0	18.8
serre01	0.42	0.58	2.7	8.3	17.8	33.4
renou01	0.91	-0.29	12.8	2.2	82.8	8.7
depen01	0.02	0.87	0.0	18.7	0.1	75.1
conso01	-0.66	-0.41	6.7	4.3	43.5	17.1
elec01	0.86	-0.23	11.6	1.3	74.7	5.1
groupe 2			33.9	34.8	43.8	27.9
serre06	0.43	0.69	3.0	12.5	18.4	47.6
renou06	0.78	-0.46	9.9	5.6	60.8	21.3
depen06	-0.05	0.83	0.0	18.2	0.3	69.7
conso06	-0.65	-0.39	6.9	4.1	42.5	15.6
elec06	0.88	-0.29	12.7	2.2	78.2	8.3
groupe 3			32.5	42.5	40.0	32.5

Coordonnées et aides à l'interprétation des variables supplémentaires

316

variable	coord1	coord2	cos1	cos2
av1960	-0.63	0.07	40.1	0.5
av1980	-0.17	-0.27	2.7	7.4
ap1980	0.77	0.15	59.0	2.2
groupe 4			30.8	3.8

Coordonnées et aides à l'interprétation des groupes actifs

317

groupe	disto	coord1	coord2	ctr1	ctr2	cos1	cos2
groupe 1	1.51	0.98	0.41	33.6	22.7	63.8	11.2
groupe 2	1.59	0.99	0.63	33.9	34.8	61.3	24.9
groupe 3	1.76	0.95	0.77	32.5	42.5	51.1	33.7
Ensemble				100.0	100.0	58.4	23.8

Coordonnées et aides à l'interprétation des groupes supplémentaires

318

groupe	disto	coord1	coord2	cos1	cos2
groupe 4	2.00	0.62	0.08	19.0	0.3
Ensemble				19.0	0.3

Rapports inertie inter/inertie totale

319

axe 1	axe 2
0.98	0.88

groupe	facteur 1	facteur 2
1	0.99	0.90
2	1.00	0.98
3	0.98	0.98

## Coordonnées et aides à l'interprétation des individus actifs

pays	groupe	poids	disto	coord1	coord2	ctr1	ctr2	cos1	cos2
Allemagne		6.67	11.97	-2.75	-1.06	17.3	4.1	63.1	9.4
	1		35.29	-2.74	-0.57	0.0	2.2	21.2	0.9
	2		36.22	-2.83	-1.21	0.3	0.2	22.2	4.0
	3		36.24	-2.68	-1.40	0.2	1.1	19.8	5.4
Autriche		6.67	12.73	3.22	-0.36	23.7	0.5	81.4	1.0
	1		38.75	3.19	-0.06	0.0	0.9	26.3	0.0
	2		37.53	3.22	-0.46	0.0	0.1	27.7	0.6
	3		38.28	3.24	-0.57	0.0	0.4	27.4	0.8
Belgique		6.67	3.01	-1.01	0.89	2.3	3.0	34.0	26.6
	1		8.60	-0.77	1.06	2.0	0.3	6.9	13.0
	2		8.73	-1.11	0.94	0.4	0.0	14.2	10.2
	3		9.75	-1.15	0.68	0.6	0.4	13.5	4.8
Danemark		6.67	10.89	0.84	-1.98	1.6	14.4	6.5	35.9
	1		28.44	0.80	-0.85	0.1	11.9	2.3	2.5
	2		27.26	0.42	-2.07	6.3	0.1	0.6	15.7
	3		42.28	1.31	-3.01	7.6	10.0	4.1	21.4
Espagne		6.67	4.42	0.09	1.32	0.0	6.4	0.2	39.3
	1		4.72	-0.42	0.52	9.1	6.0	3.8	5.7
	2		12.19	0.31	1.33	1.6	0.0	0.8	14.6
	3		22.83	0.38	2.10	3.0	5.7	0.6	19.3
Finlande		6.67	4.12	1.72	-0.35	6.8	0.5	72.1	3.0
	1		11.55	1.73	-0.20	0.0	0.2	26.0	0.4
	2		11.95	1.70	-0.29	0.0	0.0	24.2	0.7
	3		13.56	1.73	-0.57	0.0	0.4	22.1	2.4
France		6.67	4.01	-1.27	-0.98	3.7	3.5	40.1	23.9
	1		10.42	-1.08	-0.83	1.3	0.2	11.1	6.6
	2		11.43	-1.20	-0.96	0.2	0.0	12.6	8.1
	3		14.20	-1.53	-1.14	2.3	0.3	16.4	9.2
Grèce		6.67	2.11	-0.01	1.19	0.0	5.2	0.0	67.3
	1		3.59	-0.04	0.66	0.0	2.7	0.1	12.0
	2		8.63	-0.16	1.42	0.8	0.5	0.3	23.5
	3		6.78	0.17	1.49	1.2	0.9	0.5	32.9
Irlande		6.67	4.55	-0.21	1.88	0.1	13.0	1.0	77.5
	1		9.19	-0.29	0.95	0.2	8.0	0.9	9.9
	2		18.09	-0.22	2.35	0.0	2.1	0.3	30.5
	3		13.68	-0.13	2.33	0.2	1.9	0.1	39.8

pays	groupe	poids	disto	coord1	coord2	ctr1	ctr2	cos1	cos2
Italie		6.67	2.23	-0.82	0.71	1.5	1.9	29.9	22.5
	1		5.60	-0.61	0.55	1.5	0.2	6.6	5.4
	2		5.27	-0.64	0.61	1.1	0.1	7.7	7.1
	3		9.19	-1.21	0.97	5.3	0.6	15.8	10.2
Luxembourg		6.67	8.95	-1.16	1.65	3.1	10.1	15.0	30.5
	1		38.80	-1.40	2.16	2.0	2.4	5.1	12.0
	2		25.17	-1.22	0.98	0.1	4.2	5.9	3.9
	3		16.63	-0.86	1.81	3.2	0.2	4.4	19.7
Pays-Bas		6.67	2.83	-0.94	-0.50	2.0	0.9	31.4	8.7
	1		12.54	-0.90	-0.81	0.1	0.9	6.4	5.2
	2		6.51	-1.04	-0.24	0.3	0.6	16.5	0.9
	3		6.46	-0.90	-0.44	0.1	0.0	12.4	2.9
Portugal		6.67	7.18	1.81	1.49	7.5	8.2	45.8	31.0
	1		18.86	2.40	0.43	11.8	10.6	30.5	1.0
	2		25.34	2.15	1.74	4.0	0.6	18.3	12.0
	3		20.45	0.89	2.30	29.6	6.1	3.9	25.9
Royaume-Uni		6.67	10.77	-2.32	-2.10	12.3	16.3	49.9	41.0
	1		40.22	-2.50	-2.07	1.1	0.0	15.6	10.7
	2		32.46	-2.28	-2.41	0.1	0.9	15.9	17.9
	3		24.29	-2.18	-1.83	0.7	0.7	19.6	13.8
Suède		6.67	12.22	2.80	-1.81	18.0	12.0	64.3	26.7
	1		27.57	2.62	-0.94	1.1	7.0	24.9	3.2
	2		37.58	2.90	-1.74	0.3	0.0	22.3	8.0
	3		44.85	2.89	-2.74	0.3	8.2	18.6	16.8

## Individus ayant les plus fortes contributions

324

axe=1

pays	inertie	cumul
Autriche	23.7	23.7
Suède	18.0	41.7
Allemagne	17.3	59.0
Royaume-Uni	12.3	71.3
Portugal	7.5	78.8
Finlande	6.8	85.6
France	3.7	89.3
Luxembourg	3.1	92.4
Belgique	2.3	94.7
Pays-Bas	2.0	96.7
Danemark	1.6	98.4
Italie	1.5	99.9
Irlande	0.1	100.0
Espagne	0.0	100.0
Grèce	0.0	100.0

axe=2

pays	inertie	cumul
Royaume-Uni	16.3	16.3
Danemark	14.4	30.7
Irlande	13.0	43.7
Suède	12.0	55.8
Luxembourg	10.1	65.8
Portugal	8.2	74.0
Espagne	6.4	80.4
Grèce	5.2	85.7
Allemagne	4.1	89.8
France	3.5	93.3
Belgique	3.0	96.3
Italie	1.9	98.1
Pays-Bas	0.9	99.1
Autriche	0.5	99.5
Finlande	0.5	100.0

axe=1

pays	inertie	cumul
Portugal	45.4	45.4
Danemark	14.0	59.3
Espagne	13.7	73.1
Italie	7.9	81.0
Luxembourg	5.3	86.3
France	3.8	90.1
Belgique	3.0	93.0
Grèce	2.0	95.0
Royaume-Uni	1.9	96.9
Suède	1.7	98.6
Pays-Bas	0.5	99.1
Allemagne	0.4	99.5
Irlande	0.4	99.9
Autriche	0.0	100.0
Finlande	0.0	100.0

axe=2

pays	inertie	cumul
Danemark	22.0	22.0
Portugal	17.3	39.3
Suède	15.3	54.5
Irlande	12.1	66.6
Espagne	11.7	78.3
Luxembourg	6.9	85.2
Grèce	4.1	89.3
Allemagne	3.5	92.7
Royaume-Uni	1.6	94.3
Pays-Bas	1.5	95.9
Autriche	1.3	97.2
Italie	1.0	98.2
Belgique	0.7	98.9
Finlande	0.7	99.5
France	0.5	100.0

axe=1

pays	inertie	cumul
Finlande	0.0	0.0
Autriche	0.0	0.1
Irlande	0.4	0.5
Allemagne	0.4	0.9
Pays-Bas	0.5	1.4
Suède	1.7	3.1
Royaume-Uni	1.9	5.0
Grèce	2.0	7.0
Belgique	3.0	9.9
France	3.8	13.7
Luxembourg	5.3	19.0
Italie	7.9	26.9
Espagne	13.7	40.7
Danemark	14.0	54.6
Portugal	45.4	100.0

axe=2

pays	inertie	cumul
France	0.5	0.5
Finlande	0.7	1.1
Belgique	0.7	1.8
Italie	1.0	2.8
Autriche	1.3	4.1
Pays-Bas	1.5	5.7
Royaume-Uni	1.6	7.3
Allemagne	3.5	10.7
Grèce	4.1	14.8
Luxembourg	6.9	21.7
Espagne	11.7	33.4
Irlande	12.1	45.5
Suède	15.3	60.7
Portugal	17.3	78.0
Danemark	22.0	100.0

axe=1

pays	groupe	inertie	cumul
Portugal	3	29.6	29.6
Portugal	1	11.8	41.4
Espagne	1	9.1	50.4
Danemark	3	7.6	58.0
Danemark	2	6.3	64.4
Italie	3	5.3	69.6
Portugal	2	4.0	73.6
Luxembourg	3	3.2	76.8
Espagne	3	3.0	79.8
France	3	2.3	82.1
Luxembourg	1	2.0	84.2
Belgique	1	2.0	86.1
Espagne	2	1.6	87.8
Italie	1	1.5	89.3
France	1	1.3	90.6

axe=2

pays	groupe	inertie	cumul
Danemark	1	11.9	11.9
Portugal	1	10.6	22.5
Danemark	3	10.0	32.4
Suède	3	8.2	40.6
Irlande	1	8.0	48.7
Suède	1	7.0	55.7
Portugal	3	6.1	61.9
Espagne	1	6.0	67.8
Espagne	3	5.7	73.6
Luxembourg	2	4.2	77.8
Grèce	1	2.7	80.5
Luxembourg	1	2.4	82.9
Allemagne	1	2.2	85.1
Irlande	2	2.1	87.2
Irlande	3	1.9	89.1

Coordonnées et aides à l'interprétation des centres de gravité

332

variable	modalité	groupe	poids	disto	coord1	coord2	ctr1	ctr2	cos1	cos2
ancienneté	av1960		40.00	2.02	-1.32	0.12	24.1	0.3	86.8	0.7
		1		5.49	-1.25	0.26	1.2	1.1	28.4	1.2
		2		6.23	-1.34	0.02	0.1	0.6	28.8	0.0
		3		6.48	-1.39	0.08	0.8	0.1	29.6	0.1
ancienneté	av1980		20.00	2.38	-0.56	-0.73	2.2	6.0	13.4	22.6
		1		8.70	-0.66	-0.66	1.0	0.2	5.0	4.9
		2		7.20	-0.69	-0.71	1.8	0.0	6.7	7.0
		3		5.47	-0.33	-0.83	5.5	0.3	2.0	12.7
ancienneté	ap1980		40.00	2.76	1.61	0.25	35.4	1.3	93.4	2.2
		1		7.67	1.58	0.07	0.1	1.8	32.6	0.1
		2		8.94	1.69	0.33	1.4	0.4	31.9	1.3
		3		8.26	1.55	0.34	0.6	0.5	29.2	1.4

Rapports de corrélation entre les variables qualitatives et les axes

333

groupe=4

variable	axe 1	axe 2
ancienneté	61.7	7.6

Coordonnées et aides à l'interprétation des axes partiels

334

groupe=1

axe	poids	coord1	coord2	ctr1	ctr2	cos1	cos2
1	1.00	0.99	0.02	33.6	0.0	97.9	0.0
2	0.53	-0.01	0.75	0.0	16.5	0.0	56.1

groupe=2

axe	poids	coord1	coord2	ctr1	ctr2	cos1	cos2
1	1.00	0.99	0.09	33.6	0.5	97.8	0.9
2	0.65	-0.09	0.97	0.2	33.9	0.9	94.4

groupe=3

axe	poids	coord1	coord2	ctr1	ctr2	cos1	cos2
1	1.00	0.94	-0.28	30.1	4.3	87.8	7.7
2	0.78	0.29	0.94	2.3	38.0	8.5	87.5

groupe=4

axe	poids	coord1	coord2	ctr1	ctr2	cos1	cos2
1	1.00	-0.17	-0.27	0.9	4.1	2.7	7.4
2	1.00	-0.77	-0.04	20.2	0.1	59.0	0.2

### 3. La macro PLOTAFM

La macro SAS **PLOTAFM** produit des représentations graphiques associées à une analyse factorielle multiple réalisée à partir de la macro AFMULT. Elle utilise le module SAS-GRAPH (ou, à défaut, la procédure PLOT). Les graphiques peuvent représenter les individus moyens et partiels actifs et supplémentaires, les variables quantitatives et qualitatives actives et supplémentaires, les centres de gravité moyens et partiels actifs et supplémentaires, et enfin les groupes actifs et supplémentaires.

Les graphiques produits par SAS-GRAPH peuvent être sauvegardés dans différents formats (wmf, jpeg, png ...etc. avec le menu déroulant file/export as image). En se plaçant sur le graphique, on peut activer l'éditeur de graphiques et ajouter ou modifier certains éléments (texte ou autres objets).

**Cette macro utilise la table en sortie créée par la macro AFMULT (dont le nom par défaut est `_DATASOR_`).**

On dispose de deux possibilités pour utiliser la macro PLOTAFM :

- soit renseigner les paramètres de la macro dans la fenêtre éditeur de programmes, en utilisant les règles habituelles d'écriture des paramètres d'une macro (voir 3.1 *les paramètres de la macro*) :  
**%PLOTAFM(data=...,plans=... , ....)**
- soit utiliser la boîte de dialogue, qui permet de spécifier les valeurs prises par les paramètres de la macro dans des fenêtres de saisie apparaissant à l'écran (voir 3.2 *la boîte de dialogue*) :  
**%PLOTAFM(fen=o)** pour lancer la macro et ses boîtes de dialogue

#### **3.1. Les paramètres de la macro PLOTAFM**

##### **3.1.1. Paramètres généraux**

**DATA = nom de table SAS**

nom de la table SAS en sortie de la macro AFMULT.

**Ce paramètre est obligatoire.**

**TTEXT= taille des textes dans les graphiques**

**Par défaut : TTEXT = 0.8**

## **GRAPH = O ou N**

Si GRAPH = O (=oui) et si la version de SAS utilisée est munie du module SAS-GRAPH, il y a édition des graphiques avec ce module.

La macro teste systématiquement la présence de ce module dans la version de SAS utilisée : s'il est absent elle délivre un message d'avertissement et édite les graphiques avec la procédure PLOT.

**Par défaut : GRAPH = O**

## **PLANS = liste de plans factoriels**

Liste de couples d'axes définissant les plans pour lesquels on veut obtenir des graphiques. Si on définit plusieurs plans factoriels, alors tous les graphiques demandés seront édités pour chaque plan factoriel ainsi spécifié.

Exemple : PLANS = 1-2 3-4 : les graphiques seront édités pour les deux premiers plans factoriels

**Par défaut : PLANS = 1-2**

## **TYPER = S, C ou Q**

Indique le type de représentation de tous les graphiques demandés.

On peut choisir :

- **C** : on obtiendra des graphiques dans lesquels les points seront représentés par des cercles de surface **proportionnelle à la contribution de ces points à l'inertie du plan** dans lequel ils sont représentés
- **Q** : on obtiendra des graphiques dans lesquels les points seront représentés par des cercles de surface **proportionnelle à la qualité de représentation de ces points dans le plan** sur lequel ils sont représentés
- **S** : on obtiendra une **représentation simple** (i.e. la représentation habituelle)

**Ce paramètre est obligatoire.**

### **3.1.2. Paramètres spécifiant les graphiques**

Dans la suite, lorsque nous mettrons "O ou N" cela voudra dire : soit l'on met O comme oui et on obtiendra un graphique, soit l'on met un caractère quelconque différent de O (N ou bien encore rien) et on n'aura pas le graphique correspondant.

## **TGRAPH = Taille du cercle des corrélations et du graphique des groupes**

**Par défaut : TGRAPH = 12**

### **Graphiques représentant les individus**

#### ***Types de points à représenter***

## **IAM = O ou N**

Si IAM = O, alors seront représentés les **individus actifs moyens**

**Par défaut : IAM = N**

**ISM = O ou N**

Si ISM = O, alors seront représentés les **individus supplémentaires moyens**

**Par défaut : ISM = N**

**IP = O ou N**

Si IP = O alors seront représentés les **individus partiels correspondant aux catégories sélectionnées parmi les deux précédentes.**

Et donc, si IAM = N et ISM = N alors l'instruction IP = O sera sans effet.

**Par défaut : IP = N**

**Filtres :****FIAM = C xx ou Q xx , où xx est un entier**

- Si FIAM = C xx : ne seront représentés que les xx individus actifs moyens ayant la plus forte contribution à l'inertie du plan factoriel dans lequel ils sont représentés.
- Si FIAM = Q xx : ne seront représentés que les xx individus actifs moyens ayant la meilleure qualité de représentation dans le plan factoriel dans lequel ils sont représentés.

**Par défaut : pas de filtre**

**FISM = xx , où xx est un entier**

Si FISM = xx : ne seront représentés que les xx individus supplémentaires moyens ayant la meilleure qualité de représentation dans le plan factoriel dans lequel ils sont représentés.

**Par défaut : pas de filtre**

**Type de données :****TYPED = THEM, TEMP ou TEMPM**

Le paramètre TYPED n'est effectif que si l'on demande la représentation des individus partiels (IP = O). Il provoque alors le tracé des segments entre les différents points, moyens ou partiels, tracé dont la justification dépend du type de données sur lesquelles porte l'analyse.

- THEM (données **thématiques**) : les individus partiels sont reliés par des segments à l'individu moyen auquel ils sont associés.
- TEMP (données **temporelles**) : les individus partiels sont reliés par des segments dans l'ordre des groupes, et les individus moyens ne sont pas représentés.
- TEMPM : (données **temporelles** + points **moyens**) : les individus partiels sont reliés par des segments dans l'ordre des groupes, et les individus moyens sont représentés.

**Graphiques représentant les variables****Types de points à représenter :**

**VAN = O ou N**

Si VAN = O, alors seront représenté les **variables actives numériques**

**Par défaut : VAN = N**

**VAQ = O ou N**

Si VAQ = O alors seront représentées les **variables actives qualitatives**

**Par défaut : VAQ = N**

**VSN = O ou N**

Si VSN = O alors seront représentées les **variables supplémentaires numériques**

**Par défaut : VSN = N**

**VSQ = O ou N**

Si VSQ = O alors seront représentées les **variables supplémentaires qualitatives**

**Par défaut : VSQ = N**

**Filtres :**

**FVAN = C xx ou Q xx , où xx est un entier**

- Si FVAN = C xx : ne seront représentées que les xx variables actives numériques ayant la plus forte contribution à l'inertie du plan factoriel dans lequel elles sont représentées.
- Si FVAN = Q xx : ne seront représentées que les xx variables actives numériques ayant la meilleure qualité de représentation dans le plan factoriel dans lequel elles sont représentées.

**Par défaut : pas de filtre**

**FVAQ = C xx ou Q xx , où xx est un entier**

- Si FVAQ = C xx : ne seront représentées que les xx variables actives qualitatives ayant la plus forte contribution à l'inertie du plan factoriel dans lequel elles sont représentées.
- Si FVAQ = Q xx : ne seront représentées que les xx variables actives qualitatives ayant la meilleure qualité de représentation dans le plan factoriel dans lequel elles sont représentées.

**Par défaut : pas de filtre**

**FVSN = xx , où xx est un entier**

Si FVSN = xx : ne seront représentées que les xx variables supplémentaires numériques ayant la meilleure qualité de représentation dans le plan factoriel dans lequel elles sont représentées.

**Par défaut : pas de filtre**

**FVSQ = xx , où xx est un entier**

Si FVSQ = xx : ne seront représentées que les xx variables supplémentaires qualitatives ayant la meilleure qualité de représentation dans le plan factoriel dans lequel elles sont représentées.

**Par défaut : pas de filtre**

### **Graphiques représentant les centres de gravité**

***Types de points à représenter :***

**CDGAM = O ou N**

Si CDGAM = O, alors seront représentés les **centres de gravité actifs moyens**

**Par défaut : CDGAM = N**

**CDGSM = O ou N**

Si CDGSM = O alors seront représentés les **centres de gravité supplémentaires moyens**

**Par défaut : CDGSM = N**

**CDGP = O ou N**

Si CDGP = O alors seront représentés les **centres de gravité partiels correspondant aux catégories sélectionnées parmi les deux précédentes**. Et donc, si CDGAM = N et CDGSM = N alors l'instruction CDGP = O sera sans effet.

**Par défaut : CDGP = N**

***Filtres :***

**FCDGAM = C xx ou Q xx , où xx est un entier**

- Si FCDGAM = C xx : ne seront représentés que les xx centres de gravité actifs moyens ayant la plus forte contribution à l'inertie du plan factoriel dans lequel ils sont représentés.
- Si FCDGAM = Q xx : ne seront représentés que les xx centres de gravité actifs moyens ayant la meilleure qualité de représentation dans le plan factoriel dans lequel ils sont représentés.

**Par défaut : pas de filtre**

**FCDGSM = xx , où xx est un entier**

Si FCDGSM = xx : ne seront représentés que les xx centres de gravité supplémentaires moyens ayant la meilleure qualité de représentation dans le plan factoriel dans lequel ils sont représentés.

**Par défaut : pas de filtre**

**Type de données :**

**TYPEDCDG = THEM, TEMP ou TEMPM**

Le paramètre TYPEDCDG n'est effectif que si l'on demande la représentation des individus partiels (CDGP = 0). Il provoque alors le tracé des segments entre les différents points, moyens ou partiels, tracé dont la justification dépend du type de données sur lesquelles porte l'analyse.

- THEM (données **thématiques**) : les centres de gravité partiels sont reliés par des segments au centre de gravité auquel ils sont associés.
- TEMP (données **temporelles**) : les centres de gravité partiels sont reliés par des segments dans l'ordre des groupes, et les centres de gravité moyens ne sont pas représentés.
- TEMPM : (données **temporelles** + points **moyens**) : les centres de gravité partiels sont reliés par des segments dans l'ordre des groupes, et les centres de gravité moyens sont représentés.

**Graphiques représentant les groupes**

**GRA = 0 ou N**

Si GRA = 0 alors seront représentés les **groupes actifs**

**Par défaut : GRA = N**

**GRS = 0 ou N**

Si GRS = 0 alors seront représentés les **groupes supplémentaires**

**Par défaut : GRS = N**

### **3.1.3. Exemple de programmation : AFM temporelle**

On utilise le même exemple qu'avec la macro AFMULT : sur les pays de l'Union européenne.

Pour appeler la macro %AFMULT sans les boîtes de dialogue (sans écran de saisie) par exemple :

```
%plotafm(plans=1-2, iam=0, ip=0, fiam=C 9,  
typer=C, typed=TEMP, van=0, vsq=0, cdgsm=0, cdgp=0,  
typedcdg=TEMP, gra=0, grs=0, data=_DATASOR_);
```

## 3.2. La boîte de dialogue de la macro PLOTAFM

La macro PLOTAFM est dotée d'une boîte de dialogue qui facilite la saisie des différents paramètres de la macro. Cette boîte de dialogue comporte une seule fenêtre.

L'ouverture de la boîte de dialogue s'effectue en donnant la valeur "o" (pour "oui") au paramètre fen (pour "fenêtre") de la macro %PLOTAFM. Il est alors inutile de renseigner les autres paramètres de la macro.

Si la librairie qui contient le catalogue de macros est appelé compil par exemple

```
options sasstore=compil mstored;
```

```
%plotafm(fen=o)
```

- Les valeurs affichées sont celles données par défaut aux paramètres :



Si on souhaite obtenir les mêmes graphiques que ceux obtenus avec le programme (3.1.3. Exemple de programmation : AFM temporelle), on saisira à l'écran les valeurs suivantes :

```

GRAPH
Command ==>

*** EDITION DE GRAPHIQUES ILLUSTRANT LES RESULTATS D'UNE ANALYSE FACTORIELLE MULTIPLE ***

Nom de la table SAS en sortie de la macro AFMULT : datasor
Souhaitez-vous utiliser le module SAS-GRAPH ? : o

TYPES DE POINTS A REPRESENTER (o = oui), FILTRES (*), et TYPE DE DONNEES (**):

INDIVIDUS - actifs moyens : o Filtre : c 9
           - supplémentaires moyens : - Filtre : -
           - partiels correspondants : o
Type de données : TEMP

VARIABLES - actives numériques : o Filtre : ---
           - actives qualitatives : - Filtre : ---
           - supplémentaires numériques : - Filtre : ---
           - supplémentaires qualitatives : o Filtre : ---

CENTRES DE GRAVITE - actifs moyens : - Filtre : ---
                   - supplémentaires moyens : o Filtre : ---
                   - partiels correspondants : o
Type de données : TEMP

GROUPES - actifs : o
         - supplémentaires : o

Liste des plans factoriels : (ex : 1-2 1-3) 1-2

Type de représentation : c où s = simple,
                        c = proportionnelle à la contribution à l'inertie
                        q = proportionnelle à la qualité de la représentation

Taille du texte dans les graphiques : 0.8
Taille du cercle des corrélations et du graphique des groupes de variables: 12

(*) FILTRES: (optionnels)
. Pour les éléments actifs:
  ex: - c 5 sélectionne les 5 points ayant la plus forte contribution à l'inertie du plan
      - q 7 sélectionne les 7 points ayant la meilleure qualité de représentation sur le plan
. Pour les éléments supplémentaires : 6 sélectionne les 6 points ayant la meilleure qualité

(**) TYPE DE DONNEES: (nécessaire pour tracer les segments en présence de points "partiels") :
    them = thématiques, temp = temporelles ou tempm = temporelles + individus moyens

Pour quitter l'écran, positionnez le curseur ici ---> █
Si vous souhaitez : - EXECUTER la macro : tapez e
                  - Mettre fin au dialogue : tapez f
Tout autre caractère vous renverra en haut de l'écran.

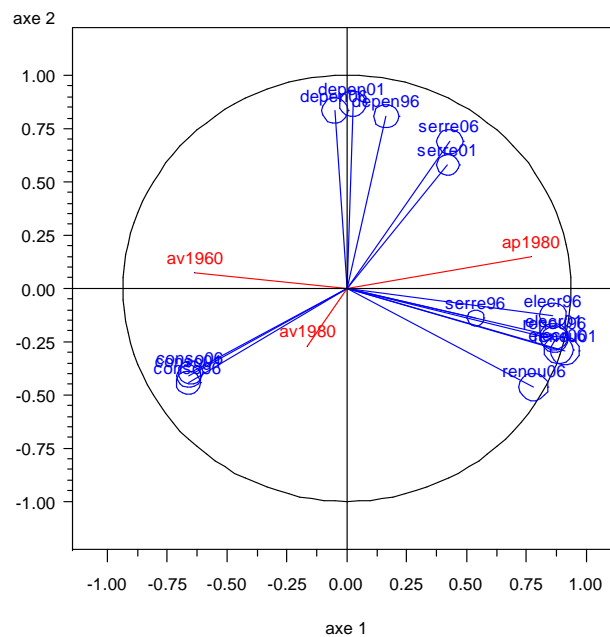
```

### 3.3. Exemple de graphiques obtenus avec PLOTAFM

On rappelle qu'il est possible d'entrer dans un graphique pour déplacer des éléments ou en ajouter. Ainsi, on peut déplacer des étiquettes qui se chevauchent.

Un cercle des corrélations sans mise en forme complémentaire :

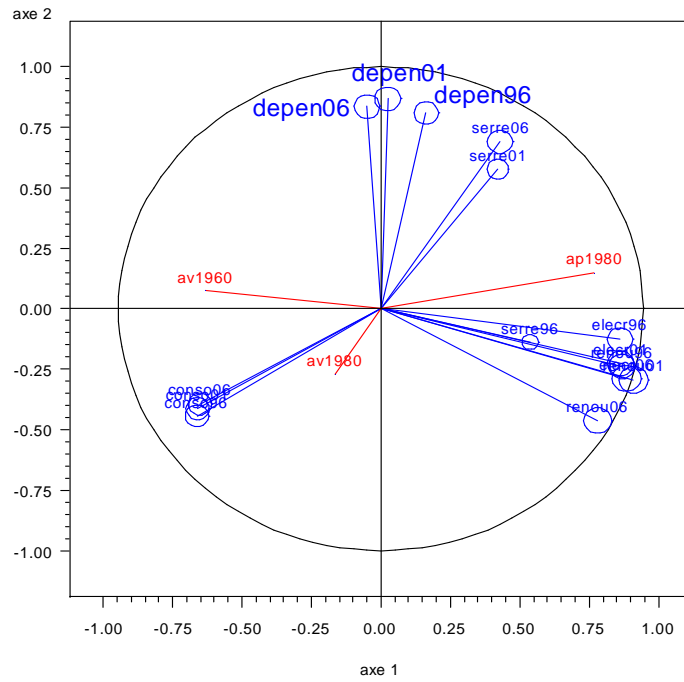
Représentation des variables numériques act - qualitatives sup -  
proportionnelle à la contribution à l'inertie expliquée par le plan



Variables supplémentaires en rouge

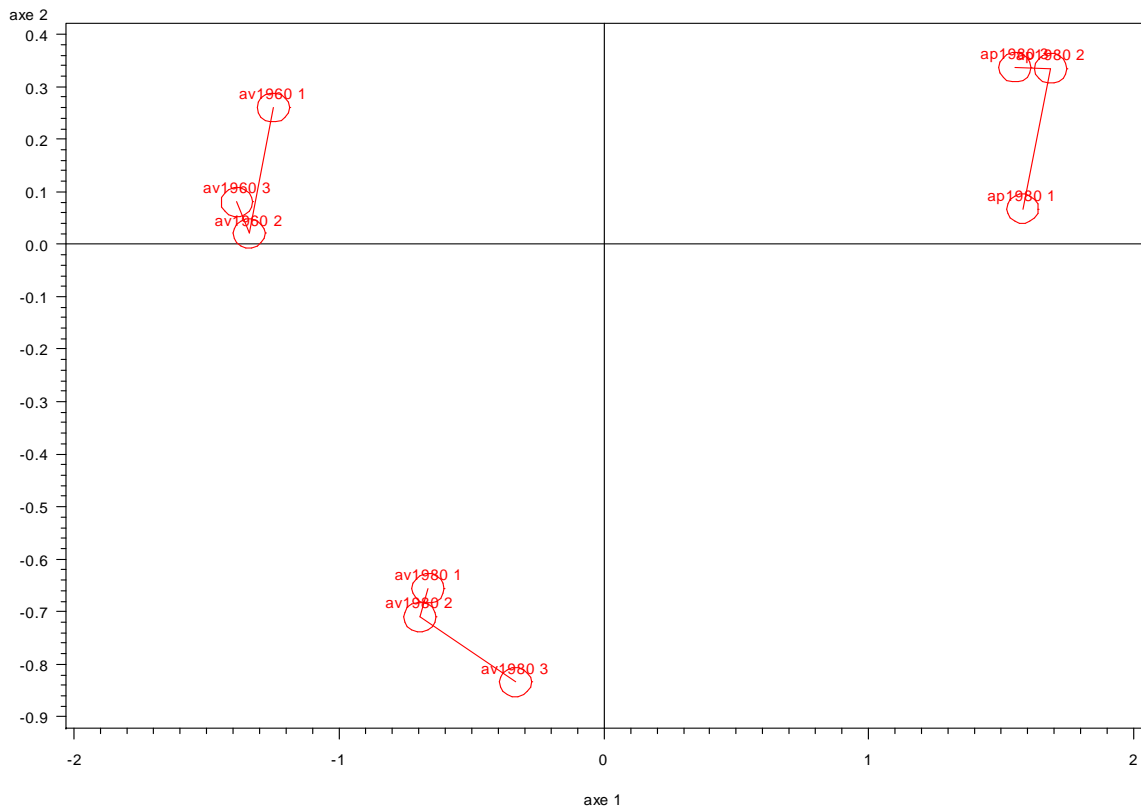
Et ci-dessous le même cercle après avoir choisi une taille de graphique plus grande et après avoir déplacé et grossi les étiquettes relatives à la variable depenXX.

Représentation des variables numériques act - qualitatives sup -  
proportionnelle à la contribution à l'inertie expliquée par le plan

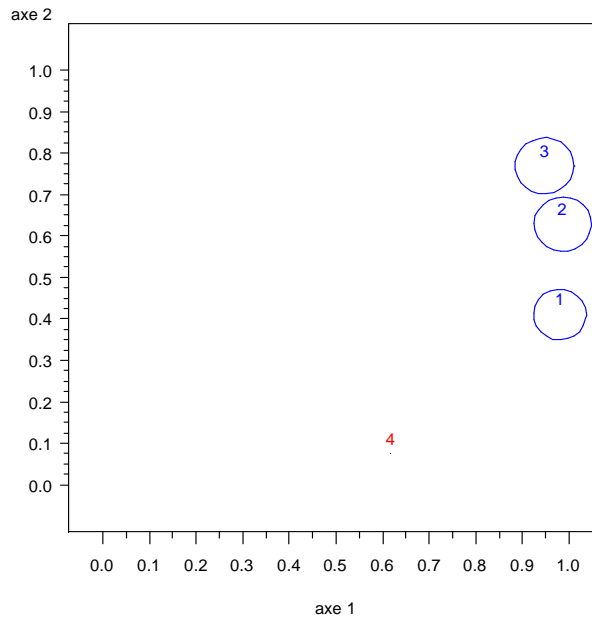


Variables supplémentaires en rouge

Représentation des centres partiels supplémentaires  
proportionnelle à la contribution à l'inertie expliquée par le plan

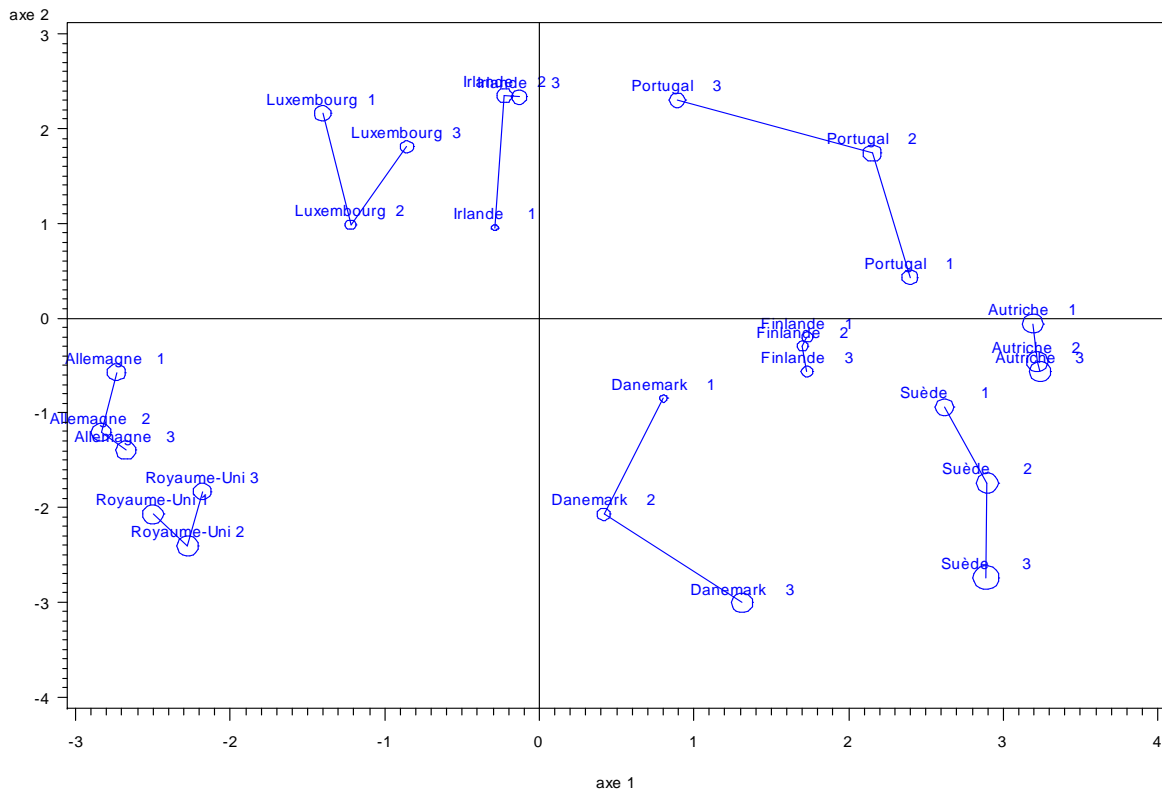


### Représentation des groupes actifs et supplémentaires proportionnelle à la contribution à l'inertie expliquée par le plan



groupes supplémentaires en rouge

### Représentation des individus partiels actifs proportionnelle à la contribution à l'inertie expliquée par le plan



9 individus actifs ayant la plus forte CTR

## Bibliographie

[1] Escofier B., Pagès J., *Analyses factorielles simples et multiples*, 4<sup>ème</sup> édition, 2008, Dunod.

[2] Gelein B., Sautory O., De nouvelles macros SAS d'analyse des données à l'Insee : comment réaliser une Analyse Factorielle Multiple, VIII<sup>èmes</sup> Journées de Méthodologie Statistique, 2002

[3] Groupe d'étude et de réflexion interrégional - GERI , *L'analyse des données évolutives : méthodes et applications*, 1996, Technip.

[4] Lebart L., Morineau A., Piron M. , *Statistique exploratoire multidimensionnelle*, 3<sup>ème</sup> édition, 2000, Dunod.

[5] SPAD TM, *Analyse des tableaux multiples*, 1997, CISIA.