

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

DIRECTION GÉNÉRALE

18, boulevard Adolphe Pinard - 75675 PARIS CEDEX 14

Les macros SAS

d'analyse des données

Document n° F 9405

15 mars 1994

Michel ISNARD, Olivier SAUTORY

**Série des documents de travail
de la Direction des Statistiques Démographiques et Sociales**

GÉNÉRALITÉS	4
AVERTISSEMENT	5
LA MACRO ACP	6
1. SYNTAXE DE LA MACRO	6
1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse.....	6
1.2. Paramètres spécifiant les sorties imprimées.....	8
1.3. Paramètres relatifs aux variables définissant des barycentres d'individus.....	9
1.4. Paramètres spécifiant la table en sortie	10
2. LES SORTIES IMPRIMÉES.....	11
3. LA TABLE SAS EN SORTIE	15
LA MACRO PLOTACP	17
LA MACRO AIDEAFC	19
1. SYNTAXE DE LA MACRO	19
2. LES SORTIES IMPRIMÉES.....	21
LA MACRO AIDEACM	24
1. SYNTAXE DE LA MACRO	24
2. LES SORTIES IMPRIMÉES.....	30
LA MACRO PLOTCOR	34
LA MACRO CAHNUM	36
1. SYNTAXE DE LA MACRO	36
1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse.....	36
1.2. Paramètres spécifiant les sorties imprimées.....	37
1.3. Paramètres divers	38
2. LES SORTIES IMPRIMÉES.....	39
3. LA TABLE SAS EN SORTIE	40
LA MACRO CAHQUAL	41
1. SYNTAXE DE LA MACRO	44
1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse.....	44
1.2. Paramètres spécifiant les sorties imprimées.....	45
1.3. Autre paramètre	46
2. LES SORTIES IMPRIMÉES.....	47
3. LA TABLE SAS EN SORTIE	48
LA MACRO PARTNUM	49
1. SYNTAXE DE LA MACRO	49
1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse.....	49
1.2. Paramètres caractérisant la partition à réaliser	50
1.3. Paramètres divers	51
2. LES SORTIES IMPRIMÉES.....	53
3. LA TABLE EN SORTIE	55

LA MACRO PARTQUAL	56
1. SYNTAXE DE LA MACRO	57
1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse.....	57
1.2. Paramètres caractérisant la partition à réaliser	57
1.3. Paramètres divers	58
2. LES SORTIES IMPRIMÉES.....	61
3. LA TABLE EN SORTIE	65
LA MACRO DESNUM	66
1. SYNTAXE DE LA MACRO	67
2. LES SORTIES IMPRIMÉES.....	69
LA MACRO DESQUAL	70
1. SYNTAXE DE LA MACRO	71
2. LES SORTIES IMPRIMÉES.....	73
EXEMPLES.....	76
1. ANALYSE EN COMPOSANTES PRINCIPALES	76
<i>Exemple 1 : Les vins français.....</i>	76
<i>Exemple 2 : Les coefficients budgétaires des ménages.....</i>	83
2. ANALYSE FACTORIELLE DES CORRESPONDANCES	90
<i>Exemple 1 : AFC sur un tableau de contingence (instruction VAR).....</i>	90
<i>Exemple 2 : AFC sur un tableau individus-variables (instruction TABLES)</i>	96
3. ANALYSE DES CORRESPONDANCES MULTIPLES	100
<i>Exemple 1 : ACM sur un tableau individus-variables (instruction TABLES)</i>	101
<i>Exemple 2 : ACM avec construction du tableau disjonctif complet</i>	105
<i>Exemple 3 : ACM sur un tableau disjonctif complet (instruction VAR)</i>	109
4. CLASSIFICATION ASCENDANTE HIÉRARCHIQUE	114
<i>Exemple 1 : tableau individus x variables numériques (les vins français).....</i>	114
<i>Exemple 2 : tableau de contingence (étrangers par CS)</i>	125
<i>Exemple 3 : tableaux de contingence juxtaposés (étrangers par CS et activité).....</i>	133
<i>Exemple 4 : tableau disjonctif complet (les chiens).....</i>	143
ANNEXE : LISTE DES PARAMÈTRES DES MACROS.....	156

Généralités

Les macros SAS présentées dans ce document permettent de réaliser des analyses factorielles et des classifications ascendantes hiérarchiques, avec édition d'aides à l'interprétation.

La liste des macros est la suivante :

ACP	réalise une analyse en composantes principales
PLOTACP	produit des représentations graphiques associées à une analyse en composantes principales
AIDEAFC	édite les aides à l'interprétation des résultats d'une analyse factorielle des correspondances (cette macro doit être précédée d'une procédure CORRESP)
AIDEACM	édite les aides à l'interprétation des résultats d'une analyse des correspondances multiples (cette macro doit être précédée d'une procédure CORRESP)
PLOTCOR	produit des représentations graphiques associées à une analyse factorielle des correspondances ou à une analyse des correspondances multiples
CAHNUM	réalise une CAH sur un tableau individus x variables numériques
CAHQVAL	réalise une CAH sur un tableau de contingence ou un tableau disjonctif complet (donc en utilisant des variables qualitatives)
PARTNUM	réalise une partition d'un ensemble d'individus à partir des résultats d'une CAH effectuée à l'aide de la macro CAHNUM, et édite une description de cette partition
PARTQUAL	réalise une partition d'un ensemble à partir des résultats d'une CAH effectuée à l'aide de la macro CAHQVAL, et édite une description de cette partition

Les deux macros suivantes peuvent être utilisées sans qu'il y ait eu une classification préalable :

DESNM	édite la description d'une partition d'un ensemble d'observations, issue ou non d'une CAH, à l'aide de variables numériques
DESQUAL	édite la description d'une partition d'un ensemble d'observations, issue ou non d'une CAH, à l'aide de variables qualitatives

Comment écrire les paramètres

Voici quelques règles relatives à l'écriture des paramètres :

- l'ordre dans lequel sont données les valeurs des paramètres n'a pas d'importance
- les paramètres doivent être séparés par des virgules (et non des points-virgules)
- certains paramètres prennent des valeurs par défaut (ces valeurs sont spécifiées dans la documentation) : ils peuvent donc être omis
- certains paramètres sont indiqués dans la documentation comme étant obligatoires : leur absence provoque l'arrêt de la macro, voire du programme SAS
- les paramètres mis à valeur manquante sont à proscrire
- l'écriture des paramètres est en format "libre" (on peut mettre des blancs où l'on veut), mais il ne faut jamais mettre de virgule dans la valeur d'un paramètre.

Certaines erreurs dans l'écriture des paramètres provoquent l'impression de messages d'erreur **sur la LOG** : ces messages commencent par le mot "Erreur"...

Les titres

Si l'on veut faire apparaître des titres en haut des pages contenant les sorties d'une macro, ces titres doivent **précéder** l'appel de la macro. D'autre part, les titres de niveaux 3 et suivants sont utilisés par les macros. L'utilisateur ne peut donc spécifier que des instructions TITLE ou TITLE2 : plus exactement, les titres figurant dans des instructions de type TITLE3, TITLE4... risquent à un moment d'être "écrasés" par ceux figurant dans les macros.

La macro ACP

La macro SAS **ACP** réalise une analyse en composantes principales "à la française" : elle utilise la procédure SAS PRINCOMP, mais les résultats sont édités sous la forme classique rencontrée dans les principaux logiciels d'analyse des données français (en particulier la macro édite des aides à l'interprétation). De plus, cette macro permet d'introduire des "points supplémentaires", qu'il s'agisse d'individus ou de variables. Les résultats de l'analyse peuvent être stockés dans une table SAS, et réutilisés par la suite par une autre procédure, ou par la macro PLOTACP.

Note : dans cette version de la macro ACP, le programme centre et réduit automatiquement les variables : on ne peut donc pas réaliser d'ACP "non réduite".

1. Syntaxe de la macro

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.

Remarques :

- on peut utiliser la clause WHERE, ou les options FIRSTOBS ou OBS, pour définir ces deux tables. Par exemple, on peut écrire :

```
DATAACT=A ( WHERE=( SEXE=" 2 " ) ) , DATASUP=B ( OBS=20 )
```

- il existe une deuxième possibilité pour définir des individus supplémentaires, en utilisant le paramètre POIDS (voir plus loin).

VARACT = liste de variables

liste des variables actives **numériques**, séparées par des blancs (**obligatoire**).

VARSUP = liste de variables

liste des variables supplémentaires **numériques**, séparées par des blancs.

ID = variable

variable servant à identifier les individus dans les éditions et récupérée dans la table en sortie (paramètre OUT) sous le nom `_LABEL_`.

Il s'agit obligatoirement d'une variable caractère.

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

Si la variable est de longueur supérieure à 8, elle est tronquée à 8 caractères.

Si ce paramètre est absent, les individus ne sont pas identifiés dans les éditions, et ne pourront pas l'être sur les graphiques ultérieurs.

Remarque : on rappelle que pour créer une variable caractère à partir d'une variable numérique, il suffit d'écrire une instruction de la forme suivante (X désignant ici une variable numérique prenant des valeurs entières comprises entre 0 et 999) :

```
Y=PUT(X,3.);
```

IDSUP = liste de variables

variables récupérées dans la table en sortie (pouvant en particulier servir d'identifiant pour les individus dans les graphiques).

Il s'agit obligatoirement de variables caractères.

Ces variables doivent figurer dans la (ou les deux) table(s) en entrée (paramètres DATAACT et DATASUP).

Si les variables sont de longueur supérieure à 8, elles sont tronquées à 8 caractères.

POIDS = variable

nom d'une variable numérique servant à pondérer les individus (les valeurs de cette variable peuvent ne pas être des entiers).

Un individu prenant une valeur négative, nulle ou manquante sur cette variable est traité comme individu supplémentaire.

Par défaut : tous les individus ont le même poids.

1.2. Paramètres spécifiant les sorties imprimées

IMPRESS = O ou N

si IMPRESS vaut O (=oui), il y a impression des sorties de la procédure SAS PRINCOMP : statistiques simples sur les variables, matrice des corrélations, tableau des valeurs propres, tableau des vecteurs propres (le tout en anglais).

Par défaut : pas d'impression (IMPRESS = N).

CORR = O ou N

si CORR vaut O (=oui), il y a impression de la matrice des corrélations entre les variables **actives** (si l'on souhaite obtenir les corrélations avec les variables supplémentaires, il faut utiliser une procédure CORR).

Par défaut : CORR = O.

VECP= n ou MAX

provoque l'édition des n premiers vecteurs propres de l'analyse, ou de tous les vecteurs propres si VECP = MAX.

Par défaut : VECP = 0 (pas d'édition).

IOA = n

les aides à l'interprétation relatives aux individus actifs¹ concernent les n premiers axes de l'analyse (coordonnées, CTR, CO2, QLT).

Par défaut : IOA = 0.

IOS = n

les aides à l'interprétation relatives aux individus supplémentaires concernent les n premiers axes de l'analyse (coordonnées, CO2, QLT).

Par défaut : IOS = 0.

IVA = n

les aides à l'interprétation relatives aux variables actives concernent les n premiers axes de l'analyse (coordonnées, CTR, CO2, QLT).

Par défaut : IVA = 0.

¹ Signification de ces paramètres et des suivants : I pour interprétation, O pour observations et V pour variables (dans la terminologie SAS), A pour actives et S pour supplémentaires.

IVS = n

les aides à l'interprétation relatives aux variables supplémentaires concernent les n premiers axes de l'analyse (coordonnées, CO2, QLT).

Par défaut : IVS = 0.

PARTIEL = n

dans les tableaux d'aides à l'interprétation **relatives aux individus**, un individu (actif ou supplémentaire) n'est édité que s'il figure parmi les n premiers points (au sens des CTR si le point est actif, au sens des CO2 si le point est supplémentaire) sur au moins un des axes édités.

Les tableaux d'aides à l'interprétation **relatives aux variables** sont toujours édités dans leur intégralité.

Il est recommandé d'utiliser ce paramètre lorsque le nombre d'individus est élevé.

Par défaut : PARTIEL = 0, ce qui signifie que tous les points sont édités.

LSIZE = longueur

valeur du paramètre "linesize" de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LSIZE = 132.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

1.3. Paramètres relatifs aux variables définissant des barycentres d'individus

CLASSES = liste de variables

liste de variables servant à définir des barycentres d'individus, séparées par des blancs.

Pour chacune de ces variables, on calcule pour chaque modalité le barycentre des individus prenant cette modalité ; ces barycentres sont des "individus supplémentaires" pour lesquels on édite les coordonnées sur les axes et des valeurs-tests (ainsi que les probabilités associées) jugeant de la significativité des points.

Sont également édités des tests de significativité globale de chaque variable.

Note : en présence d'une variable-POIDS et de ce paramètre CLASSES, la macro ne produit des résultats pertinents que si tous les individus actifs ont un poids égal à 1.

FMTCLAS = liste de formats

liste de formats associés aux variables de CLASSES, séparés par des blancs.

Ces formats, utilisés à l'édition et pour les graphiques ultérieurs éventuels, doivent être donnés dans l'ordre correspondant à l'ordre des variables de CLASSES (le nombre de formats peut être inférieur au nombre de variables).

AXECLAS = n

les éditions relatives aux variables de CLASSES concernent les n premiers axes de l'analyse.

Ce paramètre est obligatoire en présence du paramètre CLASSES.

1.4. Paramètres spécifiant la table en sortie

OUT = nom de table SAS

nom de la table SAS en sortie contenant les résultats de l'analyse : coordonnées des points sur les axes, CTR, CO2, QLT (voir § 3.).

NAXER = n

nombre d'axes récupérés dans la table OUT.

Ce paramètre est obligatoire en présence du paramètre OUT.

FILL = VAR et/ou IND et/ou BARY et/ou ALL

type de données récupérées dans la table OUT :

VAR données relatives aux variables (actives et supplémentaires)

IND données relatives aux individus (actifs et supplémentaires)

BARY données relatives aux barycentres d'individus

ALL données relatives aux variables, aux individus et aux barycentres.

Lorsque l'on veut récupérer plusieurs types de données, ces différents types doivent être séparés par des blancs ; par exemple :

FILL = VAR BARY

Par défaut : FILL = VAR.

2. Les sorties imprimées

La macro édite :

- un tableau récapitulatif donnant **les caractéristiques de l'analyse** :
 - le nombre de variables actives
 - le nombre de variables supplémentaires
 - le nombre de variables de classes
 - le nombre d'individus actifs : individus de la table-DATAACT n'ayant pas de valeur manquante sur les variables actives, et pour lesquels la variable de POIDS éventuelle est strictement positive
 - le nombre d'individus actifs éliminés : individus de la table-DATAACT ayant une valeur manquante sur l'une au moins des variables actives (quelle que soit la valeur de la variable de POIDS éventuelle)
 - le nombre d'individus supplémentaires : individus de la table-DATAACT n'ayant pas de valeur manquante sur les variables actives, et pour lesquels la variable de POIDS éventuelle est négative ou nulle, et individus de la table-DATASUP n'ayant pas de valeur manquante sur les variables actives
 - le nombre d'individus supplémentaires éliminés : individus de la table-DATASUP ayant une valeur manquante sur l'une au moins des variables actives
 - le nom de la variable de pondération éventuelle
 - la valeur du paramètre PARTIEL si ce paramètre a été renseigné
 - les valeurs des paramètres IVA, IVS, IOA, IOS, AXECLAS
 - la liste des variables actives
 - la liste éventuelle des variables supplémentaires
 - la liste éventuelle des variables de classes
- si CORR = 0 : la **matrice des corrélations** entre les variables actives
- si VECP = n : les n premiers **vecteurs propres** de l'analyse (dénommés PRIN1, PRIN2 ...)
- si VECP = MAX, tous les vecteurs propres sont imprimés
- **l'inertie totale**
- le tableau des **valeurs propres**, donnant, pour chaque axe :
 - la valeur propre, i.e. l'inertie expliquée par l'axe (**Val.Pr.**)
 - la différence entre la valeur propre et la valeur propre précédente (**Diff.**)
 - le pourcentage d'inertie expliqué par l'axe (**Pct**)
 - le pourcentage d'inertie cumulé, i.e. la part d'inertie expliquée par le sous-espace engendré par l'axe et les axes précédents (**Cum.**)

et un histogramme des valeurs propres.

Individus actifs

Si IOA > 0, la macro édite pour chaque individu actif (sélectionné, si PARTIEL > 0) :

- son **identifiant**, i.e. la valeur de la variable spécifiée dans le paramètre ID ; si ce paramètre est absent, l'identifiant est à blanc
- **CONTR** : contribution relative à l'inertie totale du nuage (en %)
- **POIDS** : poids dans le nuage (en %)
- pour chacun des n premiers axes (n est la valeur du paramètre IOA) :

COORD	coordonnée sur l'axe
CTR	contribution relative à l'inertie expliquée par l'axe (en %)
RCTR	rang du point, les points étant classés par CTR décroissante
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

Individus supplémentaires

Si IOS > 0, la macro édite pour chaque individu supplémentaire (sélectionné, si PARTIEL > 0) :

- son **identifiant**, i.e. la valeur de la variable spécifiée dans le paramètre ID ; si ce paramètre est absent, l'identifiant est à blanc
- **DIST** : carré de la distance entre le point et l'origine
- pour chacun des n premiers axes (n est la valeur du paramètre IOS) :

COORD	coordonnée sur l'axe
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
RCO2	rang du point, les points étant classés par CO2 décroissant
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

Barycentres d'individus

Si le paramètre CLASSES est renseigné, la macro édite pour chaque variable spécifiée dans ce paramètre, et pour chacun des n premiers axes (n est la valeur du paramètre AXECLAS) :

- le rapport de corrélation (**RCOR**), i.e. le rapport variance interclasse / variance totale²
- la statistique du test de Fisher (**FISH**) : il s'agit du test d'analyse de variance à 1 facteur, qui permet de juger si les coordonnées sur l'axe des barycentres d'individus (définis par les modalités de la variable) sont significativement distinctes ou non les unes des autres
- le niveau de significativité du test de Fisher (**PROB**) : une valeur faible de PROB signifie que l'hypothèse d'égalité des coordonnées des barycentres doit être rejetée
- pour chaque modalité de la variable :
 - son "nom", formé par la concaténation de la modalité, du "blanc souligné", et du nom de la variable, le tout étant tronqué à 8 caractères (par exemple, la modalité 1 de la variable SEXE se voit attribuer le nom 1_SEXE), ou bien sa valeur formatée si l'on a associé un format à la variable avec le paramètre FMTCLAS
 - son **EFFECTIF**, i.e. le nombre d'individus actifs prenant cette modalité
 - son **POIDS**, i.e. la proportion (en %) d'individus actifs prenant cette modalité
 - la coordonnée sur l'axe (**COORD**)
 - la statistique du test de nullité de la coordonnée (**VALTEST**) qui permet de juger si le point est significativement ou non différent de l'origine
 - le niveau de significativité du test (**PROB**) : une valeur faible de PROB signifie que l'on doit rejeter l'hypothèse de nullité de la coordonnée.

Variables actives

Si IVA > 0, la macro édite pour chaque variable active :

- son **identifiant**, i.e. le nom de la variable
- **CONTR** : contribution relative à l'inertie totale du nuage (en %)
- **POIDS** : poids dans le nuage (en %)
- pour chacun des n premiers axes (n est la valeur du paramètre IVA) :

COORD coordonnée sur l'axe

CTR contribution relative à l'inertie expliquée par l'axe (en %)

RCTR rang du point, les points étant classés par CTR décroissante

² La variance inter-classe est la variance des moyennes par modalité des coordonnées sur l'axe, la variance totale des coordonnées est égale à l'inertie expliquée par l'axe.

CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

Variables supplémentaires

Si $IVS > 0$, la macro édite pour chaque variable supplémentaire :

- son **identifiant**, i.e. le nom de la variable
- **DIST** : carré de la distance entre le point et l'origine
- pour chacun des n premiers axes (n est la valeur du paramètre IVS) :

COORD	coordonnée sur l'axe
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
RCO2	rang du point, les points étant classés par CO2 décroissant
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

3. La table SAS en sortie

Si le paramètre OUT est renseigné, la macro crée une table SAS, contenant les variables suivantes :

- la variable **_TYPE_** donne le type de l'observation de la table.

Les valeurs de cette variable sont les suivantes :

- l'observation **VALP** contient les valeurs propres
- les observations **AXEUNI** contiennent les coordonnées de points permettant de dessiner sur les graphiques les projections des "axes unitaires" (et donc de "reconstituer" les données par projection des individus sur ces axes)³
- les observations **VARACT** contiennent les coordonnées, les CTR, les CO2, les QLT, et la distance à l'origine pour les variables actives (si FILL = VAR ou ALL)
- les observations **VARSUP** contiennent les coordonnées, les CO2, les QLT, et la distance à l'origine pour les variables supplémentaires (si FILL = VAR ou ALL)
- les observations **OBSACT** contiennent les coordonnées, les CTR, les CO2, les QLT, et la distance à l'origine pour les individus actifs (si FILL = IND ou ALL)
- les observations **OBSSUP** contiennent les coordonnées, les CO2, les QLT, et la distance à l'origine pour les individus supplémentaires (si FILL = IND ou ALL), et les coordonnées seulement pour les barycentres d'individus générés par le paramètre CLASSES (si FILL = BARY ou ALL)
- la variable **_LABEL_** identifie l'observation ; les valeurs de cette variable dépendent du type de l'observation : si **_TYPE_** vaut

VALP **_LABEL_** est à blanc

AXEUNI **_LABEL_** contient le nom de la variable (définissant l'axe unitaire considéré) précédé d'un signe +

VARACT **_LABEL_** contient le nom de la variable

VARSUP **_LABEL_** contient le nom de la variable

OBSACT **_LABEL_** contient l'identifiant de l'individu, i.e. la valeur de la variable spécifiée dans le paramètre ID ; si ce paramètre est absent, **_LABEL_** est à blanc

OBSSUP **_LABEL_** contient l'identifiant de l'individu, i.e. la valeur de la variable spécifiée dans le paramètre ID ; si ce paramètre est absent, **LABEL** est à blanc.

Pour un "individu supplémentaire" généré par le paramètre CLASSES, i.e. associé à une modalité d'une variable spécifiée dans ce paramètre, **_LABEL_** contient le "nom" de cette modalité tel qu'il figure à l'impression (i.e. formé par la concaténation de la modalité, du signe **_**, et de la variable), ou bien sa valeur formatée

³ Les coordonnées de ces points n'ont pas de signification, seuls les axes sont significatifs.

- les variables **COORD1 ... COORD_n** (où **n** est la valeur du paramètre **NAXER**) contiennent les coordonnées des points sur les axes 1...**n**, sauf pour l'observation **_TYPE_=VALP** : ces variables contiennent dans ce cas les **n** premières valeurs propres
- les variables **CTR_1 ... CTR__n** contiennent les contributions relatives à l'inertie pour les points actifs (i.e. pour les observations **_TYPE_=VARACT** ou **_TYPE_=OBSACT**) ; elles sont à valeur manquante pour les autres observations
- les variables **CO2_1 ... CO2__n** contiennent les cosinus carrés ; elles sont à valeur manquante pour les observations **_TYPE_=VALP**, **_TYPE_=AXEUNI**, et pour les observations **_TYPE_=OBSSUP** relatives aux "individus supplémentaires" générés par le paramètre **CLASSES**
- les variables **QLT_1 ... QLT__n** contiennent les qualités de représentation sur les sous-espaces ; elles sont à valeur manquante pour les observations **_TYPE_=VALP**, **_TYPE_=AXEUNI**, et pour les observations **_TYPE_=OBSSUP** relatives aux "individus supplémentaires" générés par le paramètre **CLASSES**
- la variable **DIST** contient le carré de la distance du point à l'origine ; elle est à valeur manquante pour les observations **_TYPE_=VALP**, **_TYPE_=AXEUNI**, et pour les observations **_TYPE_=OBSSUP** relatives aux "individus supplémentaires" générés par le paramètre **CLASSES**
- les variables spécifiées le cas échéant dans le paramètre **IDSUP**.

La macro PLOTACP

La macro SAS **PLOTACP** produit des représentations graphiques associées à une analyse en composantes principales réalisée à partir de la macro ACP. Elle utilise la procédure SAS PLOT. Les graphiques peuvent représenter les individus actifs, les individus supplémentaires (y compris les barycentres d'individus), les variables actives, les variables supplémentaires, et les "axes unitaires", i.e. les projections des axes de départ de l'espace des individus (axes représentant les variables actives de l'analyse) sur le plan factoriel considéré.

Cette macro utilise la table en sortie créée par une macro ACP.

Syntaxe de la macro

DATA = nom de table SAS

nom de la table SAS en sortie créée par la macro ACP (paramètre OUT).

Par défaut : nom de la dernière table SAS créée ou utilisée.

AXEH = n

numéro de l'axe horizontal (**obligatoire**).

AXEV = n

numéro de l'axe vertical (**obligatoire**).

POINTS = OBSACT et/ou OBSSUP et/ou VARSUP et/ou VARACT et/ou AXEUNI

indique les points représentés sur le graphique (**obligatoire**) :

OBSACT	individus actifs
OBSSUP	individus supplémentaires (y compris les barycentres d'individus générés par le paramètre CLASSES de la macro ACP)
VARACT	variables actives
VARSUP	variables supplémentaires

AXEUNI "axes unitaires"

i.e. les projections des axes de départ de l'espace des individus (axes représentant les variables actives de l'analyse) sur le plan factoriel considéré.

Note : ne sont représentés sur les graphiques que des points permettant de tracer ces axes, identifiés par les noms des variables, précédés d'un signe "+".

Lorsque que l'on demande la représentation simultanée de plusieurs types de points, ces différents types doivent être séparés par des blancs ; par exemple :

```
POINTS = OBSACT OBSSUP AXEUNI
```

IDSUP = variable

variable servant à identifier les points sur les graphiques.

Si on ne spécifie pas ce paramètre, les points sont identifiés sur les graphiques de la façon suivante :

- les individus actifs et supplémentaires par les valeurs de la variable spécifiée dans le paramètre ID de la macro ACP⁴
- les barycentres d'individus par les modalités (ou leurs formats associés) des variables de classe correspondantes
- les variables actives et supplémentaires par leurs noms
- les axes unitaires par les noms des variables correspondantes précédés d'un "+".

Remarque : c'est la variable `_LABEL_` de la table en sortie créée par la macro ACP qui contient tous ces identifiants.

Si on spécifie une variable, celle-ci doit être l'une des variables récupérées dans la table en sortie créée par la macro ACP, par le paramètre IDSUP. Dans ce cas, les individus actifs et supplémentaires sont représentés par les valeurs qu'ils prennent pour la variable spécifiée ; les autres points sont représentés comme indiqué ci-dessus.

CERCLE = O ou N

Si CERCLE vaut O (= oui), le cercle des corrélations (i.e. le cercle de rayon 1) est tracé sur le graphique.

Par défaut : CERCLE = O lorsque l'on représente des variables (i.e. si VARACT ou VARSUP figurent dans la liste du paramètre POINTS) ; **sinon CERCLE = N.**

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

⁴ Si le paramètre ID de la macro ACP n'est pas renseigné, les individus n'apparaissent pas sur le graphique.

La macro AIDEAFC

La macro SAS AIDEAFC édite les aides à l'interprétation des résultats d'une analyse factorielle des correspondances (AFC) sous la forme classique rencontrée dans les principaux logiciels français d'analyse des données.

Cette macro n'effectue pas une AFC : elle doit être précédée d'une procédure CORRESP. La macro utilise la table-OUTC créée par cette procédure, et qui contient les résultats de l'AFC.

On peut demander les aides à l'interprétation pour les d premiers axes (au plus), où d est le paramètre spécifié dans l'option DIMENS de la PROC CORRESP.

La syntaxe de la procédure CORRESP est décrite dans le document suivant : **La procédure CORRESP (Version 6.07), Document n° F 9116 (3ème version)**, disponible auprès du secrétariat de l'Unité Méthodes Statistiques (DG Insee, bureau 730, tél. : 01.41.17.53.76).

1. Syntaxe de la macro

DATA = nom de table SAS

nom de la table SAS contenant les résultats de l'AFC réalisée auparavant à l'aide de la procédure CORRESP : il s'agit de la table créée par l'option OUTC = . **Il est donc nécessaire d'utiliser cette option dans l'instruction PROC CORRESP.**

Ce paramètre est obligatoire.

ID = variable

paramètre à renseigner **dans le cas où une instruction ID figure dans la procédure CORRESP** : il faut indiquer ici le nom de la variable spécifiée dans cette instruction.

IOA = n

les aides à l'interprétation relatifs aux modalités-lignes actives⁵ concernent les n premiers axes de l'analyse (coordonnées, CTR, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles _TYPE_ = OBS.

Par défaut : IOA = 0.

⁵ Signification de ces paramètres et des suivants : I pour interprétation, O pour observations et V pour variables (dans la terminologie SAS), A pour actives et S pour supplémentaires.

IOS = n

les aides à l'interprétation relatifs aux modalités-lignes supplémentaires concernent les n premiers axes de l'analyse (coordonnées, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles `_TYPE_ = SUPOBS`.

Par défaut : IOS = 0.

IVA = n

les aides à l'interprétation relatifs aux modalités-colonnes actives concernent les n premiers axes de l'analyse (coordonnées, CTR, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles `_TYPE_ = VAR`.

Par défaut : IVA = 0.

IVS = n

les aides à l'interprétation relatifs aux modalités-colonnes supplémentaires concernent les n premiers axes de l'analyse (coordonnées, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles `_TYPE_ = SUPVAR`.

Par défaut : IVS = 0.

PARTIEL = n

dans les tableaux d'aides à l'interprétation, un point (modalité-ligne ou modalité-colonne, actif ou supplémentaire) n'est édité que s'il figure parmi les n premiers points (au sens des CTR si le point est actif, au sens des CO2 si le point est supplémentaire) sur au moins un des axes édités.

Par défaut : PARTIEL = 0, ce qui signifie que tous les points sont édités.

LISTE = O ou N

si LISTE vaut O (=oui), la macro édite les listes des modalités (lignes et colonnes, actives et supplémentaires)

Par défaut : LISTE = O.

LSIZE = longueur

valeur du paramètre "linesize" de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LSIZE = 132.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

2. Les sorties imprimées

La macro édite :

- un tableau récapitulatif donnant **les caractéristiques de l'analyse** :
 - le nombre de modalités-colonnes actives
 - le nombre de modalités-colonnes supplémentaires
 - le nombre de modalités-lignes actives
 - le nombre de modalités-lignes supplémentaires
 - la valeur du paramètre PARTIEL si ce paramètre a été renseigné
 - les valeurs des paramètres IVA, IVS, IOA, IOS

si le paramètre LISTE vaut 0 :

- la liste des modalités-lignes actives
 - la liste éventuelle des modalités-lignes supplémentaires
 - la liste des modalités-colonnes actives
 - la liste éventuelle des modalités-colonnes supplémentaires
- **l'inertie totale**
 - le tableau des **valeurs propres**, donnant, pour chacun des d premiers axes (où d est le paramètre spécifié dans l'option DIMENS de la PROC CORRESP) :
 - la valeur propre, i.e. l'inertie expliquée par l'axe (**Val.Pr.**)
 - la différence entre la valeur propre et la valeur propre précédente (**Diff.**)
 - le pourcentage d'inertie expliqué par l'axe (**Pct**)
 - le pourcentage d'inertie cumulé, i.e. la part d'inertie expliquée par le sous-espace engendré par l'axe et les axes précédents (**Cum.**)

et un histogramme des valeurs propres.

Modalités-lignes actives

Si IOA > 0, la macro édite pour chaque modalité-ligne active (sélectionnée, si PARTIEL > 0) :

- son **identifiant**, i.e. :
 - la modalité de la variable-ligne, éventuellement formatée, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction TABLES (analyse d'une table sous la forme d'un tableau individus-variables)
 - la valeur de la variable-ID dans le cas où l'on a utilisé, dans la procédure CORRESP, les instructions VAR et ID (analyse d'une table déjà sous la forme d'un tableau de contingence)

- **CONTR** : contribution relative à l'inertie totale du nuage (en %)
- **POIDS** : poids dans le nuage (en %)
- pour chacun des n premiers axes (n est la valeur du paramètre IOA)

COORD	coordonnée sur l'axe
CTR	contribution relative à l'inertie expliquée par l'axe (en %)
RCTR	rang du point, les points étant classés par CTR décroissante
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

Modalités-lignes supplémentaires

Si IOS > 0, la macro édite pour chaque modalité-ligne supplémentaire (sélectionnée, si PARTIEL > 0) :

- son **identifiant**, i.e. :
 - la modalité de la variable-ligne, éventuellement formatée, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction TABLES (analyse d'une table sous la forme d'un tableau individus-variables)
 - la valeur de la variable-ID dans le cas où l'on a utilisé, dans la procédure CORRESP, les instructions VAR et ID (analyse d'une table déjà sous la forme d'un tableau de contingence)
- pour chacun des n premiers axes (n est la valeur du paramètre IOS)

COORD	coordonnée sur l'axe
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
RCO2	rang du point, les points étant classés par CO2 décroissant
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %)

Modalités-colonnes actives

Si IVA > 0, la macro édite pour chaque modalité-colonne active (sélectionnée, si PARTIEL > 0) :

- son **identifiant**, i.e. :
 - la modalité de la variable-colonne, éventuellement formatée, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction TABLES (analyse d'une table sous la forme d'un tableau individus-variables)

- le nom de la variable SAS attribué à cette modalité, dans le cas où l'on a utilisé, dans la procédure CORRESP, les instructions VAR et ID (analyse d'une table déjà sous la forme d'un tableau de contingence)
- **CONTR** : contribution relative à l'inertie totale du nuage (en %)
- **POIDS** : poids dans le nuage (en %)
- pour chacun des n premiers axes (n est la valeur du paramètre IVA)

COORD	coordonnée sur l'axe
CTR	contribution relative à l'inertie expliquée par l'axe (en %)
RCTR	rang du point, les points étant classés par CTR décroissante
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

Modalités-colonnes supplémentaires

Si $IVS > 0$, la macro édite pour chaque modalité-colonne supplémentaire (sélectionnée, si $PARTIEL > 0$) :

- son **identifiant**, i.e. :
 - la modalité de la variable-colonne, éventuellement formatée, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction TABLES (analyse d'une table sous la forme d'un tableau individus-variables)
 - le nom de la variable SAS attribué à cette modalité, dans le cas où l'on a utilisé, dans la procédure CORRESP, les instructions VAR et ID (analyse d'une table déjà sous la forme d'un tableau de contingence)
- pour chacun des n premiers axes (n est la valeur du paramètre IVS)

COORD	coordonnée sur l'axe
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
RCO2	rang du point, les points étant classés par CO2 décroissant
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

La procédure CORRESP édite également la plupart de ces statistiques, mais sous une présentation différente... et plus étendue. Pour éviter des listings trop longs, il est recommandé d'utiliser l'option **SHORT** de cette procédure, qui provoque l'édition des valeurs propres et des coordonnées des points sur les axes, associée avec les options RP et CP, qui fournissent les tableaux des profils des lignes et des profils des colonnes (pour les variables actives et supplémentaires), dont on ne dira jamais assez l'utilité pour l'interprétation d'une AFC...

La macro AIDEACM

La macro SAS **AIDEACM** édite les aides à l'interprétation des résultats d'une analyse des correspondances multiples (ACM) sous la forme classique rencontrée dans les principaux logiciels français d'analyse des données.

Cette macro n'effectue pas une ACM : elle doit être précédée d'une procédure CORRESP. La macro utilise la table-OUTC créée par cette procédure, et qui contient les résultats de l'ACM.

On peut demander les aides à l'interprétation pour les *d* premiers axes (au plus), où *d* est le paramètre spécifié dans l'option DIMENS de la PROC CORRESP.

La syntaxe de la procédure CORRESP est décrite dans le document suivant : **La procédure CORRESP (Version 6.07), Document n° F 9116 (3ème version)**, disponible auprès du secrétariat de l'Unité Méthodes Statistiques (DG Insee, bureau 730, tél. : 01.41.17.53.76).

1. Syntaxe de la macro

DATA = nom de table SAS

nom de la table SAS contenant les résultats de l'ACM réalisée auparavant à l'aide de la procédure CORRESP : il s'agit de la table créée par l'option OUTC = . **Il est donc nécessaire d'utiliser cette option dans l'instruction PROC CORRESP**

Ce paramètre est obligatoire.

DATAINIT = nom de table SAS

nom de la table SAS sur laquelle a opéré la procédure CORRESP : il faut indiquer ici le nom de la table spécifiée dans le paramètre DATA de l'instruction PROC CORRESP, avec le cas échéant les options FIRSTOBS, OBS ... et la clause WHERE.

Ce paramètre est obligatoire.

ID = variable

paramètre à renseigner **dans le cas où une instruction ID figure dans la procédure CORRESP** : il faut indiquer ici le nom de la variable spécifiée dans cette instruction.

IOA = n

les aides à l'interprétation relatives aux individus actifs⁶ concernent les n premiers axes de l'analyse (coordonnées, CTR, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles _TYPE_ = OBS.

Par défaut : IOA = 0.

IOS = n

les aides à l'interprétation relatives aux individus supplémentaires concernent les n premiers axes de l'analyse (coordonnées, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles _TYPE_ = SUPOBS.

Par défaut : IOS = 0.

IVA = n

les aides à l'interprétation relatives aux variables actives concernent les n premiers axes de l'analyse (coordonnées, CTR, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles _TYPE_ = VAR.

Par défaut : IVA = 0.

IVS = n

les aides à l'interprétation relatives aux variables supplémentaires concernent les n premiers axes de l'analyse (coordonnées, CO2, QLT) ; il s'agit des observations de la table-OUTC pour lesquelles _TYPE_ = SUPVAR.

Par défaut : IVS = 0

WEIGHT = nom de variable

si une instruction WEIGHT a été utilisée dans la procédure CORRESP, il faut indiquer le nom de la variable de pondération utilisée (qui a pu servir en particulier à définir des individus supplémentaires).

PARTIEL = n

dans les tableaux d'aides à l'interprétation **relatives aux individus**, un individu (actif ou supplémentaire) n'est édité que s'il figure parmi les n premiers points (au sens des CTR si le point est actif, au sens des CO2 si le point est supplémentaire) sur au moins un des axes édités.

Les tableaux d'aides à l'interprétation **relatives aux variables** sont toujours édités dans leur intégralité.

Il est recommandé d'utiliser ce paramètre lorsque le nombre d'individus est élevé.

Par défaut : PARTIEL = 0, ce qui signifie que tous les individus sont édités.

⁶ Signification de ces paramètres et des suivants : I pour interprétation, O pour observations et V pour variables (dans la terminologie SAS), A pour actives et S pour supplémentaires.

MISSING = OUI ou NON

il faut spécifier MISSING = OUI lorsque l'on a utilisé l'option MISSING dans la procédure CORRESP précédant l'appel de la macro.

Par défaut : MISSING = NON.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

LSIZE = longueur

valeur du paramètre "linesize" de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LSIZE = 132

Si on a utilisé une instruction TABLES dans la procédure CORRESP

Lorsque l'on a spécifié une instruction TABLES dans la procédure CORRESP précédant l'appel de la macro, les deux paramètres suivants doivent être **obligatoirement** renseignés, et les deux suivants le cas échéant.

ANALYSE = TABLES

indique le type d'analyse réalisé avec la procédure CORRESP.

VARACT = liste de noms de variables

liste des noms des **variables actives** séparés par des blancs ; l'ordre dans lequel sont donnés ces noms doit être le même que l'ordre dans lequel ont été spécifiées les variables actives dans l'instruction TABLES.

Par exemple, si on a spécifié les instructions suivantes dans la procédure CORRESP :

```
TABLES AGE TAILLE SEXE POIDS NIVEAU;  
SUPPLEMENTARY AGE SEXE;
```

le paramètre VARACT s'écrit :

```
VARACT = TAILLE POIDS NIVEAU
```

VARSUP = liste de noms de variables

liste des noms des **variables supplémentaires** séparés par des blancs ; l'ordre dans lequel sont donnés ces noms doit être le même que l'ordre dans lequel ont été spécifiées les variables supplémentaires **dans l'instruction TABLES**.

Par exemple, si on a spécifié les instructions suivantes dans la procédure CORRESP :

```
TABLES AGE TAILLE SEXE POIDS NIVEAU;  
SUPPLEMENTARY SEXE AGE;
```

le paramètre VARSUP s'écrit :

```
VARSUP = AGE SEXE
```

FORMAT = instruction format

si une instruction FORMAT a été utilisée dans la procédure CORRESP, il faut répéter ici cette instruction à **l'identique** (sans le mot-clé FORMAT et sans le point-virgule).

Si on a utilisé une instruction VAR dans la procédure CORRESP

Lorsque l'on a spécifié une instruction VAR dans la procédure CORRESP précédant l'appel de la macro (i.e. lorsque la procédure opère sur le tableau disjonctif complet préalablement construit), les trois paramètres suivants doivent être **obligatoirement** renseignés, ainsi que les deux derniers en présence de variables supplémentaires.

ANALYSE = VAR

indique le type d'analyse réalisé avec la procédure CORRESP.

VARACT = liste de noms de variables

liste des noms des **variables actives** séparés par des blancs.

On donne ici des **noms** aux variables qui seront utilisés dans les tableaux d'aides à l'interprétation. Ces noms ne doivent pas dépasser 8 caractères ; la liste doit comporter autant de noms qu'il y a de variables actives (on ne peut pas utiliser d'écriture simplifiée de la forme X1-X8 ou A--Z).

Ces noms peuvent être différents des noms des variables-SAS correspondantes figurant quelque part dans une table SAS (par exemple dans la table-DATAINIT, lorsque cette table contient ces variables et les variables indicatrices associées qui définissent le tableau disjonctif complet).

L'ordre dans lequel sont donnés ces noms doit être le même que l'ordre dans lequel ont été spécifiées les modalités de ces variables dans l'instruction VAR de la procédure CORRESP.

Il est donc nécessaire que dans l'instruction VAR de la procédure CORRESP les modalités apparaissent regroupées par variable.

Par exemple, si la variable X a deux modalités X1 et X2, la variable Y a trois modalités Y1, Y2 et Y3, et la variable Z deux modalités Z1 et Z2, il faut avoir spécifié dans la procédure CORRESP une instruction VAR de la forme :

```
VAR X1 X2 Y1-Y3 Z1 Z2;
```

Le paramètre VARACT de la macro peut alors être de la forme :

```
VARACT = VARX Y VARZ
```

NBMODACT = nombres des modalités des variables actives

Pour chacune des variables actives de la liste VARACT, on indique le nombre de modalités⁷ de la variable ; ces nombres de modalités sont donnés dans le même ordre que dans la liste VARACT, séparés par des blancs.

Dans l'exemple précédent, on indique :

```
NBMODACT = 2 3 2
```

⁷ Il s'agit des modalités prises par au moins un individu actif.

VARSUP = liste de noms de variables

liste des noms des **variables supplémentaires** séparés par des blancs.

On donne ici des **noms** aux variables qui seront utilisés dans les tableaux d'aides à l'interprétation. Ces noms ne doivent pas dépasser 8 caractères ; la liste doit comporter autant de noms qu'il y a de variables actives (on ne peut pas utiliser d'écriture simplifiée de la forme X1-X8 ou A--Z).

Ces noms peuvent être différents des noms des variables-SAS correspondantes figurant quelque part dans une table SAS (par exemple dans la table-DATAINIT, lorsque cette table contient ces variables et les variables indicatrices associées qui définissent le tableau disjonctif complet).

L'ordre dans lequel sont donnés ces noms doit être le même que l'ordre dans lequel ont été spécifiées les modalités de ces variables dans l'instruction VAR de la procédure CORRESP.

Il est donc nécessaire que dans l'instruction VAR de la procédure CORRESP les modalités apparaissent regroupées par variable.

Par exemple, si la variable X a deux modalités X1 et X2, la variable Y a trois modalités Y1, Y2 et Y3, la variable Z deux modalités Z1 et Z2, la variable T deux modalités T1 et T2, la variable U trois modalités U1 et U2, et si on a spécifié les instructions suivantes dans la procédure CORRESP :

```
VAR X1 X2 Y1-Y3 T1 T2 Z1 Z2 U1 U2;  
SUPPLEMENTARY T1 T2 U1 U2;
```

le paramètre VARSUP de la macro peut alors être de la forme :

```
VARSUP = T U
```

NBMODSUP = nombres des modalités des variables supplémentaires

Pour chacune des variables supplémentaires de la liste VARSUP, on indique le nombre de modalités⁸ de la variable ; ces nombres de modalités sont donnés dans le même ordre que dans la liste VARSUP, séparés par des blancs.

Dans l'exemple précédent, on indique :

```
NBMODSUP = 2 2
```

Remarque

L'exemple suivant montre le type de syntaxe à utiliser dans le cas où l'analyse porte sur un tableau disjonctif complet et où certaines, mais pas l'ensemble, des modalités d'une variable sont mises en supplémentaires (ce qui n'est d'ailleurs pas tout à fait légitime pour une ACM, puisque le tableau analysé n'est plus dans ce cas un tableau disjonctif "complet").

Si on a spécifié les instructions suivantes dans la procédure CORRESP :

```
VAR X1-X3 Y1-Y4 Z1-Z3 T1-T6;  
SUPPLEMENTARY Z1-Z3 T5 T6;
```

on spécifie dans la macro :

```
VARACT = X Y T,  
NBMODACT = 3 4 4,  
VARSUP = Z T,  
NBMODSUP = 3 2
```

⁸ Il s'agit des modalités prises par au moins un individu actif.

2. Les sorties imprimées

La macro édite :

- un tableau récapitulatif donnant **les caractéristiques de l'analyse** :
 - le "type" de l'analyse : TABLES ou VAR
 - le nombre de variables actives
 - le nombre de modalités actives
 - le nombre de variables supplémentaires
 - le nombre de modalités supplémentaires
 - la somme des poids des individus actifs, si le paramètre WEIGHT est renseigné, ou le nombre d'individus actifs dans le cas contraire
 - de plus, si la procédure CORRESP a permis d'obtenir des résultats sur les individus, la macro édite le nombre d'individus actifs (non pondérés, même si WEIGHT est renseigné) et le nombre d'individus supplémentaires
 - le nom de la variable de pondération éventuelle
 - la valeur du paramètre PARTIEL si ce paramètre a été renseigné
 - les valeurs des paramètres IVA, IVS, IOA, IOS
 - la liste des modalités actives
 - la liste éventuelle des modalités supplémentaires
 - **l'inertie totale**
 - le tableau des **valeurs propres**, donnant, pour chacun des d premiers axes (où d est le paramètre spécifié dans l'option DIMENS de la PROC CORRESP) :
 - la valeur propre, i.e. l'inertie expliquée par l'axe (**Val.Pr.**)
 - la différence entre la valeur propre et la valeur propre précédente (**Diff.**)
 - le pourcentage d'inertie expliqué par l'axe (**Pct**)
 - le pourcentage d'inertie cumulé, i.e. la part d'inertie expliquée par le sous-espace engendré par l'axe et les axes précédents (**Cum.**)
- et un histogramme des valeurs propres.

Variables actives

Si $IVA > 0$, pour chaque variable active, la macro édite pour chacune de ses modalités :

- son **identifiant**, i.e. :
 - la modalité de la variable, éventuellement formatée, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction TABLES (analyse d'une table sous la forme d'un tableau individus-variables)
 - le nom de la "variable"-SAS désignant la modalité, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction VAR (analyse d'une table déjà sous la forme d'un tableau disjonctif complet)
- **CONTR** : contribution relative à l'inertie totale du nuage (en %)
- **POIDS** : poids dans le nuage (en %)
- pour chacun des n premiers axes (n est la valeur du paramètre IVA)

COORD coordonnée sur l'axe

CTR contribution relative à l'inertie expliquée par l'axe (en %)

RCTR rang du point, les points étant classés par CTR décroissante

CO2 qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)

QLT qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %)

V.TEST valeur-test sur l'axe.

Ces valeurs-test sont en principe réservées aux modalités supplémentaires, elles sont données ici à titre indicatif.

De plus, pour chaque variable active, la macro édite le cumul des CONTR, POIDS et CTR de toutes ses modalités.

Variables supplémentaires

Si $IVS > 0$, pour chaque variable supplémentaire, la macro édite pour chacune de ses modalités :

- son **identifiant**, i.e. :
 - la modalité de la variable, éventuellement formatée, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction TABLES (analyse d'une table sous la forme d'un tableau individus-variables)
 - le nom de la "variable"-SAS désignant la modalité, dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction VAR (analyse d'une table déjà sous la forme d'un tableau disjonctif complet)

- **POIDS** : le poids de la modalité dans la population, i.e. la proportion d'individus dans la population prenant la modalité (en %)⁹
- pour chacun des n premiers axes (n est la valeur du paramètre IVS)

COORD	coordonnée sur l'axe
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)
RCO2	rang du point, les points étant classés par CO2 décroissant
QLT	qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %)
V.TEST	valeur-test sur l'axe.

Individus actifs

Si IOA > 0, la macro édite pour chaque individu actif (sélectionné, si PARTIEL > 0) :

- son **identifiant**, i.e. :
 - la valeur de la 1ère variable de l'instruction TABLES, lorsque l'on a utilisé cette instruction sous la forme

TABLES IDENT, X Y Z ;

 pour construire le tableau disjonctif complet associé aux variables X, Y et Z et obtenir des résultats sur les individus (identifiés ici par la variable IDENT)
 - la valeur de la variable-ID dans le cas où l'on a utilisé, dans la procédure CORRESP, les instructions VAR et ID (analyse d'une table déjà sous la forme d'un tableau disjonctif complet)
 - Row 1, Row 2 ... dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction VAR sans l'instruction ID.
- **CONTR** : contribution relative à l'inertie totale du nuage (en %)
- **POIDS** : poids dans le nuage (en %)
- pour chacun des n premiers axes (n est la valeur du paramètre IOA)

COORD	coordonnée sur l'axe
CTR	contribution relative à l'inertie expliquée par l'axe (en %)
RCTR	rang du point, les points étant classés par CTR décroissante
CO2	qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)

⁹ Le POIDS d'une modalité d'une variable supplémentaire n'a donc pas la même définition que le POIDS d'une modalité d'une variable active.

QLT qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %).

Individus supplémentaires

Si $IOS > 0$, la macro édite pour chaque individu supplémentaire (sélectionné si $PARTIEL > 0$) :

- son **identifiant**, i.e.
 - la valeur de la variable-ID dans le cas où l'on a utilisé, dans la procédure CORRESP, les instructions VAR et ID (analyse d'une table déjà sous la forme d'un tableau disjonctif complet)
 - Rowx, ... dans le cas où l'on a utilisé, dans la procédure CORRESP, l'instruction VAR sans l'instruction ID
- pour chacun des n premiers axes (n est la valeur du paramètre IOS)

COORD coordonnée sur l'axe

CO2 qualité de représentation sur l'axe (ou cosinus carré de l'angle formé par le point et l'axe) (en %)

RCO2 rang du point, les points étant classés par CO2 décroissant

QLT qualité de représentation sur le sous-espace engendré par l'axe et les axes précédents (ou cosinus carré de l'angle formé par le point et le sous-espace) (en %)

La procédure CORRESP édite également la plupart de ces statistiques, mais sous une présentation différente... et plus étendue. Pour éviter des listings trop longs, il est recommandé d'utiliser l'option **NOPRINT** de cette procédure, qui supprime toute édition.

La macro PLOTCOR

La macro SAS **PLOTCOR** produit des représentations graphiques associées à une analyse factorielle des correspondances (AFC) ou à une analyse des correspondances multiples (ACM) réalisée à partir de la procédure CORRESP. Elle utilise la procédure SAS PLOT.

Dans le cadre d'une AFC, les graphiques peuvent représenter les modalités-lignes actives, les modalités-lignes supplémentaires, les modalités-colonnes actives et les modalités-colonnes supplémentaires sur le plan factoriel considéré.

Dans le cadre d'une ACM, les graphiques peuvent représenter les individus actifs, les individus supplémentaires, les modalités actives et les modalités supplémentaires sur le plan factoriel considéré.

Syntaxe de la macro

DATA = nom de table SAS

nom de la table SAS en sortie de la procédure CORRESP créée par l'option OUTC = (**obligatoire**).

AXEH = n

numéro de l'axe horizontal (**obligatoire**)

AXEV = n

numéro de l'axe vertical (**obligatoire**)

POINTS = OBSACT et/ou OBSSUP et/ou VARSUP et/ou VARACT

indique les points représentés sur le graphique (**obligatoire**) :

OBSACT	modalités-lignes actives dans le cas d'une AFC, individus actifs dans le cas d'une ACM (observations de la table-OUTC pour lesquelles <code>_TYPE_ = OBS</code>)
OBSSUP	modalités-lignes supplémentaires dans le cas d'une AFC, individus supplémentaires dans le cas d'une AFC (observations de la table-OUTC pour lesquelles <code>_TYPE_ = SUPOBS</code>)
VARACT	modalités-colonnes actives dans le cas d'une AFC, modalités actives dans le cas d'une ACM (observations de la table-OUTC pour lesquelles <code>_TYPE_ = VAR</code>)

VARSUP modalités-colonnes supplémentaires dans le cas d'une AFC, modalités supplémentaires dans le cas d'une ACM

(observations de la table-OUTC pour lesquelles `_TYPE_ = SUPVAR`)

Lorsque que l'on demande la représentation simultanée de plusieurs types de points, ces différents types doivent être séparés par des blancs ; par exemple :

`POINTS = OBSACT VARACT`

ID = variable

paramètre à renseigner **dans le cas où une instruction ID figure dans la procédure CORRESP** : il faut indiquer ici le nom de la variable spécifiée dans cette instruction,

Remarque : les points sont identifiés sur les graphiques par les valeurs de la variable de la table-OUTC qui porte le nom `_NAME_` en l'absence de l'instruction ID, et le nom de la variable-ID en présence de l'instruction ID.

La macro CAHNUM

La macro SAS **CAHNUM** permet de réaliser une classification ascendante hiérarchique, à l'aide des procédures CLUSTER et TREE de SAS, sur les tableaux du type : individus (en lignes) × variables numériques (en colonnes).

La macro réalise une classification sur les **individus**. Les variables peuvent être réduites (i.e. divisées par leurs écart-types) ou non. On peut utiliser comme stratégie d'agrégation la méthode de WARD (maximisation de l'inertie interclasse du nuage des individus), ou toute autre méthode proposée dans la procédure CLUSTER.

La macro édite le tableau des nœuds issu de la classification, et à la demande l'arbre de classification. Les résultats de la classification sont stockés dans une table SAS, et peuvent être réutilisés par la suite par les macros PARTNUM, DESNUM..., si l'on souhaite réaliser et analyser une partition de l'ensemble des individus

Note : dans la suite, conformément à la terminologie habituelle de SAS, les individus seront appelés "observations", et les variables numériques "variables".

1. Syntaxe de la macro

1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse

DATA = nom de table SAS

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

Remarque : on peut utiliser la clause WHERE, ou les options FIRSTOBS ou OBS, pour définir cette table. Par exemple, on peut écrire :

```
DATA=A(WHERE=(SEXE="2")) ou DATA=B(OBS=20)
```

VAR = liste de variables

liste des variables **numériques**, séparées par des blancs (**obligatoire**).

ID = variable

variable servant à identifier les observations dans la table en sortie (paramètre TABN). Cette variable doit figurer dans la table en entrée (paramètre DATA).

Il s'agit obligatoirement d'une variable caractère.

Si ce paramètre est absent, la table en sortie contient une variable identifiant de nom `_NAME_` : à chaque observation est affecté son rang dans la table en entrée (OB1 pour la première, OB102 pour la 102ème...).

TABN = nom de table SAS

nom de la table SAS en sortie, appelée "table des nœuds", contenant les résultats de la classification ; elle contient un certain nombre de renseignements pour chaque nœud (i.e. chaque classe) de la hiérarchie, y compris les classes de départ réduites à un seul élément (singletons) ; elle permet en particulier de construire l'arbre de classification et de partitionner l'ensemble des observations.

Par défaut : TABN = `_TABN_`.

POIDS = variable

nom d'une variable numérique servant à pondérer les observations. Les valeurs de cette variable doivent être des entiers ; si une valeur n'est pas un entier, c'est la partie entière de cette valeur qui est utilisée.

Une observation prenant une valeur négative, nulle ou manquante sur cette variable est traitée comme observation supplémentaire et ne participe pas à la classification. Elle figure néanmoins dans la table en sortie (paramètre TABN), avec un numéro de nœud égal à ".".

Par défaut : toutes les observations ont le même poids.

1.2. Paramètres spécifiant les sorties imprimées

ARBRE = OUI ou NON

si ARBRE vaut OUI, la macro édite l'arbre de classification.

Remarque : il s'agit d'un arbre "américain" de nature quelque peu différente des arbres "français", et qui ne s'apprécie donc pas de la même façon...

Par défaut : pas d'impression (ARBRE = NON).

CCC = OUI ou NON

si CCC vaut OUI, il y a impression d'un graphique montrant l'évolution du Cubic Clustering Criterion¹⁰

Par défaut : pas d'impression (CCC = NON).

NOEUDS = n ou MAX

provoque l'édition de tableaux relatifs aux n derniers nœuds de la classification (i.e. des n classes de plus haut niveau d'agrégation), ou à tous les nœuds si NOEUDS = MAX.

Si NOEUDS = 0, aucun tableau n'est imprimé.

Par défaut : NOEUDS = 15.

¹⁰ Le CCC est une statistique censée faciliter le choix du nombre de classes "optimal".

LSIZE = longueur

valeur du paramètre "linesize" de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LSIZE = 132.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

1.3. Paramètres divers

METHODE = méthode utilisée

méthode, ou stratégie d'agrégation, utilisée pour la classification. L'utilisateur peut spécifier toute méthode proposée dans la procédure CLUSTER¹¹

Par défaut : METHODE = WARD, i.e. la méthode la plus utilisée : à chaque étape de l'algorithme sont agrégées les deux classes qui provoquent la plus petite diminution de l'inertie interclasse.

REDUC = OUI ou NON

si REDUC = OUI, les variables analysées sont préalablement réduites (i.e. divisées par leurs écarts-types).

Par défaut : REDUC = OUI.

¹¹ Le lecteur se reportera à la documentation "SAS/STAT User's Guide, volume 1" pour obtenir la liste de ces stratégies d'agrégation.

2. Les sorties imprimées

Si **NOEUDS > 0**, la macro édite deux tableaux décrivant les n derniers nœuds de la hiérarchie (i.e. les n classes de plus haut niveau d'agrégation), où n est la valeur du paramètre NOEUDS. Ces tableaux décrivent tous les nœuds si **NOEUDS = MAX**.

Le premier tableau, appelé **tableau des nœuds de la hiérarchie**, donne l'inertie totale, et pour chacun des nœuds sélectionnés :

- le nom du nœud (**Noeud**)
- l'effectif pondéré du nœud, i.e. la somme des poids des observations constituant le nœud (**Effectif pondéré**)
- les noms des deux classes qui ont été agrégées pour constituer le nœud (**Classes jointes**)
- la perte d'inertie interclasse provoquée par cette agrégation (**Perte d'inertie inter**) :
 - rapportée à l'inertie totale, en millièmes (**0/00**)
 - cumulée depuis la racine (**cum.**) : pour le nœud de nom CLj, cette quantité est égale au rapport inertie interclasse / inertie totale pour la partition en j + 1 classes
 - la différence entre cette perte d'inertie et la perte d'inertie précédente (**dif**)

et un histogramme des pertes d'inertie.

Le deuxième tableau, appelé **décomposition de la distance entre les centres de gravité des classes jointes**, donne pour chacun des nœuds sélectionnés :

- le nom du nœud (**Noeud**)
- l'effectif pondéré du nœud, i.e. la somme des poids des observations constituant le nœud (**Effectif pondéré**)
- les noms des deux classes qui ont été agrégées pour constituer le nœud (**Classes jointes**)
- la perte d'inertie interclasse provoquée par cette agrégation (**Perte d'inertie inter**)
- la décomposition de la distance entre les centres de gravité des deux classes jointes selon les variables de l'analyse : on lit dans ce tableau la contribution relative (en millièmes) de chaque variable à la distance entre les classes.

Si **ARBRE vaut OUI**, la macro édite l'**arbre de classification**.

Si **CCC vaut OUI**, la macro édite un graphique représentant l'évolution du "**Cubic Clustering Criterion**" lors de l'agrégation des différents nœuds.

3. La table SAS en sortie

La macro crée une table SAS dont le nom est spécifié dans le paramètre TABN : il s'agit de la table créée par la procédure CLUSTER. Elle contient un certain nombre de renseignements pour chaque nœud (i.e. chaque classe) de la hiérarchie, y compris les classes de départ réduites à un seul élément (singletons) ; elle permet en particulier de construire l'arbre de classification et de partitionner l'ensemble des observations.

Le lecteur intéressé pourra se reporter à la documentation "SAS/STAT User's Guide, volume 1" pour obtenir le contenu détaillé de cette table SAS.

La macro CAHQUAL

La macro SAS **CAHQUAL** permet de réaliser une classification ascendante hiérarchique, à l'aide des procédures CLUSTER et TREE de SAS, sur les types de tableaux suivants :

- tableau de contingence
- tableaux de contingence juxtaposés
- tableau disjonctif complet

La macro réalise une classification sur les **observations** (i.e. les lignes du tableau). La distance utilisée pour mesurer les proximités entre les observations est la **distance du chi-deux**. On peut utiliser comme stratégie d'agrégation la méthode de WARD (fondée sur la décomposition de l'inertie du nuage des individus), ou toute autre méthode proposée dans la procédure CLUSTER.

La macro édite le tableau des nœuds issu de la classification, et à la demande l'arbre de classification. Les résultats de la classification sont stockés dans une table SAS, et peuvent être réutilisés par la suite par les macros PARTQUAL, DESQUAL..., si l'on souhaite réaliser et analyser une partition de l'ensemble des observations.

Les tables SAS en entrée des macros CAHQUAL, PARTQUAL, DESQUAL

Les tables SAS auxquelles ces macros s'appliquent sont généralement de l'un des trois types suivants :

1. **tableau de contingence** : les observations sont les modalités d'une variable qualitative, et les variables-SAS sont les modalités d'une variable qualitative.

Exemple : la table SAS a la forme suivante

<u>IDENT</u>	<u>B1</u>	<u>B2</u>
A1	n11	n12
A2	n21	n22
A3	n31	n32

où A1, A2, A3 sont les modalités d'une variable qualitative A (stockées dans une variable identifiant IDENT), et B1, B2 sont les modalités d'une variable qualitative B ; n_{ij} est le nombre d'individus prenant simultanément les modalités A_i et B_j .

2. **tableaux de contingence juxtaposés** : les observations sont les modalités d'une variable qualitative, et les variables-SAS sont les modalités de **plusieurs** variables qualitatives.

Exemple : la table SAS a la forme suivante

IDENT	B1	B2	C1	C2	C3
A1	n11	n12	m11	m12	m13
A2	n21	n22	m21	m22	m23
A3	n31	n32	m31	m32	m33

où A1, A2, A3 sont les modalités d'une variable qualitative A (stockées dans une variable identifiant IDENT), B1, B2 sont les modalités d'une variable qualitative B, et C1, C2, C3 sont les modalités d'une variable qualitative C ; n_{ij} est le nombre d'individus prenant simultanément les modalités A_i et B_j , m_{ik} est le nombre d'individus prenant simultanément les modalités A_i et C_k

Les colonnes B1 et B2 constituent le **sous-tableau** associé à la variable B, les colonnes C1, C2 et C3 constituent le sous-tableau associé à la variable C ; les totaux de ces deux sous-tableaux peuvent ne pas être égaux.

3. **tableau disjonctif complet** : les observations sont des individus, les variables-SAS sont les variables indicatrices des modalités de plusieurs variables qualitatives.

Exemple : la table SAS a la forme suivante

IDENT	A1	A2	A3	B1	B2	C1	C2
alpha	1	0	0	1	0	0	1
beta	0	1	0	1	0	0	1
gamma	1	0	0	0	1	1	0
delta	0	0	1	1	0	0	1
eta	0	0	1	1	0	1	0

où A1, A2, A3 sont les modalités d'une variable qualitative A, B1, B2 sont les modalités d'une variable qualitative B, et C1, C2 sont les modalités d'une variable qualitative C. La variable IDENT contient les identifiants des individus. L'individu "alpha" prend les modalités A1 de A, B1 de B, et C2 de C.

Les colonnes A1 et A2 constituent le **sous-tableau** associé à la variable A, les colonnes B1, B2 et B3 constituent le sous-tableau associé à la variable B, etc. Si le tableau est bien un tableau disjonctif **complet**, les totaux de ces sous-tableaux sont tous égaux au nombre d'individus.

Notes :

- dans la suite, conformément à la terminologie habituelle de SAS, les lignes du tableau, qui selon le cas sont des modalités de variable qualitative ou des individus, seront appelées "observations", et les colonnes du tableau, qui sont des modalités de variables qualitatives, seront appelées "variables"
- le(s) tableau(x) de contingence, ou le tableau disjonctif complet, sur lequel opèrent ces macros doit avoir été **préalablement construit par l'utilisateur**
- lorsque la classification opère sur des tableaux de contingence juxtaposés, il n'est pas possible de faire apparaître la même modalité (i.e. la même variable-SAS) dans plusieurs variables qualitatives-colonnes

Ainsi, si dans la table SAS en entrée figurent les variables MOINS20, MOINS40, _2039, _40PLUS, donnant respectivement le nombre d'individus de moins de 20 ans, de moins de 40 ans, ayant entre 20 et 39 ans, de 40 ans et plus, ces variables permettent de définir deux variables qualitatives "âge" : AGE1 en 3 classes (MOINS20, _2039, _40PLUS), AGE2 en 2 classes (MOINS40, _40PLUS). L'utilisateur voulant réaliser une classification sur les tableaux de contingence juxtaposés AxAGE1 et AxAGE2¹², où A désigne la variable qualitative dont les modalités sont représentées par les observations de la table, doit au préalable définir une nouvelle variable-SAS égale à _40PLUS.

Le problème est identique dans le cas d'un tableau disjonctif complet

¹² classification dont il assure seul la pertinence...

1. Syntaxe de la macro

1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse

DATA = nom de table SAS

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

Remarque : on peut utiliser la clause WHERE, ou les options FIRSTOBS ou OBS, pour définir cette table. Par exemple, on peut écrire :

```
DATA=A(WHERE=(SEXE="2")) ou DATA=B(OBS=20)
```

VAR = liste de variables

liste des variables **numériques**, séparées par des blancs (**obligatoire**).

ID = variable

variable servant à identifier les observations dans la table en sortie (paramètre TABN). Cette variable doit figurer dans la table en entrée (paramètre DATA).

Il s'agit obligatoirement d'une variable caractère.

Si ce paramètre est absent, la table en sortie contient une variable identifiant de nom `_NAME_` : à chaque observation est affecté son rang dans la table en entrée (OB1 pour la première, OB102 pour la 102ème...).

TABN = nom de table SAS

nom de la table SAS en sortie, appelée "table des nœuds", contenant les résultats de la classification ; elle contient un certain nombre de renseignements pour chaque nœud (i.e. chaque classe) de la hiérarchie, y compris les classes de départ réduites à un seul élément (singletons) ; elle permet en particulier de construire l'arbre de classification et de partitionner l'ensemble des observations.

Par défaut : `TABN = _TABN_`.

POIDS = variable

nom d'une variable numérique servant à pondérer les observations. Les valeurs de cette variable doivent être des entiers ; si une valeur n'est pas un entier, c'est la partie entière de cette valeur qui est utilisée.

Une observation prenant une valeur négative, nulle ou manquante sur cette variable est traitée comme observation supplémentaire et ne participe pas à la classification. Elle figure néanmoins dans la table en sortie (paramètre TABN), avec un numéro de nœud égal à ".".

Par défaut : toutes les observations ont le même poids.

1.2. Paramètres spécifiant les sorties imprimées

ARBRE = OUI ou NON

si ARBRE vaut OUI, la macro édite l'arbre de classification.

Remarque : il s'agit d'un arbre "américain" de nature quelque peu différente des arbres "français", et qui ne s'apprécie donc pas de la même façon...

Par défaut : pas d'impression (ARBRE = NON).

CCC = OUI ou NON

si CCC vaut OUI, il y a impression d'un graphique montrant l'évolution du Cubic Clustering Criterion¹³.

Par défaut : pas d'impression (CCC = NON).

NOEUDS = n ou MAX

provoque l'édition de tableaux relatifs aux n derniers nœuds de la classification (i.e. des n classes de plus haut niveau d'agrégation), ou à tous les nœuds si NOEUDS = MAX.

Si NOEUDS = 0, aucun tableau n'est imprimé.

Par défaut : NOEUDS = 15.

LSIZE = longueur

valeur du paramètre "linesize" de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LSIZE = 132.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

¹³ Le CCC est une statistique censée faciliter le choix du nombre de classes "optimal".

1.3. Autre paramètre

METHODE = méthode utilisée

méthode, ou stratégie d'agrégation, utilisée pour la classification. L'utilisateur peut spécifier toute méthode proposée dans la procédure CLUSTER¹⁴.

Par défaut : METHODE = WARD, i.e. la méthode la plus utilisée : à chaque étape de l'algorithme sont agrégées les deux classes qui provoquent la plus petite diminution de l'inertie interclasse.

¹⁴ Le lecteur se reportera à la documentation "SAS/STAT User's Guide, volume 1" pour obtenir la liste de ces stratégies d'agrégation.

2. Les sorties imprimées

Si **NOEUDS** > 0, la macro édite deux tableaux décrivant les n derniers nœuds de la hiérarchie (i.e. les n classes de plus haut niveau d'agrégation), où n est la valeur du paramètre NOEUDS. Ces tableaux décrivent tous les nœuds si NOEUDS = MAX.

Le premier tableau, appelé **tableau des nœuds de la hiérarchie**, donne l'inertie totale, et pour chacun des nœuds sélectionnés :

- le nom du nœud (**Noeud**)
- l'effectif pondéré du nœud, i.e. la somme des poids des observations constituant le nœud (**Effectif pondéré**)
- les noms des deux classes qui ont été agrégées pour constituer le nœud (**Classes jointes**)
- la perte d'inertie interclasse provoquée par cette agrégation (**Perte d'inertie inter**) :
 - rapportée à l'inertie totale, en millièmes (**0/00**)
 - cumulée depuis la racine (**cum.**) : pour le nœud de nom CLj, cette quantité est égale au rapport inertie interclasse / inertie totale pour la partition en j + 1 classes
 - la différence entre cette perte d'inertie et la perte d'inertie précédente (**dif**)

et un histogramme des pertes d'inertie.

Le deuxième tableau, appelé **décomposition de la distance entre les centres de gravité des classes jointes**, donne pour chacun des nœuds sélectionnés :

- le nom du nœud (**Noeud**)
- l'effectif pondéré du nœud, i.e. la somme des poids des observations constituant le nœud (**Effectif pondéré**)
- les noms des deux classes qui ont été agrégées pour constituer le nœud (**Classes jointes**)
- la perte d'inertie interclasse provoquée par cette agrégation (**Perte d'inertie inter**)
- la décomposition de la distance entre les centres de gravité des deux classes jointes selon les modalités de l'analyse : on lit dans ce tableau la contribution relative (en millièmes) de chaque modalité à la distance entre les classes.

Si **ARBRE** vaut OUI, la macro édite l'**arbre de classification**.

Si **CCC** vaut OUI, la macro édite un graphique représentant l'évolution du "**Cubic Clustering Criterion**" lors de l'agrégation des différents nœuds.

3. La table SAS en sortie

La macro crée une table SAS dont le nom est spécifié dans le paramètre TABN : il s'agit de la table créée par la procédure CLUSTER. Elle contient un certain nombre de renseignements pour chaque nœud (i.e. chaque classe) de la hiérarchie, y compris les classes de départ réduites à un seul élément (singletons) ; elle permet en particulier de construire l'arbre de classification et de partitionner l'ensemble des observations.

Le lecteur intéressé pourra se reporter à la documentation "SAS/STAT User's Guide, volume 1" pour obtenir le contenu détaillé de cette table SAS.

La macro PARTNUM

La macro SAS **PARTNUM** permet de réaliser une partition d'un ensemble d'individus à partir des résultats d'une classification ascendante hiérarchique effectuée à l'aide de la macro **CAHNUM** (i.e. sur un tableau individus x variables numériques).

Elle édite des aides à l'interprétation de la partition : décomposition de l'inertie, des distances entre les centres des classes et l'origine, des distances entre les centres des classes... Elle permet également d'obtenir la composition des classes, et la description des classes à l'aide des variables numériques analysées.

Cette macro doit être précédée de l'appel de la macro CAHNUM. Elle utilise la table-TABN créée par CAHNUM, qui contient les résultats de la classification.

Note : dans la suite, conformément à la terminologie habituelle de SAS, les individus seront appelés "observations", et les variables numériques "variables".

1. Syntaxe de la macro

1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse

TABN = nom de table SAS

nom de la table SAS contenant les résultats de la CAH réalisée auparavant à l'aide de la macro CAHNUM : il s'agit de la table spécifiée dans le paramètre TABN de CAHNUM.

Par défaut : **TABN = _TABN_**, qui est aussi la valeur par défaut du paramètre TABN de la macro CAHNUM.

DATA = nom de table SAS

nom de la table SAS sur laquelle a opéré la macro CAHNUM ; c'est la valeur du paramètre DATA de CAHNUM (**obligatoire**).

VAR = liste des variables

liste des variables utilisées pour réaliser la classification ; cette liste est celle du paramètre VAR de CAHNUM (**obligatoire**).

ID = variable

paramètre à renseigner **dans le cas où le paramètre ID figure dans la macro CAHNUM** : il faut indiquer ici le nom de la variable spécifiée dans ce paramètre, qui permet d'identifier les observations.

POIDS = nom de variable

paramètre à renseigner **dans le cas où le paramètre POIDS figure dans la macro CAHNUM** : il faut indiquer ici le nom de la variable spécifiée dans ce paramètre, qui permet de pondérer les observations.

1.2. Paramètres caractérisant la partition à réaliser

NCL= nombre de classes

indique le nombre de classes de la partition (**obligatoire**).

LISTE = OUI ou NON

si LISTE vaut OUI, la macro édite le contenu de chaque classe de la partition créée par la macro PARTNUM.

Par défaut : LISTE = NON.

AFFECT = OUI ou NON

si AFFECT vaut OUI, la macro crée une table SAS contenant toutes les variables et toutes les observations de la table en entrée (paramètre DATA), et une variable de classe attribuant à chaque observation le numéro de la classe à laquelle elle appartient.

Par défaut : AFFECT = NON.

OUTPART = nom de table SAS

si AFFECT vaut OUI, ce paramètre est obligatoire : il indique le nom de la table SAS créée par la macro, contenant les données initiales et la variable de classe.

Remarque : les paramètres DATA et OUTPART peuvent prendre la même valeur ; cela signifie que la variable de classe est ajoutée au tableau initial à la fin de la macro PARTNUM.

VARCLASS = nom de variable

si AFFECT vaut OUI, nom de la variable de la table-OUTPART donnant, pour chaque observation, le numéro de la classe à laquelle elle appartient.

Par défaut : VARCLASS = CLASSn, où n est le nombre de classes de la partition (paramètre NCL).

1.3. Paramètres divers

DESC = OUI ou NON

si DESC vaut OUI, la macro édite la description de la partition¹⁵ à l'aide des variables analysées.

Par défaut : DESC = OUI.

ININTER = OUI ou NON

si ININTER vaut OUI, la macro édite un tableau donnant la décomposition de l'inertie interclasse selon les variables de l'analyse et les classes de la partition.

Par défaut : ININTER = NON.

LSIZE = longueur

valeur du paramètre "linesize".

Par défaut : LSIZE = 132.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

CONSOLID = OUI ou NON

si CONSOLID vaut OUI, la macro utilise l'algorithme des **centres mobiles**¹⁶ pour "consolider" la partition en NCL classes donnée par la CAH : partant des centres de gravité des classes obtenues, un certain nombre de réaffectations des observations dans les classes sont réalisées de façon à augmenter l'inertie interclasse de la partition.

Par défaut : CONSOLID = NON.

¹⁵ Cette description est réalisée par l'appel de la macro **DESNUM**.

¹⁶ La macro utilise la procédure SAS **FASTCLUS**.

OBSSUP = OUI ou NON

si OBSSUP vaut OUI, la macro affecte les observations de la table-DATA éliminées de la classification à la classe dont elles sont le plus proches (au sens de la distance¹⁷ utilisée pour la CAH). Cette affectation, qui n'a pas d'influence sur la description de la partition par la macro PARTNUM, a pour conséquence l'attribution à ces observations d'une valeur de la variable-VARCLASS dans la table-OUTPART : la macro DESNUM permet alors de décrire la partition en prenant en compte ces "observations supplémentaires".

Par défaut : OBSSUP = NON.

REDUC = OUI ou NON

ce paramètre doit prendre la même valeur que le paramètre REDUC de la macro CAHNUM.

Par défaut : REDUC = OUI.

¹⁷ Si l'observation a été éliminée en raison de valeur(s) manquante(s) sur des variables analysées, la distance est calculée à partir des valeurs non manquantes.

2. Les sorties imprimées

La macro édite :

- **si le paramètre CONSOLID vaut OUI, les résultats de la consolidation :**
 - un tableau donnant l'inertie interclasse avant et après la consolidation, ainsi que le gain d'inertie
 - un tableau croisant les observations (non pondérées) selon les deux variables de classe (avant et après la consolidation)
- **si le paramètre LISTE vaut OUI, la composition de chaque classe de la partition.** Les observations éliminées de la classification (à cause d'un poids négatif ou nul, ou de la présence d'une valeur manquante dans les variables analysées par exemple) sont listées à part.
- un tableau donnant la **décomposition de l'inertie totale** du nuage relative à la partition considérée :
 - inertie totale
 - inertie interclasse
 - décomposition par classe de l'inertie interclasse
 - inertie intraclasse
 - décomposition par classe de l'inertie intraclasse.

Pour chaque classe, le tableau donne de plus **l'effectif absolu**, i.e. le nombre d'observations constituant la classe, et **l'effectif pondéré**, i.e. la somme des poids¹⁸ des observations constituant la classe (en millièmes).

- un tableau donnant, pour chaque classe de la partition, **la décomposition de la distance entre le centre de la classe et le centre de gravité du nuage** (ou RHO2) selon les variables de l'analyse. Cette décomposition, donnée en millièmes, est signée : un signe - (resp. l'absence de signe -) indique que la moyenne de la variable dans la classe est plus faible (resp. plus élevée) que sa moyenne dans la population.
- un tableau donnant, pour chaque couple de classes de la partition, **la décomposition de la distance entre les centres des deux classes** selon les variables de l'analyse. Cette décomposition, donnée en millièmes, est signée : un signe - (resp. l'absence de signe -) indique que la moyenne de la variable dans la première classe est plus faible (resp. plus élevée) que sa moyenne dans la deuxième classe. Ce tableau donne également la perte d'inertie interclasse que provoquerait l'agrégation des deux classes considérées.
- **si le paramètre ININTER vaut OUI**, un tableau donnant la **décomposition de l'inertie interclasse selon les variables de l'analyse et les classes de la partition.**

¹⁸ Le poids d'une observation est égal à la valeur de la variable-POIDS ; il vaut 1 si ce paramètre n'est pas renseigné.

Si le paramètre DESC vaut OUI, la macro édite la description de la partition à l'aide des variables analysées : il s'agit d'un tableau donnant, pour chaque variable ayant participé à l'analyse :

- sa moyenne et son écart-type (**Ec.type**) dans la population totale (**Ensemble**) et dans chaque classe
- sur la ligne **V.test**, les valeurs des statistiques de test suivantes :
 - dans la colonne **Ensemble**, la statistique du test de Fisher de comparaison de moyennes (ou test d'analyse de la variance) mesurant la "liaison" entre la variable numérique considérée et la partition : plus grands sont les écarts entre les moyennes de la variable dans les différentes classes, plus la statistique est élevée
 - dans chacune des autres colonnes, la statistique du test comparant la moyenne de la variable dans la classe et sa moyenne dans la population totale : plus grand est l'écart entre ces moyennes, plus la statistique est élevée
- sur la ligne suivante (**Proba**), les niveaux de significativité des tests précédents : une valeur faible de Proba signifie que l'on doit rejeter l'hypothèse testée (égalité de toutes les moyennes par classe pour le test de Fisher, égalité entre la moyenne dans la classe et la moyenne générale pour les tests suivants).

3. La table en sortie

Si le paramètre AFFECT vaut OUI, la macro crée une table SAS, dont le nom est spécifié dans le paramètre OUTPART. Cette table contient toutes les variables et toutes les observations de la table en entrée (paramètre DATA), et une variable supplémentaire, dont le nom est spécifié dans le paramètre VARCLASS. Cette variable attribue à chaque observation le numéro de la classe à laquelle elle appartient : les modalités de cette variable (qui est **caractère**) sont 1, 2, ..., n, où n est le nombre de classes de la partition.

Si OBSSUP vaut NON, la variable-VARCLASS prend la valeur " " pour les observations exclues de la classification.

Si OBSSUP vaut OUI, pour chaque observation exclue, la variable-VARCLASS prend comme valeur le numéro de la classe dont le centre de gravité est le plus proche de l'observation (au sens de la distance utilisée par la CAH). La table en sortie contient alors une autre variable supplémentaire, de nom CLINITn (où n est la valeur du paramètre NCL), qui prend la valeur de la variable-VARCLASS pour une observation ayant pris part à la classification, et la valeur " " pour une observation exclue.

Remarque : lorsque la table spécifiée dans le paramètre OUTPART existe déjà, la table créée par la macro remplace la table préexistante.

La macro PARTQUAL

La macro SAS **PARTQUAL** permet de réaliser une partition à partir des résultats d'une classification ascendante hiérarchique effectuée à l'aide de la macro **CAHQUAL** sur un tableau de contingence, sur des tableaux de contingence juxtaposés, ou sur un tableau disjonctif complet. Dans les deux premiers cas, les unités partitionnées sont les modalités de la variable qualitative définissant les lignes du tableau ; dans le dernier cas, les unités partitionnées sont les individus.

La macro édite des aides à l'interprétation de la partition : décomposition de l'inertie, des distances entre les centres des classes et l'origine, des distances entre les centres des classes... Elle permet également d'obtenir la composition des classes, et la description des classes à l'aide des modalités de la (des) variable(s) analysée(s).

Cette macro doit être précédée de l'appel de la macro CAHQUAL. Elle utilise la table-TABN créée par CAHQUAL, qui contient les résultats de la classification.

Le § "Les tables SAS en entrée des macros CAHQUAL, PARTQUAL et DESQUAL" de la description de la macro CAHQUAL, présente les tableaux sur lesquels on peut utiliser la macro.

Notes :

- Dans la suite, conformément à la terminologie habituelle de SAS, les lignes du tableau, qui selon le cas sont des modalités de variable qualitative ou des individus, seront appelées "observations", et les colonnes du tableau, qui sont des modalités de variables qualitatives, seront appelées "variables". Une "variable qualitative" fait donc référence à un groupe de colonnes (ou "variables") du tableau analysé.
- Quelques précisions sur les poids des observations :
 - Dans le cas d'un tableau de contingence, le poids d'une observation, i.e. d'une modalité-ligne, est égal à l'effectif de cette modalité (i.e. la somme de la ligne dans le tableau), éventuellement multiplié par la valeur de la variable-POIDS si ce paramètre est renseigné.
 - Dans le cas de tableaux de contingence juxtaposés, le poids d'une observation, i.e. d'une modalité-ligne, **dans un sous-tableau** associé à une des variables qualitatives en colonne, est égal à l'effectif de cette modalité (i.e. la somme de la ligne dans le sous-tableau), éventuellement multiplié par la valeur de la variable-POIDS si ce paramètre est renseigné. Le poids d'une observation dans le tableau complet est égal à la somme des poids de l'observation dans chacun des sous-tableaux.
 - Dans le cas d'un tableau disjonctif complet, le poids d'une observation, i. e. d'un individu, **dans un sous-tableau** associé à une des variables qualitatives en colonne, est égal à 1 (i. e. la somme de la ligne dans le sous-tableau), éventuellement multiplié par la valeur de la variable-POIDS si ce paramètre est renseigné. Le poids d'une observation dans le tableau complet est égal à la somme des poids de l'observation dans chacun des sous-tableaux.

1. Syntaxe de la macro

1.1. Paramètres spécifiant les tables SAS et les variables de l'analyse

TABN = nom de table SAS

nom de la table SAS contenant les résultats de la CAH réalisée auparavant à l'aide de la macro CAHQUAL : il s'agit de la table spécifiée dans le paramètre TABN de CAHQUAL.

Par défaut : TABN = _TABN_, qui est aussi la valeur par défaut du paramètre TABN de la macro CAHQUAL.

DATA = nom de table SAS

nom de la table SAS sur laquelle a opéré la macro CAHQUAL ; c'est la valeur du paramètre DATA de CAHQUAL (**obligatoire**).

VAR = liste des variables

liste des variables utilisées pour réaliser la classification ; cette liste est celle du paramètre VAR de CAHQUAL (**obligatoire**).

ID = variable

paramètre à renseigner **dans le cas où le paramètre ID figure dans la macro CAHQUAL** : il faut indiquer ici le nom de la variable spécifiée dans ce paramètre, qui permet d'identifier les observations.

POIDS = nom de variable

paramètre à renseigner **dans le cas où le paramètre POIDS figure dans la macro CAHQUAL** : il faut indiquer ici le nom de la variable spécifiée dans ce paramètre, qui permet de pondérer les observations.

1.2. Paramètres caractérisant la partition à réaliser

NCL= nombre de classes

indique le nombre de classes de la partition (**obligatoire**).

LISTE = OUI ou NON

si LISTE vaut OUI, la macro édite le contenu de chaque classe de la partition créée par la macro PARTQUAL.

Par défaut : LISTE = NON.

AFFECT = OUI ou NON

si AFFECT vaut OUI, la macro crée une table SAS contenant toutes les variables et toutes les observations de la table en entrée (paramètre DATA), et une variable de classe attribuant à chaque observation le numéro de la classe à laquelle elle appartient.

Par défaut : AFFECT = NON.

OUTPART = nom de table SAS

si AFFECT vaut OUI, ce paramètre est obligatoire : il indique le nom de la table SAS créée par la macro, contenant les données initiales et la variable de classe.

Remarque : les paramètres DATA et OUTPART peuvent prendre la même valeur ; cela signifie que la variable de classe est ajoutée au tableau initial à la fin de la macro PARTQUAL.

VARCLASS = nom de variable

si AFFECT vaut OUI, nom de la variable de la table-OUTPART donnant, pour chaque observation, le numéro de la classe à laquelle elle appartient.

Par défaut : VARCLASS = CLASSn, où n est le nombre de classes de la partition (paramètre NCL).

1.3. Paramètres divers

DESC = OUI ou NON

si DESC vaut OUI, la macro édite la description de la partition¹⁹ à l'aide des variables analysées.

Par défaut : DESC = OUI.

ININTER = OUI ou NON

si ININTER vaut OUI, la macro édite un tableau donnant la décomposition de l'inertie interclasse selon les variables de l'analyse et les classes de la partition.

Par défaut : ININTER = NON.

LSIZE = longueur

valeur du paramètre "linesize" de SAS en vigueur au moment de l'appel de la macro.

Par défaut : LSIZE = 132.

¹⁹ Cette description est réalisée par l'appel de la macro **DESQUAL**.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

CONSOLID = OUI ou NON

si CONSOLID vaut OUI, la macro utilise l'algorithme **des centres mobiles**²⁰ pour "consolider" la partition en NCL classes donnée par la CAH : partant des centres de gravité des classes obtenues, un certain nombre de réaffectations des observations dans les classes sont réalisées de façon à augmenter l'inertie interclasse de la partition.

Par défaut : CONSOLID = NON.

OBSSUP = OUI ou NON

si OBSSUP vaut OUI, la macro affecte les observations de la table-DATA éliminées de la classification à la classe dont elles sont le plus proches (au sens de la distance²¹ utilisée pour la CAH). Cette affectation, qui n'a pas d'influence sur la description de la partition par la macro PARTQUAL, a pour conséquence l'attribution à ces observations d'une valeur de la variable-VARCLASS dans la table-OUTPART : la macro DESQUAL permet alors de décrire la partition en prenant en compte ces "observations supplémentaires".

Par défaut : OBSSUP = NON.

LIBVAR = liste de libellés

liste des libellés des **variables qualitatives** séparés par des blancs (les blancs ne sont pas autorisés dans les libellés). Ces libellés seront utilisés dans les tableaux de description des classes.

Lorsque l'on opère sur plusieurs variables qualitatives, l'ordre dans lequel sont donnés ces noms doit être le même que l'ordre dans lequel ont été spécifiées les modalités de ces variables dans le paramètre VAR. **Il est nécessaire que dans ce paramètre VAR les modalités apparaissent regroupées par variable** (même si l'on utilise les libellés par défaut indiqués plus loin).

On se reportera à l'exemple présenté plus loin.

Par défaut : les libellés des variables sont Var001, Var002...

²⁰ La macro utilise la procédure SAS **FASTCLUS**.

²¹ Si l'observation a été éliminée en raison de valeur(s) manquante(s) sur des variables analysées, la distance est calculée à partir des valeurs non manquantes.

NBMODA = liste des nombres de modalités

pour chacune des variables qualitatives, on indique le nombre de modalités²² de la variable ; ces nombres de modalités sont donnés dans le même ordre que dans la liste LIBVAR, séparés par des blancs.

La somme de ces nombres des modalités **doit être égal** au nombre de variables SAS du paramètre VAR.

Ce paramètre est **obligatoire**, sauf dans le cas d'un tableau de contingence (une seule variable qualitative).

Exemples :

- dans le cas d'un tableau de contingence, où la variable qualitative en colonne est la variable "état matrimonial" à 4 modalités, on peut écrire par exemple :

```
VAR = CELIBAT MARIÉS VEUFS DIVORCÉS, LIBVAR = état_matrimonial
```

ou

```
VAR = CELIBAT MARIÉS VEUFS DIVORCÉS
```

- dans le cas d'un tableau disjonctif complet à trois variables qualitatives X, Y et Z, telles que la variable X a deux modalités X1 et X2, la variable Y trois modalités Y1, Y2 et Y3, et la variable Z deux modalités Z1 et Z2, on peut écrire par exemple :

```
VAR = X1 X2 Y1-Y3 Z1 Z2, LIBVAR = VarX VarY VarZ, NBMODA = 2 3 2
```

Note : le paramètre LIBVAR peut comporter moins de libellés que le nombre de valeurs spécifiées dans la liste NBMODA : les libellés "manquants" sont alors VARn,VARn + 1, ...

PROBA = probabilité

ce paramètre permet de limiter la liste des modalités caractérisant les différentes classes. Seules sont listées les modalités pour lesquelles le niveau de significativité du test est inférieur ou égal à ce paramètre.

Par défaut : PROBA = 1 (toutes les modalités sont listées).

²² Il s'agit des modalités prises par au moins une observation.

2. Les sorties imprimées

La macro édite :

- **si le paramètre CONSOLID vaut OUI, les résultats de la consolidation :**
 - un tableau donnant l'inertie interclasse avant et après la consolidation, ainsi que le gain d'inertie
 - un tableau croisant les observations (non pondérées) selon les deux variables de classe (avant et après la consolidation)
- **si le paramètre LISTE vaut OUI, la composition de chaque classe de la partition.** Les observations éliminées de la classification (à cause d'un poids négatif ou nul, ou de la présence d'une valeur manquante dans les variables analysées par exemple) sont listées à part
- un tableau donnant la **décomposition de l'inertie totale** du nuage relative à la partition considérée :
 - inertie totale
 - inertie interclasse
 - décomposition par classe de l'inertie interclasse
 - inertie intraclasse
 - décomposition par classe de l'inertie intraclasse

Pour chaque classe, le tableau donne de plus l'**effectif absolu**, i.e. le nombre d'observations constituant la classe, et l'**effectif pondéré**, i.e. la somme des poids **dans le tableau complet** des observations constituant la classe (en millièmes)

- un tableau donnant, pour chaque classe de la partition, **la décomposition de la distance entre le centre de la classe et le centre de gravité du nuage** (ou RHO2) selon les modalités de l'analyse. Cette décomposition, donnée en millièmes, est signée : un signe - (resp. l'absence de signe -) indique que la fréquence de la modalité dans la classe est plus faible (resp. plus élevée) que sa fréquence dans la population
- un tableau donnant, pour chaque couple de classes de la partition, **la décomposition de la distance entre les centres des deux classes** selon les modalités de l'analyse. Cette décomposition, donnée en millièmes, est signée : un signe - (resp. l'absence de signe -) indique que la fréquence de la modalité dans la première classe est plus faible (resp. plus élevée) que sa fréquence dans la deuxième classe. Ce tableau donne également la perte d'inertie interclasse que provoquerait l'agrégation des deux classes considérées
- **si le paramètre ININTER vaut OUI**, un tableau donnant la **décomposition de l'inertie interclasse selon les modalités de l'analyse et les classes de la partition.**

Si le paramètre DESC vaut OUI, la macro édite la description de la partition à l'aide des variables qualitatives (et de leurs modalités) analysées : il s'agit d'un tableau donnant, pour chaque variable ayant participé à l'analyse :

- le(s) **tableau(x) de contingence** croisant la variable de classe et chaque variable qualitative
- le tableau des **CHI2 des tableaux de contingence** ; ce tableau donne, pour chaque variable qualitative :
 - son **nombre de modalités**
 - l'**effectif pondéré du tableau**, i.e. l'effectif total du tableau de contingence ; c'est aussi la somme des poids des observations dans le (sous-)tableau associé à la variable qualitative
 - le nombre de cases du tableau de contingence ayant un effectif inférieur à 5 (**Nbre de cases d'effectif < 5**)
 - la proportion dans le tableau de contingence de cases d'effectif inférieur à 5 (**Nbre de cases d'effectif < 5 (en %)**)
 - la valeur de la statistique du chi-deux permettant de réaliser le test d'indépendance entre la variable de classe et la variable qualitative (**CHI2**)
 - le nombre de degrés de liberté associé à la statistique du chi-deux, égal à $(p-1)(q-1)$, où p est le nombre de modalités de la variable de classe et q le nombre de modalités de la variable qualitative (**Degrés de liberté**)
 - le niveau de significativité du test d'indépendance du chi-deux (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'indépendance entre la variable de classe et la variable qualitative doit être rejetée, autrement dit que la variable qualitative n'a pas la même répartition dans les différentes classes.

Dans ce tableau, les variables qualitatives sont classées selon les valeurs croissantes de PROBA²³.

- un tableau donnant la **caractérisation de chaque classe de la partition par les variables analysées** ; ce tableau donne, pour chaque modalité de la variable de classe et pour chaque variable qualitative (identifiée par le libellé donné dans le paramètre LIBVAR) :
 - l'**effectif pondéré** de la classe, i.e. la somme des poids des observations constituant la classe, dans le sous-tableau associé à la variable
 - la distance du chi-deux (au sens de distance entre deux lois de probabilité) entre la distribution de la variable dans la population et la distribution de la variable dans la classe (**CHI2**)
 - le nombre de degrés de liberté associé à cette distance, égal à $q-1$, où q est le nombre de modalités de la variable qualitative (**Degrés de liberté**)
 - le niveau de significativité du test permettant de juger de la significativité des écarts entre les deux distributions (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'égalité des deux distributions doit être rejetée, autrement dit que la variable qualitative n'a pas la même répartition dans la classe et dans la population.

²³ Le classement relatif de variables ayant des valeurs de PROBA nulles et des degrés de liberté associés différents est sans signification.

Dans ce tableau, les variables qualitatives sont rangées, au sein de chaque classe, selon les valeurs croissantes de PROBA²⁴.

- un tableau donnant la **caractérisation de la partition par les modalités des variables analysées** ; ce tableau donne, pour chaque modalité de chaque variable qualitative :
 - le **label** de la modalité de la variable qualitative (qui est le nom de la variable-SAS si aucun label ne lui a été attribué)
 - le libellé de la **variable** qualitative
 - l'**effectif pondéré** de la modalité, i.e. le total (éventuellement pondéré) de la colonne représentant la modalité
 - la distance du chi-deux (au sens de distance entre deux lois de probabilité) entre la distribution de la variable de classe dans la population et la distribution de cette variable de classe parmi les observations prenant la modalité considérée (**CHI2**)
 - le nombre de degrés de liberté associé à cette distance, égal à $p-1$, où p est le nombre de modalités de la variable de classe (**Degrés de liberté**)
 - le niveau de significativité du test permettant de juger de la significativité des écarts entre les deux distributions (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'égalité des deux distributions doit être rejetée, autrement dit que la variable de classe n'a pas la même répartition dans la population entière et dans la sous-population définie par la modalité considérée.

Dans ce tableau, les modalités des variables qualitatives sont rangées selon les valeurs croissantes de PROBA.

- un tableau donnant la **caractérisation de chaque classe de la partition par les modalités des variables analysées** ; ce tableau donne, pour chaque modalité de la variable de classe et pour chaque modalité de chaque variable qualitative :
 - le **label** de la modalité de la variable qualitative (qui est le nom de la variable-SAS si aucun label ne lui a été attribué)
 - le libellé de la **variable** qualitative
 - l'**effectif pondéré** de la modalité dans la classe, i.e. la somme des poids des observations de la classe prenant la modalité considérée
 - la fréquence de la modalité dans la classe (**Fréquence dans la classe (%)**) : c'est le rapport entre l'effectif pondéré de la modalité dans la classe et l'effectif pondéré de la classe
 - la fréquence de la modalité dans la population (**Fréquence dans la pop. (%)**) : c'est le rapport entre l'effectif pondéré de la modalité et l'effectif pondéré du sous-tableau associé à la variable

²⁴ Le classement relatif de variables ayant des valeurs de PROBA nulles et des degrés de liberté associés différents est sans signification.

- le niveau de significativité du test permettant de juger de la significativité de l'écart entre ces deux fréquences²⁵ (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'égalité des deux fréquences doit être rejetée.
- la valeur de la variable aléatoire normale centrée réduite "équivalente" (**Val Test**) conduisant au même niveau de significativité (pour un test bilatéral).

Dans ce tableau, les modalités des variables qualitatives sont rangées, au sein de chaque classe, par valeur-test décroissante, et regroupées en deux catégories : les modalités sur-représentées (resp. sous-représentées) sont telles que la fréquence dans la classe est supérieure (resp. inférieure) à la fréquence dans la population.

²⁵ La statistique du test dépend de la taille de l'échantillon, car elle suit, sous l'hypothèse d'égalité des deux fréquences, une loi hypergéométrique : pour les grands échantillons, on transforme cette statistique pour se ramener à une loi normale centrée réduite.

3. La table en sortie

Si le paramètre AFFECT vaut OUI, la macro crée une table SAS, dont le nom est spécifié dans le paramètre OUTPART. Cette table contient toutes les variables et toutes les observations de la table en entrée (paramètre DATA), et une variable supplémentaire, dont le nom est spécifié dans le paramètre VARCLASS. Cette variable attribue à chaque observation le numéro de la classe à laquelle elle appartient : les modalités de cette variable (qui est **caractère**) sont 1, 2, ..., n, où n est le nombre de classes de la partition.

Si **OBSSUP** vaut **NON**, la variable-VARCLASS prend la valeur " " pour les observations exclues de la classification.

Si **OBSSUP** vaut **OUI**, pour chaque observation exclue, la variable-VARCLASS prend comme valeur le numéro de la classe dont le centre de gravité est le plus proche de l'observation (au sens de la distance utilisée par la CAH). La table en sortie contient alors une autre variable supplémentaire, de nom CLINITn (où n est la valeur du paramètre NCL), qui prend la valeur de la variable-VARCLASS pour une observation ayant pris part à la classification, et la valeur " " pour une observation exclue.

Remarque : lorsque la table spécifiée dans le paramètre OUTPART existe déjà, la table créée par la macro remplace la table préexistante.

La macro DESNUM

La macro SAS **DESNUM** permet de décrire la partition d'un ensemble d'observations à l'aide d'une ou plusieurs variables numériques "explicatives". Elle opère sur une table SAS contenant (au moins) :

- une "variable de classe" : cette variable (numérique ou caractère) à p modalités définit une partition de la population en p classes, une classe étant constituée par l'ensemble des observations prenant une même modalité
- les variables numériques explicatives.

La macro édite un tableau donnant les moyennes et les écarts-types des variables dans chaque classe, ainsi que des tests statistiques permettant de caractériser les classes de la partition par ces variables.

Note : la macro DESNUM produit le même type de sorties que celles de la macro PARTNUM lorsque le paramètre DESC de PARTNUM vaut OUI. Cette macro peut donc être utilisée comme aide à la description d'une partition lorsque l'on désire l'expliquer avec des variables "supplémentaires" n'ayant pas pris part à la classification. On opère alors de la façon suivante :

- on réalise une CAH avec CAHNUM
- on construit une partition avec PARTNUM ; si le paramètre DESC vaut OUI, on obtient la description de la partition par les variables **analysées**
- on utilise DESNUM, avec la table-OUTPART et la variable-VARCLASS créées par PARTNUM, pour décrire la partition avec des variables numériques quelconques (pouvant inclure les variables analysées)
- on peut également utiliser DESQUAL si l'on souhaite décrire la partition avec des variables qualitatives.

1. Syntaxe de la macro

DATA = nom de table SAS

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

Remarque : on peut utiliser la clause WHERE, ou les options FIRSTOBS ou OBS, pour définir cette table. Par exemple, on peut écrire :

```
DATA=A(WHERE=(SEXE="2")) ou DATA=B(OBS=20)
```

VAR = liste de variables

liste des variables **numériques** explicatives, séparées par des blancs (**obligatoire**).

POIDS = variable

nom d'une variable numérique servant à pondérer les observations. Les valeurs de cette variable doivent être des entiers ; si une valeur n'est pas un entier, c'est la partie entière de cette valeur qui est utilisée.

Une observation prenant une valeur négative, nulle ou manquante sur cette variable ne participe pas à la description.

Note : lorsque l'on utilise DESNUM à la suite des macros CAHNUM et PARTNUM, ce paramètre POIDS doit prendre la même valeur que dans les deux autres macros si l'on veut décrire la partition créée par PARTNUM ; en revanche, si l'on veut décrire une autre partition, incluant en particulier des "observations supplémentaires" (générées lorsque le paramètre OBSSUP de PARTNUM vaut OUI), ce paramètre POIDS doit être changé (ou laissé à blanc le cas échéant).

Par défaut : toutes les observations ont le même poids.

VARCLASS = nom de variable

nom de la variable de classe (**obligatoire**).

MISSING = OUI ou NON

si MISSING vaut OUI, les observations ayant une valeur manquante pour la variable de classe constituent une classe à part entière.

Par défaut : MISSING = NON.

NCL = n

nombre de modalités de la variable de classe.

Si NCL n'est pas renseigné, la macro calcule ce nombre.

Note : lorsque l'on utilise DESNUM à la suite de la macro PARTNUM, ce paramètre NCL doit prendre la même valeur que dans cette macro, et doit être augmenté d'une unité si MISSING vaut OUI.

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N.

2. Les sorties imprimées

La macro édite un tableau donnant, pour chaque variable explicative :

- sa **moyenne** et son écart-type (**Ec.type**) dans la population totale (**Ensemble**) et dans chaque classe
- sur la ligne **V.test**, les valeurs des statistiques de test suivantes :
 - dans la colonne **Ensemble**, la statistique du test de Fisher de comparaison de moyennes (ou test d'analyse de la variance) mesurant la "liaison" entre la variable numérique considérée et la partition : plus grands sont les écarts entre les moyennes de la variable dans les différentes classes, plus la statistique est élevée
 - dans chacune des autres colonnes, la statistique du test comparant la moyenne de la variable dans la classe et sa moyenne dans la population totale : plus grand est l'écart entre ces moyennes, plus la statistique est élevée
- sur la ligne suivante (**Proba**), les niveaux de significativité des tests précédents : une valeur faible de Proba signifie que l'on doit rejeter l'hypothèse testée (égalité de toutes les moyennes par classe pour le test de Fisher, égalité entre la moyenne dans la classe et la moyenne générale pour les tests suivants).

La macro DESQUAL

La macro SAS **DESQUAL** permet de décrire la partition d'un ensemble d'observations à l'aide d'une ou plusieurs variables qualitatives "explicatives". Elle opère sur une table SAS contenant (au moins) :

- une "variable de classe" : cette variable (numérique ou caractère) à p modalités définit une partition de la population en p classes, une classe étant constituée par l'ensemble des observations prenant une même modalité
- pour chaque variable qualitative²⁶ à q modalités, q variables numériques représentant les q modalités.

La macro édite les tableaux de contingence croisant la variable de classe avec chaque variable qualitative, ainsi que des tableaux et des tests statistiques permettant de caractériser les classes de la partition par les modalités des variables explicatives.

Le § "Les tables SAS en entrée des macros CAHQQUAL, PARTQUAL et DESQUAL" de la description de la macro CAHQQUAL, présente les tableaux sur lesquels on peut utiliser la macro.

Note : la macro DESQUAL produit le même type de sorties que celles de la macro PARTQUAL lorsque le paramètre DESC de PARTQUAL vaut OUI. Cette macro peut donc être utilisée comme aide à la description d'une partition lorsque l'on désire l'expliquer avec des variables "supplémentaires" n'ayant pas pris part à la classification. On opère alors de la façon suivante :

- on réalise une CAH avec CAHQQUAL
- on construit une partition avec PARTQUAL ; si le paramètre DESC vaut OUI, on obtient la description de la partition par les variables **analysées**
- on utilise DESQUAL, avec la table-OUTPART et la variable-VARCLASS créées par PARTQUAL, pour décrire la partition avec des variables qualitatives quelconques (pouvant inclure les variables analysées)
- on peut également utiliser DESNUM si l'on souhaite décrire la partition avec des variables numériques.

²⁶ La variable qualitative peut ne pas être une variable SAS de la table.

1. Syntaxe de la macro

DATA = nom de table SAS

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

Remarque : on peut utiliser la clause WHERE, ou les options FIRSTOBS ou OBS, pour définir cette table. Par exemple, on peut écrire :

```
DATA=A(WHERE=(SEXE="2")) ou DATA=B(OBS=20)
```

VAR = liste de variables

liste des variables **numériques**, séparées par des blancs (**obligatoire**) : ces variables-SAS représentent les modalités des variables qualitatives explicatives.

POIDS = variable

nom d'une variable numérique servant à pondérer les observations. Les valeurs de cette variable doivent être des entiers ; si une valeur n'est pas un entier, c'est la partie entière de cette valeur qui est utilisée.

Une observation prenant une valeur négative, nulle ou manquante sur cette variable ne participe pas à la description.

Note : lorsque l'on utilise DESQUAL à la suite des macros CAHQUAL et PARTQUAL, ce paramètre POIDS doit prendre la même valeur que dans les deux autres macros si l'on veut décrire la partition créée par PARTQUAL ; en revanche, si l'on veut décrire une autre partition, incluant en particulier des "observations supplémentaires" (générées lorsque le paramètre OBSSUP de PARTQUAL vaut OUI), ce paramètre POIDS doit être changé (ou laissé à blanc le cas échéant).

Par défaut : toutes les observations ont le même poids.

VARCLASS = nom de variable

nom de la variable de classe (**obligatoire**).

MISSING = OUI ou NON

si MISSING vaut OUI, les observations ayant une valeur manquante pour la variable de classe constituent une classe à part entière.

Par défaut : MISSING = NON.

NCL = n

nombre de modalités de la variable de classe.

Si NCL n'est pas renseigné, la macro calcule ce nombre.

Note : lorsque l'on utilise DESQUAL à la suite de la macro PARTQUAL, ce paramètre NCL doit prendre la même valeur que dans cette macro, et doit être augmenté d'une unité si MISSING vaut OUI

NOTES = O ou N

si NOTES vaut N (= non), les notes produites par SAS durant l'exécution de la macro ne sont pas éditées (sur la Log).

Par défaut : NOTES = N

LIBVAR = liste de libellés

liste des libellés des **variables qualitatives** séparés par des blancs (les blancs ne sont pas autorisés dans les libellés). Ces libellés seront utilisés dans les tableaux de description des classes.

Lorsque l'on opère sur plusieurs variables qualitatives, l'ordre dans lequel sont donnés ces noms doit être le même que l'ordre dans lequel ont été spécifiées les modalités de ces variables dans le paramètre VAR. **Il est nécessaire que dans ce paramètre VAR les modalités apparaissent regroupées par variable** (même si l'on utilise les libellés par défaut indiqués plus loin).

On se reportera à l'exemple présenté plus loin.

Par défaut : les libellés des variables sont Var001, Var002...

NBMODA = liste des nombres de modalités

pour chacune des variables qualitatives, on indique le nombre de modalités²⁷ de la variable ; ces nombres de modalités sont donnés dans le même ordre que dans la liste LIBVAR, séparés par des blancs.

La somme de ces nombres des modalités **doit être égal** au nombre de variables SAS du paramètre VAR.

Ce paramètre est **obligatoire**, sauf dans le cas d'un tableau de contingence (une seule variable qualitative).

Exemples :

- dans le cas d'un tableau de contingence, où la variable qualitative en colonne est la variable "état matrimonial" à 4 modalités, on peut écrire par exemple :

```
VAR = CELIBAT MARIÉS VEUFS DIVORCÉS, LIBVAR = état_matrimonial  
OU
```

```
VAR = CELIBAT MARIÉS VEUFS DIVORCÉS
```

- dans le cas d'un tableau disjonctif complet à trois variables qualitatives X, Y et Z, telles que la variable X a deux modalités X1 et X2, la variable Y trois modalités Y1, Y2 et Y3, et la variable Z deux modalités Z1 et Z2, on peut écrire par exemple :

```
VAR = X1 X2 Y1-Y3 Z1 Z2, LIBVAR = VarX VarY VarZ, NBMODA = 2 3 2
```

Note : le paramètre LIBVAR peut comporter moins de libellés que le nombre de valeurs spécifiées dans la liste NBMODA : les libellés "manquants" sont alors VARn, VARn + 1,...

PROBA = probabilité

ce paramètre permet de limiter la liste des modalités caractérisant les différentes classes. Seules sont listées les modalités pour lesquelles le niveau de significativité du test est inférieur ou égal à ce paramètre.

Par défaut : PROBA = 1 (toutes les modalités sont listées).

²⁷ Il s'agit des modalités prises par au moins une observation.

2. Les sorties imprimées

La macro édite :

- le(s) **tableau(x) de contingence** croisant la variable de classe et chaque variable qualitative
- le tableau des **CHI2 des tableaux de contingence** ; ce tableau donne, pour chaque variable qualitative :
 - son **nombre de modalités**
 - **l'effectif pondéré du tableau**, i.e. l'effectif total du tableau de contingence ; c'est aussi la somme des poids des observations dans le (sous-)tableau associé à la variable qualitative
 - le nombre de cases du tableau de contingence ayant un effectif inférieur à 5 (**Nbre de cases d'effectif < 5**)
 - la proportion dans le tableau de contingence de cases d'effectif inférieur à 5 (**Nbre de cases d'effectif < 5 (en %)**)
 - la valeur de la statistique du chi-deux permettant de réaliser le test d'indépendance entre la variable de classe et la variable qualitative (**CHI2**)
 - le nombre de degrés de liberté associé à la statistique du chi-deux, égal à $(p-1)(q-1)$, où p est le nombre de modalités de la variable de classe et q le nombre de modalités de la variable qualitative (**Degrés de liberté**)
 - le niveau de significativité du test d'indépendance du chi-deux (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'indépendance entre la variable de classe et la variable qualitative doit être rejetée, autrement dit que la variable qualitative n'a pas la même répartition dans les différentes classes.

Dans ce tableau, les variables qualitatives sont classées selon les valeurs croissantes de PROBA²⁸.

- un tableau donnant **la caractérisation de chaque classe de la partition par les variables explicatives** ; ce tableau donne, pour chaque modalité de la variable de classe et pour chaque variable qualitative (identifiée par le libellé donné dans le paramètre LIBVAR) :
 - **l'effectif pondéré** de la classe, i.e. la somme des poids des observations constituant la classe, dans le sous-tableau associé à la variable
 - la distance du chi-deux (au sens de distance entre deux lois de probabilité) entre la distribution de la variable dans la population et la distribution de la variable dans la classe (**CHI2**)
 - le nombre de degrés de liberté associé à cette distance, égal à $q-1$, où q est le nombre de modalités de la variable qualitative (**Degrés de liberté**)
 - le niveau de significativité du test permettant de juger de la significativité des écarts entre les deux distributions (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'égalité des deux distributions doit être rejetée, autrement dit que la variable qualitative n'a pas la même répartition dans la classe et dans la population.

²⁸ Le classement relatif de variables ayant des valeurs de PROBA nulles et des degrés de liberté associés différents est sans signification.

Dans ce tableau, les variables qualitatives sont rangées, au sein de chaque classe, selon les valeurs croissantes de PROBA²⁹.

- un tableau donnant la **caractérisation de la partition par les modalités des variables explicatives** ; ce tableau donne, pour chaque modalité de chaque variable qualitative :
 - le **label** de la modalité de la variable qualitative (qui est le nom de la variable-SAS si aucun label ne lui a été attribué)
 - le libellé de la **variable** qualitative
 - l'**effectif pondéré** de la modalité, i.e. le total (éventuellement pondéré) de la colonne représentant la modalité
 - la distance du chi-deux (au sens de distance entre deux lois de probabilité) entre la distribution de la variable de classe dans la population et la distribution de cette variable de classe parmi les observations prenant la modalité considérée (**CHI2**)
 - le nombre de degrés de liberté associé à cette distance, égal à $p-1$, où p est le nombre de modalités de la variable de classe (**Degrés de liberté**)
 - le niveau de significativité du test permettant de juger de la significativité des écarts entre les deux distributions (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'égalité des deux distributions doit être rejetée, autrement dit que la variable de classe n'a pas la même répartition dans la population entière et dans la sous-population définie par la modalité considérée.

Dans ce tableau, les modalités des variables qualitatives sont rangées selon les valeurs croissantes de PROBA.

- un tableau donnant la **caractérisation de chaque classe de la partition par les modalités des variables explicatives** ; ce tableau donne, pour chaque modalité de la variable de classe et pour chaque modalité de chaque variable qualitative :
 - le **label** de la modalité de la variable qualitative (qui est le nom de la variable-SAS si aucun label ne lui a été attribué)
 - le libellé de la **variable** qualitative
 - l'**effectif pondéré** de la modalité dans la classe, i.e. la somme des poids des observations de la classe prenant la modalité considérée
 - la fréquence de la modalité dans la classe (**Fréquence dans la classe (%)**) : c'est le rapport entre l'effectif pondéré de la modalité dans la classe et l'effectif pondéré de la classe
 - la fréquence de la modalité dans la population (**Fréquence dans la pop. (%)**) : c'est le rapport entre l'effectif pondéré de la modalité et l'effectif pondéré du sous-tableau associé à la variable
 - le niveau de significativité du test permettant de juger de la significativité de l'écart entre ces deux fréquences³⁰ (**PROBA**) : une valeur faible de PROBA signifie que l'hypothèse d'égalité des deux fréquences doit être rejetée.

²⁹ Le classement relatif de variables ayant des valeurs de PROBA nulles et des degrés de liberté associés différents est sans signification.

³⁰ La statistique du test dépend de la taille de l'échantillon, car elle suit, sous l'hypothèse d'égalité des deux fréquences, une loi hypergéométrique : pour les grands échantillons, on transforme cette statistique pour se ramener à une loi normale centrée réduite.

- la valeur de la variable aléatoire normale centrée réduite "équivalente" (**Val.Test**) conduisant au même niveau de significativité (pour un test bilatéral).

Dans ce tableau, les modalités des variables qualitatives sont rangées, au sein de chaque classe, par valeur-test décroissante, et regroupées en deux catégories : les modalités sur-représentées (resp. sous-représentées) sont telles que la fréquence dans la classe est supérieure (resp. inférieure) à la fréquence dans la population.

Exemples

1. Analyse en composantes principales

Exemple 1 : Les vins français

```
LIBNAME COMPIL ...;
OPTIONS SASMSTORE=COMPIL MSTORED NODATE PAGESIZE=66;

DATA A;
INPUT ANNEE $ BORD_ROU BORD_BLA BOUR_ROU BOUR_BLA BEAUJOL ALSACE COT_RHON
      PAY_LOIR TEMP_MOY TEMP_MAX TEMP_MIN PLUIE;

CARDS;
69 1 2 6 5 4 3 4 3 11.7 32.8 -8.1 618
70 6 5 3 3 4 3 4 3 11.6 31.2 -6.1 631
71 5 5 5 3 4 6 3 3 11.8 31.6 -8.6 508
72 2 3 4 2 3 1 3 1 11.1 30.1 -7.2 740
73 3 3 2 4 4 4 2 3 11.6 32.7 -4.2 576
74 2 3 2 2 2 3 1 1 12.0 34.0 -0.4 668
75 6 1 1 3 3 4 1 3 11.7 35.7 -4.5 658
76 4 3 5 3 6 5 4 4 12.4 35.4 -6.2 417
77 3 3 2 2 2 3 1 1 11.7 29.0 -3.7 717
78 4 5 6 5 5 5 6 5 10.9 30.1 -10.3 743
79 4 3 4 4 4 4 5 4 11.0 32.1 -12.7 729
80 2 4 2 3 1 3 3 2 11.2 31.3 -5.1 690
81 5 5 3 4 3 4 3 5 11.8 31.9 -3.1 746
82 6 5 4 3 3 4 4 5 12.4 32.8 -6.6 700
83 4 4 3 4 5 6 4 4 12.3 33.4 -4.2 623
84 2 3 2 2 1 2 2 4 11.7 34.8 -3.3 745
85 5 5 5 4 5 5 3 5 10.8 32.0 -15.9 501
;
PROC PRINT;
TITLE 'Notes (de 1 à 6) données aux vins français de 1969 à 1985';
RUN;

TITLE 'Analyse en composantes principales sur les vins français';
%ACP(DATAACT=A,
      VARACT=BORD_ROU--PAY_LOIR,
      VARSUP=TEMP_MOY--PLUIE,
      ID=ANNEE,
      VECP=5,
      IOA=3,
      IVA=3,
      IVS=3,
      OUT=SOR,
      NAXER=3,
      FILL=ALL);

PROC PRINT DATA=SOR;
TITLE2 "Table en sortie créée par la macro";
RUN;

OPTIONS PAGESIZE=95;

TITLE2 "Les variables dans le plan 1-2";
%PLOTACP(AXEH=1,AXEV=2,POINTS=VARACT VARSUP);

OPTIONS PAGESIZE=66;

TITLE2 "Les individus dans le plan 1-2";
%PLOTACP(AXEH=1,AXEV=2,POINTS=OBSACT);

TITLE2 "Les individus et les axes unitaires dans le plan 1-2";
%PLOTACP(AXEH=1,AXEV=2,POINTS=OBSACT AXEUNI);
```

Notes (de 1 à 6) données aux vins français de 1969 à 1985

OBS	ANNEE	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR	TEMP_MOY	TEMP_MAX	TEMP_MIN	PLUIE
1	69	1	2	6	5	4	3	4	3	11.7	32.8	-8.1	618
2	70	6	5	3	3	4	3	4	3	11.6	31.2	-6.1	631
3	71	5	5	5	3	4	6	3	3	11.8	31.6	-8.6	508
4	72	2	3	4	2	3	1	3	1	11.1	30.1	-7.2	740
5	73	3	3	2	4	4	4	2	3	11.6	32.7	-4.2	576
6	74	2	3	2	2	2	3	1	1	12.0	34.0	-0.4	668
7	75	6	1	1	3	3	4	1	3	11.7	35.7	-4.5	658
8	76	4	3	5	3	6	5	4	4	12.4	35.4	-6.2	417
9	77	3	3	2	2	2	3	1	1	11.7	29.0	-3.7	717
10	78	4	5	6	5	5	5	6	5	10.9	30.1	-10.3	743
11	79	4	3	4	4	4	4	5	4	11.0	32.1	-12.7	729
12	80	2	4	2	3	1	3	3	2	11.2	31.3	-5.1	690
13	81	5	5	3	4	3	4	3	5	11.8	31.9	-3.1	746
14	82	6	5	4	3	3	4	4	5	12.4	32.8	-6.6	700
15	83	4	4	3	4	5	6	4	4	12.3	33.4	-4.2	623
16	84	2	3	2	2	1	2	2	4	11.7	34.8	-3.3	745
17	85	5	5	5	4	5	5	3	5	10.8	32.0	-15.9	501

Analyse en composantes principales sur les vins français

```

*****
*          Caractéristiques de l'analyse          *
*****
*   Nombre de variables actives           =    8   *
*   Nombre de variables supplémentaires =    4   *
*   Nombre de variables de classes       =    0   *
*                                           *
*   Nombre d'individus actifs            =   17   *
*   Nombre d'individus actifs éliminés  =    0   *
*   Nombre d'individus supplémentaires  =    0   *
*   Nombre d'individus supp. éliminés   =    0   *
*                                           *
*   Variable de pondération              =         *
*                                           *
*   Edition des aides à l'interprétation :   *
*     variables actives          sur 3 axes   *
*     variables supplémentaires sur 3 axes   *
*     individus actifs          sur 3 axes   *
*     individus supplémentaires sur 0 axes   *
*     barycentres d'individus   sur 0 axes   *
*****

```

Liste des variables actives :

BORD_ROU BORD_BLA BOUR_ROU BOUR_BLA BEAUJOL ALSACE COT_RHON PAY_LOIR

Liste des variables supplémentaires :

TEMP_MOY TEMP_MAX TEMP_MIN PLUIE

Matrice des corrélations

VARIABLE	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR
BORD_ROU	1.00000	0.43393	0.02227	0.12580	0.38159	0.53507	0.17917	0.53298
BORD_BLA	0.43393	1.00000	0.35818	0.19549	0.20989	0.38124	0.46131	0.46517
BOUR_ROU	0.02227	0.35818	1.00000	0.56005	0.66184	0.34588	0.74774	0.42185
BOUR_BLA	0.12580	0.19549	0.56005	1.00000	0.60995	0.51754	0.64888	0.61143
BEAUJOL	0.38159	0.20989	0.66184	0.60995	1.00000	0.67378	0.59560	0.49080
ALSACE	0.53507	0.38124	0.34588	0.51754	0.67378	1.00000	0.34427	0.56348
COT_RHON	0.17917	0.46131	0.74774	0.64888	0.59560	0.34427	1.00000	0.58179
PAY_LOIR	0.53298	0.46517	0.42185	0.61143	0.49080	0.56348	0.58179	1.00000

Analyse en composantes principales sur les vins français

Les 5 premiers vecteurs propres

NAME	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR
PRIN1	0.24658	0.27551	0.35791	0.37312	0.39706	0.36317	0.39199	0.39086
PRIN2	0.65806	0.30065	-0.42963	-0.28058	-0.10993	0.27868	-0.29819	0.18597
PRIN3	-0.07882	0.71340	0.18984	-0.23877	-0.38425	-0.38630	0.30486	0.04451
PRIN4	0.07118	0.16192	0.36830	-0.49695	0.43927	0.20868	-0.07536	-0.58646
PRIN5	0.50376	-0.40751	0.11761	-0.28198	0.20274	-0.57689	0.28936	0.17082

Valeurs propres

Inertie totale : 8

N°	Val.Pr.	Diff.	Pct	Cum	!
1	4.2397	.	53.00	53.00	!*****
2	1.3624	2.8774	17.03	70.03	!*****
3	0.8886	0.4737	11.11	81.13	!*****
4	0.5667	0.3219	7.08	88.22	!*****
5	0.4082	0.1585	5.10	93.32	!*****
6	0.2375	0.1708	2.97	96.29	!****
7	0.1700	0.0675	2.12	98.41	!***
8	0.1270	0.0430	1.59	100.0	!**

Aides à l'interprétation pour les individus actifs

! Individus actifs !	! AXE1 !					! AXE2 !					! AXE3 !						
! Ident.	CONTR	POIDS	COORD	CTR	RCTR	COORD	CTR	RCTR	COORD	CTR	RCTR	COORD	CTR	RCTR	COORD	CTR	RCTR
! 69	8.93	5.88	! 0.54	0.4	16	2.4	2.4	! -3.27	46.1	1	87.9	90.2	! -0.67	3.0	10	3.7	94.0
! 70	3.37	5.88	! 0.53	0.4	17	6.2	6.2	! 1.06	4.8	6	24.5	30.6	! 1.00	6.6	5	21.8	52.5
! 71	4.49	5.88	! 1.41	2.7	11	32.4	32.4	! 0.93	3.7	8	14.1	46.5	! 0.19	0.2	14	0.6	47.0
! 72	8.28	5.88	! -2.43	8.2	5	52.4	52.4	! -1.54	10.3	3	21.1	73.5	! 0.96	6.1	7	8.2	81.7
! 73	2.15	5.88	! -0.55	0.4	15	10.3	10.3	! -0.07	0.0	15	0.2	10.5	! -1.17	9.1	3	46.9	57.4
! 74	8.21	5.88	! -3.21	14.3	2	92.1	92.1	! -0.02	0.0	17	0.0	92.1	! -0.05	0.0	16	0.0	92.2
! 75	9.15	5.88	! -1.74	4.2	9	24.4	24.4	! 1.57	10.6	2	19.7	44.1	! -2.35	36.5	1	44.3	88.5
! 76	4.66	5.88	! 1.65	3.8	10	43.0	43.0	! -0.46	0.9	11	3.4	46.4	! -0.97	6.2	6	14.9	61.3
! 77	7.44	5.88	! -3.05	12.9	3	91.9	91.9	! 0.40	0.7	12	1.6	93.5	! -0.10	0.1	15	0.1	93.6
! 78	11.35	5.88	! 3.71	19.1	1	89.1	89.1	! -1.05	4.7	7	7.1	96.2	! 0.62	2.5	12	2.5	98.7
! 79	2.44	5.88	! 1.23	2.1	13	45.8	45.8	! -0.74	2.4	9	16.6	62.4	! -0.27	0.5	13	2.2	64.5
! 80	5.12	5.88	! -2.01	5.6	7	58.1	58.1	! -0.28	0.3	14	1.1	59.2	! 1.06	7.4	4	16.0	75.2
! 81	3.15	5.88	! 1.04	1.5	14	25.4	25.4	! 1.13	5.5	5	29.7	55.0	! 0.62	2.6	11	9.1	64.1
! 82	4.20	5.88	! 1.34	2.5	12	31.3	31.3	! 1.34	7.8	4	31.5	62.8	! 1.17	9.1	2	24.1	86.8
! 83	4.05	5.88	! 1.79	4.4	8	58.2	58.2	! 0.37	0.6	13	2.5	60.7	! -0.89	5.3	8	14.5	75.1
! 84	7.74	5.88	! -2.63	9.6	4	65.7	65.7	! 0.03	0.0	16	0.0	65.7	! 0.84	4.7	9	6.8	72.4
! 85	5.26	5.88	! 2.38	7.8	6	79.0	79.0	! 0.61	1.6	10	5.2	84.2	! 0.02	0.0	17	0.0	84.2

Analyse en composantes principales sur les vins français

Aides à l'interprétation pour les variables actives

! Variables actives !	! AXE1 !					! AXE2 !					! AXE3 !						
! Ident.	CONTR	POIDS	COORD	CTR	RCTR	COORD	CTR	RCTR	COORD	CTR	RCTR	COORD	CTR	RCTR	COORD	CTR	RCTR
! BORD_ROU	12.50	12.50	! 0.51	6.1	8	25.8	25.8	! 0.77	43.3	1	59.0	84.8	! -0.07	0.6	7	0.6	85.3
! BORD_BLA	12.50	12.50	! 0.57	7.6	7	32.2	32.2	! 0.35	9.0	3	12.3	44.5	! 0.67	50.9	1	45.2	89.7
! BOUR_ROU	12.50	12.50	! 0.74	12.8	6	54.3	54.3	! -0.50	18.5	2	25.1	79.5	! 0.18	3.6	6	3.2	82.7
! BOUR_BLA	12.50	12.50	! 0.77	13.9	4	59.0	59.0	! -0.33	7.9	5	10.7	69.7	! -0.23	5.7	5	5.1	74.8
! BEAUJOL	12.50	12.50	! 0.82	15.8	1	66.8	66.8	! -0.13	1.2	8	1.6	68.5	! -0.36	14.8	3	13.1	81.6
! ALSACE	12.50	12.50	! 0.75	13.2	5	55.9	55.9	! 0.33	7.8	6	10.6	66.5	! -0.36	14.9	2	13.3	79.8
! COT_RHON	12.50	12.50	! 0.81	15.4	2	65.1	65.1	! -0.35	8.9	4	12.1	77.3	! 0.29	9.3	4	8.3	85.5
! PAY_LOIR	12.50	12.50	! 0.80	15.3	3	64.8	64.8	! 0.22	3.5	7	4.7	69.5	! 0.04	0.2	8	0.2	69.7

Analyse en composantes principales sur les vins français

Aides à l'interprétation pour les variables supplémentaires

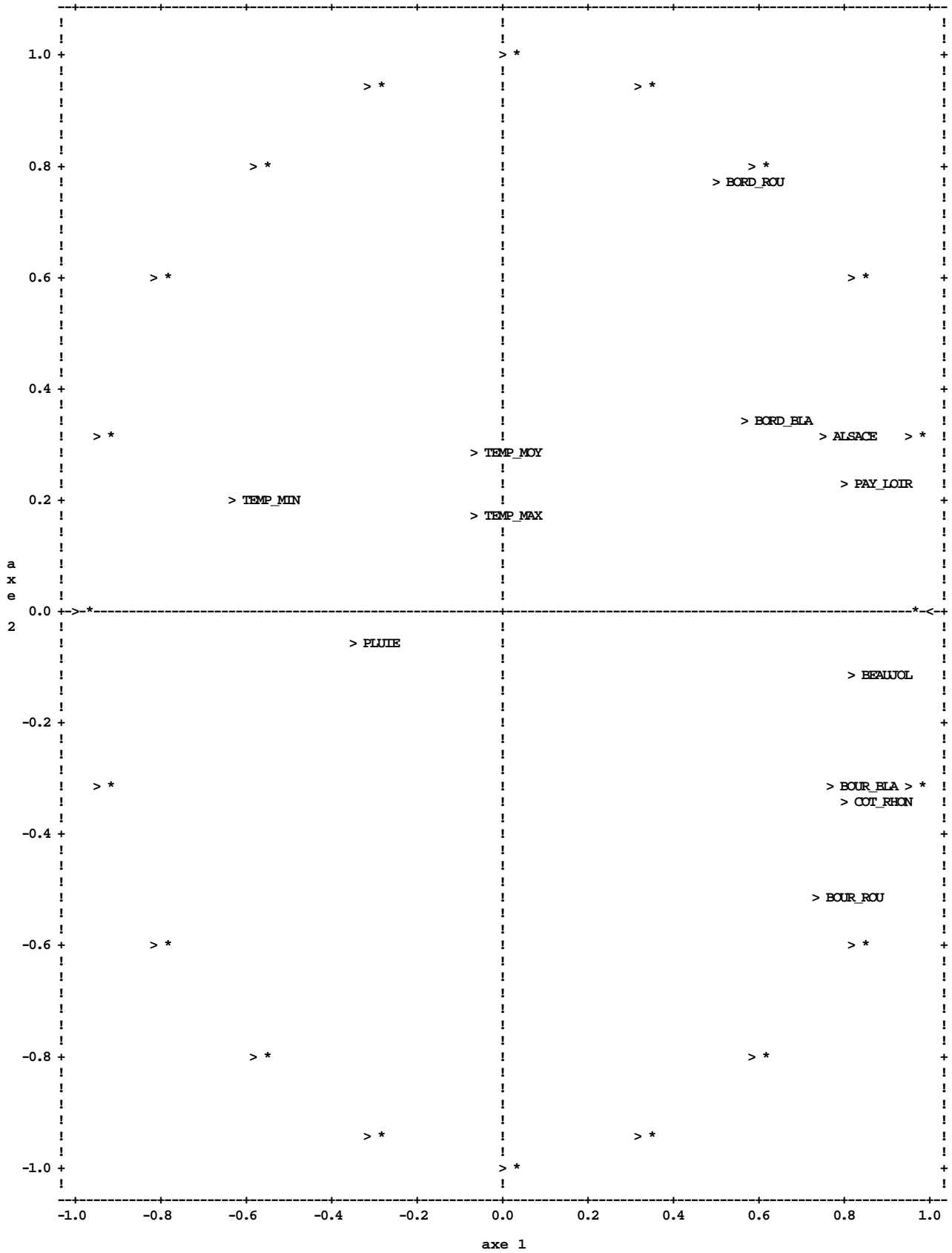
! Variables supplém.	! _____ AXE1 _____ !	! _____ AXE2 _____ !	! _____ AXE3 _____ !
! Ident.	DIST POIDS ! COORD CIR CO2 RCO2 QLT !	COORD CIR CO2 RCO2 QLT !	COORD CIR CO2 RCO2 QLT !
! TEMP_MOY	1.00 . ! -0.06 . 0.4 3 0.4 !	0.29 . 8.7 1 9.1 !	-0.23 . 5.1 3 14.2 !
! TEMP_MAX	1.00 . ! -0.06 . 0.4 4 0.4 !	0.16 . 2.5 3 2.9 !	-0.55 . 29.7 1 32.6 !
! TEMP_MIN	1.00 . ! -0.63 . 40.2 1 40.2 !	0.21 . 4.6 2 44.8 !	-0.08 . 0.6 4 45.4 !
! PLUIE	1.00 . ! -0.36 . 12.7 2 12.7 !	-0.07 . 0.4 4 13.1 !	0.42 . 18.0 2 31.1 !

Analyse en composantes principales sur les vins français
Table en sortie créée par la macro

OBS	COORD1	COORD2	COORD3	LABEL	CO2_1	CO2_2	CO2_3	CIR_1	CIR_2	CIR_3	QLT_1	QLT_2	QLT_3	DIST	TYPE
1	4.23973	1.36236	0.88861	VALP
2	0.98631	2.63226	-0.31527	+BORD_RO	AXEUNI
3	1.10203	1.20260	2.85361	+BORD_EL	AXEUNI
4	1.43164	-1.71853	0.75937	+BOUR_RO	AXEUNI
5	1.49248	-1.12231	-0.95507	+BOUR_EL	AXEUNI
6	1.58826	-0.43971	-1.53699	+BEAUJOL	AXEUNI
7	1.45270	1.11474	-1.54522	+ALSACE	AXEUNI
8	1.56794	-1.19278	1.21943	+COT_RHO	AXEUNI
9	1.56343	0.74387	0.17803	+PAY_LOI	AXEUNI
10	0.50772	0.76809	-0.07430	BORD_ROU	25.7779	58.9969	0.5520	6.0801	43.3049	0.6212	25.7779	84.7748	85.3268	1.0000	VARACT
11	0.56729	0.35092	0.67250	BORD_ELA	32.1815	12.3144	45.2251	7.5905	9.0390	50.8941	32.1815	44.4959	89.7210	1.0000	VARACT
12	0.73696	-0.50147	0.17896	BOUR_ROU	54.3106	25.1469	3.2026	12.8099	18.4584	3.6040	54.3106	79.4575	82.6601	1.0000	VARACT
13	0.76828	-0.32749	-0.22508	BOUR_ELA	59.0247	10.7250	5.0660	13.9218	7.8723	5.7010	59.0247	69.7497	74.8157	1.0000	VARACT
14	0.81758	-0.12831	-0.36222	BEAUJOL	66.8435	1.6463	13.1200	15.7660	1.2084	14.7646	66.8435	68.4898	81.6097	1.0000	VARACT
15	0.74780	0.32528	-0.36415	ALSACE	55.9204	10.5808	13.2608	13.1896	7.7665	14.9231	55.9204	66.5011	79.7620	1.0000	VARACT
16	0.80712	-0.34805	0.28738	COT_RHON	65.1445	12.1141	8.2586	15.3652	8.8920	9.2938	65.1445	77.2585	85.5171	1.0000	VARACT
17	0.80480	0.21706	0.04196	PAY_LOIR	64.7701	4.7116	0.1760	15.2769	3.4584	0.1981	64.7701	69.4818	69.6578	1.0000	VARACT
18	-0.06268	0.29426	-0.22602	TEMP_MOY	0.3929	8.6591	5.1084	.	.	.	0.3929	9.0520	14.1604	1.0000	VARSUP
19	-0.06052	0.15916	-0.54535	TEMP_MAX	0.3663	2.5331	29.7407	.	.	.	0.3663	2.8993	32.6400	1.0000	VARSUP
20	-0.63403	0.21404	-0.07828	TEMP_MIN	40.1995	4.5815	0.6127	.	.	.	40.1995	44.7810	45.3937	1.0000	VARSUP
21	-0.35658	-0.06528	0.42398	PLUIE	12.7148	0.4262	17.9762	.	.	.	12.7148	13.1410	31.1171	1.0000	VARSUP
22	0.53825	-3.26612	-0.67184	69	2.3861	87.8560	3.7174	0.4020	46.0600	2.9879	2.3861	90.2420	93.9594	12.1421	OBSACT
23	0.53217	1.05859	1.00014	70	6.1823	24.4625	21.8357	0.3929	4.8386	6.6216	6.1823	30.6448	52.4805	4.5810	OBSACT
24	1.40596	0.92639	0.18548	71	32.4061	14.0692	0.5640	2.7426	3.7055	0.2277	32.4061	46.4753	47.0394	6.0998	OBSACT
25	-2.42949	-1.54121	0.96329	72	52.3840	21.0812	8.2354	8.1892	10.2562	6.1426	52.3840	73.4652	81.7006	11.2676	OBSACT
26	-0.54986	-0.07448	-1.17245	73	10.3235	0.1894	46.9371	0.4195	0.0240	9.0998	10.3235	10.5129	57.4500	2.9287	OBSACT
27	-3.20740	-0.02144	-0.05390	74	92.1358	0.0041	0.0260	14.2731	0.0020	0.0192	92.1358	92.1399	92.1660	11.1655	OBSACT
28	-1.74291	1.56654	-2.34914	75	24.4053	19.7159	44.3355	4.2147	10.5960	36.5306	24.4053	44.1212	88.4567	12.4470	OBSACT
29	1.65223	-0.46096	-0.97146	76	43.0477	3.3507	14.8818	3.7875	0.9175	6.2472	43.0477	46.3984	61.2803	6.3415	OBSACT
30	-3.04873	0.40200	-0.10461	77	91.8621	1.5971	0.1082	12.8959	0.6978	0.0724	91.8621	93.4592	93.5674	10.1182	OBSACT
31	3.70872	-1.04716	0.61705	78	89.1338	7.1059	2.4674	19.0837	4.7346	2.5205	89.1338	96.2397	98.7071	15.4314	OBSACT
32	1.23340	-0.74188	-0.26825	79	45.7884	16.5661	2.1658	2.1107	2.3765	0.4763	45.7884	62.3545	64.5204	3.3224	OBSACT
33	-2.01239	-0.28126	1.05595	80	58.1133	1.1352	16.0006	5.6187	0.3416	7.3811	58.1133	59.2485	75.2491	6.9686	OBSACT
34	1.04342	1.12803	0.62392	81	25.3755	29.6572	9.0729	1.5105	5.4941	2.5769	25.3755	55.0327	64.1056	4.2905	OBSACT
35	1.33712	1.34040	1.17263	82	31.2997	31.4537	24.0726	2.4806	7.7576	9.1025	31.2997	62.7535	86.8260	5.7121	OBSACT
36	1.79005	0.36726	-0.89202	83	58.2284	2.4511	14.4596	4.4457	0.5824	5.2673	58.2284	60.6795	75.1391	5.5029	OBSACT
37	-2.62861	0.03456	0.84472	84	65.6557	0.0113	6.7802	9.5866	0.0052	4.7235	65.6557	65.6671	72.4473	10.5240	OBSACT
38	2.37806	0.61076	0.02050	85	79.0197	5.2123	0.0059	7.8462	1.6106	0.0028	79.0197	84.2320	84.2379	7.1567	OBSACT

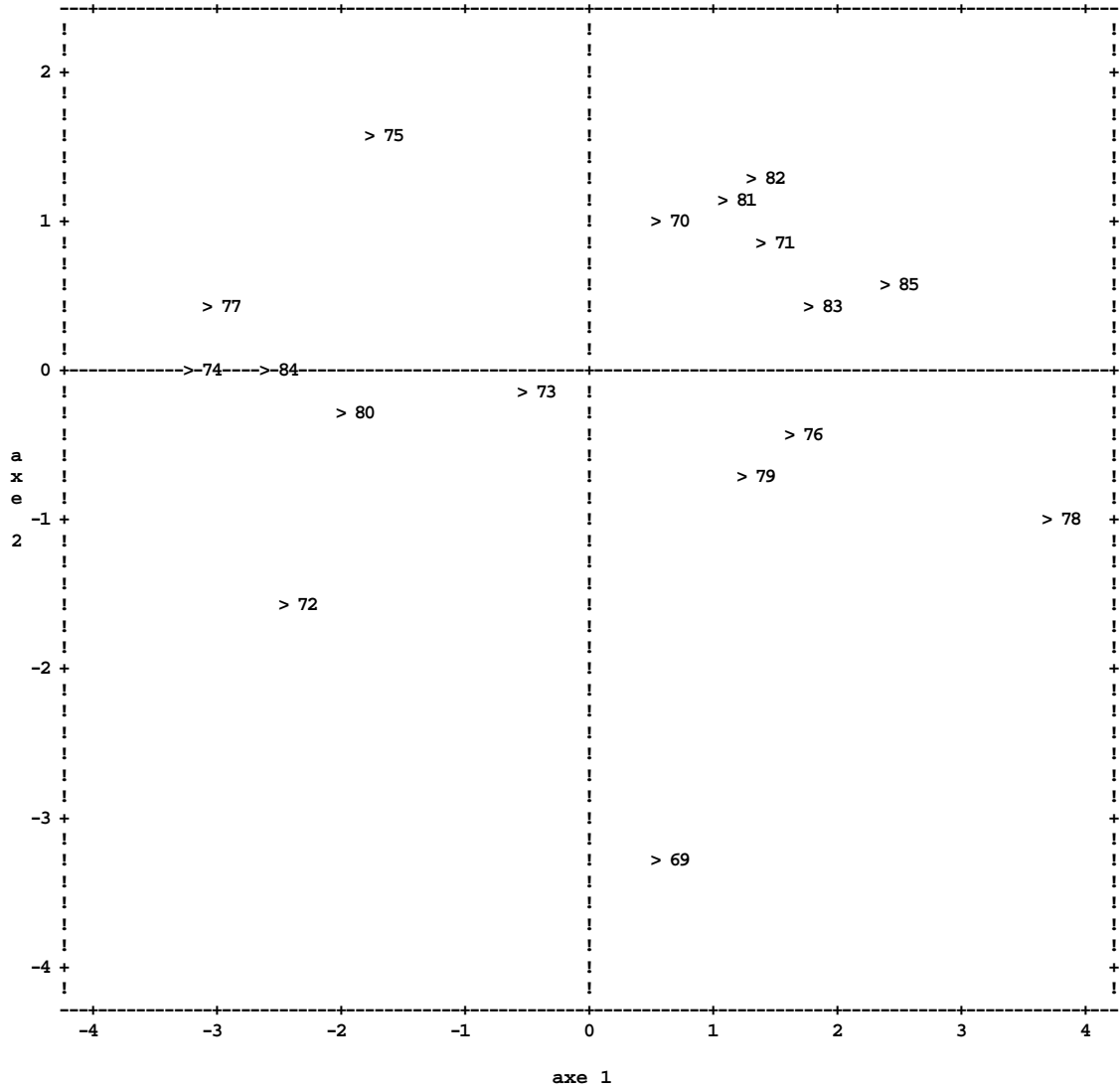
Analyse en composantes principales sur les vins français
 Les variables dans le plan 1-2

Plot of COORD2*COORD1\$ LABEL_. Symbol points to label.



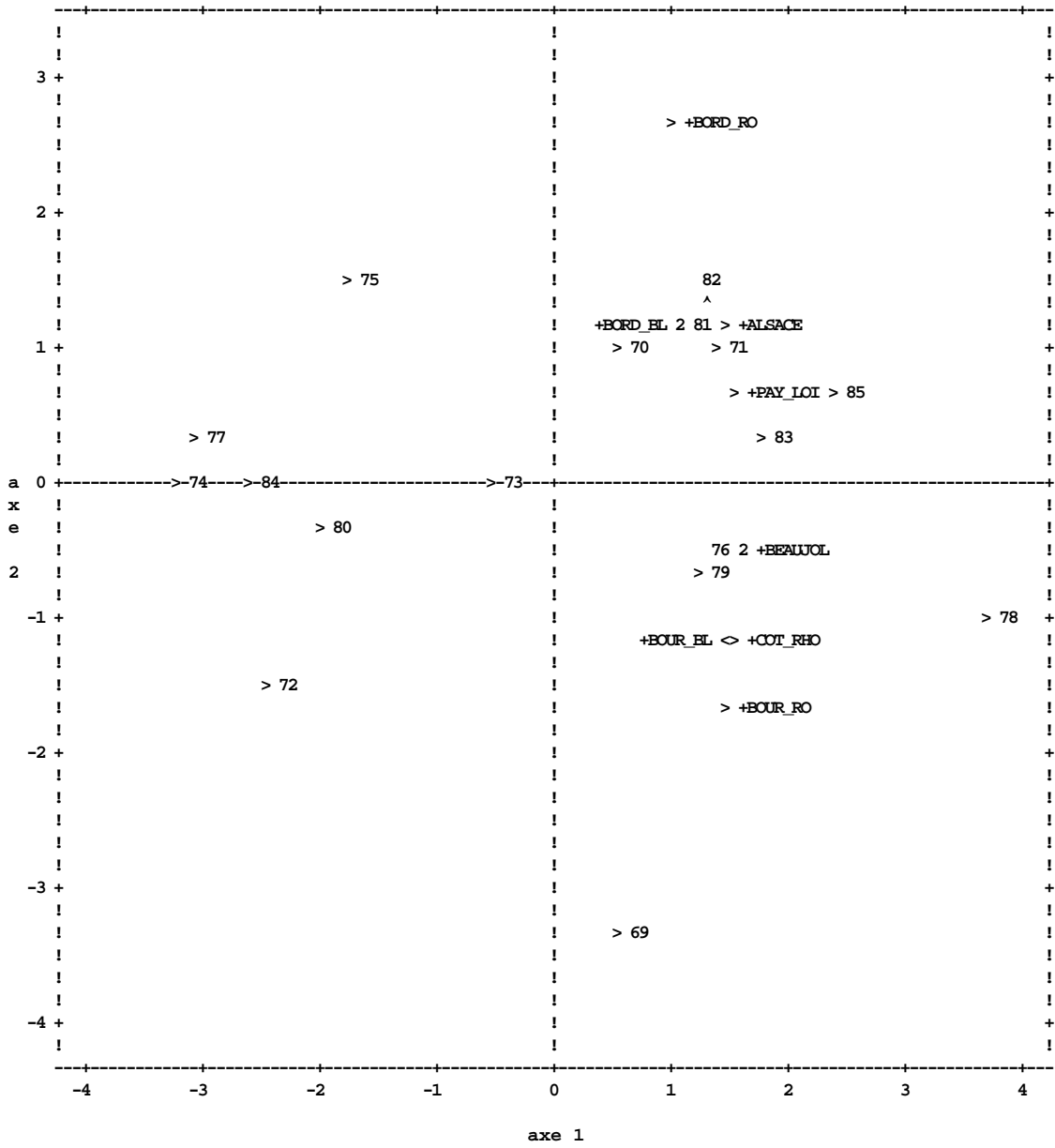
Analyse en composantes principales sur les vins français
Les individus dans le plan 1-2

Plot of COORD2*COORD1\$ LABEL_. Symbol points to label.



Analyse en composantes principales sur les vins français
 Les individus et les axes unitaires dans le plan 1-2

Plot of COORD2*COORD1\$ LABEL_. Symbol points to label.



Exemple 2 : Les coefficients budgétaires des ménages

```
LIBNAME COMPIL ...;

OPTIONS SASMSTORE=COMPIL MSTORED NODATE PAGESIZE=66;

PROC FORMAT;
VALUE $AGE 1='<25 ans'
           2='25-34'
           3='35-44'
           4='45-54'
           5='55-64'
           6='65-74'
           7='>75 ans';
VALUE $CCOM 0='C.rurale'
            1='Petit.uu'
            2='Moyen.uu'
            3='Gross.uu'
            4='Agg.Pari'
            5='Paris';
VALUE $CS 1='Agricult'
          2='Art.comm'
          3='Cadr.sup'
          4='Pro.inte'
          5='Employés'
          6='Ouvriers'
          7='Retraité'
          8='Aut.inac';

TITLE'Analyse en composantes principales sur les coefficients budgétaires des ménages';

%ACP(DATAACT=BUDGET,
      VARACT=ALIMENTA--IMP_EPAR HYGIENE TELECOM,
      IVA=4,
      CLASSES=AGE_CHEF CAT_COMM CS_CHEF,
      AXECLAS=4,
      FMTCLAS=$AGE. $CCOM. $CS.,
      OUT=SOR,
      NAXER=4,
      FILL=VAR BARY);

OPTIONS PAGESIZE=95;

TITLE2 "Les variables dans le plan 1-2";
%PLOTACP(AXEH=1,AXEV=2,POINTS=VARACT);

OPTIONS PAGESIZE=66;

TITLE2 "Les barycentres d'individus dans le plan 1-2";
%PLOTACP(AXEH=1,AXEV=2,POINTS=OBSSUP);

OPTIONS PAGESIZE=95;

TITLE2 "Les variables dans le plan 3-4";
%PLOTACP(AXEH=3,AXEV=4,POINTS=VARACT);

OPTIONS PAGESIZE=66;

TITLE2 "Les barycentres d'individus dans le plan 3-4";
%PLOTACP(AXEH=3,AXEV=4,POINTS=OBSSUP);
```

Analyse en composantes principales sur les coefficients budgétaires des ménages

```

*****
*          Caractéristiques de l'analyse          *
*****
*   Nombre de variables actives           =   10   *
*   Nombre de variables supplémentaires   =    0   *
*   Nombre de variables de classes       =    3   *
*                                           *
*   Nombre d'individus actifs            =  9038  *
*   Nombre d'individus actifs éliminés   =    0   *
*   Nombre d'individus supplémentaires   =    0   *
*   Nombre d'individus supp. éliminés    =    0   *
*                                           *
*   Variable de pondération              =         *
*                                           *
*   Edition des aides à l'interprétation :      *
*     variables actives sur 4 axes             *
*     variables supplémentaires sur 0 axes     *
*     individus actifs sur 0 axes             *
*     individus supplémentaires sur 0 axes    *
*     barycentres d'individus sur 4 axes     *
*****

```

Liste des variables actives :

ALIMENTA HABILLEM LOG_ENER EQUI_LOG SANTE TRANSPOR CULT_LOI IMP_EPAR HYGIENE TELECOM

Liste des variables de classes (définissant des barycentres d'individus) :

AGE_CHEF CAT_COMM CS_CHEF

Matrice des corrélations

VARIABLE	ALIMENTA	HABILLEM	LOG_ENER	EQUI_LOG	SANTE	TRANSPOR	CULT_LOI	IMP_EPAR	HYGIENE	TELECOM
ALIMENTA	1.00000	-0.08134	-0.14303	-0.12095	-0.09383	-0.33495	-0.13553	-0.24203	-0.01656	0.11077
HABILLEM	-0.08134	1.00000	-0.18890	-0.02619	-0.09238	-0.11129	0.04371	-0.05542	0.10559	-0.04876
LOG_ENER	-0.14303	-0.18890	1.00000	-0.09980	-0.15719	-0.31410	-0.23129	-0.20922	-0.07755	0.07596
EQUI_LOG	-0.12095	-0.02619	-0.09980	1.00000	-0.07456	-0.13377	-0.07404	-0.08832	0.00424	-0.05860
SANTE	-0.09383	-0.09238	-0.15719	-0.07456	1.00000	-0.16485	-0.12891	-0.12498	-0.02720	-0.03224
TRANSPOR	-0.33495	-0.11129	-0.31410	-0.13377	-0.16485	1.00000	-0.10060	-0.12037	-0.10235	-0.16696
CULT_LOI	-0.13553	0.04371	-0.23129	-0.07404	-0.12891	-0.10060	1.00000	-0.03253	0.01266	-0.03640
IMP_EPAR	-0.24203	-0.05542	-0.20922	-0.08832	-0.12498	-0.12037	-0.03253	1.00000	-0.04374	-0.06487
HYGIENE	-0.01656	0.10559	-0.07755	0.00424	-0.02720	-0.10235	0.01266	-0.04374	1.00000	0.00034
TELECOM	0.11077	-0.04876	0.07596	-0.05860	-0.03224	-0.16696	-0.03640	-0.06487	0.00034	1.00000

Valeurs propres

Inertie totale : 10

N°	Val.Pr.	Diff.	Pct	Cum	!
1	1.5633	.	15.63	15.63	!*****
2	1.3113	0.2520	13.11	28.75	!*****
3	1.1571	0.1542	11.57	40.32	!*****
4	1.1130	0.0441	11.13	51.45	!*****
5	1.0949	0.0181	10.95	62.40	!*****
6	1.0113	0.0836	10.11	72.51	!*****
7	0.9721	0.0392	9.72	82.23	!*****
8	0.9009	0.0712	9.01	91.24	!*****
9	0.8760	0.0249	8.76	100.0	!*****
10	0.0000	0.8760	0.00	100.0	!

Test d'égalité des coordonnées (sur chacun des 4 premiers axes)
des centres de gravité, des classes définies par les variables AGE_CHEF CAT_COMM CS_CHEF
(test F d'analyse de la variance de FISHER-SNEDECOR)
RCOR = rapport de corrélation, FISH = Statistique de FISHER, PROB = Probabilité de dépassement

VARIABLE	RCOR1	RCOR2	RCOR3	RCOR4	FISH1	FISH2	FISH3	FISH4	PROB1	PROB2	PROB3	PROB4
AGE_CHEF	0.09	0.01	0.02	0.02	158	20.7	33.7	25.5	0.000	0.000	0.000	0.000
CAT_COMM	0.00	0.01	0.03	0.00	6.85	21.4	46.8	4.77	0.000	0.000	0.000	0.000
CS_CHEF	0.11	0.01	0.06	0.01	156	17.4	81.3	10.1	0.000	0.000	0.000	0.000

Analyse en composantes principales sur les coefficients budgétaires des ménages

Test de nullité des coordonnées (sur chacun des 4 premiers axes)
des centres de gravité des classes définies par les variables AGE_CHEF CAT_COMM CS_CHEF
COORD = Coordonnée , VALTEST = Statistique du test, PROB = Probabilité de dépassement

----- VARIABLE=AGE_CHEF -----														
MODALITE	EFFECTIF	POIDS	COORD1	COORD2	COORD3	COORD4	VALTEST1	VALTEST2	VALTEST3	VALTEST4	PROB1	PROB2	PROB3	PROB4
<25 ans	350	3.9	-0.43	0.25	0.05	-0.01	-6.55	4.20	0.81	-0.09	0.000	0.000	0.419	0.926
25-34	1717	19.0	-0.17	-0.22	0.11	0.18	-6.44	-8.80	4.68	7.65	0.000	0.000	0.000	0.000
35-44	2062	22.8	-0.18	-0.07	0.14	-0.01	-7.37	-2.95	6.85	-0.46	0.000	0.003	0.000	0.644
45-54	1432	15.8	-0.34	0.03	0.08	-0.21	-11.3	1.23	3.05	-8.33	0.000	0.220	0.002	0.000
55-64	1447	16.0	-0.00	0.06	-0.04	-0.09	-0.01	2.23	-1.70	-3.50	0.990	0.026	0.090	0.000
65-74	1099	12.2	0.40	0.15	-0.22	-0.03	11.35	4.65	-7.24	-1.00	0.000	0.000	0.000	0.316
>75 ans	931	10.3	0.93	0.13	-0.33	0.20	24.05	3.61	-9.84	6.13	0.000	0.000	0.000	0.000

----- VARIABLE=CAT_COMM -----														
MODALITE	EFFECTIF	POIDS	COORD1	COORD2	COORD3	COORD4	VALTEST1	VALTEST2	VALTEST3	VALTEST4	PROB1	PROB2	PROB3	PROB4
C.rurale	2577	28.5	0.10	-0.07	-0.24	-0.05	4.95	-3.85	-13.1	-3.04	0.000	0.000	0.000	0.002
Petit.uu	1530	16.9	0.03	-0.14	-0.02	0.02	0.96	-5.42	-0.75	0.72	0.338	0.000	0.455	0.474
Moyen.uu	1173	13.0	-0.05	-0.01	0.06	0.05	-1.50	-0.34	1.89	1.80	0.133	0.733	0.059	0.072
Gross.uu	2530	28.0	-0.05	0.10	0.10	0.03	-2.18	5.26	5.66	1.53	0.029	0.000	0.000	0.127
Agg.Pari	888	9.8	-0.13	0.02	0.19	0.05	-3.37	0.50	5.54	1.47	0.001	0.617	0.000	0.143
Paris	340	3.8	-0.04	0.44	0.42	-0.18	-0.56	7.24	7.26	-3.28	0.576	0.000	0.000	0.001

----- VARIABLE=CS_CHEF -----														
MODALITE	EFFECTIF	POIDS	COORD1	COORD2	COORD3	COORD4	VALTEST1	VALTEST2	VALTEST3	VALTEST4	PROB1	PROB2	PROB3	PROB4
Agricult	321	3.6	0.06	-0.03	-0.36	-0.26	0.94	-0.53	-6.05	-4.45	0.345	0.600	0.000	0.000
Art.comm	449	5.0	-0.24	-0.13	0.30	-0.25	-4.26	-2.46	6.05	-5.25	0.000	0.014	0.000	0.000
Cadr.sup	805	8.9	-0.80	0.03	0.58	-0.01	-19.0	0.68	16.06	-0.39	0.000	0.499	0.000	0.700
Pro.inte	1309	14.5	-0.50	-0.04	0.20	0.01	-15.6	-1.31	7.31	0.31	0.000	0.189	0.000	0.757
Employés	1023	11.3	-0.02	-0.01	0.14	0.14	-0.46	-0.16	4.44	4.39	0.644	0.872	0.000	0.000
Ouvriers	2091	23.1	-0.00	-0.17	-0.11	0.01	-0.06	-7.84	-5.29	0.60	0.955	0.000	0.000	0.549
Retraité	2588	28.6	0.44	0.12	-0.26	0.03	21.04	6.50	-14.3	1.83	0.000	0.000	0.000	0.067
Aut.inac	452	5.0	0.61	0.32	-0.01	-0.11	10.57	6.06	-0.20	-2.33	0.000	0.000	0.843	0.020

Aides à l'interprétation pour les variables actives

! Variables actives ! _____ AXE1 _____ ! _____ AXE2 _____ ! _____ AXE3 _____ !																	
! Ident.	CONTR	POIDS	! COORD	CTR	RCTR	CO2	QLT	! COORD	CTR	RCTR	CO2	QLT	! COORD	CTR	RCTR	CO2	QLT
! ALIMENTA	10.00	10.00	! 0.57	21.1	3	33.1	33.1	! 0.35	9.4	6	12.3	45.4	! -0.39	13.3	4	15.4	60.8
! HABILLEM	10.00	10.00	! -0.19	2.3	7	3.5	3.5	! 0.58	25.3	1	33.2	36.8	! 0.11	1.0	8	1.2	37.9
! LOG_ENER	10.00	10.00	! 0.58	21.4	2	33.4	33.4	! -0.49	18.3	2	24.0	57.4	! 0.47	19.2	3	22.2	79.6
! EQUI_LOG	10.00	10.00	! -0.03	0.0	9	0.1	0.1	! 0.05	0.2	8	0.3	0.4	! 0.16	2.2	7	2.6	2.9
! SANTE	10.00	10.00	! 0.09	0.5	8	0.8	0.8	! -0.05	0.2	9	0.2	1.0	! -0.60	31.0	1	35.9	36.9
! TRANSPOR	10.00	10.00	! -0.65	27.4	1	42.8	42.8	! -0.43	14.2	5	18.7	61.5	! -0.33	9.3	5	10.8	72.2
! CULT_LOI	10.00	10.00	! -0.30	5.7	6	9.0	9.0	! 0.45	15.2	4	20.0	28.9	! 0.19	3.0	6	3.5	32.4
! IMP_EPAR	10.00	10.00	! -0.35	8.1	5	12.6	12.6	! -0.01	0.0	10	0.0	12.6	! 0.48	19.5	2	22.6	35.2
! HYGIENE	10.00	10.00	! 0.00	0.0	10	0.0	0.0	! 0.46	16.5	3	21.6	21.6	! 0.09	0.6	10	0.7	22.4
! TELECOM	10.00	10.00	! 0.46	13.5	4	21.1	21.1	! 0.09	0.6	7	0.8	21.9	! 0.09	0.7	9	0.8	22.7

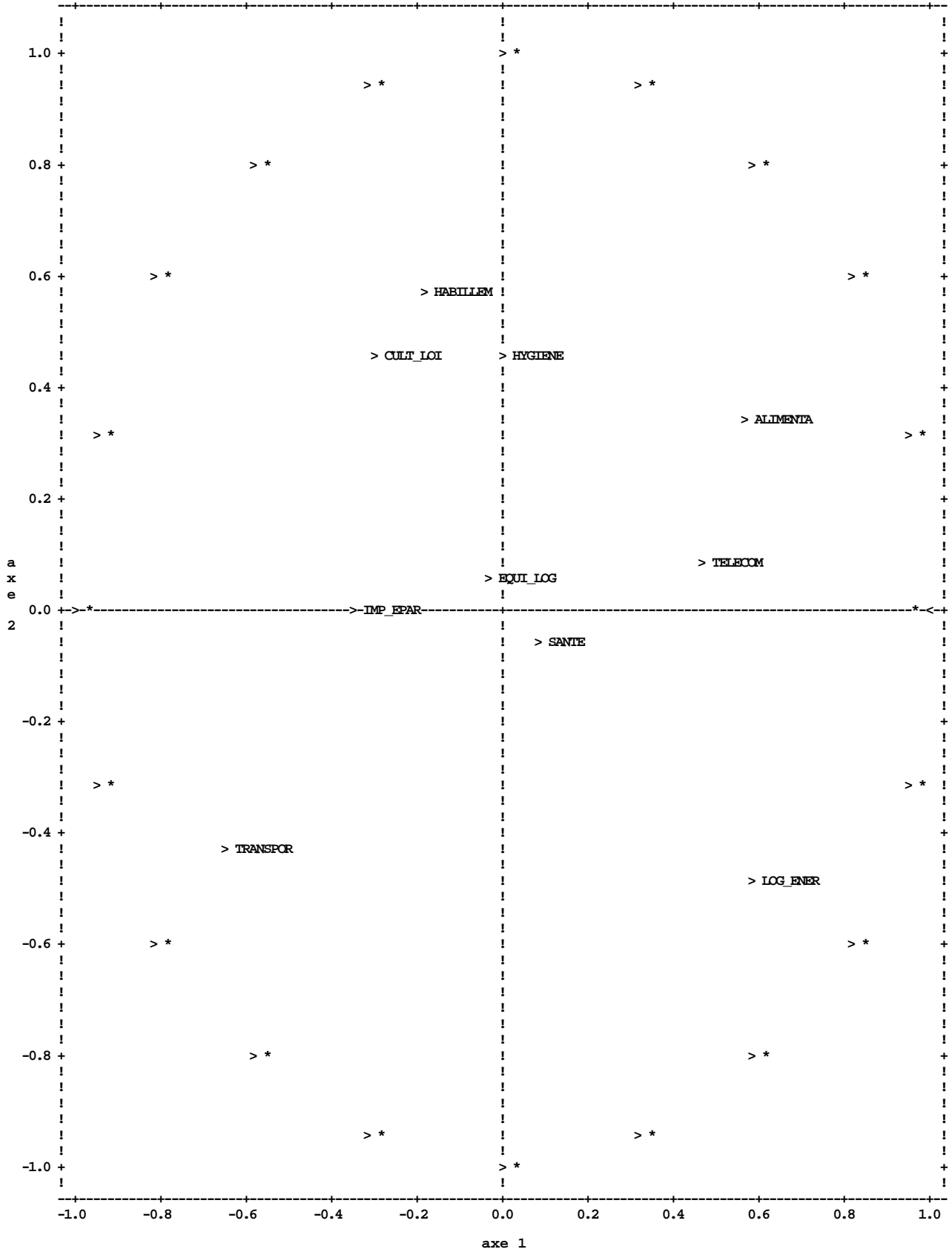
Aides à l'interprétation pour les variables actives

! Variables actives ! _____ AXE4 _____ !							
! Ident.	CONTR	POIDS	! COORD	CTR	RCTR	CO2	QLT
! ALIMENTA	10.00	10.00	! -0.33	9.7	3	10.8	71.6
! HABILLEM	10.00	10.00	! 0.14	1.8	8	1.9	39.9
! LOG_ENER	10.00	10.00	! 0.09	0.7	10	0.8	80.4
! EQUI_LOG	10.00	10.00	! 0.70	44.3	1	49.3	52.2
! SANTE	10.00	10.00	! 0.41	15.0	2	16.7	53.6
! TRANSPOR	10.00	10.00	! -0.22	4.4	7	4.9	77.1
! CULT_LOI	10.00	10.00	! -0.30	7.9	5	8.8	41.3
! IMP_EPAR	10.00	10.00	! -0.14	1.7	9	1.9	37.2
! HYGIENE	10.00	10.00	! 0.27	6.3	6	7.0	29.4
! TELECOM	10.00	10.00	! -0.30	8.1	4	9.1	31.8

Analyse en composantes principales sur les coefficients budgétaires des ménages

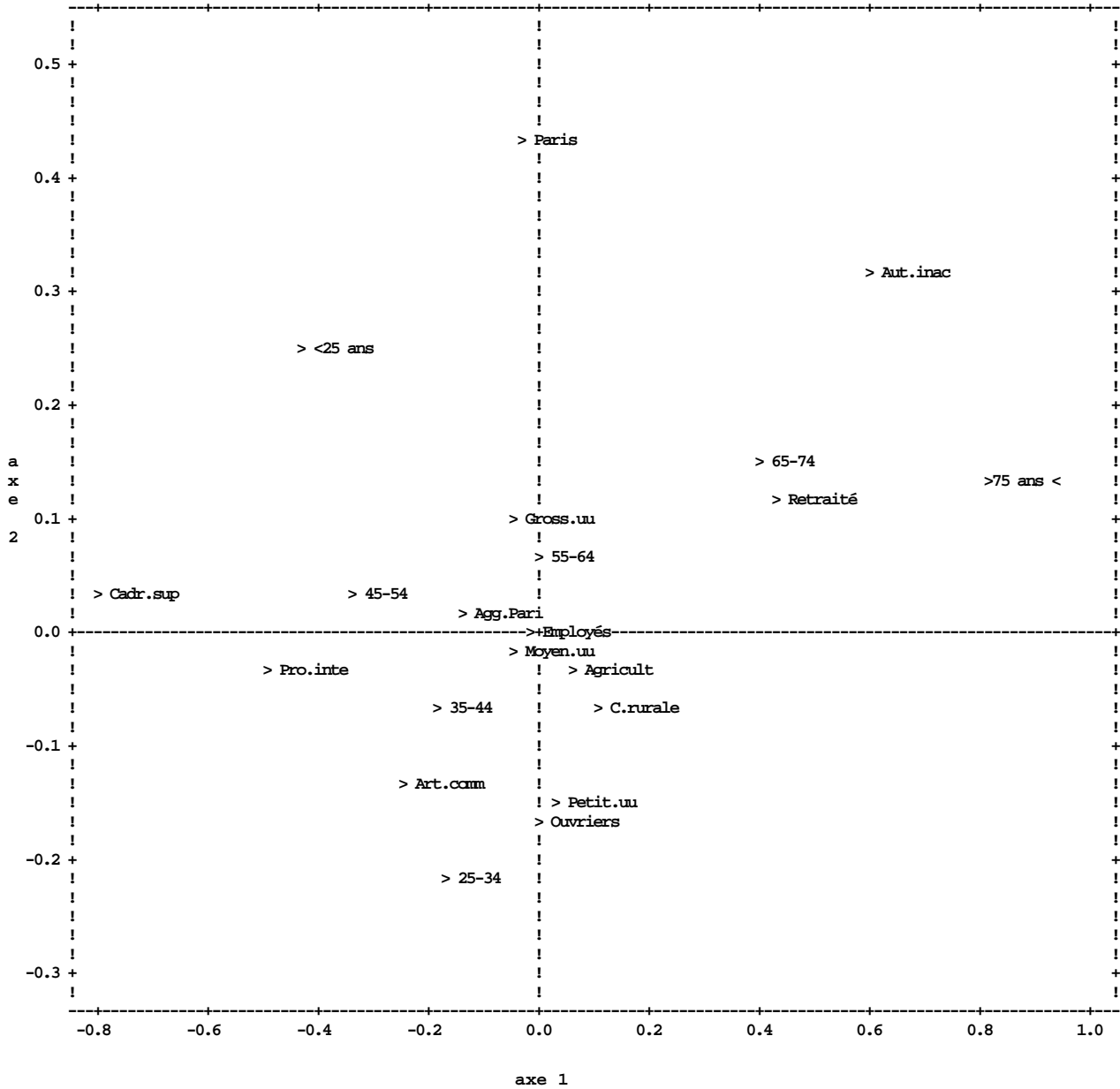
Les variables dans le plan 1-2

Plot of COORD2*COORD1\$ LABEL_. Symbol points to label.



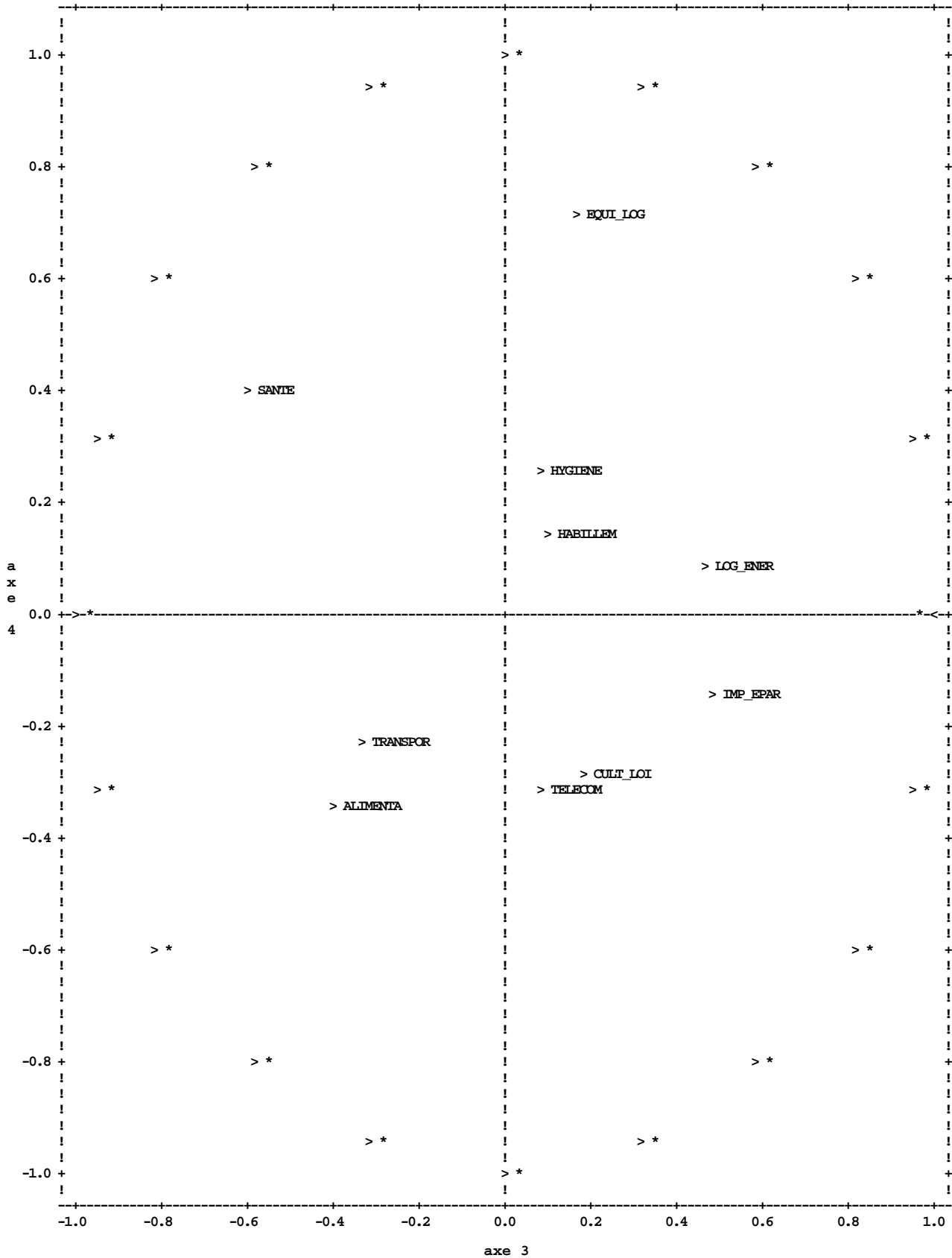
Analyse en composantes principales sur les coefficients budgétaires des ménages
 Les barycentres d'individus dans le plan 1-2

Plot of COORD2*COORD1\$_LABEL_. Symbol points to label.



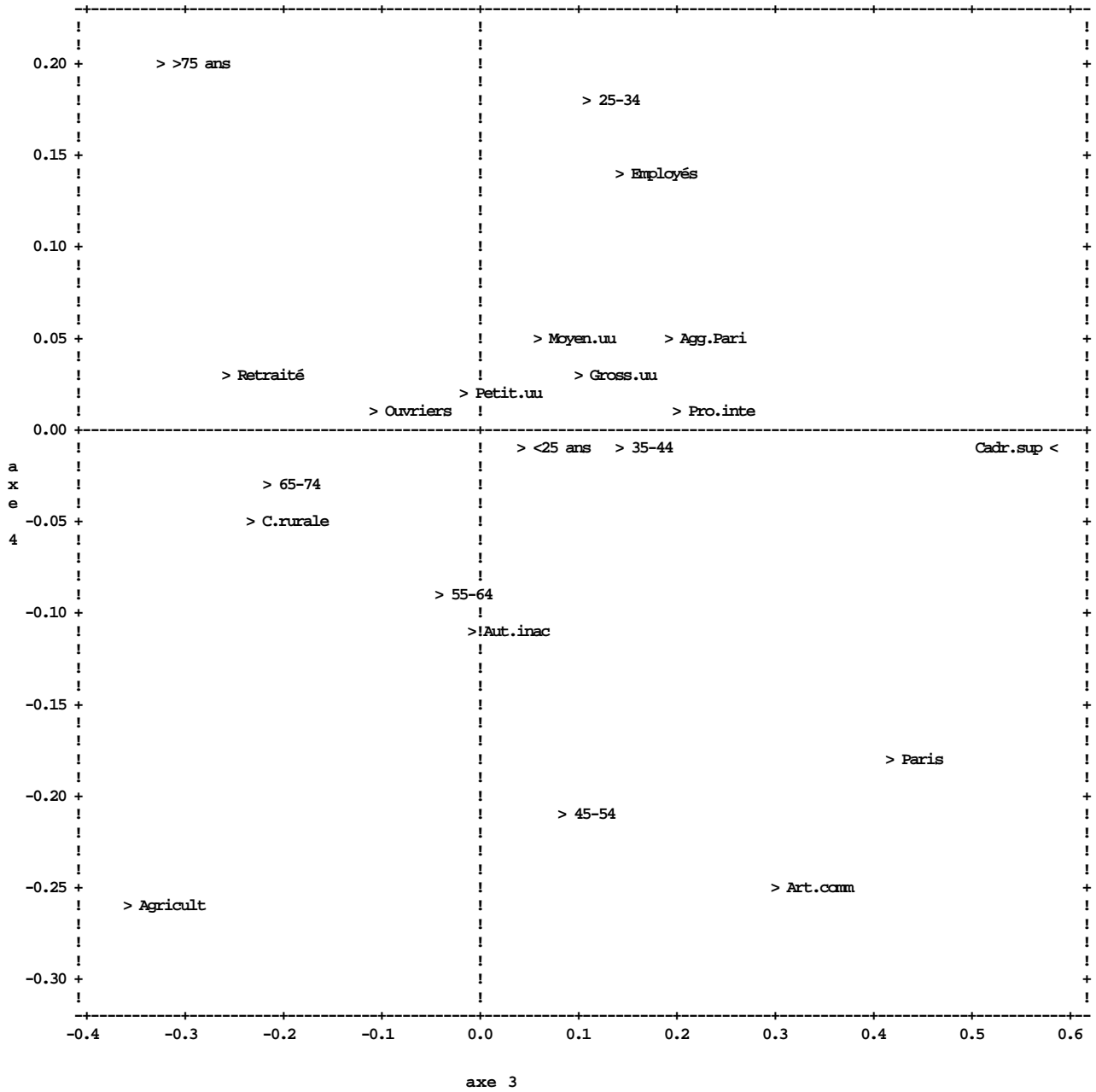
Analyse en composantes principales sur les coefficients budgétaires des ménages
 Les variables dans le plan 3-4

Plot of COORD4*COORD3\$_LABEL_. Symbol points to label.



Analyse en composantes principales sur les coefficients budgétaires des ménages
 Les barycentres d'individus dans le plan 3-4

Plot of COORD4*COORD3\$ LABEL_. Symbol points to label.



2. Analyse factorielle des correspondances

Exemple 1 : AFC sur un tableau de contingence (instruction VAR)

```
LIBNAME COMPIL ...;
OPTIONS SASMSTORE=COMPIL MSTORED NODATE PAGESIZE=66;
FILENAME ENT ...;

DATA A; INFILE ENT(ETRANGER);
INPUT AGRI ARTC CSUP PINT EMPL SERV OQ ONQ OAGR RETR / INAC NAT $13.;
IF NAT="Français" THEN SUPPLE=-1; ELSE SUPPLE=1;
LABEL AGRI="Agricult"
      ARTC="Art,com"
      CSUP="Cad.sup"
      PINT="Pr.inter"
      EMPL="Employés"
      SERV="Per.serv"
      OQ  ="Ouv.qual"
      ONQ ="Ouv.nqua"
      OAGR="Ouv.agri"
      RETR="Retraité"
      INAC="Aut.inac";

PROC PRINT LABEL DATA=A;
ID NAT;
TITLE "Répartition des étrangers par catégorie socioprofessionnelle en 1982";
TITLE2 "Tableau de données";

PROC CORRESP DATA=A OUTC=SOR SHORT RP CP DIMENS=6;
VAR AGRI ARTC CSUP PINT EMPL SERV OQ ONQ OAGR RETR INAC;
SUPPLEMENTARY RETR INAC;
ID NAT;
WEIGHT SUPPLE;
TITLE "AFC sur le tableau nationalités x catégories socioprofessionnelles";
TITLE2 "Sorties de la procédure CORRESP";
RUN;

TITLE2 "Sorties de la macro AIDEAFC";
%AIDEAFC(IOA=3,IOS=3,IVA=3,IVS=3,DATA=SOR,ID=NAT);

TITLE2 "Les points actifs et supplémentaires dans le plan 1-2";
%PLOTGOR(AXEH=1,AXEV=2,POINTS=OBSACT OBSSUP VARACT VARSUP,DATA=SOR,ID=NAT);

TITLE2 "Les points actifs et supplémentaires dans le plan 3-4";
%PLOTGOR(AXEH=3,AXEV=4,POINTS=OBSACT OBSSUP VARACT VARSUP,DATA=SOR,ID=NAT);
```

Répartition des étrangers par catégorie socioprofessionnelle en 1982
Tableau de données

NAT	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri	Retraité	Aut.inac	SUPPLE
Italiens	3176	15268	4228	13836	12952	6892	45960	39660	2668	68096	127572	1
Espagnols	1356	6608	2284	8464	13512	18112	39160	38736	9376	53936	135612	1
Portugais	420	5456	948	9128	30456	44972	111200	171680	11188	8080	373776	1
Aut.CEE	3592	6256	15872	14116	10108	2700	7120	4880	1052	30636	63676	1
Algériens	104	11708	2236	8560	33328	11220	84516	153336	1904	16224	481980	1
Marocains	212	3544	1400	2780	11524	6256	37140	83772	19212	2636	272832	1
Tunisiens	76	3044	1460	3112	7836	3712	20148	33092	2856	2072	113392	1
Aut.étrangers	1508	9380	22316	20540	34972	15276	46472	84612	3636	60884	382604	1
Français	1460460	1768524	1847128	3858160	5212188	776032	3647048	2855108	242728	7159616	21754420	-1

AFC sur le tableau nationalités x catégories socioprofessionnelles
Sorties de la procédure CORRESP

The Correspondence Analysis Procedure

Row Profiles

	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri
Italiens	0.021958	0.105559	0.029231	0.095658	0.089546	0.047649	0.317754	0.274198	0.018446
Espagnols	0.009854	0.048020	0.016598	0.061508	0.098192	0.131620	0.284576	0.281495	0.068136
Portugais	0.001090	0.014155	0.002459	0.023682	0.079015	0.116675	0.288495	0.445404	0.029026
Aut.CEE	0.054676	0.095226	0.241598	0.214868	0.153860	0.041098	0.108378	0.074282	0.016013
Algériens	0.000339	0.038148	0.007285	0.027891	0.108591	0.036558	0.275375	0.499609	0.006204
Marocains	0.001278	0.021370	0.008442	0.016763	0.069489	0.037723	0.223951	0.505137	0.115847
Tunisiens	0.001009	0.040406	0.019380	0.041308	0.104014	0.049273	0.267442	0.439259	0.037910
Aut.étrangers	0.006317	0.039294	0.093485	0.086045	0.146503	0.063993	0.194678	0.354452	0.015232

Supplementary Row Profiles

	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri
Français	0.067404	0.081622	0.085249	0.178063	0.240555	0.035816	0.168320	0.131770	0.011202

Column Profiles

	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri
Italiens	0.304098	0.249217	0.083320	0.171799	0.083730	0.063148	0.117330	0.065041	0.051414
Espagnols	0.129835	0.107861	0.045010	0.105096	0.087350	0.165952	0.099970	0.063526	0.180683
Portugais	0.040214	0.089057	0.018682	0.113341	0.196887	0.412058	0.283879	0.281550	0.215602
Aut.CEE	0.343930	0.102115	0.312786	0.175276	0.065344	0.024739	0.018176	0.008003	0.020273
Algériens	0.009958	0.191107	0.044064	0.106288	0.215453	0.102804	0.215758	0.251466	0.036692
Marocains	0.020299	0.057848	0.027589	0.034519	0.074498	0.057321	0.094814	0.137383	0.370230
Tunisiens	0.007277	0.049687	0.028772	0.038641	0.050657	0.034011	0.051435	0.054270	0.055037
Aut.étrangers	0.144389	0.153108	0.439776	0.255041	0.226081	0.139967	0.118637	0.138761	0.070069

Supplementary Column Profiles

	Retraité	Aut.inac
Italiens	0.280734	0.065373
Espagnols	0.222358	0.069493
Portugais	0.033311	0.191538
Aut.CEE	0.126301	0.032630
Algériens	0.066885	0.246986
Marocains	0.010867	0.139810
Tunisiens	0.008542	0.058107
Aut.étrangers	0.251002	0.196062

Inertia and Chi-Square Decomposition

Singular Values	Principal Inertias	Chi-Squares	Percents	14	28	42	56	70
0.42111	0.17733	269579	68.14%	*****				
0.17913	0.03209	48779.5	12.33%	****				
0.16462	0.02710	41197.3	10.41%	****				
0.14382	0.02068	31441.9	7.95%	***				
0.05326	0.00284	4312.96	1.09%					
0.01406	0.00020	300.546	0.08%					
0.00401	0.00002	24.4776	0.01%					
	0.26025	395635	(Degrees of Freedom = 56)					

Row Coordinates

	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6
Italiens	0.26058	-0.10170	0.32142	0.25056	0.00173	-0.01807
Espagnols	-0.00987	0.19955	0.27403	-0.11214	0.08200	0.02455
Portugais	-0.27692	-0.02764	0.05622	-0.16120	-0.05010	-0.00672
Aut.CEE	1.60227	0.13243	-0.05462	-0.02593	-0.12049	0.01692
Algériens	-0.20073	-0.20536	-0.10239	0.10647	-0.00649	0.01678
Marocains	-0.31263	0.40108	-0.15746	0.16285	-0.01669	-0.00411
Tunisiens	-0.11579	-0.01602	-0.05241	0.08208	0.03762	-0.01134
Aut.étrangers	0.36579	-0.05478	-0.17088	-0.09568	0.07380	-0.01216

Supplementary Row Coordinates

	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6
Français	0.976643	-.037251	0.234801	0.161932	-.009712	0.540205

Column Coordinates

	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6
Agricult	1.57091	0.20928	0.53167	0.25994	-0.40842	0.10574
Art,com	0.46695	-0.10042	0.31372	0.36535	0.06522	-0.01285
Cad.sup	1.56095	0.10540	-0.37926	-0.17955	-0.03745	-0.01997
Pr.inter	0.83084	0.00562	0.11475	0.02888	0.02120	-0.01543
Employés	0.19337	-0.08617	-0.10164	-0.03233	0.10364	0.02855
Per.serv	-0.12098	0.06837	0.25716	-0.41841	0.00146	-0.00101
Ouv.qual	-0.13158	-0.06155	0.12199	0.02251	-0.00570	-0.00422
Ouv.nqua	-0.23219	-0.03166	-0.12291	0.02715	-0.02610	-0.00129
Ouv.agri	-0.28368	0.91436	0.00084	0.13456	0.04666	0.00137

Supplementary Column Coordinates

	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6
Retraité	0.802896	0.046681	0.572467	0.155331	0.376679	0.016296
Aut.inac	-0.030173	-.000398	-.211470	0.083027	0.093944	0.022649

AFC sur le tableau nationalités x catégories socioprofessionnelles
Sorties de la macro AIDEAFC

```
*****
*      Caractéristiques de l'analyse      *
*****
*      Nombre de modalités-colonnes actives = 9      *
*      Nombre de modalités-colonnes supp.   = 2      *
*      *                                         *
*      Nombre de modalités-lignes actives  = 8      *
*      Nombre de modalités-lignes supp.    = 1      *
*      *                                         *
*      Edition des aides à l'interprétation :      *
*      modalités-colonnes actives sur 3 axes      *
*      modalités-colonnes supp. sur 3 axes      *
*      modalités-lignes actives sur 3 axes      *
*      modalités-lignes supp. sur 3 axes      *
*****
```

Liste des modalités-lignes actives :

Italiens Espagnols Portugais Aut.CEE Algériens Marocains Tunisiens Aut.étrangers

Liste des modalités-lignes supplémentaires :

Français

Liste des modalités-colonnes actives :

Agricult Art,com Cad.sup Pr.inter Employés Per.serv Ouv.qual Ouv.nqua Ouv.agri

Liste des modalités-colonnes supplémentaires :

Retraité Aut.inac

Valeurs propres

Inertie totale : 0.26025

N° Val.Pr. Diff. Pct Cum !

N°	Val.Pr.	Diff.	Pct	Cum	!
1	0.1773	.	68.14	68.14	!*****
2	0.0321	0.1452	12.33	80.47	!*****
3	0.0271	0.0050	10.41	90.88	!*****
4	0.0207	0.0064	7.95	98.83	!*****
5	0.0028	0.0178	1.09	99.92	!*
6	0.0002	0.0026	0.08	99.99	!

Aides à l'interprétation pour les modalités-lignes actives

! Mod. lignes actives !	AXE1					AXE2					AXE3					!		
! Ident.	CONTR	POIDS	COORD	CTR	RCR	CO2	QLT	COORD	CTR	RCR	CO2	QLT	COORD	CTR	RCR	CO2	QLT	!
! Italiens	8.94	9.51	0.26	3.6	6	27.8	27.8	-0.10	3.1	4	4.2	32.0	0.32	36.3	1	42.2	74.2	!
! Espagnol	4.69	9.05	-0.01	0.0	8	0.1	0.1	0.20	11.2	3	29.5	29.6	0.27	25.1	2	55.7	85.2	!
! Portugai	10.63	25.36	-0.28	11.0	3	70.3	70.3	-0.03	0.6	7	0.7	71.0	0.06	3.0	6	2.9	73.9	!
! Aut.CEE	43.23	4.32	1.60	62.6	1	98.6	98.6	0.13	2.4	5	0.7	99.3	-0.05	0.5	8	0.1	99.4	!
! Algérien	8.12	20.19	-0.20	4.6	5	38.5	38.5	-0.21	26.5	2	40.3	78.8	-0.10	7.8	5	10.0	88.9	!
! Marocain	13.00	10.91	-0.31	6.0	4	31.5	31.5	0.40	54.7	1	51.9	83.4	-0.16	10.0	4	8.0	91.4	!
! Tunisien	0.48	4.96	-0.12	0.4	7	53.7	53.7	-0.02	0.0	8	1.0	54.7	-0.05	0.5	7	11.0	65.7	!
! Aut.étra	10.91	15.70	0.37	11.8	2	74.0	74.0	-0.05	1.5	6	1.7	75.7	-0.17	16.9	3	16.2	91.8	!

AFC sur le tableau nationalités x catégories socioprofessionnelles
Sorties de la macro AIDEAFC

Aides à l'interprétation pour les modalités-lignes supplémentaires

! Mod. lignes supplé.	AXE1					AXE2					AXE3					!		
! Ident.	CONTR	POIDS	COORD	CTR	RCO2	QLT	COORD	CTR	RCO2	QLT	COORD	CTR	RCO2	QLT	!			
! Français	.	.	0.98	.	68.8	1	68.8	-0.04	.	0.1	1	68.9	0.23	.	4.0	1	72.9	!

Aides à l'interprétation pour les modalités-colonnes actives

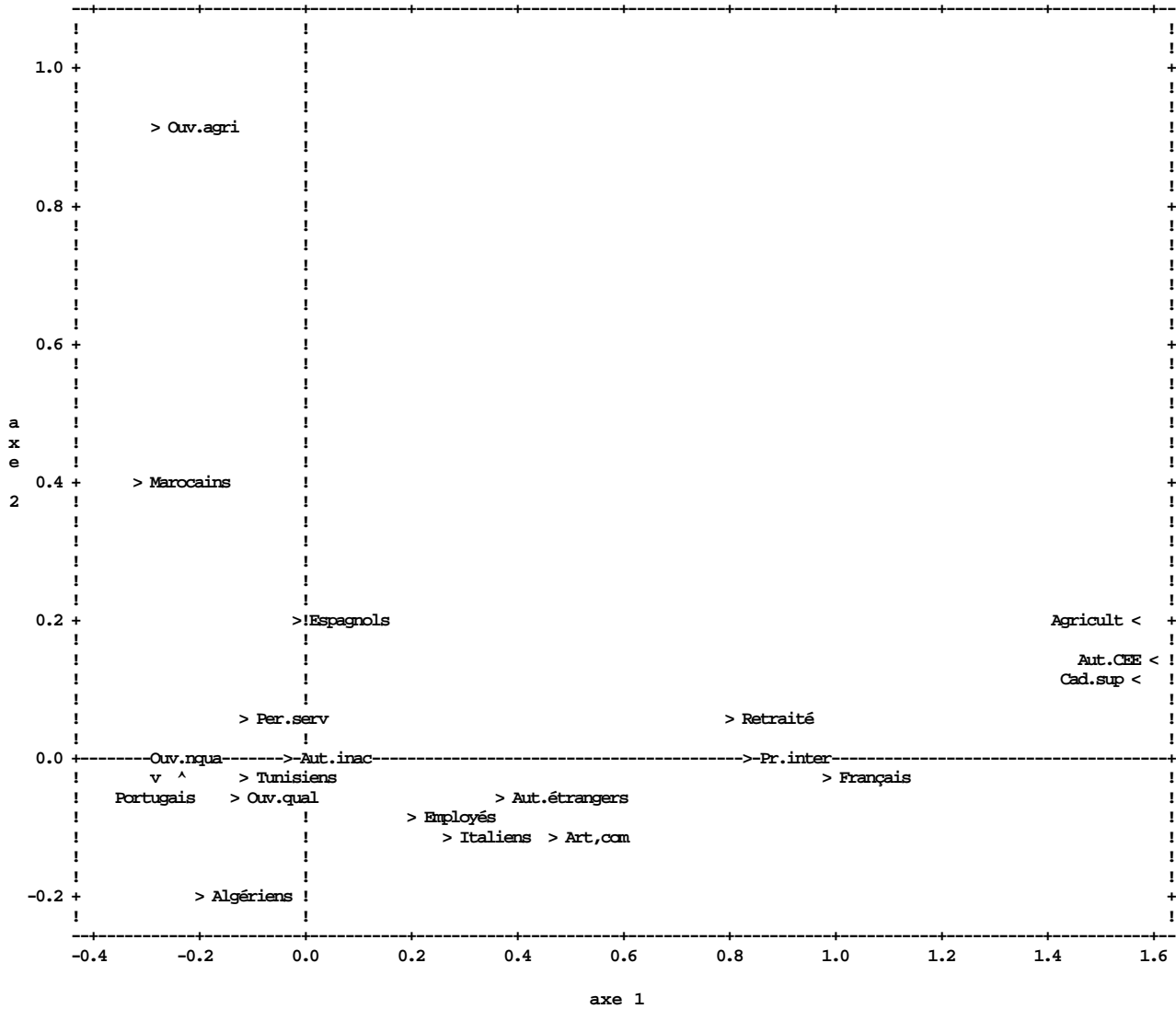
! Mod. col. actives !	AXE1					AXE2					AXE3					!		
! Ident.	CONTR	POIDS	COORD	CTR	RCR	CO2	QLT	COORD	CTR	RCR	CO2	QLT	COORD	CTR	RCR	CO2	QLT	!
! Agricult	8.02	0.69	1.57	9.6	4	81.2	81.2	0.21	0.9	8	1.4	82.6	0.53	7.2	6	9.3	91.9	!
! Art,com	7.19	4.03	0.47	5.0	5	46.9	46.9	-0.10	1.3	4	2.2	49.1	0.31	14.6	4	21.2	70.3	!
! Cad.sup	33.68	3.34	1.56	45.9	1	92.8	92.8	0.11	1.2	6	0.4	93.2	-0.38	17.7	2	5.5	98.7	!
! Pr.inter	14.35	5.30	0.83	20.6	2	97.9	97.9	0.01	0.0	9	0.0	97.9	0.11	2.6	8	1.9	99.8	!
! Employés	2.65	10.18	0.19	2.1	7	55.2	55.2	-0.09	2.4	3	11.0	66.2	-0.10	3.9	7	15.2	81.4	!
! Per.serv	7.19	7.18	-0.12	0.6	9	5.6	5.6	0.07	1.0	7	1.8	7.4	0.26	17.5	3	25.4	32.8	!
! Ouv.qual	3.62	25.77	-0.13	2.5	6	47.3	47.3	-0.06	3.0	2	10.4	57.7	0.12	14.2	5	40.7	98.4	!
! Ouv.nqua	11.01	40.11	-0.23	12.2	3	75.5	75.5	-0.03	1.3	5	1.4	76.9	-0.12	22.4	1	21.1	98.0	!
! Ouv.agri	12.29	3.41	-0.28	1.5	8	8.6	8.6	0.91	88.9	1	89.2	97.8	0.00	0.0	9	0.0	97.8	!

Aides à l'interprétation pour les modalités-colonnes supplémentaires

! Mod. col. supplé.	AXE1					AXE2					AXE3					!		
! Ident.	CONTR	POIDS	COORD	CTR	RCO2	QLT	COORD	CTR	RCO2	QLT	COORD	CTR	RCO2	QLT	!			
! Retraité	.	.	0.80	.	54.9	1	54.9	0.05	.	0.2	1	55.1	0.57	.	27.9	2	83.0	!
! Aut.inac	.	.	-0.03	.	1.5	2	1.5	-0.00	.	0.0	2	1.5	-0.21	.	72.3	1	73.8	!

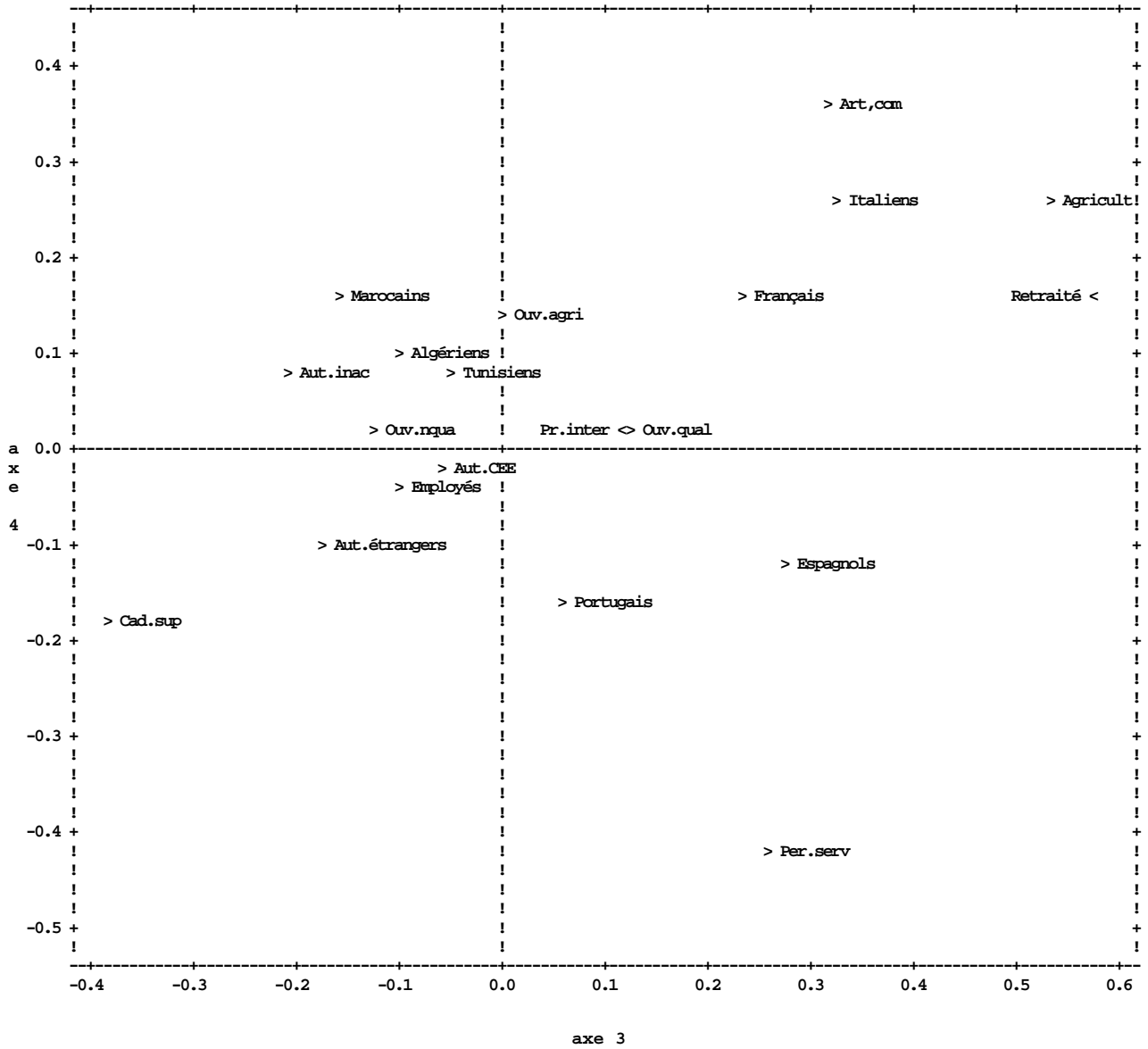
AFC sur le tableau nationalités x catégories socioprofessionnelles
 Les points actifs et supplémentaires dans le plan 1-2

Plot of DIM2*DIM1\$NAT. Symbol points to label.



AFC sur le tableau nationalités x catégories socioprofessionnelles
 Les points actifs et supplémentaires dans le plan 3-4

Plot of DIM4*DIM3\$NAT. Symbol points to label.



Exemple 2 : AFC sur un tableau individus-variables (instruction TABLES)

```
LIBNAME COMPIL ...;
OPTIONS SASMSTORE=COMPIL MSTORED NODATE PAGESIZE=66;
FILENAME ENT ...;

PROC FORMAT;
VALUE TAILFOR 1="Petit"
              2="T.moyen"
              3="Grand";
VALUE VELOFOR 1="Lent"
              2="V.moyen"
              3="Rapide";
VALUE FONCFOR 1="Compagnie"
              2="Chasse"
              3="Utilité";

DATA A; INFILE ENT(CHIENS);
INPUT CHIEN $CHAR15. +1 NOM $4. +1 TAILLE +2 VELOCITE à34 FONCTION;
IF NOM="SUPP" THEN SUPPLE=-1;ELSE SUPPLE=1;

PROC PRINT DATA=A;
TITLE "Caractéristiques de races de chiens";
TITLE2 "Tableau de données";

PROC CORRESP DATA=A DIMENS=2 OUTC=SOR SHORT RP CP;
TABLES TAILLE,VELOCITE FONCTION;
SUPPLEMENTARY FONCTION;
WEIGHT SUPPLE;
FORMAT TAILLE TAILFOR. VELOCITE VELOFOR. FONCTION FONCFOR.;
TITLE "AFC sur le tableau TAILLE x VELOCITE construit par la procédure CORRESP";
TITLE2 "Sorties de la procédure CORRESP";
RUN;

TITLE2 "Sorties de la macro AIDEAFC";
%AIDEAFC(IOA=2,IOS=2,IVA=2,IVS=2,DATA=SOR);

TITLE2 "Les points actifs et supplémentaires dans le plan 1-2";
%PLOTGOR(AXEH=1,AXEV=2,POINTS=OBSACT OBSSUP VARACT VARSUP,DATA=SOR);
```

Caractéristiques de races de chiens
Tableau de données

OBS	CHIEN	NOM	TAILLE	VELOCITE	FONCTION	SUPPLE
1	BEAUCERON	BEAU	3	3	3	1
2	BASSET	BASS	1	1	2	1
3	BERGER ALLEMAND	BERG	3	3	3	1
4	BOXER	BOXE	2	2	1	1
5	BULL-DOG	BULD	1	1	1	1
6	BULL-MASTIFF	BULM	3	1	3	1
7	CANICHE	CANI	1	2	1	1
8	CHIHUAHUA	CHIU	1	1	1	1
9	COCKER	COCK	2	1	1	1
10	COLLEY	COLL	3	3	1	1
11	DALMATIEN	DALM	2	2	1	1
12	DOBERMANN	DOBE	3	3	3	1
13	DOGUE ALLEMAND	DOGU	3	3	3	1
14	EPAGNE BRETON	EPAB	2	2	2	1
15	EPAGNE FRANCAIS	EPAF	3	2	2	1
16	FOX-HOUND	FOXH	3	3	2	1
17	FOX TERRIER	FOXT	1	2	1	1
18	GRAND BLEU	GRBL	3	2	2	1
19	LABRADOR	LABR	2	2	2	1
20	LEVRIER	LEVR	3	3	2	1
21	MASTIFF	MAST	3	1	3	1
22	PEKINOIS	PEKI	1	1	1	1
23	POINTER	POIN	3	3	2	1
24	SAINT-BERNARD	STBE	3	1	3	1
25	SETTER	SETT	3	3	2	1
26	TECKEL	TECK	1	1	1	1
27	TERRE-NEUVE	TERN	3	1	3	1
28	SUPPLEME	SUPP	1	3	2	-1

AFC sur le tableau TAILLE x VELOCITE construit par la procédure CORRESP
Sorties de la procédure CORRESP

The Correspondence Analysis Procedure

Row Profiles

	Lent	Rapide	V.moyen
Grand	0.266667	0.600000	0.133333
Petit	0.714286	0.000000	0.285714
T.moyen	0.200000	0.000000	0.800000

Column Profiles

	Lent	Rapide	V.moyen
Grand	0.40000	1.00000	0.25000
Petit	0.50000	0.00000	0.25000
T.moyen	0.10000	0.00000	0.50000

Supplementary Column Profiles

	Chasse	Compagni	Utilité
Grand	0.66667	0.10000	1.00000
Petit	0.11111	0.60000	0.00000
T.moyen	0.22222	0.30000	0.00000

Inertia and Chi-Square Decomposition

Singular Values	Principal Inertias	Chi-Squares	Percents	14	28	42	56	70
0.65218	0.42534	11.4842	72.27%	*****				
0.40402	0.16323	4.4072	27.73%	*****				
	0.58857	15.8914 (Degrees of Freedom = 4)						

Row Coordinates

	Dim1	Dim2
Grand	-.571706	0.071777
Petit	0.550409	-.591702
T.moyen	0.944546	0.613051

Column Coordinates

	Dim1	Dim2
Lent	0.216161	-.509470
Rapide	-.876606	0.177658
V.moyen	0.715980	0.436972

Supplementary Column Coordinates

	Dim1	Dim2
Chasse	-.168790	0.292909
Compagni	0.853196	-.405745
Utilité	-.876606	0.177658

AFC sur le tableau TAILLE x VELOCITE construit par la procédure CORRESP
Sorties de la macro AIDEAFC

```
*****
*      Caractéristiques de l'analyse      *
*****
*      Nombre de modalités-colonnes actives = 3 *
*      Nombre de modalités-colonnes supp.   = 3 *
*                                          *
*      Nombre de modalités-lignes actives  = 3 *
*      Nombre de modalités-lignes supp.    = 0 *
*                                          *
*      Edition des aides à l'interprétation : *
*      modalités-colonnes actives sur 2 axes *
*      modalités-colonnes supp. sur 2 axes  *
*      modalités-lignes actives sur 2 axes  *
*      modalités-lignes supp. sur 2 axes    *
*****
```

Liste des modalités-lignes actives :
Grand Petit T.moyen
Liste des modalités-colonnes actives :
Lent Rapide V.moyen
Liste des modalités-colonnes supplémentaires :
Chasse Compagni Utilité

Valeurs propres

Inertie totale : 0.58857

N°	Val.Pr.	Diff.	Pct	Cum	!
1	0.4253	.	72.27	72.27	!*****
2	0.1632	0.2621	27.73	100.0	!*****

Aides à l'interprétation pour les modalités-lignes actives

! Mod. lignes actives !	! _____ AXE1 _____ !					! _____ AXE2 _____ !				
! Ident.	COORD	POIDS	COORD	CIR	RCIR	COORD	CIR	RCIR	COORD	QLT
! Grand	31.34	55.56	-0.57	42.7	1	98.4	98.4	0.07	1.8	3
! Petit	28.77	25.93	0.55	18.5	3	46.4	46.4	-0.59	55.6	1
! T.moyen	39.90	18.52	0.94	38.8	2	70.4	70.4	0.61	42.6	2

Aides à l'interprétation pour les modalités-colonnes actives

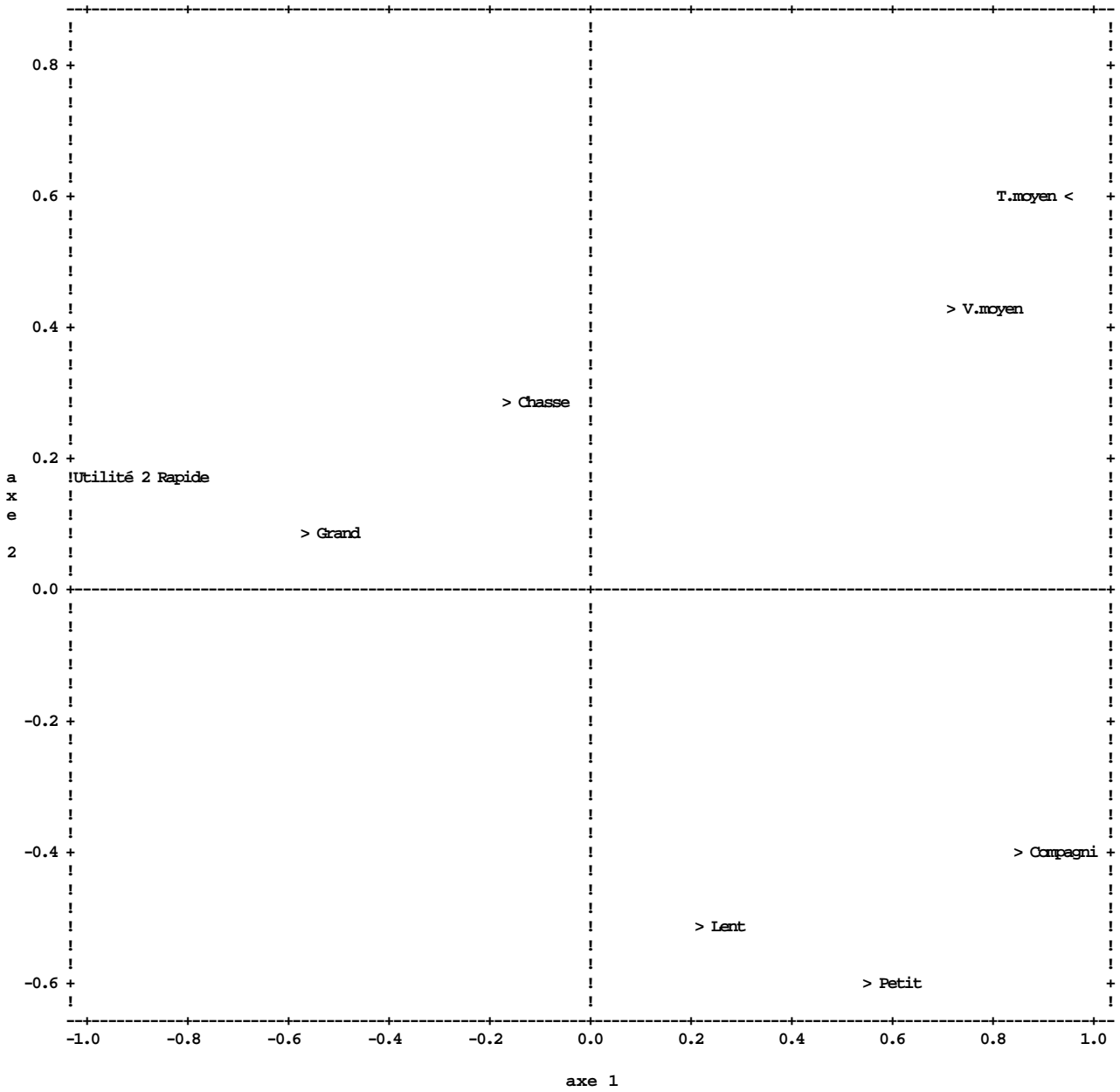
! Mod. col. actives !	! _____ AXE1 _____ !					! _____ AXE2 _____ !				
! Ident.	COORD	CIR	RCIR	COORD	QLT	COORD	CIR	RCIR	COORD	QLT
! Lent	19.27	37.04	0.22	4.1	3	15.3	15.3	-0.51	58.9	1
! Rapide	45.31	33.33	-0.88	60.2	1	96.1	96.1	0.18	6.4	3
! V.moyen	35.42	29.63	0.72	35.7	2	72.9	72.9	0.44	34.7	2

Aides à l'interprétation pour les modalités-colonnes supplémentaires

! Mod. col. supplé. !	! _____ AXE1 _____ !					! _____ AXE2 _____ !				
! Ident.	COORD	CIR	RCO2	COORD	QLT	COORD	CIR	RCO2	COORD	QLT
! Chasse	.	.	-0.17	24.9	3	24.9	0.29	75.1	1	100.0
! Compagni	.	.	0.85	81.6	2	81.6	-0.41	18.4	2	100.0
! Utilité	.	.	-0.88	96.1	1	96.1	0.18	3.9	3	100.0

AFC sur le tableau TAILLE x VELOCITE construit par la procédure CORRESP
Les points actifs et supplémentaires dans le plan 1-2

Plot of DIM2*DIM1\$_NAME_. Symbol points to label.



3. Analyse des correspondances multiples

```

PROC FORMAT;
VALUE FONCFOR 1="Compagnie"
                2="Chasse"
                3="Utilité";
VALUE TAILFOR 1="Petit"
                2="Tail.moy"
                3="Grand";
VALUE POIDFOR 1="P.L,ger"
                2="Poid.moy"
                3="P.lourd";
VALUE VELOFOR 1="Lent"
                2="Vite.moy"
                3="Rapide";
VALUE INTEFOR 1="Peu inte"
                2="Inte.moy"
                3="Très int";
VALUE AGREFOR 1="Peu agre"
                2="Agressif";
VALUE AFFEFOR 1="Peu affe"
                2="Affectueux";

OPTIONS NODATE PAGESIZE=66;

FILENAME ENT ...;

DATA A; INFILE ENT(CHIENS);
INPUT CHIEN $CHAR15. +1 NOM $4. +1 TAILLE POIDS VELOCITE INTELLIG AFFECTIO
        AGRESSIV FONCTION;
IF NOM="SUPP" THEN SUPPLE=-1;ELSE SUPPLE=1;

PROC PRINT DATA=A;
TITLE "Caractéristiques de races de chiens";
TITLE2 "Tableau de données";

```

Caractéristiques de races de chiens
Tableau de données

OBS	CHIEN	NOM	TAILLE	POIDS	VELOCITE	INTELLIG	AFFECTIO	AGRESSIV	FONCTION	SUPPLE
1	BEAUCERON	BEAU	3	2	3	2	2	2	3	1
2	BASSET	BASS	1	1	1	1	1	2	2	1
3	BERGER ALLEMAND	BERG	3	2	3	3	2	2	3	1
4	BOXER	BOXE	2	2	2	2	2	2	1	1
5	BULL-DOG	BULD	1	1	1	2	2	1	1	1
6	BULL-MASTIFF	BULM	3	3	1	3	1	2	3	1
7	CANICHE	CANI	1	1	2	3	2	1	1	1
8	CHIHUAHUA	CHIU	1	1	1	1	2	1	1	1
9	COCKER	COCK	2	1	1	2	2	2	1	1
10	COLLEY	COLL	3	2	3	2	2	1	1	1
11	DALMATIEN	DALM	2	2	2	2	2	1	1	1
12	DOBERMANN	DOBE	3	2	3	3	1	2	3	1
13	DOGUE ALLEMAND	DOGU	3	3	3	1	1	2	3	1
14	EPAGNE BRETON	EPAB	2	2	2	3	2	1	2	1
15	EPAGNE FRANCAIS	EPAF	3	2	2	2	1	1	2	1
16	FOX-HOUND	FOXH	3	2	3	1	1	2	2	1
17	FOX TERRIER	FOXT	1	1	2	2	2	2	1	1
18	GRAND ELEU	GREL	3	2	2	1	1	2	2	1
19	LABRADOR	LABR	2	2	2	2	2	1	2	1
20	LEVRIER	LEVR	3	2	3	1	1	1	2	1
21	MASTIFF	MAST	3	3	1	1	1	2	3	1
22	PEKINOIS	PEKI	1	1	1	1	2	1	1	1
23	POINIER	FOIN	3	2	3	3	1	1	2	1
24	SAINT-BERNARD	STBE	3	3	1	2	1	2	3	1
25	SETTER	SEIT	3	2	3	2	1	1	2	1
26	TECKEL	TECK	1	1	1	2	2	1	1	1
27	TERRE-NEUVE	TERN	3	3	1	2	1	1	3	1
28	SUPPLEME	SUPP	1	2	3	1	2	1	2	-1

Exemple 1 : ACM sur un tableau individus-variables (instruction TABLES)

```
LIBNAME COMPIL ...;
OPTIONS SASSTORE=COMPIL MSTORED;

PROC CORRESP NOPRINT DIMENS=6 DATA=A MCA OUTC=SOR;
TABLES TAILLE POIDS VELOCITE INTELLIG AFFECTIO AGRESSIV FONCTION;
SUPPLEMENTARY FONCTION;
WEIGHT SUPPLE;
FORMAT FONCTION FONCFOR. TAILLE TAILFOR. POIDS POIDFOR. VELOCITE VELOFOR.
        INTELLIG INTEFOR. AFFECTIO AFFEFOR. AGRESSIV AGREFOR.;

TITLE "Analyse des correspondances multiples sur le tableau des caractéristiques
des races de chiens";
TITLE2
"Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES";
%AIDEACM(DATA = SOR, DATAINIT = A,
        ANALYSE = TABLES,
        VARACT = TAILLE POIDS VELOCITE INTELLIG AFFECTIO AGRESSIV,
        VARSUP = FONCTION,
        IVA = 4, IVS = 4,
        FORMAT = FONCTION FONCFOR. TAILLE TAILFOR. POIDS POIDFOR.
                VELOCITE VELOFOR. INTELLIG INTEFOR. AFFECTIO AFFEFOR.
                AGRESSIV AGREFOR.,
        WEIGHT = SUPPLE);

TITLE2 "Les modalités actives et supplémentaires dans le plan 1-2";
%PLOTGOR(AXEH=1, AXEV=2, POINTS=VARACT VARSUP, DATA=SOR);
```

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES

```
*****
*          Caractéristiques de l'analyse          *
*****
*   Type d'analyse           =  TABLES   *
*
*   Nombre de variables actives       =    6   *
*   Nombre de modalités actives      =   16   *
*   Nombre de variables supplémentaires =    1   *
*   Nombre de modalités supplémentaires =    3   *
*
*   Somme des poids des individus actifs =   27   *
*
*   Variable de pondération           =  SUPPLE *
*
*   Edition des aides à l'interprétation :      *
*   variables actives                 sur 4 axes *
*   variables supplémentaires sur 4 axes      *
*****
```

Liste des modalités actives :

```
pour la variable TAILLE      : Grand Petit Tail.moy
pour la variable POIDS       : P.lourd P.Léger Poid.moy
pour la variable VELOCITE    : Lent Rapide Vite.moy
pour la variable INTELLIG    : Inte.moy Peu inte Très int
pour la variable AFFECTIO    : Affectue Peu affe
pour la variable AGRESSIV    : Agressif Peu agre
```

Liste des modalités supplémentaires :

```
pour la variable FONCTION    : Chasse Compagni Utilité
```

Valeurs propres

Inertie totale : 1.67

N°	Val.Pr.	Diff.	Pct	Cum	!
1	0.4816	.	28.90	28.90	!*****
2	0.3847	0.0969	23.08	51.98	!*****
3	0.2110	0.1738	12.66	64.64	!*****
4	0.1576	0.0534	9.45	74.09	!*****
5	0.1501	0.0074	9.01	83.10	!*****
6	0.1233	0.0268	7.40	90.50	!*****

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES

Aides à l'interprétation pour les modalités actives

Modalités actives		AXE1								AXE2					
Variable	Modalité	COVIR	POIDS	COORD	CTR	RCIR	CO2	QLT	V.TEST	COORD	CTR	RCIR	CO2	QLT	V.TEST
AFFECTIO	Affectue	4.81	8.64	0.78	10.8	5	64.8	64.8	4.10	-0.27	1.6	13	7.7	72.4	-1.41
	Peu affe	5.19	8.02	-0.84	11.6	4	64.8	64.8	-4.10	0.29	1.7	11	7.7	72.4	1.41
AFFECTIO		10.00	16.67		22.4						3.3				
AGRESSIV	Agressif	5.19	8.02	-0.43	3.1	10	17.3	17.3	-2.12	0.21	0.9	14	4.1	21.4	1.03
	Peu agre	4.81	8.64	0.40	2.9	11	17.3	17.3	2.12	-0.19	0.8	15	4.1	21.4	-1.03
AGRESSIV		10.00	16.67		6.0						1.8				
INTELLIG	Inte.moy	5.19	8.02	0.37	2.3	12	12.7	12.7	1.82	-0.29	1.7	12	7.6	20.2	-1.40
	Peu inte	7.04	4.94	-0.35	1.2	15	5.1	5.1	-1.15	0.81	8.4	7	27.5	32.7	2.68
	Très int	7.78	3.70	-0.34	0.9	16	3.2	3.2	-0.91	-0.46	2.0	9	6.0	9.2	-1.25
INTELLIG		20.00	16.67		4.4						12.1				
POIDS	P.lourd	8.15	3.09	-1.02	6.6	7	23.4	23.4	-2.47	0.97	7.6	8	21.6	45.0	2.37
	P.Léger	7.04	4.94	1.17	14.0	1	57.5	57.5	3.87	0.82	8.7	6	28.6	86.1	2.73
	Poid.moy	4.81	8.64	-0.31	1.7	13	10.0	10.0	-1.62	-0.82	15.1	2	72.2	82.3	-4.33
POIDS		20.00	16.67		22.3						31.4				
TAILLE	Grand	4.44	9.26	-0.84	13.5	2	87.5	87.5	-4.77	-0.02	0.0	16	0.1	87.6	-0.12
	Petit	7.41	4.32	1.18	12.6	3	49.1	49.1	3.57	0.92	9.6	5	29.9	79.0	2.79
	Tail.moy	8.15	3.09	0.85	4.6	8	16.5	16.5	2.07	-1.23	12.2	3	34.5	50.9	-2.99
TAILLE		20.00	16.67		30.7						21.8				
VELOCITE	Lent	6.30	6.17	0.32	1.3	14	6.0	6.0	1.25	1.04	17.5	1	64.2	70.2	4.09
	Rapide	6.67	5.56	-0.89	9.2	6	39.8	39.8	-3.22	-0.37	2.0	10	6.9	46.7	-1.34
	Vite.moy	7.04	4.94	0.60	3.7	9	15.3	15.3	2.00	-0.89	10.1	4	33.2	48.5	-2.94
VELOCITE		20.00	16.67		14.2						29.6				

Aides à l'interprétation pour les modalités actives

Modalités actives		AXE3							AXE4						
Variable	Modalité	CONTN	POIDS	COORD	CTR	RCTR	CO2	QLT	V.TEST	COORD	CTR	RCTR	CO2	QLT	V.TEST
AFFECTIO	Affectue	4.81	8.64	-0.06	0.2	15	0.4	72.8	-0.32	0.08	0.3	13	0.6	73.5	0.41
	Peu affe	5.19	8.02	0.07	0.2	14	0.4	72.8	0.32	-0.08	0.4	12	0.6	73.5	-0.41
AFFECTIO		10.00	16.67		0.3						0.7				
AGRESSIV	Agressif	5.19	8.02	0.33	4.2	8	10.3	31.7	1.64	0.55	15.5	3	28.2	59.9	2.71
	Peu agre	4.81	8.64	-0.31	3.9	9	10.3	31.7	-1.64	-0.51	14.4	4	28.2	59.9	-2.71
AGRESSIV		10.00	16.67		8.2						29.8				
INTELLIG	Inte.moy	5.19	8.02	0.49	9.3	4	22.6	42.8	2.42	-0.60	18.5	2	33.8	76.6	-2.97
	Peu inte	7.04	4.94	-0.35	2.9	12	5.2	37.9	-1.16	0.02	0.0	16	0.0	37.9	0.08
	Très int	7.78	3.70	-0.60	6.3	6	10.3	19.5	-1.64	1.28	38.2	1	46.5	66.0	3.48
INTELLIG		20.00	16.67		18.5						56.8				
POIDS	P.lourd	8.15	3.09	1.22	21.8	1	33.9	78.9	2.97	0.07	0.1	15	0.1	79.0	0.16
	P.Léger	7.04	4.94	-0.36	3.0	10	5.4	91.6	-1.19	0.16	0.9	9	1.1	92.7	0.55
	Poid.moy	4.81	8.64	-0.23	2.2	13	5.8	88.0	-1.22	-0.12	0.8	10	1.5	89.5	-0.63
POIDS		20.00	16.67		27.0						1.7				
TAILLE	Grand	4.44	9.26	-0.05	0.1	16	0.3	87.9	-0.29	-0.17	1.7	8	3.6	91.5	-0.97
	Petit	7.41	4.32	-0.62	7.8	5	13.3	92.3	-1.86	0.12	0.4	11	0.5	92.8	0.36
	Tail.moy	8.15	3.09	1.02	15.1	3	23.5	74.4	2.47	0.34	2.3	6	2.7	77.1	0.83
TAILLE		20.00	16.67		23.0						4.4				
VELOCITE	Lent	6.30	6.17	0.40	4.7	7	9.5	79.7	1.57	-0.08	0.3	14	0.4	80.1	-0.31
	Rapide	6.67	5.56	-0.76	15.3	2	29.1	75.8	-2.75	-0.24	2.0	7	2.9	78.7	-0.86
	Vite.moy	7.04	4.94	0.36	3.0	11	5.3	53.9	1.18	0.37	4.3	5	5.8	59.7	1.23
VELOCITE		20.00	16.67		23.0						6.6				

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES

Aides à l'interprétation pour les modalités supplémentaires

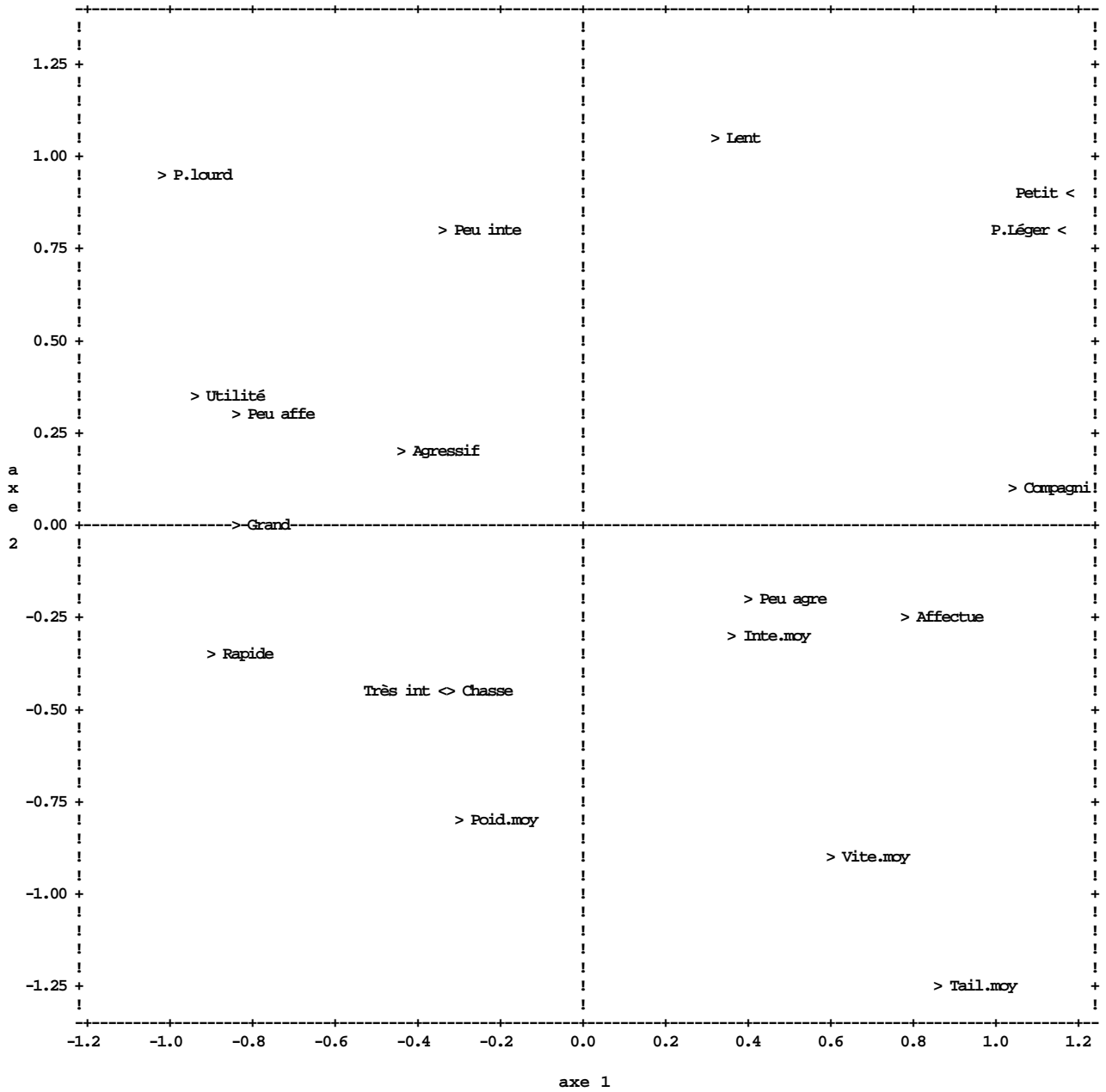
Modalités supplémentaires		AXE1							AXE2						
Variable	Modalité	CONTN	POIDS	COORD	CTR	CO2	RCO2	QLT	V.TEST	COORD	CTR	CO2	RCO2	QLT	V.TEST
FONCTION	Chasse	. 33.33		-0.32	. 5.2	3	5.2	-1.16	-0.43	. 9.3	1	14.5	-1.56		
	Compagni	. 37.04		1.04	. 63.5	1	63.5	4.06	0.10	. 0.5	3	64.1	0.37		
	Utilité	. 29.63		-0.94	. 36.9	2	36.9	-3.10	0.37	. 5.7	2	42.6	1.22		

Aides à l'interprétation pour les modalités supplémentaires

Modalités supplémentaires		AXE3							AXE4						
Variable	Modalité	CONTN	POIDS	COORD	CTR	CO2	RCO2	QLT	V.TEST	COORD	CTR	CO2	RCO2	QLT	V.TEST
FONCTION	Chasse	. 33.33		-0.35	. 6.1	2	20.6	-1.26	-0.18	. 1.7	2	22.3	-0.66		
	Compagni	. 37.04		-0.07	. 0.3	3	64.3	-0.27	-0.09	. 0.5	3	64.8	-0.35		
	Utilité	. 29.63		0.48	. 9.6	1	52.2	1.58	0.32	. 4.3	1	56.5	1.05		

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
 Les modalités actives et supplémentaires dans le plan 1-2

Plot of DIM2*DIM1\$_NAME_. Symbol points to label.



Exemple 2 : ACM avec construction du tableau disjonctif complet

```
LIBNAME COMPIL ...; /* uniquement à l'INSEE */
OPTIONS SASMSTORE=COMPIL MSTORED;

PROC CORRESP NOPRINT DIMENS=6 DATA=A OUTC=SOR;
TABLES CHIEN,TAILLE POIDS VELOCITE INTELLIG AFFECTIO AGRESSIV FONCTION;
SUPPLEMENTARY FONCTION;
WEIGHT SUPPLE;
FORMAT FONCTION FONCFOR. TAILLE TAILFOR. POIDS POIDFOR. VELOCITE VELOFOR.
        INTELLIG INTEFOR. AFFECTIO AFFEFOR. AGRESSIV AGREFOR.;

TITLE "Analyse des correspondances multiples sur le tableau des caractéristiques
des races de chiens";
TITLE2
"Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES et
construction du tableau disjonctif complet";
%AIDEACM(DATA = SOR, DATAINIT = A,
        ANALYSE = TABLES,
        VARACT = TAILLE POIDS VELOCITE INTELLIG AFFECTIO AGRESSIV,
        VARSUP = FONCTION,
        IVA = 2, IVS = 2, IOA = 2, IOS = 2,
        FORMAT = FONCTION FONCFOR. TAILLE TAILFOR. POIDS POIDFOR.
                VELOCITE VELOFOR. INTELLIG INTEFOR. AFFECTIO AFFEFOR.
                AGRESSIV AGREFOR.,
        WEIGHT = SUPPLE);

TITLE2 "Les points actifs et supplémentaires dans le plan 1-2";
%PLOTGOR(AXEH=1,AXEV=2,POINTS=VARACT VARSUP OBSACT,DATA=SOR);
```

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES et construction du tableau disjonctif complet

```
*****
*          Caractéristiques de l'analyse          *
*****
*      Type d'analyse                = TABLES *
*
*      Nombre de variables actives    =      6 *
*      Nombre de modalités actives    =     16 *
*      Nombre de variables supplémentaires =    1 *
*      Nombre de modalités supplémentaires =    3 *
*
*      Somme des poids des individus actifs =    27 *
*      Nombre d'individus actifs      =    27 *
*      Nombre d'individus supplémentaires =    0 *
*
*      Variable de pondération        = SUPPLE *
*
*      Edition des aides à l'interprétation : *
*      variables actives              sur 2 axes *
*      variables supplémentaires sur 2 axes *
*      individus actifs              sur 2 axes *
*      individus supplémentaires sur 2 axes *
*****
```

Liste des modalités actives :

```
pour la variable TAILLE      : Grand Petit Tail.moy
pour la variable POIDS       : P.lourd P.Léger Poid.moy
pour la variable VELOCITE    : Lent Rapide Vite.moy
pour la variable INTELLIG    : Inte.moy Peu inte Très int
pour la variable AFFECTIO    : Affectue Peu affe
pour la variable AGRESSIV    : Agressif Peu agre
```

Liste des modalités supplémentaires :

```
pour la variable FONCTION    : Chasse Compagni Utilité
```

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES et construction du tableau disjonctif complet

Valeurs propres

Inertie totale : 1.67

N°	Val.Pr.	Diff.	Pct	Cum	!
1	0.4816	.	28.90	28.90	!*****
2	0.3847	0.0969	23.08	51.98	!*****
3	0.2110	0.1738	12.66	64.64	!*****
4	0.1576	0.0534	9.45	74.09	!*****
5	0.1501	0.0074	9.01	83.10	!*****
6	0.1233	0.0268	7.40	90.50	!*****

Aides à l'interprétation pour les individus actifs

Individus actifs Ident.	AXE1						AXE2					
	CONTR	POIDS	COORD	CTR	RCTR	CO2	QLT	COORD	CTR	RCTR	CO2	QLT
BASSET !	4.24	3.70	0.25	0.5	25	3.4	3.4	1.10	11.7	1	63.5	66.9
BEAUCERO !	2.52	3.70	-0.32	0.8	24	8.9	8.9	-0.42	1.7	19	15.4	24.2
BERGER A !	3.42	3.70	-0.49	1.8	20	15.4	15.4	-0.46	2.1	17	14.0	29.4
BOXER !	3.99	3.70	0.45	1.5	22	11.1	11.1	-0.88	7.5	6	43.3	54.4
BULL-DOG !	3.65	3.70	1.01	7.9	2	62.4	62.4	0.55	2.9	10	18.4	80.8
BULL-MAS !	4.65	3.70	-0.75	4.4	11	27.1	27.1	0.55	2.9	12	14.3	41.4
CANICHE !	4.80	3.70	0.91	6.4	4	38.5	38.5	-0.02	0.0	27	0.0	38.5
CHIHUAHU !	4.13	3.70	0.84	5.4	8	38.0	38.0	0.84	6.9	7	38.3	76.3
COCKER !	4.28	3.70	0.73	4.1	12	27.9	27.9	0.08	0.1	25	0.3	28.2
COLLEY !	2.47	3.70	-0.12	0.1	27	1.2	1.2	-0.53	2.7	13	24.9	26.2
DALMATIE !	3.94	3.70	0.65	3.2	15	23.6	23.6	-0.99	9.4	3	55.3	78.9
DOBERMAN !	3.47	3.70	-0.87	5.9	7	48.8	48.8	-0.32	1.0	21	6.4	55.1
DOGUE AL !	4.34	3.70	-1.05	8.4	1	56.1	56.1	0.51	2.5	15	13.1	69.2
EPAGNE B !	4.84	3.70	0.48	1.8	21	10.5	10.5	-1.04	10.4	2	49.4	59.9
EPAGNE F !	2.66	3.70	-0.14	0.2	26	1.8	1.8	-0.52	2.6	14	22.2	24.0
FOX-HOUN !	3.06	3.70	-0.88	5.9	6	55.8	55.8	0.03	0.0	26	0.0	55.9
FOX TERR !	3.96	3.70	0.88	6.0	5	43.6	43.6	0.14	0.2	22	1.1	44.7
GRAND EL !	3.20	3.70	-0.52	2.1	18	18.6	18.6	-0.11	0.1	23	0.9	19.5
LABRADOR !	3.94	3.70	0.65	3.2	15	23.6	23.6	-0.99	9.4	3	55.3	78.9
LEVRIER !	3.00	3.70	-0.68	3.5	13	33.9	33.9	-0.08	0.1	24	0.5	34.4
MASTIFF !	4.23	3.70	-0.76	4.4	10	30.0	30.0	0.89	7.6	5	41.4	71.4
PEKINOIS !	4.13	3.70	0.84	5.4	8	38.0	38.0	0.84	6.9	7	38.3	76.3
POINTER !	3.42	3.70	-0.67	3.5	14	29.5	29.5	-0.42	1.7	18	11.7	41.1
SAINTE-BE !	3.75	3.70	-0.58	2.6	17	20.2	20.2	0.59	3.4	9	20.9	41.0
SETTER !	2.52	3.70	-0.50	2.0	19	22.4	22.4	-0.38	1.4	20	12.5	34.9
TECKEL !	3.65	3.70	1.01	7.9	2	62.4	62.4	0.55	2.9	10	18.4	80.8
TERRE-NE !	3.70	3.70	-0.38	1.1	23	8.8	8.8	0.49	2.3	16	14.2	23.0

Aides à l'interprétation pour les modalités actives

Modalités actives		AXE1							AXE2						
Variable	Modalité	CONTR	POIDS	COORD	CIR	RCIR	CO2	QLT	V.TEST	COORD	CIR	RCIR	CO2	QLT	V.TEST
AFFECTIO	Affectue	4.81	8.64	0.78	10.8	5	64.8	64.8	4.10	-0.27	1.6	13	7.7	72.4	-1.41
	Peu affe	5.19	8.02	-0.84	11.6	4	64.8	64.8	-4.10	0.29	1.7	11	7.7	72.4	1.41
AFFECTIO		10.00	16.67		22.4						3.3				
AGRESSIV	Agressif	5.19	8.02	-0.43	3.1	10	17.3	17.3	-2.12	0.21	0.9	14	4.1	21.4	1.03
	Peu agre	4.81	8.64	0.40	2.9	11	17.3	17.3	2.12	-0.19	0.8	15	4.1	21.4	-1.03
AGRESSIV		10.00	16.67		6.0						1.8				
INTELLIG	Inte.moy	5.19	8.02	0.37	2.3	12	12.7	12.7	1.82	-0.29	1.7	12	7.6	20.2	-1.40
	Peu inte	7.04	4.94	-0.35	1.2	15	5.1	5.1	-1.15	0.81	8.4	7	27.5	32.7	2.68
	Très int	7.78	3.70	-0.34	0.9	16	3.2	3.2	-0.91	-0.46	2.0	9	6.0	9.2	-1.25
INTELLIG		20.00	16.67		4.4						12.1				
POIDS	P.lourd	8.15	3.09	-1.02	6.6	7	23.4	23.4	-2.47	0.97	7.6	8	21.6	45.0	2.37
	P.Léger	7.04	4.94	1.17	14.0	1	57.5	57.5	3.87	0.82	8.7	6	28.6	86.1	2.73
	Poid.moy	4.81	8.64	-0.31	1.7	13	10.0	10.0	-1.62	-0.82	15.1	2	72.2	82.3	-4.33
POIDS		20.00	16.67		22.3						31.4				
TAILLE	Grand	4.44	9.26	-0.84	13.5	2	87.5	87.5	-4.77	-0.02	0.0	16	0.1	87.6	-0.12
	Petit	7.41	4.32	1.18	12.6	3	49.1	49.1	3.57	0.92	9.6	5	29.9	79.0	2.79
	Tail.moy	8.15	3.09	0.85	4.6	8	16.5	16.5	2.07	-1.23	12.2	3	34.5	50.9	-2.99
TAILLE		20.00	16.67		30.7						21.8				
VELOCITE	Lent	6.30	6.17	0.32	1.3	14	6.0	6.0	1.25	1.04	17.5	1	64.2	70.2	4.09
	Rapide	6.67	5.56	-0.89	9.2	6	39.8	39.8	-3.22	-0.37	2.0	10	6.9	46.7	-1.34
	Vite.moy	7.04	4.94	0.60	3.7	9	15.3	15.3	2.00	-0.89	10.1	4	33.2	48.5	-2.94
VELOCITE		20.00	16.67		14.2						29.6				

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens

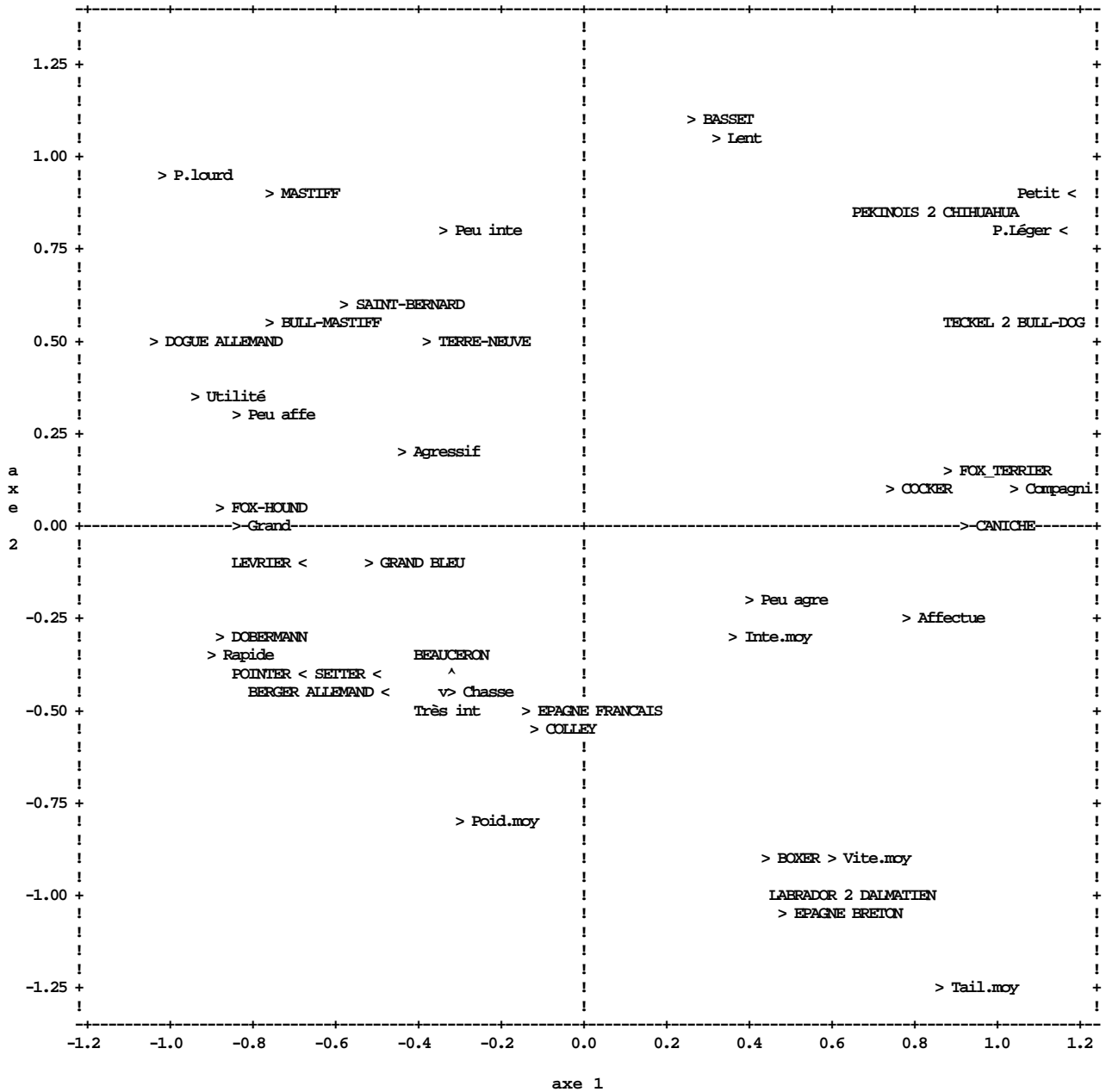
Sorties de la MACRO AIDEACM après une utilisation de l'instruction TABLES et construction du tableau disjonctif complet

Aides à l'interprétation pour les modalités supplémentaires

Modalités supplémentaires		AXE1							AXE2						
Variable	Modalité	CONTR	POIDS	COORD	CIR	CO2	RCO2	QLT	V.TEST	COORD	CIR	CO2	RCO2	QLT	V.TEST
FONCTION	Chasse	. 33.33		-0.32	. 5.2	3	5.2		-1.16	-0.43	. 9.3	1	14.5		-1.56
	Compagni	. 37.04		1.04	. 63.5	1	63.5		4.06	0.10	. 0.5	3	64.1		0.37
	Utilité	. 29.63		-0.94	. 36.9	2	36.9		-3.10	0.37	. 5.7	2	42.6		1.22

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
 Les points actifs et supplémentaires dans le plan 1-2

Plot of DIM2*DIM1\$_NAME_. Symbol points to label.



Exemple 3 : ACM sur un tableau disjonctif complet (instruction VAR)

```
LIBNAME COMPIL ...; /* uniquement à l'INSEE */
OPTIONS SASMSTORE=COMPIL MSTORED;

DATA B;SET A;

TAI1=(TAILLE=1);
TAI2=(TAILLE=2);
TAI3=(TAILLE=3);

POI1=(POIDS=1);
POI2=(POIDS=2);
POI3=(POIDS=3);

VEL1=(VELOCITE=1);
VEL2=(VELOCITE=2);
VEL3=(VELOCITE=3);

INT1=(INTELLIG=1);
INT2=(INTELLIG=2);
INT3=(INTELLIG=3);

AFF1=(AFFECTIO=1);
AFF2=(AFFECTIO=2);

AGR1=(AGRESSIV=1);
AGR2=(AGRESSIV=2);

COMP=(FONCTION=1);
CHAS=(FONCTION=2);
UTIL=(FONCTION=3);

PROC CORRESP DATA=B NOPRINT DIMENS=6 OUTC=SOR;
ID CHIEN;
VAR TAI1--UTIL;
SUPPLEMENTARY COMP--UTIL;
WEIGHT SUPPLE;

TITLE "Analyse des correspondances multiples sur le tableau des caractéristiques
des races de chiens";
TITLE2
"Sorties de la MACRO AIDEACM après une utilisation de l'instruction VAR (sur le
tableau disjonctif complet)";
%AIDEACM(DATA = SOR, DATAINIT = B,
        ANALYSE = VAR,
        ID = CHIEN,
        VARACT = TAILLE POIDS VITESSE INTELLO AFFECT AGRESSIV,
        NBMODACT = 3 3 3 3 2 2,
        VARSUP = FONCTION,
        NBMODSUP = 3,
        IVA = 2, IVS = 2, IOA = 2, IOS = 2,
        WEIGHT = SUPPLE);

TITLE2 "Les points actifs et supplémentaires dans le plan 1-2";
%PLOTGOR(AXEH=1,AXEV=2,POINTS=VARACT VARSUP OBSACT OBSSUP,DATA=SOR,ID=CHIEN);
```

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction VAR (sur le tableau disjonctif complet)

```
*****
*          Caractéristiques de l'analyse          *
*****
*      Type d'analyse                =  VAR      *
*
*      Nombre de variables actives      =    6    *
*      Nombre de modalités actives     =   16    *
*      Nombre de variables supplémentaires =    1    *
*      Nombre de modalités supplémentaires =    3    *
*
*      Somme des poids des individus actifs =   27   *
*      Nombre d'individus actifs       =   27   *
*      Nombre d'individus supplémentaires =    1    *
*
*      Variable de pondération          = SUPPLE  *
*
*      Edition des aides à l'interprétation :    *
*      variables actives                 sur 2 axes *
*      variables supplémentaires sur 2 axes *
*      individus actifs                 sur 2 axes *
*      individus supplémentaires sur 2 axes *
*****
```

Liste des modalités actives :

```
pour la variable TAILLE      : TAI1  TAI2  TAI3
pour la variable POIDS      : POI1  POI2  POI3
pour la variable VITESSE    : VEL1  VEL2  VEL3
pour la variable INTELL    : INT1  INT2  INT3
pour la variable AFFECT     : AFF1  AFF2
pour la variable AGRESSIV   : AGR1  AGR2
```

Liste des modalités supplémentaires :

```
pour la variable FONCTION   : COMP  CHAS  UTIL
```

Valeurs propres

Inertie totale : 1.67

N°	Val.Pr.	Diff.	Pct	Cum	!
1	0.4816	.	28.90	28.90	!*****
2	0.3847	0.0969	23.08	51.98	!*****
3	0.2110	0.1738	12.66	64.64	!*****
4	0.1576	0.0534	9.45	74.09	!*****
5	0.1501	0.0074	9.01	83.10	!*****
6	0.1233	0.0268	7.40	90.50	!*****

Aides à l'interprétation pour les individus actifs

Ident.	Individus actifs			AXEL					AXE2				
	CONTR	POIDS	COORD	CTR	RCTR	CO2	QLT	COORD	CTR	RCTR	CO2	QLT	
BEAUCERO !	2.52	3.70	-0.32	0.8	24	8.9	8.9	-0.42	1.7	19	15.4	24.2	
BASSET !	4.24	3.70	0.25	0.5	25	3.4	3.4	1.10	11.7	1	63.5	66.9	
BERGER A !	3.42	3.70	-0.49	1.8	20	15.4	15.4	-0.46	2.1	17	14.0	29.4	
BOXER !	3.99	3.70	0.45	1.5	22	11.1	11.1	-0.88	7.5	6	43.3	54.4	
BULL-DOG !	3.65	3.70	1.01	7.9	2	62.4	62.4	0.55	2.9	10	18.4	80.8	
BULL-MAS !	4.65	3.70	-0.75	4.4	11	27.1	27.1	0.55	2.9	12	14.3	41.4	
CANICHE !	4.80	3.70	0.91	6.4	4	38.5	38.5	-0.02	0.0	27	0.0	38.5	
CHIHUAHU !	4.13	3.70	0.84	5.4	8	38.0	38.0	0.84	6.9	7	38.3	76.3	
COCKER !	4.28	3.70	0.73	4.1	12	27.9	27.9	0.08	0.1	25	0.3	28.2	
COLLEY !	2.47	3.70	-0.12	0.1	27	1.2	1.2	-0.53	2.7	13	24.9	26.2	
DALMATIE !	3.94	3.70	0.65	3.2	15	23.6	23.6	-0.99	9.4	3	55.3	78.9	
DOBERMAN !	3.47	3.70	-0.87	5.9	7	48.8	48.8	-0.32	1.0	21	6.4	55.1	
DOGUE AL !	4.34	3.70	-1.05	8.4	1	56.1	56.1	0.51	2.5	15	13.1	69.2	
EPAGNE B !	4.84	3.70	0.48	1.8	21	10.5	10.5	-1.04	10.4	2	49.4	59.9	
EPAGNE F !	2.66	3.70	-0.14	0.2	26	1.8	1.8	-0.52	2.6	14	22.2	24.0	
FOX-HOUN !	3.06	3.70	-0.88	5.9	6	55.8	55.8	0.03	0.0	26	0.0	55.9	
FOX TERR !	3.96	3.70	0.88	6.0	5	43.6	43.6	0.14	0.2	22	1.1	44.7	
GRAND EL !	3.20	3.70	-0.52	2.1	18	18.6	18.6	-0.11	0.1	23	0.9	19.5	
LABRADOR !	3.94	3.70	0.65	3.2	15	23.6	23.6	-0.99	9.4	3	55.3	78.9	
LEVRIER !	3.00	3.70	-0.68	3.5	13	33.9	33.9	-0.08	0.1	24	0.5	34.4	
MASTIFF !	4.23	3.70	-0.76	4.4	10	30.0	30.0	0.89	7.6	5	41.4	71.4	
PEKINOIS !	4.13	3.70	0.84	5.4	8	38.0	38.0	0.84	6.9	7	38.3	76.3	
POINTER !	3.42	3.70	-0.67	3.5	14	29.5	29.5	-0.42	1.7	18	11.7	41.1	
SAINTE-BE !	3.75	3.70	-0.58	2.6	17	20.2	20.2	0.59	3.4	9	20.9	41.0	
SETTER !	2.52	3.70	-0.50	2.0	19	22.4	22.4	-0.38	1.4	20	12.5	34.9	
TECKEL !	3.65	3.70	1.01	7.9	2	62.4	62.4	0.55	2.9	10	18.4	80.8	
TERRE-NE !	3.70	3.70	-0.38	1.1	23	8.8	8.8	0.49	2.3	16	14.2	23.0	

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
Sorties de la MACRO AIDEACM après une utilisation de l'instruction VAR (sur le tableau disjonctif complet)

Aides à l'interprétation pour les individus supplémentaires

Individus supplément.		AXE1						AXE2					
Ident.	CONTR	POIDS	COORD	CTR	CO2	RCO2	QLT	COORD	CTR	CO2	RCO2	QLT	V.TEST
SUPLEME	.	.	0.20	.	2.3	1	2.3	0.02	.	0.0	1	2.3	.

Aides à l'interprétation pour les modalités actives

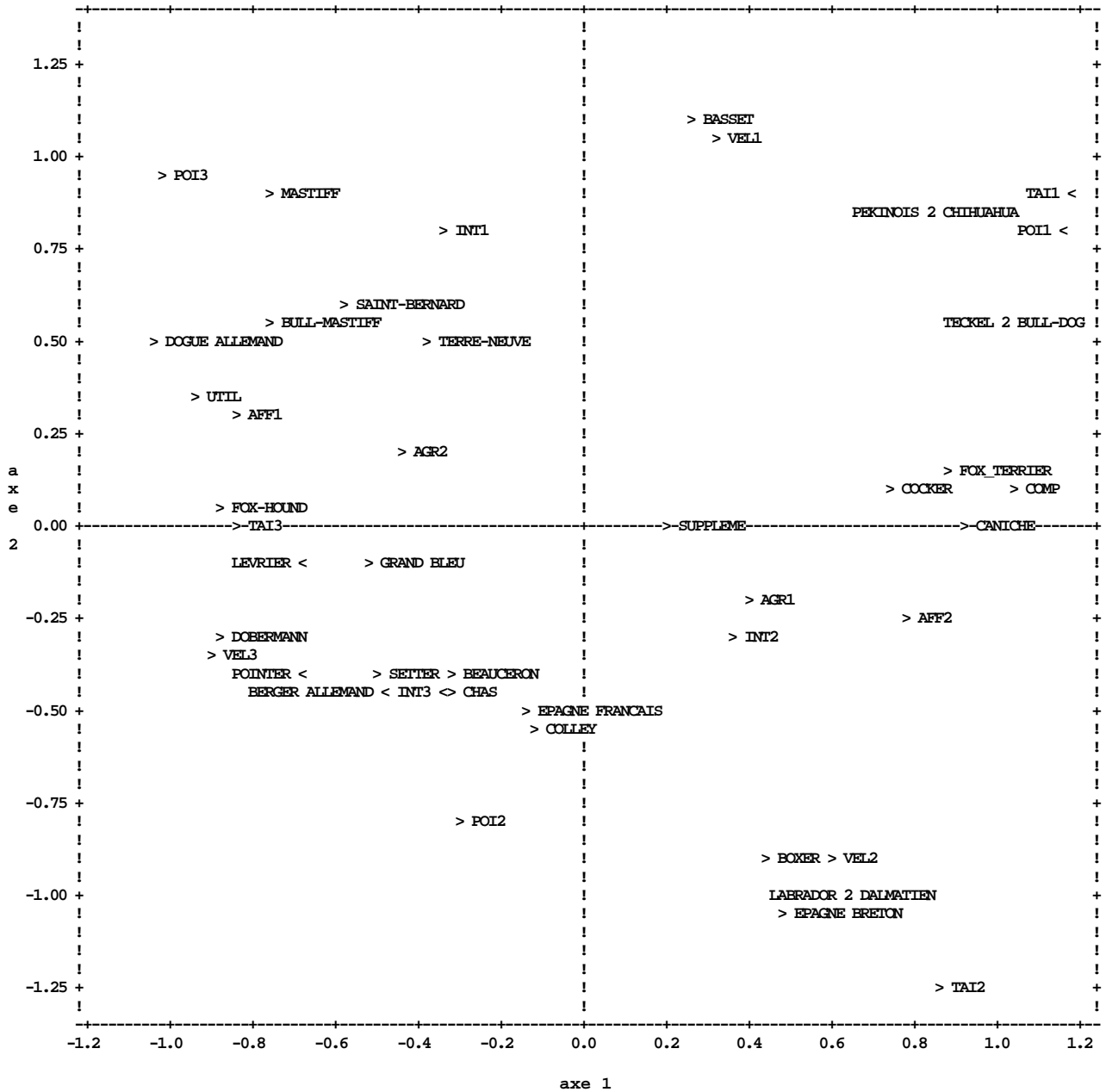
Modalités actives		AXE1						AXE2							
Variable	Modalité	CONTR	POIDS	COORD	CTR	RCIR	CO2	QLT	V.TEST	COORD	CTR	RCIR	CO2	QLT	V.TEST
AFFECT	AFF1	5.19	8.02	-0.84	11.6	4	64.8	64.8	-4.10	0.29	1.7	11	7.7	72.4	1.41
	AFF2	4.81	8.64	0.78	10.8	5	64.8	64.8	4.10	-0.27	1.6	13	7.7	72.4	-1.41
AFFECT		10.00	16.67		22.4						3.3				
AGRESSIV	AGR1	4.81	8.64	0.40	2.9	11	17.3	17.3	2.12	-0.19	0.8	15	4.1	21.4	-1.03
	AGR2	5.19	8.02	-0.43	3.1	10	17.3	17.3	-2.12	0.21	0.9	14	4.1	21.4	1.03
AGRESSIV		10.00	16.67		6.0						1.8				
INTELLO	INT1	7.04	4.94	-0.35	1.2	15	5.1	5.1	-1.15	0.81	8.4	7	27.5	32.7	2.68
	INT2	5.19	8.02	0.37	2.3	12	12.7	12.7	1.82	-0.29	1.7	12	7.6	20.2	-1.40
	INT3	7.78	3.70	-0.34	0.9	16	3.2	3.2	-0.91	-0.46	2.0	9	6.0	9.2	-1.25
INTELLO		20.00	16.67		4.4						12.1				
POIDS	POI1	7.04	4.94	1.17	14.0	1	57.5	57.5	3.87	0.82	8.7	6	28.6	86.1	2.73
	POI2	4.81	8.64	-0.31	1.7	13	10.0	10.0	-1.62	-0.82	15.1	2	72.2	82.3	-4.33
	POI3	8.15	3.09	-1.02	6.6	7	23.4	23.4	-2.47	0.97	7.6	8	21.6	45.0	2.37
POIDS		20.00	16.67		22.3						31.4				
TAILLE	TAI1	7.41	4.32	1.18	12.6	3	49.1	49.1	3.57	0.92	9.6	5	29.9	79.0	2.79
	TAI2	8.15	3.09	0.85	4.6	8	16.5	16.5	2.07	-1.23	12.2	3	34.5	50.9	-2.99
	TAI3	4.44	9.26	-0.84	13.5	2	87.5	87.5	-4.77	-0.02	0.0	16	0.1	87.6	-0.12
TAILLE		20.00	16.67		30.7						21.8				
VITESSE	VEL1	6.30	6.17	0.32	1.3	14	6.0	6.0	1.25	1.04	17.5	1	64.2	70.2	4.09
	VEL2	7.04	4.94	0.60	3.7	9	15.3	15.3	2.00	-0.89	10.1	4	33.2	48.5	-2.94
	VEL3	6.67	5.56	-0.89	9.2	6	39.8	39.8	-3.22	-0.37	2.0	10	6.9	46.7	-1.34
VITESSE		20.00	16.67		14.2						29.6				

Aides à l'interprétation pour les modalités supplémentaires

Modalités supplémentaires		AXE1						AXE2							
Variable	Modalité	CONTR	POIDS	COORD	CTR	CO2	RCO2	QLT	V.TEST	COORD	CTR	CO2	RCO2	QLT	V.TEST
FONCTION	CHAS	.	33.33	-0.32	.	5.2	3	5.2	-1.16	-0.43	.	9.3	1	14.5	-1.56
	COMP	.	37.04	1.04	.	63.5	1	63.5	4.06	0.10	.	0.5	3	64.1	0.37
	UTIL	.	29.63	-0.94	.	36.9	2	36.9	-3.10	0.37	.	5.7	2	42.6	1.22

Analyse des correspondances multiples sur le tableau des caractéristiques des races de chiens
 Les points actifs et supplémentaires dans le plan 1-2

Plot of DIM2*DIM1\$CHIEN. Symbol points to label.



4. Classification ascendante hiérarchique

Exemple 1 : tableau individus x variables numériques (les vins français)

```
LIBNAME COMPIL ...; /* uniquement à l'INSEE */
OPTIONS MSTORED SASMSTORE=COMPIL NODATE PAGESIZE=66;

DATA A;
INPUT ANNEE $ BORD_ROU BORD_BLA BOUR_ROU BOUR_BLA BEAUJOL ALSACE COT_RHON
      PAY_LOIR TEMP_MOY TEMP_MAX TEMP_MIN PLUIE PRESID $;
IF ANNEE <='90' THEN POND=1;ELSE POND=0;

/* Les instructions suivantes sont utiles pour l'utilisation de DESQUAL */

POMP=(PRESID='Pomp'); GISC=(PRESID='Gisc');
MITT=(PRESID='Mitt'); BALL=(PRESID='Ball');
LABEL POMP=Pompidou GISC="Giscard d'Estaing"
      MITT=Mitterrand BALL=Balladur;

CARDS;
69 1 2 6 5 4 3 4 3 11.7 32.8 -8.1 618 Pomp
70 6 5 3 3 4 3 4 3 11.6 31.2 -6.1 631 Pomp
71 5 5 5 3 4 6 3 3 11.8 31.6 -8.6 508 Pomp
72 2 3 4 2 3 1 3 1 11.1 30.1 -7.2 740 Pomp
73 3 3 2 4 4 4 2 3 11.6 32.7 -4.2 576 Pomp
74 2 3 2 2 2 3 1 1 12.0 34.0 -0.4 668 Gisc
75 6 1 1 3 3 4 1 3 11.7 35.7 -4.5 658 Gisc
76 4 3 5 3 6 5 4 4 12.4 35.4 -6.2 417 Gisc
77 3 3 2 2 2 3 1 1 11.7 29.0 -3.7 717 Gisc
78 4 5 6 5 5 5 6 5 10.9 30.1 -10.3 743 Gisc
79 4 3 4 4 4 4 5 4 11.0 32.1 -12.7 729 Gisc
80 2 4 2 3 1 3 3 2 11.2 31.3 -5.1 690 Gisc
81 5 5 3 4 3 4 3 5 11.8 31.9 -3.1 746 Mitt
82 6 5 4 3 3 4 4 5 12.4 32.8 -6.6 700 Mitt
83 4 4 3 4 5 6 4 4 12.3 33.4 -4.2 623 Mitt
84 2 3 2 2 1 2 2 4 11.7 34.8 -3.3 745 Mitt
85 5 5 5 4 5 5 3 5 10.8 32.0 -15.9 501 Mitt
98 2 5 5 2 3 5 4 5 10.6 33.0 -12.9 601 Ball
99 5 1 5 4 5 5 3 1 10.4 32.7 -15.7 551 Ball
;
PROC PRINT;
TITLE 'Notes (de 1 à 6) données aux vins français';
RUN;

TITLE "CAH sur un tableau individus x variables numériques : les vins français";
TITLE2 "Les sorties de la macro CAHNUM";
%CAHNUM (DATA=A,
        VAR=BORD_ROU--PAY_LOIR,
        ID=ANNEE,
        POIDS=POND,
        ARBRE=OUI,
        NOEUDS=MAX);
RUN;
```


CAH sur un tableau individus x variables numériques : les vins français
Les sorties de la macro CAHNUM

```
*****
*      Caractéristiques de la classification      *
*****
*      Nombre de variables          =      8      *
*
*      Nombre d'observations        =      17      *
*      Nombre d'observations pondérées =      17      *
*      Nombre d'observations éliminées =      2      *
*
*      Variable de pondération      = POND        *
*
*      Variable identifiant         = ANNEE       *
*
*      Nombre de noeuds édités     = MAX        *
*
*      Méthode utilisée             = WARD       *
*****
```

Liste des variables :

BORD_ROU BORD_BLA BOUR_ROU BOUR_BLA BEAUJOL ALSACE COT_RHON PAY_LOIR

Tableau des noeuds de la hiérarchie

Inertie totale : 8.000000

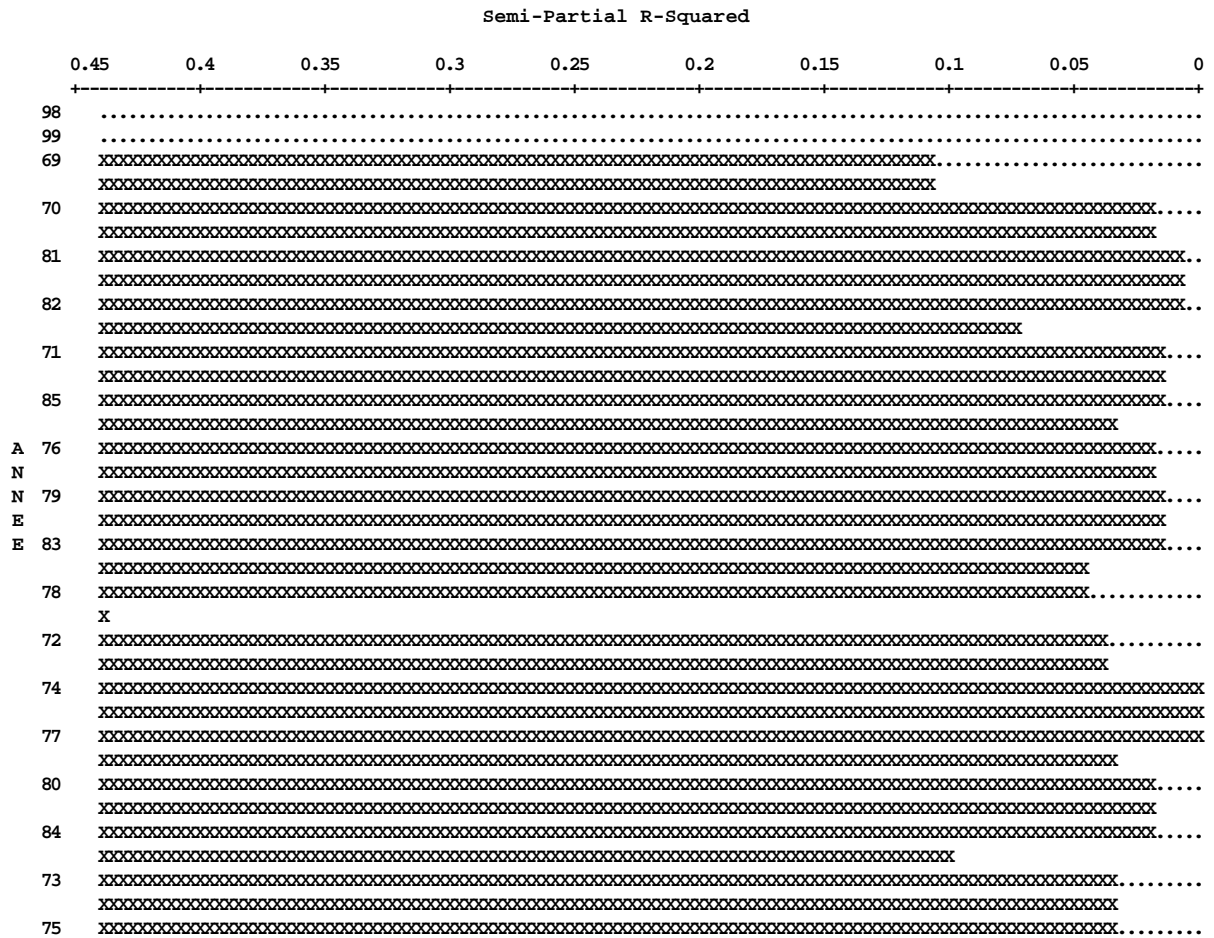
Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter- 0/00 cum. dif !	Histogramme
CL1	17.0	CL2-CL3	438 438 . !*****	
CL2	10.0	69-CL4	106 544 332 !*****	
CL3	7.0	CL6-CL9	99 643 8 !*****	
CL4	9.0	CL10-CL5	73 716 25 !*****	
CL5	6.0	CL7-78	44 760 29 !*****	
CL6	5.0	72-CL8	37 797 7 !*****	
CL7	5.0	CL14-CL12	36 833 1 !*****	
CL8	4.0	CL16-CL11	34 867 2 !*****	
CL9	2.0	73-75	34 901 0 !*****	
CL10	3.0	70-CL15	19 920 15 !***	
CL11	2.0	80-84	19 939 0 !***	
CL12	3.0	76-CL13	18 956 1 !**	
CL13	2.0	79-83	17 973 1 !**	
CL14	2.0	71-85	16 989 1 !**	
CL15	2.0	81-82	9 998 7 !*	
CL16	2.0	74-77	2 1000 8 !	

Décomposition de la distance entre les centres de gravité
des classes jointes
(Il s'agit aussi de la décomposition de la perte d'inertie interclasse)

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR
CL1	17	CL2-CL3	438	68	89	157	114	148	111	170	143
CL2	10	69-CL4	106	368	265	88	121	4	103	0	50
CL3	7	CL6-CL9	99	233	109	38	197	162	163	14	83
CL4	9	CL10-CL5	73	148	99	159	55	239	270	27	3
CL5	6	CL7-78	44	9	98	157	297	3	3	358	75
CL6	5	72-CL8	37	4	7	282	11	188	290	133	85
CL7	5	CL14-CL12	36	101	482	108	7	32	36	232	0
CL8	4	CL16-CL11	34	23	39	0	60	115	33	262	470
CL9	2	73-75	34	406	310	48	119	57	0	58	0
CL10	3	70-CL15	19	27	0	28	70	135	153	34	553
CL11	2	80-84	19	0	140	0	215	0	117	105	423
CL12	3	76-CL13	18	0	50	278	305	330	0	37	0
CL13	2	79-83	17	0	154	96	0	114	518	116	0
CL14	2	71-85	16	0	0	0	250	121	136	0	493
CL15	2	81-82	9	166	0	179	440	0	0	215	0
CL16	2	74-77	2	1000	0	0	0	0	0	0	0

CAH sur un tableau individus x variables numériques : les vins français
 Les sorties de la macro CAHNUM
 Arbre de classification

Ward's Minimum Variance Cluster Analysis



CAH sur un tableau individus x variables numériques : les vins français
 La macro PARTNUM : partition en 5 classes
 Partition en 5 classes : composition des classes

VARIABLE DE CLASSE : CLASS5

Individus non inclus dans une classe
 98 99

Classe numéro 1
 72 74 77 80 84

Classe numéro 2
 70 81 82

Classe numéro 3
 71 76 78 79 83 85

Classe numéro 4
 73 75

Classe numéro 5
 69

Partition en 5 classes : Décomposition de l'inertie totale

	CLASS5	Inertie	En 0/00 de l'inertie totale	Effectif absolu	Effectif pondéré (en 0/00)
Inertie totale		8.0000	1000	17	1000
Inertie inter		5.7256	716	.	.
	1	2.2147	277	5	294
	2	0.6316	79	3	176
	3	1.5305	191	6	353
	4	0.6346	79	2	118
	5	0.7142	89	1	59
Inertie intra		2.2744	284	.	.
	1	0.7290	91	5	294
	2	0.2263	28	3	176
	3	1.0492	131	6	353
	4	0.2699	34	2	118
	5	0.0000	0	1	59

Décomposition signée du RHO2 (distance entre le centre des classes et le centre de gravité du nuage)
 selon les variables de l'analyse
 tous les chiffres sont en millièmes, exceptés l'effectif absolu et le RHO2

Classe	Effectif absolu	Effectif pondéré	RHO2	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR
1	5	294	7.5300	-135	-19	-68	-174	-195	-161	-89	-160
2	3	176	3.5790	418	364	-2	0	-3	-4	45	163
3	6	353	4.3364	31	44	147	73	226	248	136	95
4	2	118	5.3940	42	-358	-320	9	0	3	-260	-9
5	1	59	12.1421	-261	-159	234	262	12	-33	34	-4

Décomposition signée de la distance entre les centres de gravité des classes de la partition
 tous les chiffres sont en millièmes
 exceptées la perte d'inertie inter et la distance

Classes	Distance	Perte d'inertie inter	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR
2 - 1	16.2204	1.78902	307	142	24	87	76	59	92	213
3 - 1	22.7096	3.64325	83	29	101	129	214	201	111	133
4 - 1	9.3884	0.78894	233	-109	-38	197	162	163	-14	83
5 - 1	21.6505	1.06130	-28	-47	266	396	118	10	99	36
3 - 2	4.9733	0.58510	-148	-99	159	55	239	270	27	-3
4 - 2	12.0467	0.85036	-47	-532	-124	3	1	6	-209	-80
5 - 2	23.1436	1.02104	-390	-277	137	131	10	-11	3	-41
4 - 3	14.2232	1.25499	1	-235	-313	-9	-66	-57	-268	-52
5 - 3	14.1406	0.71297	-325	-236	56	105	-26	-198	-1	-52
5 - 4	20.6124	0.80833	-246	0	437	119	6	-29	162	0

CAH sur un tableau individus x variables numériques : les vins français
 La macro PARTNUM : partition en 5 classes
 Décomposition de l'inertie inter de la partition en 5 classes
 Selon les variables de l'analyse et les classes

TABLE OF CLASS5 BY IN_INTER

CLASS5	IN_INTER								
Frequency!									
Percent !									
Row Pct !									
Col Pct	BORD_ROU	BORD_ELA	BOUR_ROU	BOUR_ELA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR	Total
1	0.2981	0.0418	0.1499	0.3854	0.4329	0.3559	0.1966	0.354	2.2147
	5.21	0.73	2.62	6.73	7.56	6.22	3.43	6.18	38.68
	13.46	1.89	6.77	17.40	19.55	16.07	8.88	15.98	
	36.26	6.15	20.08	55.79	54.86	46.54	31.60	58.05	
2	0.2643	0.2299	0.0015	0.0003	0.0018	0.0026	0.0285	0.1028	0.6316
	4.62	4.02	0.03	0.01	0.03	0.05	0.50	1.79	11.03
	41.85	36.41	0.23	0.05	0.28	0.41	4.51	16.27	
	32.15	33.79	0.20	0.04	0.22	0.34	4.58	16.85	
3	0.0472	0.0678	0.2245	0.1123	0.3457	0.3802	0.2079	0.1449	1.5305
	0.83	1.18	3.92	1.96	6.04	6.64	3.63	2.53	26.73
	3.09	4.43	14.67	7.34	22.58	24.84	13.58	9.47	
	5.75	9.97	30.08	16.26	43.81	49.72	33.41	23.76	
4	0.0263	0.2272	0.2031	0.0055	537E-7	0.0022	0.1648	0.0055	0.6346
	0.46	3.97	3.55	0.10	0.00	0.04	2.88	0.10	11.08
	4.15	35.80	32.01	0.86	0.01	0.34	25.96	0.86	
	3.20	33.39	27.22	0.79	0.01	0.29	26.48	0.90	
5	0.1862	0.1136	0.1673	0.1874	0.0087	0.0238	0.0245	0.0027	0.7142
	3.25	1.98	2.92	3.27	0.15	0.42	0.43	0.05	12.47
	26.06	15.90	23.43	26.24	1.22	3.34	3.43	0.38	
	22.64	16.70	22.42	27.12	1.10	3.12	3.94	0.45	
Total	0.82221	0.68038	0.74631	0.69091	0.78905	0.76467	0.62222	0.60989	5.72563
	14.36	11.88	13.03	12.07	13.78	13.36	10.87	10.65	100.00

Moyenne et écart-type par classe, tests de significativité :
test de Fisher dans la colonne Ensemble et test de comparaisons de moyennes

Variable	Statistique	Ensemble	_1	_2	_3	_4	_5
	Eff.pond	17.0000	5.00000	3.00000	6.00000	2.00000	1.00000
ALSACE	Moyenne	3.8235	2.40000	3.66667	5.16667	4.00000	3.00000
	Ec.type	1.2941	0.80000	0.47140	0.68718	0.00000	0.00000
	V.test	9.7480	-2.84019	-0.22444	3.06610	0.19917	-0.63636
	Proba	0.0009	0.00451	0.82241	0.00217	0.84213	0.52454
BEAUJOL	Moyenne	3.4706	1.80000	3.33333	4.83333	3.50000	4.00000
	Ec.type	1.3770	0.74833	0.47140	0.68718	0.50000	0.00000
	V.test	11.2215	-3.13244	-0.18456	2.92356	0.03120	0.38446
	Proba	0.0005	0.00173	0.85357	0.00346	0.97511	0.70064
BORD_BLA	Moyenne	3.6471	3.20000	5.00000	4.16667	2.00000	2.00000
	Ec.type	1.1853	0.40000	0.00000	0.89753	1.00000	0.00000
	V.test	6.3861	-0.97388	2.11359	1.29509	-2.02967	-1.38962
	Proba	0.0054	0.33012	0.03455	0.19529	0.04239	0.16465
BORD_ROU	Moyenne	3.7647	2.20000	5.66667	4.33333	4.50000	1.00000
	Ec.type	1.5541	0.40000	0.47140	0.47140	1.50000	0.00000
	V.test	13.8735	-2.59961	2.26610	1.08091	0.69105	-1.77898
	Proba	0.0002	0.00933	0.02345	0.27974	0.48953	0.07524
BOUR_BLA	Moyenne	3.2941	2.20000	3.33333	3.83333	3.50000	5.00000
	Ec.type	0.9558	0.40000	0.47140	0.68718	0.50000	0.00000
	V.test	6.7059	-2.95573	0.07597	1.66667	0.31463	1.78483
	Proba	0.0045	0.00312	0.93944	0.09558	0.75305	0.07429
BOUR_ROU	Moyenne	3.4706	2.40000	3.33333	4.66667	1.50000	6.00000
	Ec.type	1.4997	0.80000	0.47140	0.94281	0.50000	0.00000
	V.test	8.8253	-1.84319	-0.16946	2.35609	-1.91919	1.68660
	Proba	0.0015	0.06530	0.86543	0.01847	0.05496	0.09168
COT_RHON	Moyenne	3.1176	2.00000	3.66667	4.16667	1.50000	4.00000
	Ec.type	1.3669	0.89443	0.47140	1.06719	0.50000	0.00000
	V.test	4.9412	-2.11111	0.74370	2.26712	-1.72848	0.64550
	Proba	0.0138	0.03476	0.45706	0.02338	0.08390	0.51861
PAY_LOIR	Moyenne	3.2941	1.80000	4.33333	4.16667	3.00000	3.00000
	Ec.type	1.3619	1.16619	0.94281	0.68718	0.00000	0.00000
	V.test	4.6901	-2.83273	1.41296	1.89276	-0.31544	-0.21597
	Proba	0.0164	0.00462	0.15767	0.05839	0.75243	0.82901

CAH sur un tableau individus x variables numériques : les vins français
Table en sortie créée par la macro PARTNUM

OBS	CLASS5	ANNEE	BORD_ROU	BORD_BLA	BOUR_ROU	BOUR_BLA	BEAUJOL	ALSACE	COT_RHON	PAY_LOIR	TEMP_MOY	TEMP_MAX	TEMP_MIN	PLUIE	POND	CLINIT5
1	5	69	1	2	6	5	4	3	4	3	11.7	32.8	-8.1	618	1	5
2	2	70	6	5	3	3	4	3	4	3	11.6	31.2	-6.1	631	1	2
3	3	71	5	5	5	3	4	6	3	3	11.8	31.6	-8.6	508	1	3
4	1	72	2	3	4	2	3	1	3	1	11.1	30.1	-7.2	740	1	1
5	4	73	3	3	2	4	4	4	2	3	11.6	32.7	-4.2	576	1	4
6	1	74	2	3	2	2	2	3	1	1	12.0	34.0	-0.4	668	1	1
7	4	75	6	1	1	3	3	4	1	3	11.7	35.7	-4.5	658	1	4
8	3	76	4	3	5	3	6	5	4	4	12.4	35.4	-6.2	417	1	3
9	1	77	3	3	2	2	2	3	1	1	11.7	29.0	-3.7	717	1	1
10	3	78	4	5	6	5	5	5	6	5	10.9	30.1	-10.3	743	1	3
11	3	79	4	3	4	4	4	4	5	4	11.0	32.1	-12.7	729	1	3
12	1	80	2	4	2	3	1	3	3	2	11.2	31.3	-5.1	690	1	1
13	2	81	5	5	3	4	3	4	3	5	11.8	31.9	-3.1	746	1	2
14	2	82	6	5	4	3	3	4	4	5	12.4	32.8	-6.6	700	1	2
15	3	83	4	4	3	4	5	6	4	4	12.3	33.4	-4.2	623	1	3
16	1	84	2	3	2	2	1	2	2	4	11.7	34.8	-3.3	745	1	1
17	3	85	5	5	5	4	5	5	3	5	10.8	32.0	-15.9	501	1	3
18	3	98	2	5	5	2	3	5	4	5	10.6	33.0	-12.9	601	0	
19	3	99	5	1	5	4	5	5	3	1	10.4	32.7	-15.7	551	0	

CAH sur un tableau individus x variables numériques : les vins français
 La macro DESNUM : description de la partition en 5 classes avec des variables supplémentaires
 Moyenne et écart-type par modalité de la variable CLASS5, tests de significativité :
 test de Fisher dans la colonne Ensemble et test de comparaisons de moyennes

Variable	Statistique	Ensemble	_1	_2	_3	_4	_5
	Eff.pond	17.000	5.000	3.000	6.000	2.000	1.000
PLUIE	Moyenne	647.647	712.000	692.333	586.833	617.000	618.000
	Ec.type	95.288	29.387	47.261	121.335	41.000	0.000
	V.test	1.508	1.744	0.868	-1.885	-0.470	-0.311
	Proba	0.261	0.081	0.385	0.059	0.639	0.756
TEMP_MAX	Moyenne	32.406	31.840	31.967	32.433	34.200	32.800
	Ec.type	1.808	2.228	0.655	1.642	1.500	0.000
	V.test	0.563	-0.808	-0.450	0.045	1.450	0.218
	Proba	0.694	0.419	0.653	0.964	0.147	0.827
TEMP_MIN	Moyenne	-6.482	-3.940	-5.267	-9.650	-4.350	-8.100
	Ec.type	3.697	2.235	1.546	3.905	0.150	0.000
	V.test	2.636	1.776	0.609	-2.531	0.842	-0.438
	Proba	0.087	0.076	0.543	0.011	0.400	0.662
TEMP_MOY	Moyenne	11.629	11.540	11.933	11.533	11.650	11.700
	Ec.type	0.481	0.338	0.340	0.662	0.050	0.000
	V.test	0.319	-0.480	1.170	-0.590	0.063	0.147
	Proba	0.860	0.631	0.242	0.555	0.950	0.883

CAH sur un tableau individus x variables numériques : les vins français
 La macro DESNUM : description de la partition en 5 classes avec des observations et des variables supplémentaires
 Moyenne et écart-type par modalité de la variable CLASS5, tests de significativité :
 test de Fisher dans la colonne Ensemble et test de comparaisons de moyennes

Variable	Statistique	Ensemble	_1	_2	_3	_4	_5
	Eff.pond	19.000	5.000	3.000	8.000	2.000	1.000
ALSACE	Moyenne	3.947	2.400	3.667	5.125	4.000	3.000
	Ec.type	1.276	0.800	0.471	0.599	0.000	0.000
	V.test	12.567	-3.074	-0.404	3.339	0.060	-0.742
	Proba	0.000	0.002	0.686	0.001	0.952	0.458
BEAUJOL	Moyenne	3.526	1.800	3.333	4.625	3.500	4.000
	Ec.type	1.352	0.748	0.471	0.857	0.500	0.000
	V.test	8.853	-3.237	-0.262	2.940	-0.028	0.350
	Proba	0.001	0.001	0.793	0.003	0.977	0.726
BORD_BLA	Moyenne	3.579	3.200	5.000	3.875	2.000	2.000
	Ec.type	1.311	0.400	0.000	1.364	1.000	0.000
	V.test	2.962	-0.733	1.992	0.817	-1.753	-1.205
	Proba	0.058	0.463	0.046	0.414	0.080	0.228
BORD_ROU	Moyenne	3.737	2.200	5.667	4.125	4.500	1.000
	Ec.type	1.551	0.400	0.471	0.927	1.500	0.000
	V.test	8.951	-2.513	2.286	0.906	0.716	-1.765
	Proba	0.001	0.012	0.022	0.365	0.474	0.078
BOUR_BLA	Moyenne	3.263	2.200	3.333	3.625	3.500	5.000
	Ec.type	0.965	0.400	0.471	0.857	0.500	0.000
	V.test	4.393	-2.794	0.134	1.357	0.357	1.800
	Proba	0.017	0.005	0.894	0.175	0.721	0.072
BOUR_ROU	Moyenne	3.632	2.400	3.333	4.750	1.500	6.000
	Ec.type	1.494	0.800	0.471	0.829	0.500	0.000
	V.test	11.548	-2.090	-0.367	2.708	-2.076	1.585
	Proba	0.000	0.037	0.714	0.007	0.038	0.113
COT_RHON	Moyenne	3.158	2.000	3.667	4.000	1.500	4.000
	Ec.type	1.308	0.894	0.471	1.000	0.500	0.000
	V.test	5.146	-2.244	0.714	2.329	-1.844	0.644
	Proba	0.009	0.025	0.475	0.020	0.065	0.520
PAY_LOIR	Moyenne	3.263	1.800	4.333	3.875	3.000	3.000
	Ec.type	1.445	1.166	0.943	1.269	0.000	0.000
	V.test	2.717	-2.567	1.360	1.532	-0.265	-0.182
	Proba	0.073	0.010	0.174	0.126	0.791	0.856
PLUIE	Moyenne	640.105	712.000	692.333	584.125	617.000	618.000
	Ec.type	93.131	29.387	47.261	105.924	41.000	0.000
	V.test	2.039	1.957	1.030	-2.175	-0.361	-0.237
	Proba	0.144	0.050	0.303	0.030	0.718	0.812
TEMP_MAX	Moyenne	32.453	31.840	31.967	32.538	34.200	32.800
	Ec.type	1.716	2.228	0.655	1.435	1.500	0.000
	V.test	0.659	-0.905	-0.520	0.179	1.482	0.202
	Proba	0.630	0.365	0.603	0.858	0.138	0.840
TEMP_MIN	Moyenne	-7.305	-3.940	-5.267	-10.813	-4.350	-8.100
	Ec.type	4.265	2.235	1.546	3.998	0.150	0.000
	V.test	4.059	2.001	0.878	-2.975	1.008	-0.186
	Proba	0.022	0.045	0.380	0.003	0.313	0.852
TEMP_MOY	Moyenne	11.511	11.540	11.933	11.275	11.650	11.700
	Ec.type	0.573	0.338	0.340	0.729	0.050	0.000
	V.test	0.716	0.130	1.356	-1.487	0.354	0.331
	Proba	0.595	0.896	0.175	0.137	0.723	0.741

CAH sur un tableau individus x variables numériques : les vins français
 La macro DESQUAL : description de la partition en 5 classes avec une variable supplémentaire qualitative
 Tableau(x) de contingence croisant la variable de classe et les variables explicatives

----- Variable=PRESIDENT -----

TABLE OF CLASS5 BY MODALITE

CLASS5	MODALITE					Total
Frequency!						
Percent !						
Row Pct !						
Col Pct	!Balladur!	Giscard !	Mitterra!	Pompidou!		
	!	!d'Estain!	!nd	!	!	
1	!	0 !	3 !	1 !	1 !	5
	!	0.00 !	15.79 !	5.26 !	5.26 !	26.32
	!	0.00 !	60.00 !	20.00 !	20.00 !	
	!	0.00 !	42.86 !	20.00 !	20.00 !	
2	!	0 !	0 !	2 !	1 !	3
	!	0.00 !	0.00 !	10.53 !	5.26 !	15.79
	!	0.00 !	0.00 !	66.67 !	33.33 !	
	!	0.00 !	0.00 !	40.00 !	20.00 !	
3	!	2 !	3 !	2 !	1 !	8
	!	10.53 !	15.79 !	10.53 !	5.26 !	42.11
	!	25.00 !	37.50 !	25.00 !	12.50 !	
	!	100.00 !	42.86 !	40.00 !	20.00 !	
4	!	0 !	1 !	0 !	1 !	2
	!	0.00 !	5.26 !	0.00 !	5.26 !	10.53
	!	0.00 !	50.00 !	0.00 !	50.00 !	
	!	0.00 !	14.29 !	0.00 !	20.00 !	
5	!	0 !	0 !	0 !	1 !	1
	!	0.00 !	0.00 !	0.00 !	5.26 !	5.26
	!	0.00 !	0.00 !	0.00 !	100.00 !	
	!	0.00 !	0.00 !	0.00 !	20.00 !	
Total		2	7	5	5	19
		10.53	36.84	26.32	26.32	100.00

CHI2 des tableaux de contingence croisant la variable de classe et les variables explicatives
 (triés par niveau de significativité)

Variable	Nbre de modalités	Effectif pondéré du tableau	Nbre de cases d'effectif <5	Nbre de cases d'effectif <5 (en %)	CHI2	Degrés de liberté	PROBA
PRESIDENT	4	19	20	100	10.97	12	0.5311

Caractérisation de chaque classe de la partition par les variables explicatives
 (triées par niveau de significativité)

CLASS5	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
1	PRESIDENT	5	1.41	3	0.7042
2	PRESIDENT	3	3.33	3	0.3430
3	PRESIDENT	8	2.18	3	0.5362
4	PRESIDENT	2	1.26	3	0.7393
5	PRESIDENT	1	2.80	3	0.4235

Caractérisation de la partition par les modalités des variables explicatives
(triées par niveau de significativité)

Modalité	Label	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
POMP	Pompidou	PRESIDENT	5	3.20	4	0.5247
BALL	Balladur	PRESIDENT	2	2.75	4	0.6005
MITT	Mitterrand	PRESIDENT	5	2.73	4	0.6046
GISC	Giscard d'Estain	PRESIDENT	7	2.30	4	0.6814

CAH sur un tableau individus x variables numériques : les vins français

La macro DESQUAL : description de la partition en 5 classes avec une variable supplémentaire qualitative
Caractérisation de chaque modalité de la variable CLASS5 par les modalités des variables explicatives
(triées par niveau de significativité)

CLASS5	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
1	sur-représentées	GISC	Giscard d'Estain	PRESIDENT	3	60.0	36.8	0.2366	0.71733
1	sous-représentées	POMP	Pompidou	PRESIDENT	1	20.0	26.3	0.6026	0.26008
		MITT	Mitterrand	PRESIDENT	1	20.0	26.3	0.6026	0.26008
		BALL	Balladur	PRESIDENT	0	0.0	10.5	0.5322	0.08071

CLASS5	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
2	sur-représentées	MITT	Mitterrand	PRESIDENT	2	66.7	26.3	0.1548	1.01607
		POMP	Pompidou	PRESIDENT	1	33.3	26.3	0.6244	-0.31694
2	sous-représentées	BALL	Balladur	PRESIDENT	0	0.0	10.5	0.7018	0.52945
		GISC	Giscard d'Estain	PRESIDENT	0	0.0	36.8	0.2270	-0.74864

CLASS5	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
3	sur-représentées	BALL	Balladur	PRESIDENT	2	25.0	10.5	0.1637	0.97919
		GISC	Giscard d'Estain	PRESIDENT	3	37.5	36.8	0.6634	-0.42168
3	sous-représentées	MITT	Mitterrand	PRESIDENT	2	25.0	26.3	0.6641	0.42364
		POMP	Pompidou	PRESIDENT	1	12.5	26.3	0.2668	-0.62261

CLASS5	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
4	sur-représentées	POMP	Pompidou	PRESIDENT	1	50.0	26.3	0.4678	0.08071
		GISC	Giscard d'Estain	PRESIDENT	1	50.0	36.8	0.6140	-0.28985
4	sous-représentées	BALL	Balladur	PRESIDENT	0	0.0	10.5	0.7953	0.82503
		MITT	Mitterrand	PRESIDENT	0	0.0	26.3	0.5322	0.08071

CLASS5	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
5	sur-représentées	POMP	Pompidou	PRESIDENT	1	100.0	26.3	0.2632	0.63364
5	sous-représentées	BALL	Balladur	PRESIDENT	0	0.0	10.5	0.8947	1.25212
		MITT	Mitterrand	PRESIDENT	0	0.0	26.3	0.7368	0.63364
		GISC	Giscard d'Estain	PRESIDENT	0	0.0	36.8	0.6316	0.33604

Exemple 2 : tableau de contingence (étrangers par CS)

```
LIBNAME COMPIL ...;
OPTIONS MSTORED SASMSTORE=COMPIL NODATE PAGESIZE=66;
FILENAME ENT ...;

DATA A; INFILE ENT(ETRANGER);
INPUT AGRI ARTC CSUP PINT EMPL SERV OQ ONQ OAGR RETR / INAC NAT $13.;
IF NAT="Français" THEN POND=0;ELSE POND=1;
LABEL AGRI="Agricult"
      ARTC="Art,com"
      CSUP="Cad.sup"
      PINT="Pr.inter"
      EMPL="Employés"
      SERV="Per.serv"
      OQ  ="Ouv.qual"
      ONQ ="Ouv.nqua"
      OAGR="Ouv.agri"
      RETR="Retrait,"
      INAC="Aut.inac";

PROC PRINT LABEL DATA=A;
ID NAT;
TITLE "Répartition des étrangers par catégorie socioprofessionnelle en 1982";
TITLE2 "Tableau de données";
RUN;

TITLE "CAH sur un tableau de contingence : les étrangers par CS en 1982";
TITLE2 "Les sorties de la macro CAHQQUAL";
%CAHQQUAL(DATA=A,
          VAR=AGRI--OAGR,
          ID=NAT,
          POIDS=POND,
          ARBRE=OUI,
          NOEUDS=MAX);

RUN;

TITLE2 "La macro PARTQUAL : partition en 3 classes";
%PARTQUAL(DATA=A,
          VAR=AGRI--OAGR,
          ID=NAT,
          POIDS=POND,
          NCL=3,
          AFFECT=OUI,
          OUTPART=PART3,
          LIBVAR=Cat_socioprofessionnelle,
          OBSSUP=OUI);

RUN;

PROC PRINT DATA=PART3;
TITLE2 "Table en sortie créée par la macro PARTQUAL";
RUN;

TITLE2 "La macro DESQUAL : description de la partition en 4 classes"
      " avec des variables supplémentaires";
%DESQUAL(DATA=PART3,
          VAR=RETR INAC,
          POIDS=POND,
          VARCLASS=CLASS3,
          LIBVAR=Cat_socioprof);
```

Répartition des étrangers par catégorie socioprofessionnelle en 1982
Tableau de données

NAT	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri	Retraité	Aut.inac	POND
Italiens	3176	15268	4228	13836	12952	6892	45960	39660	2668	68096	127572	1
Espagnols	1356	6608	2284	8464	13512	18112	39160	38736	9376	53936	135612	1
Portugais	420	5456	948	9128	30456	44972	111200	171680	11188	8080	373776	1
Aut.CEE	3592	6256	15872	14116	10108	2700	7120	4880	1052	30636	63676	1
Algériens	104	11708	2236	8560	33328	11220	84516	153336	1904	16224	481980	1
Marocains	212	3544	1400	2780	11524	6256	37140	83772	19212	2636	272832	1
Tunisiens	76	3044	1460	3112	7836	3712	20148	33092	2856	2072	113392	1
Aut.étrangers	1508	9380	22316	20540	34972	15276	46472	84612	3636	60884	382604	1
Français	1460460	1768524	1847128	3858160	5212188	776032	3647048	2855108	242728	7159616	21754420	0

CAH sur un tableau de contingence : les étrangers par CS en 1982
Les sorties de la macro CAHQVAL

```

*****
*      Caractéristiques de la classification      *
*****
*      Nombre de variables                      =    9      *
*      Nombre d'observations                    =    8      *
*      Nombre d'observations pondérées          =1.52E6    *
*      Nombre d'observations éliminées         =    1      *
*      Variable de pondération                  = POND      *
*      Variable identifiant                     = NAT      *
*      Nombre de noeuds édités                 = MAX      *
*      Méthode utilisée                         = WARD      *
*****

```

Liste des variables :
AGRI ARTC CSUP PINT EMPL SERV OQ ONQ OAGR

Tableau des noeuds de la hiérarchie

Inertie totale : 0.260254

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter 0/00 cum. dif !	Histogramme
CL1	1520192.0	CL2-Aut.C	452 452 . !*****	
CL2	1454496.0	CL4-CL3	219 671 233 !*****	
CL3	1071144.0	CL5-Maroc	105 775 114 !*****	
CL4	383352.0	Itali-Aut.,	87 862 18 !*****	
CL5	905304.0	CL6-CL7	81 943 5 !*****	
CL6	523056.0	Espag-Portu	49 992 32 !*****	
CL7	382248.0	Algér-Tunis	8 1000 41 !*	

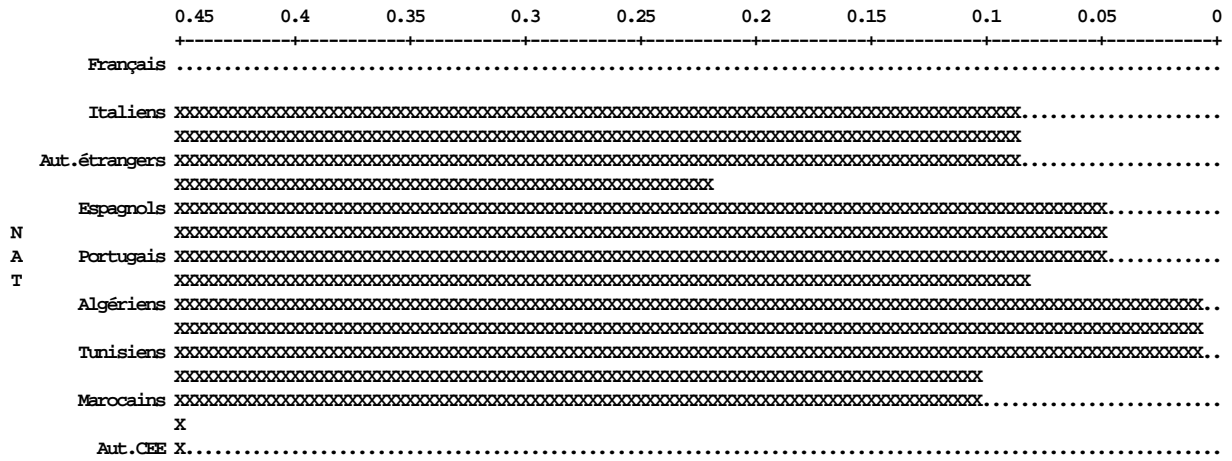
Décomposition de la distance entre les centres de gravité
des classes jointes
(Il s'agit aussi de la décomposition de la perte d'inertie interclasse)

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri
CL1	1520192	CL2-Aut.C	452	128	29	499	190	10	5	33	102	4
CL2	1454496	CL4-CL3	219	49	105	369	220	39	20	13	126	60
CL3	1071144	CL5-Maroc	105	0	6	0	15	20	111	44	38	766
CL4	383352	Itali-Aut.,	87	94	286	325	5	84	10	154	42	1
CL5	905304	CL6-CL7	81	9	41	3	1	38	635	5	125	145
CL6	523056	Espag-Portu	49	58	149	31	141	19	16	0	350	234
CL7	382248	Algér-Tunis	8	1	3	89	69	4	46	5	185	599

CAH sur un tableau de contingence : les étrangers par CS en 1982
 Les sorties de la macro CAHQVAL
 Arbre de classification

Ward's Minimum Variance Cluster Analysis

Semi-Partial R-Squared



CAH sur un tableau de contingence : les étrangers par CS en 1982
 La macro PARTQUAL : partition en 3 classes
 Partition en 3 classes : composition des classes

VARIABLE DE CLASSE : CLASS3

Individus non inclus dans une classe
 Français

Classe numéro 1
 Algériens Espagnols Marocains Portugais Tunisiens

Classe numéro 2
 Aut.étrangers Italiens

Classe numéro 3
 Aut.CEE

Partition en 3 classes : Décomposition de l'inertie totale

	CLASS3	Inertie	En 0/00 de l'inertie totale	Effectif absolu	Effectif pondéré (en 0/00)
Inertie totale		0.2603	1000	8	1000
Inertie inter		0.1746	671	.	.
	1	0.0329	127	5	705
	2	0.0291	112	2	252
	3	0.1125	432	1	43
Inertie intra		0.0857	329	.	.
	1	0.0631	243	5	705
	2	0.0226	87	2	252
	3	0.0000	0	1	43

Décomposition signée du RHO2 (distance entre le centre des classes et le centre de gravité du nuage) selon les variables de l'analyse
 tous les chiffres sont en millièmes, exceptés l'effectif absolu et le RHO2

Classe	Effectif absolu	Effectif pondéré	RHO2	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri
1	5	705	0.04676	-73	-76	-420	-215	-28	14	19	121	35
2	2	252	0.11542	36	124	334	220	46	-24	-9	-128	-79
3	1	43	2.60328	128	29	499	190	10	-5	-33	-102	-4

Décomposition signée de la distance entre les centres de gravité des classes de la partition
 tous les chiffres sont en millièmes
 exceptées la perte d'inertie inter et la distance

Classes	Distance	Perte d'inertie									
		inter	Agricult	Art,com	Cad.sup	Pr.inter	Employés	Per.serv	Ouv.qual	Ouv.nqua	Ouv.agri
2 - 1	0.30678	0.05697	49	105	369	220	39	-20	-13	-126	-60
3 - 1	3.33107	0.13564	121	33	492	194	12	-6	-31	-105	-6
3 - 2	1.70808	0.06302	154	14	521	173	5	-2	-40	-91	-0

Décomposition de l'inertie inter de la partition en 3 classes
 Selon les variables de l'analyse et les classes

TABLE OF CLASS3 BY IN_INTER

CLASS3	IN_INTER									
Frequency!										
Percent !										
Row Pct !										
Col Pct	!AGRI	!ARTIC	!CSUP	!PINT	!EMPL	!SERV	!OQ	!ONQ	!OAGR	! Total
1	! 0.0024 !	! 0.0025 !	! 0.0138 !	! 0.0071 !	! 0.0009 !	! 0.0005 !	! 0.0006 !	! 0.004 !	! 0.0011 !	! 0.0329
	! 1.38 !	! 1.43 !	! 7.93 !	! 4.05 !	! 0.53 !	! 0.27 !	! 0.36 !	! 2.28 !	! 0.66 !	! 18.88
	! 7.31 !	! 7.59 !	! 42.00 !	! 21.47 !	! 2.79 !	! 1.41 !	! 1.89 !	! 12.07 !	! 3.47 !	
	! 13.51 !	! 26.77 !	! 17.37 !	! 20.29 !	! 26.93 !	! 27.07 !	! 13.44 !	! 20.71 !	! 29.54 !	
2	! 0.001 !	! 0.0036 !	! 0.0097 !	! 0.0064 !	! 0.0013 !	! 0.0007 !	! 0.0003 !	! 0.0037 !	! 0.0023 !	! 0.0291
	! 0.60 !	! 2.06 !	! 5.57 !	! 3.67 !	! 0.77 !	! 0.39 !	! 0.15 !	! 2.13 !	! 1.32 !	! 16.67
	! 3.61 !	! 12.38 !	! 33.38 !	! 22.02 !	! 4.61 !	! 2.35 !	! 0.92 !	! 12.79 !	! 7.94 !	
	! 5.89 !	! 38.59 !	! 12.19 !	! 18.38 !	! 39.28 !	! 39.90 !	! 5.80 !	! 19.38 !	! 59.72 !	
3	! 0.0144 !	! 0.0032 !	! 0.0561 !	! 0.0214 !	! 0.0012 !	! 0.0006 !	! 0.0037 !	! 0.0115 !	! 0.0004 !	! 0.1125
	! 8.24 !	! 1.85 !	! 32.16 !	! 12.25 !	! 0.66 !	! 0.32 !	! 2.14 !	! 6.59 !	! 0.24 !	! 64.45
	! 12.78 !	! 2.88 !	! 49.89 !	! 19.00 !	! 1.02 !	! 0.50 !	! 3.32 !	! 10.23 !	! 0.37 !	
	! 80.61 !	! 34.65 !	! 70.44 !	! 61.32 !	! 33.79 !	! 33.03 !	! 80.76 !	! 59.92 !	! 10.74 !	
Total	0.01783	0.00934	0.07968	0.03486	0.00341	0.00172	0.00463	0.01921	0.00387	0.17456
	10.22	5.35	45.65	19.97	1.95	0.98	2.65	11.00	2.22	100.00

CAH sur un tableau de contingence : les étrangers par CS en 1982

La macro PARTQUAL : partition en 3 classes

Tableau(x) de contingence croisant la variable de classe et les variables analysées

----- Variable-Cat_socioprofess -----

TABLE OF CLASS3 BY MODALITE

CLASS3	MODALITE									
Frequency!										
Percent !										
Row Pct !										
Col Pct	!Agricult	!Art,com	!Cad.sup	!Employés	!Ouv.agri	!Ouv.nqua	!Ouv.qual	!Per.serv	!Pr.inter	! Total
1	! 2168 !	! 30360 !	! 8328 !	! 96656 !	! 44536 !	! 480616 !	! 292164 !	! 84272 !	! 32044 !	!1071144
	! 0.14 !	! 2.00 !	! 0.55 !	! 6.36 !	! 2.93 !	! 31.62 !	! 19.22 !	! 5.54 !	! 2.11 !	! 70.46
	! 0.20 !	! 2.83 !	! 0.78 !	! 9.02 !	! 4.16 !	! 44.87 !	! 27.28 !	! 7.87 !	! 2.99 !	
	! 20.76 !	! 49.56 !	! 16.41 !	! 62.48 !	! 85.82 !	! 78.82 !	! 74.59 !	! 77.21 !	! 39.79 !	
2	! 4684 !	! 24648 !	! 26544 !	! 47924 !	! 6304 !	! 124272 !	! 92432 !	! 22168 !	! 34376 !	! 383352
	! 0.31 !	! 1.62 !	! 1.75 !	! 3.15 !	! 0.41 !	! 8.17 !	! 6.08 !	! 1.46 !	! 2.26 !	! 25.22
	! 1.22 !	! 6.43 !	! 6.92 !	! 12.50 !	! 1.64 !	! 32.42 !	! 24.11 !	! 5.78 !	! 8.97 !	
	! 44.85 !	! 40.23 !	! 52.31 !	! 30.98 !	! 12.15 !	! 20.38 !	! 23.60 !	! 20.31 !	! 42.68 !	
3	! 3592 !	! 6256 !	! 15872 !	! 10108 !	! 1052 !	! 4880 !	! 7120 !	! 2700 !	! 14116 !	! 65696
	! 0.24 !	! 0.41 !	! 1.04 !	! 0.66 !	! 0.07 !	! 0.32 !	! 0.47 !	! 0.18 !	! 0.93 !	! 4.32
	! 5.47 !	! 9.52 !	! 24.16 !	! 15.39 !	! 1.60 !	! 7.43 !	! 10.84 !	! 4.11 !	! 21.49 !	
	! 34.39 !	! 10.21 !	! 31.28 !	! 6.53 !	! 2.03 !	! 0.80 !	! 1.82 !	! 2.47 !	! 17.53 !	
Total	10444	61264	50744	154688	51892	609768	391716	109140	80536	1520192
	0.69	4.03	3.34	10.18	3.41	40.11	25.77	7.18	5.30	100.00

CHI2 des tableaux de contingence croisant la variable de classe et les variables analysées
(triés par niveau de significativité)

Variable	Nbre de modalités	Effectif pondéré du tableau	Nbre de cases d'effectif <5	Nbre de cases d'effectif <5 (en %)	CHI2	Degrés de liberté	PROBA
Cat_socioprofess	9	1520192	0	0	265358	16	0.0000

Caractérisation de chaque classe de la partition par les variables analysées
(triées par niveau de significativité)

CLASS3	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
1	Cat_socioprofess	1071144	50086.6	8	0.0000
2	Cat_socioprofess	383352	44246.5	8	0.0000
3	Cat_socioprofess	65696	171025	8	0.0000

CAH sur un tableau de contingence : les étrangers par CS en 1982
 La macro PARTQUAL : partition en 3 classes
 Caractérisation de la partition par les modalités des variables analysées
 (triées par niveau de significativité)

Modalité	Label	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
CSUP	Cad.sup	Cat_socioprofess	50744	121136	2	0.0000
PINT	Pr.inter	Cat_socioprofess	80536	52997.5	2	0.0000
ONQ	Ouv.nqua	Cat_socioprofess	609768	29198.8	2	0.0000
AGRI	Agricult	Cat_socioprofess	10444	27111.9	2	0.0000
ARTC	Art,com	Cat_socioprofess	61264	14195.1	2	0.0000
OQ	Ouv.qual	Cat_socioprofess	391716	7036.65	2	0.0000
OAGR	Ouv.agri	Cat_socioprofess	51892	5885.03	2	0.0000
EMPL	Employés	Cat_socioprofess	154688	5187.45	2	0.0000
SERV	Per.serv	Cat_socioprofess	109140	2610.25	2	0.0000

Caractérisation de chaque classe de la partition par les modalités des variables analysées
 (triées par niveau de significativité)

CLASS3	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
1	sur-représentées	ONQ	Ouv.nqua	Cat_socioprofess	480616	44.9	40.1	0.0000	188.069
		OAGR	Ouv.agri	Cat_socioprofess	44536	4.2	3.4	0.0000	86.548
		OQ	Ouv.qual	Cat_socioprofess	292164	27.3	25.8	0.0000	66.623
		SERV	Per.serv	Cat_socioprofess	84272	7.9	7.2	0.0000	52.575
1	sous-représentées	EMPL	Employés	Cat_socioprofess	96656	9.0	10.2	0.0000	-70.395
		AGRI	Agricult	Cat_socioprofess	2168	0.2	0.7	0.0000	-102.417
		ARTC	Art,com	Cat_socioprofess	30360	2.8	4.0	0.0000	-108.148
		PINT	Pr.inter	Cat_socioprofess	32044	3.0	5.3	0.0000	-180.342
		CSUP	Cad.sup	Cat_socioprofess	8328	0.8	3.3	0.0000	-250.008

CLASS3	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
2	sur-représentées	CSUP	Cad.sup	Cat_socioprofess	26544	6.9	3.3	0.0000	128.040
		PINT	Pr.inter	Cat_socioprofess	34376	9.0	5.3	0.0000	108.218
		ARTC	Art,com	Cat_socioprofess	24648	6.4	4.0	0.0000	81.227
		EMPL	Employés	Cat_socioprofess	47924	12.5	10.2	0.0000	53.442
		AGRI	Agricult	Cat_socioprofess	4684	1.2	0.7	0.0000	42.325
2	sous-représentées	OQ	Ouv.qual	Cat_socioprofess	92432	24.1	25.8	0.0000	-27.303
		SERV	Per.serv	Cat_socioprofess	22168	5.8	7.2	0.0000	-40.029
		OAGR	Ouv.agri	Cat_socioprofess	6304	1.6	3.4	0.0000	-78.118
		ONQ	Ouv.nqua	Cat_socioprofess	124272	32.4	40.1	0.0000	-113.792

CLASS3	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
3	sur-représentées	CSUP	Cad.sup	Cat_socioprofess	15872	24.2	3.3	0.0000	175.885
		PINT	Pr.inter	Cat_socioprofess	14116	21.5	5.3	0.0000	132.125
		AGRI	Agricult	Cat_socioprofess	3592	5.5	0.7	0.0000	83.526
		ARTC	Art,com	Cat_socioprofess	6256	9.5	4.0	0.0000	59.038
		EMPL	Employés	Cat_socioprofess	10108	15.4	10.2	0.0000	41.242
3	sous-représentées	OAGR	Ouv.agri	Cat_socioprofess	1052	1.6	3.4	0.0000	-30.630
		SERV	Per.serv	Cat_socioprofess	2700	4.1	7.2	0.0000	-34.947
		OQ	Ouv.qual	Cat_socioprofess	7120	10.8	25.8	0.0000	-101.882
		ONQ	Ouv.nqua	Cat_socioprofess	4880	7.4	40.1	0.0000	-206.661

CAH sur un tableau de contingence : les étrangers par CS en 1982
 La macro DESQUAL : description de la partition en 4 classes avec des variables supplémentaires
 Tableau(x) de contingence croisant la variable de classe et les variables explicatives

----- Variable=Cat_socioprof -----

TABLE OF CLASS3 BY MODALITE

CLASS3	MODALITE		
Frequency!			
Percent !			
Row Pct !			
Col Pct !	!Aut.inac!	Retraité!	Total
1	!1377592 !	82948 !	!1460540
	! 62.79 !	3.78 !	! 66.57
	! 94.32 !	5.68 !	
	! 70.59 !	34.20 !	
2	! 510176 !	128980 !	! 639156
	! 23.25 !	5.88 !	! 29.13
	! 79.82 !	20.18 !	
	! 26.14 !	53.17 !	
3	! 63676 !	30636 !	! 94312
	! 2.90 !	1.40 !	! 4.30
	! 67.52 !	32.48 !	
	! 3.26 !	12.63 !	
Total	1951444	242564	2194008
	88.94	11.06	100.00

CHI2 des tableaux de contingence croisant la variable de classe et les variables explicatives
 (triés par niveau de significativité)

Variable	Nbre de modalités	Effectif pondéré du tableau	Nbre de cases d'effectif <5	Nbre de cases d'effectif <5 (en %)	CHI2	Degrés de liberté	PROBA
Cat_socioprof	2	2194008	0	0	141081	2	0.0000

Caractérisation de chaque classe de la partition par les variables explicatives
 (triées par niveau de significativité)

CLASS3	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
1	Cat_socioprof	1460540	42934.2	1	0.0000
2	Cat_socioprof	639156	54109.1	1	0.0000
3	Cat_socioprof	94312	44037.3	1	0.0000

Caractérisation de la partition par les modalités des variables explicatives
 (triées par niveau de significativité)

Modalité	Label	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
RETR	Retraité	Cat_socioprof	242564	125483	2	0.0000
INAC	Aut.inac	Cat_socioprof	1951444	15597.5	2	0.0000

CAH sur un tableau de contingence : les étrangers par CS en 1982
 La macro DESQUAL : description de la partition en 4 classes avec des variables supplémentaires
 Caractérisation de chaque modalité de la variable CLASS3 par les modalités des variables explicatives
 (triées par niveau de significativité)

CLASS3	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
1	sur-représentées	INAC	Aut.inac	Cat_socioprof	1377592	94.3	88.9	0.0000	338.569
1	sous-représentées	RETR	Retraité	Cat_socioprof	82948	5.7	11.1	0.0000	-338.565

CLASS3	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
2	sur-représentées	RETR	Retraité	Cat_socioprof	128980	20.2	11.1	0.0000	257.131
2	sous-représentées	INAC	Aut.inac	Cat_socioprof	510176	79.8	88.9	0.0000	-257.127

CLASS3	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
3	sur-représentées	RETR	Retraité	Cat_socioprof	30636	32.5	11.1	0.0000	167.867
3	sous-représentées	INAC	Aut.inac	Cat_socioprof	63676	67.5	88.9	0.0000	-167.860

Exemple 3 : tableaux de contingence juxtaposés (étrangers par CS et activité)

```
LIBNAME COMPIL ...;
OPTIONS SASMSTORE=COMPIL MSTORED NODATE PAGESIZE=66;

PROC PRINT DATA=ETR;
ID NAT;
TITLE "Répartition de la population active par nationalité (en ligne)"
      " et par CS et activité économique (en colonne) en 1990";
RUN;

TITLE "CAH sur deux tableaux de contingence juxtaposés : "
      " nationalité x CS et nationalité x activité économique";
TITLE2 "Les sorties de la macro CAHQVAL";
%CAHQVAL (DATA=ETR,
          VAR=AGRICULT--SERV_NMA,
          ID=NAT,
          POIDS=POND,
          NOEUDS=MAX);
RUN;

TITLE2 "La macro PARTQUAL : partition en 4 classes";
%PARTQUAL(DATA=ETR,
          VAR=AGRICULT--SERV_NMA,
          ID=NAT,
          POIDS=POND,
          NCL=4,
          LISTE=OUI,
          AFFECT=OUI,
          OUTPART=PART4,
          LIBVAR=Cat_socioprofessionnelle_Activité_économique,
          NEMODA=18 14);
RUN;
```

Répartition de la population active par nationalité (en ligne) et par CS et activité économique (en colonne) en 1990

	A	A	C	P	C	C	I	I		E	E	E	P	O		
	G	R	H	R	A	A	N	N	T	C	M	M	M	E	U	
	R	T	O	E	O	D	D	T	T	E	O	P	P	P	R	V
	I	I	M	F	F	—	—	—	C	N	—	—	—	—	—	—
	C	S	M	—	—	P	E	P	E	H	T	P	E	C	S	Q
N	U	A	E	E	L	U	N	U	N	N	R	U	N	O	E	U
A	L	N	R	N	I	B	T	B	T	I	E	B	T	M	R	A
T	T	S	C	T	B	L	R	L	R	C	M	L	R	M	V	L
Espagnol	1036	4104	1736	364	164	1464	2176	1800	2672	1677	2700	2852	4120	2720	14140	26760
Italiens	1296	8444	3744	1100	340	1868	4096	1644	4344	2008	4920	2608	4024	2580	6168	32808
Portugai	796	15008	2804	716	228	932	2396	2080	5368	3300	8020	10864	11996	11420	53332	122120
Algérien	120	5292	10292	408	516	3072	2964	5540	5292	2840	2644	9808	12208	7644	16352	82440
Marocain	596	2640	4660	180	424	3156	1752	3744	2812	1604	1236	5032	6336	4716	13328	51373
Tunisien	80	2748	2708	244	284	1204	1360	1356	2036	996	924	2288	2764	2484	5636	24573
Afr-noir	40	848	2012	112	356	2928	1960	3436	2116	1380	392	4372	4733	1848	6384	12848
Indochin	100	1120	2088	140	140	424	900	692	1128	1176	328	1212	2216	2252	4316	12392
Turcs	512	3464	968	144	56	356	352	768	608	272	404	944	1336	1360	1584	19016
Autres	4088	7374	9752	3180	3996	24288	29344	16858	17228	5364	2844	7656	18948	7584	21473	38452
Fra-nais	992525	763730	721734	163422	291226	949108	1269479	1901796	1304969	719389	528714	2296418	2199715	887633	993272	3697190
Fra-acqu	11748	35692	33208	6676	13380	33329	43044	45225	43652	22424	21252	67616	75148	36508	52852	153258

	O	A		B	B	B	B	C	T	S	L	A	O	S		
	U	G	E	I	I	I	A	O	R	E	O	S	R	E		
	O	V	R	I	E	E	E	T	M	A	R	C	S	G	R	
	U	—	I	—	E	N	N	I	M	N	V	A	U	—	V	
	V	A	C	A	R	—	—	M	E	S	—	T	R	F	—	
N	—	G	O	—	G	I	E	C	E	R	P	M	I	A	I	N
A	N	R	L	A	I	N	Q	O	N	C	O	A	O	N	N	M
T	Q	I	E	—	E	T	U	M	T	E	R	R	N	C	A	A
Espagnol	19612	4764	5100	1700	200	6580	6665	5284	16304	8248	2308	19584	476	336	660	10324
Italiens	18908	928	1960	1476	688	9448	8212	5812	24416	9168	2620	18556	260	316	728	6556
Portugai	127700	9080	9048	8828	660	29032	21456	24768	105360	27596	8840	79600	1696	812	1588	32976
Algérien	94152	1744	1208	3832	1368	20780	17180	10492	43656	20104	10808	50528	804	380	716	17572
Marocain	72048	18892	16060	3160	2320	14256	18256	6992	28612	13104	5361	29900	432	208	408	12684
Tunisien	26249	2484	2068	1472	252	4472	4048	3257	15280	7228	2597	16200	168	92	172	4744
Afr-noir	26115	248	168	1136	232	3224	5172	3443	3848	5284	3280	20964	184	173	428	8320
Indochin	16168	540	400	1320	48	2972	4152	7212	1272	5532	1380	10628	56	84	148	1584
Turcs	33448	2464	2252	1424	184	6560	4828	9696	14800	3288	868	5972	44	36	52	1952
Autres	38351	1760	5548	3084	1176	9580	14416	19347	15253	24724	8325	76501	556	920	3004	37634
Fra-nais	2486527	232000	1207870	589535	251060	1144419	1448127	1119574	1309571	2465316	1343106	4699903	70212	156476	444628	4010307
Fra-acqu	108787	7144	17916	14500	6144	45800	55744	47334	68884	94024	33250	185713	2764	4216	10912	118769

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
Les sorties de la macro CAHQUAL

```

*****
*      Caractéristiques de la classification      *
*****
*      Nombre de variables          =      32      *
*
*      Nombre d'observations        =      10      *
*      Nombre d'observations pondérées = 2.87E6   *
*      Nombre d'observations éliminées =      2      *
*
*      Variable de pondération      = POND        *
*
*      Variable identifiant         = NAT         *
*
*      Nombre de noeuds édités      = MAX        *
*
*      Méthode utilisée             = WARD       *
*****

```

Liste des variables :

```

AGRICULT ARTISANS COMMERC CHEF_ENT PROF_LIB CAD_PUBL CAD_ENTR INT_PUBL INT_ENTR TECHNIC CONTREM
EMP_PUBL EMP_ENTR EMP_COMM
PER_SERV OUV_QUAL OUV_NQ OUV_AGRIC AGRICOLE I_A_A ENERGIE BIEN_INT BIEN_EQU BIEN_COM BATIMENT COMMERCE
TRANSPOR SERV_MAR LOCATION
ASSURANC ORG_FINA SERV_NMA

```

Tableau des noeuds de la hiérarchie

Inertie totale : 0.199560

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter 0/00 cum. dif !	Histogramme
CL1	2873324.0	CL2-Autre	520 520 . !*****	*****
CL2	2394716.0	CL3-Maroc	144 664 377 !*****	*****
CL3	2048434.0	CL4-CL6	103 767 41 !*****	*****
CL4	1231106.0	CL7-CL5	66 833 38 !*****	*****
CL5	860432.0	Portu-Turcs	55 887 11 !*****	*****
CL6	817328.0	CL9-CL8	43 931 11 !****	*****
CL7	370674.0	Espag-Itali	32 963 11 !****	*****
CL8	212104.0	Afr-n-Indoc	28 991 5 !***	*****
CL9	605224.0	Algér-Tunis	9 1000 18 !*	*****

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
 Les sorties de la macro CAHQVAL
 Décomposition de la distance entre les centres de gravité des classes jointes
 (Il s'agit aussi de la décomposition de la perte d'inertie interclasse)

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter	AGRICULT	ARTISANS	COMMERCE	CHEF_ENT	PROF_LIB	CAD_PUEL	CAD_ENTR	INT_PUEL	INT_ENTR
CL1	2873324	CL2-Autre	520	19	1	5	16	31	190	235	71	55
CL2	2394716	CL3-Maroc	144	0	31	0	2	0	3	2	2	3
CL3	2048434	CL4-CL6	103	18	78	105	2	3	19	1	44	3
CL4	1231106	CL7-CL5	66	52	60	53	26	3	28	79	19	65
CL5	860432	Portu-Turcs	55	11	14	4	0	0	1	0	3	1
CL6	817328	CL9-CL8	43	0	6	2	0	3	35	15	31	4
CL7	370674	Espag-Itali	32	1	124	34	30	2	1	25	1	19
CL8	212104	Afr-n-Indoc	28	1	8	19	1	2	74	4	84	2
CL9	605224	Algér-Tunis	9	1	71	15	6	7	5	12	9	11

Noeud	TECHNIC	CONTREM	EMP_PUEL	EMP_ENTR	EMP_COMM	PER_SERV	OUV_QUAL	OUV_NQ	OUV_AGR	AGRICOLE	I_A_A	ENERGIE	BIEN_INT
CL1	4	1	0	20	0	1	58	83	16	2	2	0	15
CL2	2	17	1	1	1	15	5	21	460	298	1	35	0
CL3	4	59	29	31	4	75	0	69	34	57	2	3	3
CL4	32	94	0	12	0	11	1	291	2	16	8	6	1
CL5	2	22	9	4	4	230	1	226	15	9	0	1	21
CL6	29	5	15	17	3	25	143	0	5	3	5	5	26
CL7	1	65	2	1	1	223	15	4	160	110	2	12	21
CL8	5	0	75	15	31	0	48	3	4	2	16	2	9
CL9	2	1	33	42	1	7	4	46	257	191	9	12	101

Noeud	BIEN_EQU	BIEN_COM	BATIMENT	COMMERCE	TRANSPOR	SERV_MAR	LOCATION	ASSURANC	ORG_FIN	SERV_NMA
CL1	2	3	78	3	0	35	0	1	9	42
CL2	31	21	28	2	0	15	1	1	1	1
CL3	4	13	243	11	54	29	1	1	1	1
CL4	17	21	65	20	2	1	0	4	8	4
CL5	11	221	13	7	5	96	8	2	4	56
CL6	14	142	361	5	0	82	1	1	4	13
CL7	4	0	69	0	0	7	5	0	0	60
CL8	7	330	8	44	17	39	1	0	3	146
CL9	43	0	37	25	33	4	4	1	1	10

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
 La macro PARTQUAL : partition en 4 classes
 Partition en 4 classes : composition des classes

VARIABLE DE CLASSE : CLASS4

Individus non inclus dans une classe
 Fra-acqu Fra-nais

Classe numéro 1
 Afr-noir Algérien Indochin Tunisien

Classe numéro 2
 Espagnol Italiens Portugai Turcs

Classe numéro 3
 Marocain

Classe numéro 4
 Autres

Partition en 4 classes : Décomposition de l'inertie totale

	CLASS4	Inertie	En 0/00 de l'inertie totale	Effectif absolu	Effectif pondéré (en 0/00)
Inertie totale		0.1996	1000	10	1000
Inertie inter		0.1531	767	.	.
	1	0.0134	67	4	284
	2	0.0239	120	4	428
	3	0.0292	146	1	121
	4	0.0866	434	1	167
Inertie intra		0.0465	233	.	.
	1	0.0161	80	4	284
	2	0.0304	152	4	428
	3	0.0000	0	1	121
	4	0.0000	0	1	167

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
 La macro PARTQUAL : partition en 4 classes

Décomposition signée du RHO2 (distance entre le centre des classes et le centre de gravité du nuage)
 selon les variables de l'analyse

tous les chiffres sont en millièmes, exceptés l'effectif absolu et le RHO2

Classe	Effectif absolu	Effectif pondéré	RHO2	AGRICULT	ARTISANS	COMMERCE	CHEF_ENT	PROF_LIB	CAD_FUEL	CAD_ENIR	INT_FUEL	INT_ENIR
1	4	284	0.04711	-48	-36	68	-13	-4	-31	-76	0	-7
2	4	428	0.05587	-0	56	-56	-1	-21	-131	-91	-89	-25
3	1	121	0.24233	-2	-24	-0	-6	-2	-7	-33	-2	-14
4	1	167	0.51967	19	-1	5	16	31	190	235	71	55

Classe	TECHNIC	CONTREM	EMP_FUEL	EMP_ENIR	EMP_COMM	PER_SERV	OUV_QUAL	OUV_NQ	OUV_AGR	AGRICOLE	I_A_A	ENERGIE	BIEN_INT
1	1	-26	33	8	5	-40	31	154	-110	-155	-0	-0	1
2	-4	43	-7	-31	-1	47	31	-1	-1	-0	3	-8	11
3	-4	-12	-1	-5	-1	-10	0	48	434	262	-0	30	2
4	4	-1	-0	20	0	-1	-58	-83	-16	-2	-2	-0	-15

Classe	BIEN_EQU	BIEN_COM	BATIMENT	COMMERCE	TRANSPOR	SERV_MAR	LOCATION	ASSURANC	ORG_FINA	SERV_NMA
1	1	-8	-52	6	47	7	-0	-1	-7	-24
2	-4	7	264	-7	-20	-30	2	0	-1	-8
3	31	-22	-5	-3	-0	-28	-0	-1	-4	-9
4	-2	3	-78	3	0	35	-0	1	9	42

Décomposition signée de la distance entre les centres de gravité des classes de la partition
tous les chiffres sont en millièmes
exceptées la perte d'inertie inter et la distance

Classes	Distance	Perte d'inertie		AGRICULT	ARTISANS	COMMERC	CHEF_ENT	PROF_LIB	CAD_PUEL	CAD_ENTR	INT_PUEL	INT_ENTR
		inter										
2 - 1	0.12032	0.02057		18	78	-105	2	-3	-19	-1	-44	-3
3 - 1	0.32266	0.02731		2	-4	-12	-0	-0	-0	-3	-2	-5
4 - 1	0.65908	0.06924		33	1	-0	20	31	188	254	54	53
3 - 2	0.29662	0.02790		-2	-59	8	-3	1	7	-1	8	-1
4 - 2	0.83449	0.10009		12	-6	14	12	31	192	212	82	51
4 - 3	1.06004	0.07412		15	3	3	15	21	118	181	43	48

Classes	TECHNIC	CONTREM	EMP_PUEL	EMP_ENTR	EMP_COMM	PER_SERV	OUV_QUAL	OUV_NQ	OUV_AGRI	AGRICOLE	I_A_A	ENERGIE
2 - 1	-4	59	-29	-31	-4	75	0	-69	34	57	2	-3
3 - 1	-4	-1	-9	-9	-3	-0	-4	2	487	353	-0	24
4 - 1	2	0	-3	10	-0	1	-68	-130	-1	5	-1	0
3 - 2	-1	-36	0	0	-0	-35	-5	43	371	219	-1	38
4 - 2	5	-7	0	25	0	-6	-56	-49	-9	-1	-3	1
4 - 3	6	1	0	18	0	1	-30	-94	-164	-75	-1	-7

Classes	BIEN_INT	BIEN_EQU	BIEN_COM	BATIMENT	COMMERCE	TRANSFOR	SERV_MAR	LOCATION	ASSURANC	ORG_FINA	SERV_NMA
2 - 1	3	-4	13	243	-11	-54	-29	1	1	1	1
3 - 1	1	20	-9	1	-6	-8	-31	-0	-0	-0	-1
4 - 1	-14	-2	5	-35	1	-2	21	-0	1	11	50
3 - 2	-0	35	-28	-83	-0	3	-6	-1	-1	-2	-2
4 - 2	-15	-0	0	-125	4	2	37	-1	1	6	34
4 - 3	-11	-13	12	-26	4	0	45	-0	1	9	36

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
 La macro PARTQUAL : partition en 4 classes
 Tableau(x) de contingence croisant la variable de classe et les variables analysées

----- Variable=Cat_socioprofess -----

TABLE OF CLASS4 BY MODALITE

CLASS4	MODALITE										Total
Frequency!											
Percent !											
Row Pct !											
Col Pct	!AGRICULT!	!ARTISANS!	!CAD_ENTR!	!CAD_PUBL!	!CHEF_ENT!	!COMMERC!	!CONIREM!	!EMP_COMM!	!EMP_ENTR!		
1	340	10008	7184	7628	904	17100	4288	14228	21921	463206	
	0.02	0.64	0.46	0.49	0.06	1.09	0.27	0.91	1.40	29.52	
	0.07	2.16	1.55	1.65	0.20	3.69	0.93	3.07	4.73		
	3.92	19.61	15.19	19.22	13.72	41.95	17.57	31.90	31.92		
2	3640	31020	9020	4620	2324	9252	16044	18080	21476	652905	
	0.23	1.98	0.57	0.29	0.15	0.59	1.02	1.15	1.37	41.61	
	0.56	4.75	1.38	0.71	0.36	1.42	2.46	2.77	3.29		
	42.01	60.77	19.07	11.64	35.28	22.70	65.72	40.53	31.27		
3	596	2640	1752	3156	180	4660	1236	4716	6336	194529	
	0.04	0.17	0.11	0.20	0.01	0.30	0.08	0.30	0.40	12.40	
	0.31	1.36	0.90	1.62	0.09	2.40	0.64	2.42	3.26		
	6.88	5.17	3.70	7.95	2.73	11.43	5.06	10.57	9.23		
4	4088	7374	29344	24288	3180	9752	2844	7584	18948	258540	
	0.26	0.47	1.87	1.55	0.20	0.62	0.18	0.48	1.21	16.48	
	1.58	2.85	11.35	9.39	1.23	3.77	1.10	2.93	7.33		
	47.18	14.45	62.04	61.19	48.27	23.92	11.65	17.00	27.59		
Total	8664	51042	47300	39692	6588	40764	24412	44608	68681	1569180	
	0.55	3.25	3.01	2.53	0.42	2.60	1.56	2.84	4.38	100.00	

(Continued)

TABLE OF CLASS4 BY MODALITE

CLASS4	MODALITE										Total
Frequency!											
Percent !											
Row Pct !											
Col Pct	!EMP_PUBL!	!INT_ENTR!	!INT_PUBL!	!OUV_AGRI!	!OUV_NQ!	!OUV_QUAL!	!PER_SERV!	!PROF_LIB!	!TECHNIC!		
1	17680	10572	11024	5016	162684	132253	32688	1296	6392	463206	
	1.13	0.67	0.70	0.32	10.37	8.43	2.08	0.08	0.41	29.52	
	3.82	2.28	2.38	1.08	35.12	28.55	7.06	0.28	1.38		
	37.11	24.25	29.07	11.69	34.41	31.28	22.90	19.93	31.00		
2	17268	12992	6292	17236	199668	200704	75224	788	7257	652905	
	1.10	0.83	0.40	1.10	12.72	12.79	4.79	0.05	0.46	41.61	
	2.64	1.99	0.96	2.64	30.58	30.74	11.52	0.12	1.11		
	36.25	29.80	16.59	40.17	42.24	47.47	52.71	12.12	35.20		
3	5032	2812	3744	18892	72048	51373	13328	424	1604	194529	
	0.32	0.18	0.24	1.20	4.59	3.27	0.85	0.03	0.10	12.40	
	2.59	1.45	1.92	9.71	37.04	26.41	6.85	0.22	0.82		
	10.56	6.45	9.87	44.03	15.24	12.15	9.34	6.52	7.78		
4	7656	17228	16858	1760	38351	38452	21473	3996	5364	258540	
	0.49	1.10	1.07	0.11	2.44	2.45	1.37	0.25	0.34	16.48	
	2.96	6.66	6.52	0.68	14.83	14.87	8.31	1.55	2.07		
	16.07	39.51	44.46	4.10	8.11	9.09	15.05	61.44	26.02		
Total	47636	43604	37918	42904	472751	422782	142713	6504	20617	1569180	
	3.04	2.78	2.42	2.73	30.13	26.94	9.09	0.41	1.31	100.00	

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
 La macro PARTQUAL : partition en 4 classes
 Tableau(x) de contingence croisant la variable de classe et les variables analysées

----- Variable=Activité_économi -----

TABLE OF CLASS4 BY MODALITE

CLASS4	MODALITE								Total
Frequency!									
Percent !									
Row Pct !									
Col Pct	!AGRICOLE!	!ASSURANC!	!BATIMENT!	!BIEN_COM!	!BIEN_EQU!	!BIEN_INT!	!COMMERCE!		Total
1	3844	729	64056	24404	30552	31448	38148		354122
	0.29	0.06	4.91	1.87	2.34	2.41	2.93		27.15
	1.09	0.21	18.09	6.89	8.63	8.88	10.77		
	8.77	21.72	23.83	25.34	29.27	29.42	30.70		
2	18360	1500	160880	45560	41161	51620	48300		578201
	1.41	0.12	12.34	3.49	3.16	3.96	3.70		44.34
	3.18	0.26	27.82	7.88	7.12	8.93	8.35		
	41.91	44.68	59.85	47.31	39.43	48.29	38.87		
3	16060	208	28612	6992	18256	14256	13104		151753
	1.23	0.02	2.19	0.54	1.40	1.09	1.00		11.64
	10.58	0.14	18.85	4.61	12.03	9.39	8.64		
	36.66	6.20	10.64	7.26	17.49	13.34	10.54		
4	5548	920	15253	19347	14416	9580	24724		220068
	0.43	0.07	1.17	1.48	1.11	0.73	1.90		16.87
	2.52	0.42	6.93	8.79	6.55	4.35	11.23		
	12.66	27.41	5.67	20.09	13.81	8.96	19.89		
Total	43812	3357	268801	96303	104385	106904	124276		1304144
	3.36	0.26	20.61	7.38	8.00	8.20	9.53		100.00

(Continued)

TABLE OF CLASS4 BY MODALITE

CLASS4	MODALITE							Total
Frequency!								
Percent !								
Row Pct !								
Col Pct	!ENERGIE !	!I_A_A	!LOCATION!	!ORG_FINA!	!SERV_MAR!	!SERV_NMA!	!TRANSPOR!	Total
1	1900	7760	1212	1464	98320	32220	18065	354122
	0.15	0.60	0.09	0.11	7.54	2.47	1.39	27.15
	0.54	2.19	0.34	0.41	27.76	9.10	5.10	
	26.66	28.29	25.92	18.52	29.94	23.98	38.94	
2	1732	13428	2476	3028	123712	51808	14636	578201
	0.13	1.03	0.19	0.23	9.49	3.97	1.12	44.34
	0.30	2.32	0.43	0.52	21.40	8.96	2.53	
	24.30	48.95	52.95	38.31	37.67	38.56	31.55	
3	2320	3160	432	408	29900	12684	5361	151753
	0.18	0.24	0.03	0.03	2.29	0.97	0.41	11.64
	1.53	2.08	0.28	0.27	19.70	8.36	3.53	
	32.55	11.52	9.24	5.16	9.10	9.44	11.56	
4	1176	3084	556	3004	76501	37634	8325	220068
	0.09	0.24	0.04	0.23	5.87	2.89	0.64	16.87
	0.53	1.40	0.25	1.37	34.76	17.10	3.78	
	16.50	11.24	11.89	38.01	23.29	28.01	17.95	
Total	7128	27432	4676	7904	328433	134346	46387	1304144
	0.55	2.10	0.36	0.61	25.18	10.30	3.56	100.00

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique
 La macro PARTQUAL : partition en 4 classes
 CHI2 des tableaux de contingence croisant la variable de classe et les variables analysées
 (triés par niveau de significativité)

Variable	Nbre de modalités	Effectif pondéré du tableau	Nbre de cases d'effectif <5	Nbre de cases d'effectif <5 (en %)	CHI2	Degrés de liberté	PROBA
Activité_économi	14	1304144	0	0	116610	39	0.0000
Cat_socioprofess	18	1569180	0	0	321684	51	0.0000

Caractérisation de chaque classe de la partition par les variables analysées
 (triées par niveau de significativité)

CLASS4	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
1	Activité_économi	354122	11686.9	13	0.0000
	Cat_socioprofess	463206	25007.0	17	0.0000
2	Activité_économi	578201	23560.5	13	0.0000
	Cat_socioprofess	652905	44448.4	17	0.0000
3	Activité_économi	151753	34286.9	13	0.0000
	Cat_socioprofess	194529	49063.4	17	0.0000
4	Activité_économi	220068	47075.5	13	0.0000
	Cat_socioprofess	258540	203166	17	0.0000

Caractérisation de la partition par les modalités des variables analysées
 (triées par niveau de significativité)

Modalité	Label	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
CAD_ENVIR	CAD_ENVIR	Cat_socioprofess	47300	71543.7	3	0.0000
CAD_FUEL	CAD_FUEL	Cat_socioprofess	39692	58794.7	3	0.0000
OUV_AGRIC	OUV_AGRIC	Cat_socioprofess	42904	43266.2	3	0.0000
BATIMENT	BATIMENT	Activité_économi	268801	35897.4	3	0.0000
AGRICOLE	AGRICOLE	Activité_économi	43812	29540.0	3	0.0000
OUV_NQ	OUV_NQ	Cat_socioprofess	472751	27034.1	3	0.0000
INT_FUEL	INT_FUEL	Cat_socioprofess	37918	23920.5	3	0.0000
OUV_QUAL	OUV_QUAL	Cat_socioprofess	422782	17939.8	3	0.0000
INT_ENVIR	INT_ENVIR	Cat_socioprofess	43604	17158.9	3	0.0000
SERV_MAR	SERV_MAR	Activité_économi	328433	14058.1	3	0.0000
SERV_NMA	SERV_NMA	Activité_économi	134346	11940.4	3	0.0000
PROF_LIB	PROF_LIB	Cat_socioprofess	6504	9724.26	3	0.0000
ARTISANS	ARTISANS	Cat_socioprofess	51042	8481.28	3	0.0000
EMP_ENVIR	EMP_ENVIR	Cat_socioprofess	68681	7602.94	3	0.0000
PER_SERV	PER_SERV	Cat_socioprofess	142713	7596.09	3	0.0000
AGRICULT	AGRICULT	Cat_socioprofess	8664	7094.42	3	0.0000
COMMERC	COMMERC	Cat_socioprofess	40764	7040.18	3	0.0000
CONIREM	CONIREM	Cat_socioprofess	24412	5997.53	3	0.0000
CHEF_ENT	CHEF_ENT	Cat_socioprofess	6588	5158.59	3	0.0000
BIEN_INT	BIEN_INT	Activité_économi	106904	4810.32	3	0.0000
BIEN_EQU	BIEN_EQU	Activité_économi	104385	4391.94	3	0.0000
TRANSPOR	TRANSPOR	Activité_économi	46387	4116.55	3	0.0000
ENERGIE	ENERGIE	Activité_économi	7128	3325.44	3	0.0000
ORG_FINA	ORG_FINA	Activité_économi	7904	2657.91	3	0.0000
BIEN_COM	BIEN_COM	Activité_économi	96303	2483.22	3	0.0000
COMMERCE	COMMERCE	Activité_économi	124276	2212.26	3	0.0000
TECHNIC	TECHNIC	Cat_socioprofess	20617	1712.56	3	0.0000
EMP_FUEL	EMP_FUEL	Cat_socioprofess	47636	1393.65	3	0.0000
I_A_A	I_A_A	Activité_économi	27432	660.75	3	0.0000
ASSURANC	ASSURANC	Activité_économi	3357	342.65	3	0.0000
EMP_COMM	EMP_COMM	Cat_socioprofess	44608	225.08	3	0.0000
LOCATION	LOCATION	Activité_économi	4676	172.84	3	0.0000

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique

La macro PARTQUAL : partition en 4 classes

Caractérisation de chaque classe de la partition par les modalités des variables analysées
(triées par niveau de significativité)

CLASS4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test		
1	sur-représentées	OUV_NQ	OUV_NQ	Cat_socioprofess	162684	35.1	30.1	0.0000	87.141		
		TRANSPOR	TRANSPOR	Activité_économi	18065	5.1	3.6	0.0000	55.134		
		COMMERCE	COMMERCE	Cat_socioprofess	17100	3.7	2.6	0.0000	53.168		
		SERV_MAR	SERV_MAR	Activité_économi	98320	27.8	25.2	0.0000	41.036		
		EMP_FUEL	EMP_FUEL	Cat_socioprofess	17680	3.8	3.0	0.0000	35.773		
		OUV_QUAL	OUV_QUAL	Cat_socioprofess	132253	28.6	26.9	0.0000	29.243		
		COMMERCE	COMMERCE	Activité_économi	38148	10.8	9.5	0.0000	29.024		
		BIEN_INT	BIEN_INT	Activité_économi	31448	8.9	8.2	0.0000	17.175		
		BIEN_EQU	BIEN_EQU	Activité_économi	30552	8.6	8.0	0.0000	15.853		
		EMP_ENTR	EMP_ENTR	Cat_socioprofess	21921	4.7	4.4	0.0000	13.945		
		EMP_COMM	EMP_COMM	Cat_socioprofess	14228	3.1	2.8	0.0000	11.051		
		TECHNIC	TECHNIC	Cat_socioprofess	6392	1.4	1.3	0.0000	4.682		
		I_A_A	I_A_A	Activité_économi	7760	2.2	2.1	0.0000	4.253		
1	sous-représentées	ENERGIE	ENERGIE	Activité_économi	1900	0.5	0.5	0.1758	-0.931		
		LOCATION	LOCATION	Activité_économi	1212	0.3	0.4	0.0294	-1.890		
		INT_FUEL	INT_FUEL	Cat_socioprofess	11024	2.4	2.4	0.0273	-1.922		
		ASSURANC	ASSURANC	Activité_économi	729	0.2	0.3	0.0000	-7.315		
		BIEN_COM	BIEN_COM	Activité_économi	24404	6.9	7.4	0.0000	-13.269		
		PROF_LIB	PROF_LIB	Cat_socioprofess	1296	0.3	0.4	0.0000	-17.981		
		ORG_FINA	ORG_FINA	Activité_économi	1464	0.4	0.6	0.0000	-18.347		
		INT_ENTR	INT_ENTR	Cat_socioprofess	10572	2.3	2.8	0.0000	-25.173		
		SERV_NMA	SERV_NMA	Activité_économi	32220	9.1	10.3	0.0000	-28.066		
		CHEF_ENT	CHEF_ENT	Cat_socioprofess	904	0.2	0.4	0.0000	-31.462		
		BATIMENT	BATIMENT	Activité_économi	64056	18.1	20.6	0.0000	-44.141		
		CONTREM	CONTREM	Cat_socioprofess	4288	0.9	1.6	0.0000	-44.475		
		CAD_FUEL	CAD_FUEL	Cat_socioprofess	7628	1.6	2.5	0.0000	-48.453		
		ARTISANS	ARTISANS	Cat_socioprofess	10008	2.2	3.3	0.0000	-52.890		
		PER_SERV	PER_SERV	Cat_socioprofess	32688	7.1	9.1	0.0000	-59.404		
		AGRICULT	AGRICULT	Cat_socioprofess	340	0.1	0.6	0.0000	-68.214		
		CAD_ENTR	CAD_ENTR	Cat_socioprofess	7184	1.6	3.0	0.0000	-76.227		
		OUV_AGR	OUV_AGR	Cat_socioprofess	5016	1.1	2.7	0.0000	-93.488		
		AGRICOLE	AGRICOLE	Activité_économi	3844	1.1	3.4	0.0000	-103.622		
		2	sur-représentées	BATIMENT	BATIMENT	Activité_économi	160880	27.8	20.6	0.0000	180.564
				OUV_QUAL	OUV_QUAL	Cat_socioprofess	200704	30.7	26.9	0.0000	90.001
				PER_SERV	PER_SERV	Cat_socioprofess	75224	11.5	9.1	0.0000	88.186
				ARTISANS	ARTISANS	Cat_socioprofess	31020	4.8	3.3	0.0000	88.061
CONTREM	CONTREM			Cat_socioprofess	16044	2.5	1.6	0.0000	76.147		
BIEN_INT	BIEN_INT			Activité_économi	51620	8.9	8.2	0.0000	27.047		
BIEN_COM	BIEN_COM			Activité_économi	45560	7.9	7.4	0.0000	19.251		
I_A_A	I_A_A			Activité_économi	13428	2.3	2.1	0.0000	15.491		
LOCATION	LOCATION			Activité_économi	2476	0.4	0.4	0.0000	11.824		
OUV_NQ	OUV_NQ			Cat_socioprofess	199668	30.6	30.1	0.0000	10.464		
AGRICULT	AGRICULT			Cat_socioprofess	3640	0.6	0.6	0.2180	0.779		
ASSURANC	ASSURANC			Activité_économi	1500	0.3	0.3	0.3356	0.424		
2	sous-représentées			EMP_COMM	EMP_COMM	Cat_socioprofess	18080	2.8	2.8	0.0000	-4.686
				OUV_AGR	OUV_AGR	Cat_socioprofess	17236	2.6	2.7	0.0000	-6.122
				AGRICOLE	AGRICOLE	Activité_économi	18360	3.2	3.4	0.0000	-10.438
				CHEF_ENT	CHEF_ENT	Cat_socioprofess	2324	0.4	0.4	0.0000	-10.572
		ORG_FINA	ORG_FINA	Activité_économi	3028	0.5	0.6	0.0000	-10.899		
		TECHNIC	TECHNIC	Cat_socioprofess	7257	1.1	1.3	0.0000	-19.036		
		EMP_FUEL	EMP_FUEL	Cat_socioprofess	17268	2.6	3.0	0.0000	-24.344		
		BIEN_EQU	BIEN_EQU	Activité_économi	41161	7.1	8.0	0.0000	-33.455		
		ENERGIE	ENERGIE	Activité_économi	1732	0.3	0.5	0.0000	-35.790		
		COMMERCE	COMMERCE	Activité_économi	48300	8.4	9.5	0.0000	-41.094		
		SERV_NMA	SERV_NMA	Activité_économi	51808	9.0	10.3	0.0000	-45.296		
		INT_ENTR	INT_ENTR	Cat_socioprofess	12992	2.0	2.8	0.0000	-52.189		
		PROF_LIB	PROF_LIB	Cat_socioprofess	788	0.1	0.4	0.0000	-54.683		
		EMP_ENTR	EMP_ENTR	Cat_socioprofess	21476	3.3	4.4	0.0000	-57.519		
		TRANSPOR	TRANSPOR	Activité_économi	14636	2.5	3.6	0.0000	-57.735		
		COMMERCE	COMMERCE	Cat_socioprofess	9252	1.4	2.6	0.0000	-82.922		
		SERV_MAR	SERV_MAR	Activité_économi	123712	21.4	25.2	0.0000	-89.488		
		CAD_ENTR	CAD_ENTR	Cat_socioprofess	9020	1.4	3.0	0.0000	-108.568		
		INT_FUEL	INT_FUEL	Cat_socioprofess	6292	1.0	2.4	0.0000	-109.190		
		CAD_FUEL	CAD_FUEL	Cat_socioprofess	4620	0.7	2.5	0.0000	-138.862		

CAH sur deux tableaux de contingence juxtaposés : nationalité x CS et nationalité x activité économique

La macro PARTQUAL : partition en 4 classes

Caractérisation de chaque classe de la partition par les modalités des variables analysées
(triées par niveau de significativité)

CLASS4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test		
3	sur-représentées	OUV_AGRI	OUV_AGRI	Cat_socioprofess	18892	9.7	2.7	0.0000	151.358		
		AGRICOLE	AGRICOLE	Activité économi	16060	10.6	3.4	0.0000	127.904		
		OUV_NQ	OUV_NQ	Cat_socioprofess	72048	37.0	30.1	0.0000	69.234		
		BIEN_EQU	BIEN_EQU	Activité économi	18256	12.0	8.0	0.0000	56.559		
		ENERGIE	ENERGIE	Activité économi	2320	1.5	0.5	0.0000	43.442		
		BIEN_INT	BIEN_INT	Activité économi	14256	9.4	8.2	0.0000	17.603		
3	sous-représentées	TRANSPOR	TRANSPOR	Activité économi	5361	3.5	3.6	0.2984	-0.529		
		I A A	I A A	Activité économi	3160	2.1	2.1	0.2761	-0.595		
		LOCATION	LOCATION	Activité économi	432	0.3	0.4	0.0000	-5.333		
		OUV_QUAL	OUV_QUAL	Cat_socioprofess	51373	26.4	26.9	0.0000	-5.683		
		COMMERC	COMMERC	Cat_socioprofess	4660	2.4	2.6	0.0000	-6.080		
		ASSURANC	ASSURANC	Activité économi	208	0.1	0.3	0.0000	-11.116		
		EMP_COMM	EMP_COMM	Cat_socioprofess	4716	2.4	2.8	0.0000	-12.245		
		EMP_PUEL	EMP_PUEL	Cat_socioprofess	5032	2.6	3.0	0.0000	-12.729		
		COMMERCE	COMMERCE	Activité économi	13104	8.6	9.5	0.0000	-12.855		
		INT_PUEL	INT_PUEL	Cat_socioprofess	3744	1.9	2.4	0.0000	-15.795		
		PROF_LIB	PROF_LIB	Cat_socioprofess	424	0.2	0.4	0.0000	-16.343		
		AGRICULT	AGRICULT	Cat_socioprofess	596	0.3	0.6	0.0000	-17.551		
		BATIMENT	BATIMENT	Activité économi	28612	18.9	20.6	0.0000	-18.253		
		ORG_FINA	ORG_FINA	Activité économi	408	0.3	0.6	0.0000	-21.091		
		TECHNIC	TECHNIC	Cat_socioprofess	1604	0.8	1.3	0.0000	-22.223		
		EMP_ENTR	EMP_ENTR	Cat_socioprofess	6336	3.3	4.4	0.0000	-27.343		
		SERV_NMA	SERV_NMA	Activité économi	12684	8.4	10.3	0.0000	-27.548		
		CAD_PUEL	CAD_PUEL	Cat_socioprofess	3156	1.6	2.5	0.0000	-29.733		
		CHEF_ENT	CHEF_ENT	Cat_socioprofess	180	0.1	0.4	0.0000	-31.218		
		PER_SERV	PER_SERV	Cat_socioprofess	13328	6.9	9.1	0.0000	-38.770		
		CONIREM	CONIREM	Cat_socioprofess	1236	0.6	1.6	0.0000	-41.581		
		INT_ENTR	INT_ENTR	Cat_socioprofess	2812	1.4	2.8	0.0000	-43.394		
		BIEN_COM	BIEN_COM	Activité économi	6992	4.6	7.4	0.0000	-48.132		
		SERV_MAR	SERV_MAR	Activité économi	29900	19.7	25.2	0.0000	-54.191		
		ARTISANS	ARTISANS	Cat_socioprofess	2640	1.4	3.3	0.0000	-59.378		
		CAD_ENTR	CAD_ENTR	Cat_socioprofess	1752	0.9	3.0	0.0000	-72.542		
		CLASS4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
		4	sur-représentées	CAD_ENTR	CAD_ENTR	Cat_socioprofess	29344	11.3	3.0	0.0000	208.538
CAD_PUEL	CAD_PUEL			Cat_socioprofess	24288	9.4	2.5	0.0000	187.280		
INT_PUEL	INT_PUEL			Cat_socioprofess	16858	6.5	2.4	0.0000	121.509		
INT_ENTR	INT_ENTR			Cat_socioprofess	17228	6.7	2.8	0.0000	110.117		
SERV_MAR	SERV_MAR			Activité économi	76501	34.8	25.2	0.0000	108.808		
SERV_NMA	SERV_NMA			Activité économi	37634	17.1	10.3	0.0000	104.973		
PROF_LIB	PROF_LIB			Cat_socioprofess	3996	1.5	0.4	0.0000	75.108		
EMP_ENTR	EMP_ENTR			Cat_socioprofess	18948	7.3	4.4	0.0000	72.509		
AGRICULT	AGRICULT			Cat_socioprofess	4088	1.6	0.6	0.0000	62.197		
CHEF_ENT	CHEF_ENT			Cat_socioprofess	3180	1.2	0.4	0.0000	55.856		
ORG_FINA	ORG_FINA			Activité économi	3004	1.4	0.6	0.0000	42.616		
COMMERCE	COMMERCE			Cat_socioprofess	9752	3.8	2.6	0.0000	38.123		
TECHNIC	TECHNIC			Cat_socioprofess	5364	2.1	1.3	0.0000	33.884		
COMMERCE	COMMERCE			Activité économi	24724	11.2	9.5	0.0000	29.017		
BIEN_COM	BIEN_COM			Activité économi	19347	8.8	7.4	0.0000	26.807		
ASSURANC	ASSURANC			Activité économi	920	0.4	0.3	0.0000	14.803		
TRANSPOR	TRANSPOR			Activité économi	8325	3.8	3.6	0.0000	6.215		
EMP_COMM	EMP_COMM	Cat_socioprofess	7584	2.9	2.8	0.0012	3.026				
4	sous-représentées	ENERGIE	ENERGIE	Activité économi	1176	0.5	0.5	0.2039	-0.828		
		EMP_PUEL	EMP_PUEL	Cat_socioprofess	7656	3.0	3.0	0.0078	-2.417		
		LOCATION	LOCATION	Activité économi	556	0.3	0.4	0.0000	-9.714		
		ARTISANS	ARTISANS	Cat_socioprofess	7374	2.9	3.3	0.0000	-12.875		
		PER_SERV	PER_SERV	Cat_socioprofess	21473	8.3	9.1	0.0000	-15.517		
		CONIREM	CONIREM	Cat_socioprofess	2844	1.1	1.6	0.0000	-21.878		
		AGRICOLE	AGRICOLE	Activité économi	5548	2.5	3.4	0.0000	-25.256		
		I A A	I A A	Activité économi	3084	1.4	2.1	0.0000	-27.159		
		BIEN_EQU	BIEN_EQU	Activité économi	14416	6.6	8.0	0.0000	-28.558		
		BIEN_INT	BIEN_INT	Activité économi	9580	4.4	8.2	0.0000	-80.378		
		OUV_AGRI	OUV_AGRI	Cat_socioprofess	1760	0.7	2.7	0.0000	-88.617		
		OUV_QUAL	OUV_QUAL	Cat_socioprofess	38452	14.9	26.9	0.0000	-163.022		
		OUV_NQ	OUV_NQ	Cat_socioprofess	38351	14.8	30.1	0.0000	-201.099		
		BATIMENT	BATIMENT	Activité économi	15253	6.9	20.6	0.0000	-201.974		

Exemple 4 : tableau disjonctif complet (les chiens)

```
LIBNAME COMPIL ...;
OPTIONS MSTORED SASMSTORE=COMPIL NODATE PAGESIZE=66;
FILENAME ENT ...;

DATA A; INFILE ENT(CHIENS);
INPUT CHIEN $CHAR15. +1 NOM $4. +1 TAILLE POIDS VELOCITE INTELLIG AFFECTIO
      AGRESSIV FONCTION;
      IF NOM="SUPP" THEN DELETE;

TAI1=(TAILLE=1);
TAI2=(TAILLE=2);
TAI3=(TAILLE=3);

POI1=(POIDS=1);
POI2=(POIDS=2);
POI3=(POIDS=3);

VEL1=(VELOCITE=1);
VEL2=(VELOCITE=2);
VEL3=(VELOCITE=3);

INT1=(INTELLIG=1);
INT2=(INTELLIG=2);
INT3=(INTELLIG=3);

AFF1=(AFFECTIO=1);
AFF2=(AFFECTIO=2);

AGR1=(AGRESSIV=1);
AGR2=(AGRESSIV=2);

COMP=(FONCTION=1);
CHAS=(FONCTION=2);
UTIL=(FONCTION=3);

PROC PRINT DATA=A;
TITLE "Caractéristiques de races de chiens";
TITLE2 "Tableau de données";
RUN;

TITLE "CAH sur un tableau disjonctif complet : les races de chiens";
TITLE2 "Les sorties de la macro CAHQVAL";
%CAHQVAL(DATA=A,
        VAR=TAI1--AGR2,
        ID=NOM,
        ARBRE=OUI,
        NOEUDS=MAX);
RUN;

TITLE2 "La macro PARTQUAL : partition en 4 classes";
%PARTQUAL(DATA=A,
        NCL=4,
        VAR=TAI1--AGR2,
        ID=NOM,
        AFFECT=OUI,
        OUTPART=PART4,
        VARCLASS=CL4,
        LIBVAR=Taille Poids Vitesse Intelligence Affection Agressivité,
        NBMODA=3 3 3 2 2);
PROC PRINT DATA=PART4;
TITLE2 "Table en sortie créée par la macro PARTQUAL";
RUN;

TITLE2 "La macro DESQUAL : description de la partition en 4 classes"
      " avec des variables supplémentaires";
%DESQUAL(DATA=PART4,
        VAR=COMP CHAS UTIL,
        NCL=4,
        VARCLASS=CL4,
        LIBVAR=Fonction);
```


CAH sur un tableau disjonctif complet : les races de chiens
 Les sorties de la macro CAHQVAL
 Tableau des noeuds de la hiérarchie

Inertie totale : 1.666667

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter !		Histogramme
			0/00	cum. dif !	
CL1	162.0	CL3-CL2	260	260	. !*****
CL2	72.0	CL5-CL7	165	425	94 !*****
CL3	90.0	CL4-CL8	137	562	29 !*****
CL4	60.0	CL6-CL16	51	613	86 !*****
CL5	42.0	CL9-CL12	47	660	4 !*****
CL6	42.0	CL10-CL15	46	706	2 !*****
CL7	30.0	CL13-COCK	42	748	4 !*****
CL8	30.0	CL14-CL18	30	777	12 !*****
CL9	30.0	CL11-CL24	28	805	1 !*****
CL10	24.0	CL23-CL17	24	830	4 !*****
CL11	18.0	BASS-CL25	20	850	5 !*****
CL12	12.0	CANI-FOXT	20	869	0 !*****
CL13	24.0	CL19-EPAB	20	889	0 !*****
CL14	18.0	BULM-CL20	19	908	1 !*****
CL15	18.0	CL21-GREL	18	926	0 !****
CL16	18.0	CL22-POIN	12	938	6 !***
CL17	12.0	EPAF-SEIT	12	950	1 !**
CL18	12.0	DOGU-MAST	11	960	1 !**
CL19	18.0	BOXE-CL26	10	970	1 !**
CL20	12.0	STEBE-TERN	7	978	2 !**
CL21	12.0	FOXH-LEVR	7	985	0 !**
CL22	12.0	BERG-DOBE	7	993	0 !**
CL23	12.0	BEAU-COLL	7	1000	0 !**
CL24	12.0	BULD-TECK	0	1000	7 !
CL25	12.0	CHIU-PEKI	0	1000	0 !
CL26	12.0	DALM-LAER	0	1000	0 !

Décomposition de la distance entre les centres de gravité
 des classes jointes
 (Il s'agit aussi de la décomposition de la perte d'inertie interclasse)

Noeud	Effectif pondéré	Classes jointes	Perte d'inertie inter	TAI1	TAI2	TAI3	FOI1	FOI2	FOI3	VEL1	VEL2	VEL3	INT1	INT2	INT3	AFF1	AFF2	AGR1	AGR2
CL1	162	CL3-CL2	260	125	89	171	143	20	57	14	43	103	2	7	4	101	94	13	14
CL2	72	CL5-CL7	165	252	353	0	141	81	0	47	58	0	40	19	1	3	3	2	2
CL3	90	CL4-CL8	137	0	0	0	0	174	488	156	12	98	3	0	4	17	16	16	17
CL4	60	CL6-CL16	51	0	0	0	0	0	0	0	42	37	95	103	687	1	1	17	18
CL5	42	CL9-CL12	47	0	0	0	0	0	0	301	377	0	136	2	126	9	9	19	21
CL6	42	CL10-CL15	46	0	0	0	0	0	0	0	3	3	469	289	0	72	67	47	50
CL7	30	CL13-COCK	42	0	0	0	240	137	0	192	240	0	0	9	20	0	0	77	83
CL8	30	CL14-CL18	30	0	0	0	0	0	0	101	0	112	506	138	75	0	0	32	35
CL9	30	CL11-CL24	28	0	0	0	0	0	0	0	0	0	532	327	0	36	34	34	36
CL10	24	CL23-CL17	24	0	0	0	0	0	0	0	128	114	0	0	0	315	292	73	79
CL11	18	BASS-CL25	20	0	0	0	0	0	0	0	0	0	0	0	0	259	241	241	259
CL12	12	CANI-FOXT	20	0	0	0	0	0	0	0	0	0	0	196	425	0	0	182	196
CL13	24	CL19-EPAB	20	0	0	0	0	0	0	0	0	0	0	296	641	0	0	31	33
CL14	18	BULM-CL20	19	0	0	0	0	0	0	0	0	0	0	274	594	0	0	64	69
CL15	18	CL21-GREL	18	0	0	0	0	0	0	0	458	407	0	0	0	0	0	65	70
CL16	18	CL22-POIN	12	0	0	0	0	0	0	0	0	0	0	0	0	104	96	385	415
CL17	12	EPAF-SEIT	12	0	0	0	0	0	0	0	529	471	0	0	0	0	0	0	0
CL18	12	DOGU-MAST	11	0	0	0	0	0	0	474	0	526	0	0	0	0	0	0	0
CL19	18	BOXE-CL26	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	481	519
CL20	12	STEBE-TERN	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	481	519
CL21	12	FOXH-LEVR	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	481	519
CL22	12	BERG-DOBE	7	0	0	0	0	0	0	0	0	0	0	0	0	519	481	0	0
CL23	12	BEAU-COLL	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	481	519
CL24	12	BULD-TECK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CL25	12	CHIU-PEKI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CL26	12	DALM-LAER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

CAH sur un tableau disjonctif complet : les races de chiens
 La macro PARTQUAL : partition en 4 classes
 Partition en 4 classes : composition des classes

VARIABLE DE CLASSE : CL4

Classe numéro 1
 BOXE COCK DALM EPAB LABR

Classe numéro 2
 BASS BULD CANI CHIU FOXT PEKI TECK

Classe numéro 3
 BEAU BERG COLL DOBE EPAF FOXH GRBL LEVR POIN SETT

Classe numéro 4
 BULM DOGU MAST STBE TERN

Partition en 4 classes : Décomposition de l'inertie totale

	CL4	Inertie	En 0/00 de l'inertie totale	Effectif absolu	Effectif pondéré (en 0/00)
Inertie totale		1.6667	1000	27	1000
Inertie inter		0.9367	562	.	.
	1	0.2315	139	5	185
	2	0.2848	171	7	259
	3	0.1863	112	10	370
	4	0.2341	140	5	185
Inertie intra		0.7300	438	.	.
	1	0.1183	71	5	185
	2	0.1916	115	7	259
	3	0.3096	186	10	370
	4	0.1105	66	5	185

Décomposition signée du RHO2 (distance entre le centre des classes et le centre de gravité du nuage)
 selon les variables de l'analyse
 tous les chiffres sont en millièmes, exceptés l'effectif absolu et le RHO2

Classe	Effectif		RHO2	Effectif															
	absolu	pondéré		TAI1	TAI2	TAI3	POI1	POI2	POI3	VEL1	VEL2	VEL3	INT1	INT2	INT3	AFF1	AFF2	AGR1	AGR2
1	5	185	1.25028	-35	478	-74	-4	20	-25	-10	114	-44	-39	28	-0	-64	60	2	-2
2	7	259	1.09854	321	-28	-84	254	-79	-28	48	-0	-51	9	-1	-4	-36	34	11	-12
3	10	370	0.50288	-86	-61	118	-98	148	-61	-123	-10	217	0	-5	9	33	-31	-0	0
4	5	185	1.26393	-34	-24	47	-39	-68	473	66	-39	-7	5	-2	-0	74	-68	-26	28

Décomposition signée de la distance entre les centres de gravité des classes de la partition
 tous les chiffres sont en millièmes
 exceptées la perte d'inertie inter et la distance

Classes	Distance	Perte d'inertie																
		inter	TAI1	TAI2	TAI3	POI1	POI2	POI3	VEL1	VEL2	VEL3	INT1	INT2	INT3	AFF1	AFF2	AGR1	AGR2
2 - 1	2.55223	0.27570	252	-353	0	141	-81	0	47	-58	0	40	-19	-1	3	-3	2	-2
3 - 1	2.22316	0.27446	0	-405	135	-10	6	0	-8	-91	144	23	-25	3	76	-71	-1	2
4 - 1	3.78999	0.35093	0	-237	79	-6	-54	237	43	-95	5	24	-15	0	91	-85	-14	15
3 - 2	2.64649	0.40360	-243	0	113	-213	121	0	-87	-2	121	-4	-0	7	41	-38	-6	6
4 - 2	3.14481	0.33972	-204	0	95	-179	0	286	1	-15	6	-0	-0	1	81	-75	-27	29
4 - 3	1.84522	0.22780	0	0	0	0	-174	488	156	-12	-98	3	0	-4	17	-16	-16	17

CAH sur un tableau disjonctif complet : les races de chiens
 La macro PARTQUAL : partition en 4 classes
 Tableau(x) de contingence croisant la variable de classe et les variables analysées

----- Variable=Taille -----

TABLE OF CL4 BY MODALITE

CL4	MODALITE			
Frequency!				
Percent !				
Row Pct !				
Col Pct	!TAI1	!TAI2	!TAI3	Total
1	0	5	0	5
	0.00	18.52	0.00	18.52
	0.00	100.00	0.00	
	0.00	100.00	0.00	
2	7	0	0	7
	25.93	0.00	0.00	25.93
	100.00	0.00	0.00	
	100.00	0.00	0.00	
3	0	0	10	10
	0.00	0.00	37.04	37.04
	0.00	0.00	100.00	
	0.00	0.00	66.67	
4	0	0	5	5
	0.00	0.00	18.52	18.52
	0.00	0.00	100.00	
	0.00	0.00	33.33	
Total	7	5	15	27
	25.93	18.52	55.56	100.00

----- Variable=Poids -----

TABLE OF CL4 BY MODALITE

CL4	MODALITE			
Frequency!				
Percent !				
Row Pct !				
Col Pct	!POI1	!POI2	!POI3	Total
1	1	4	0	5
	3.70	14.81	0.00	18.52
	20.00	80.00	0.00	
	12.50	28.57	0.00	
2	7	0	0	7
	25.93	0.00	0.00	25.93
	100.00	0.00	0.00	
	87.50	0.00	0.00	
3	0	10	0	10
	0.00	37.04	0.00	37.04
	0.00	100.00	0.00	
	0.00	71.43	0.00	
4	0	0	5	5
	0.00	0.00	18.52	18.52
	0.00	0.00	100.00	
	0.00	0.00	100.00	
Total	8	14	5	27
	29.63	51.85	18.52	100.00

CAH sur un tableau disjonctif complet : les races de chiens
 La macro PARTQUAL : partition en 4 classes
 Tableau(x) de contingence croisant la variable de classe et les variables analysées

----- Variable=Vélocité -----

TABLE OF CL4 BY MODALITE

CL4	MODALITE			Total
Frequency!				
Percent !				
Row Pct !				
Col Pct	!VEL1	!VEL2	!VEL3	Total
1	1	4	0	5
	3.70	14.81	0.00	18.52
	20.00	80.00	0.00	
	10.00	50.00	0.00	
2	5	2	0	7
	18.52	7.41	0.00	25.93
	71.43	28.57	0.00	
	50.00	25.00	0.00	
3	0	2	8	10
	0.00	7.41	29.63	37.04
	0.00	20.00	80.00	
	0.00	25.00	88.89	
4	4	0	1	5
	14.81	0.00	3.70	18.52
	80.00	0.00	20.00	
	40.00	0.00	11.11	
Total	10	8	9	27
	37.04	29.63	33.33	100.00

----- Variable=Intelligence -----

TABLE OF CL4 BY MODALITE

CL4	MODALITE			Total
Frequency!				
Percent !				
Row Pct !				
Col Pct	!INT1	!INT2	!INT3	Total
1	0	4	1	5
	0.00	14.81	3.70	18.52
	0.00	80.00	20.00	
	0.00	30.77	16.67	
2	3	3	1	7
	11.11	11.11	3.70	25.93
	42.86	42.86	14.29	
	37.50	23.08	16.67	
3	3	4	3	10
	11.11	14.81	11.11	37.04
	30.00	40.00	30.00	
	37.50	30.77	50.00	
4	2	2	1	5
	7.41	7.41	3.70	18.52
	40.00	40.00	20.00	
	25.00	15.38	16.67	
Total	8	13	6	27
	29.63	48.15	22.22	100.00

CAH sur un tableau disjonctif complet : les races de chiens
 La macro PARTQUAL : partition en 4 classes
 Tableau(x) de contingence croisant la variable de classe et les variables analysées

----- Variable=Affection -----

TABLE OF CLA BY MODALITE

CLA	MODALITE		
	!AFF1	!AFF2	! Total
Frequency!			
Percent !			
Row Pct !			
Col Pct	!AFF1	!AFF2	! Total
1	0	5	5
	0.00	18.52	18.52
	0.00	100.00	
	0.00	35.71	
2	1	6	7
	3.70	22.22	25.93
	14.29	85.71	
	7.69	42.86	
3	7	3	10
	25.93	11.11	37.04
	70.00	30.00	
	53.85	21.43	
4	5	0	5
	18.52	0.00	18.52
	100.00	0.00	
	38.46	0.00	
Total	13	14	27
	48.15	51.85	100.00

----- Variable=Agressivité -----

TABLE OF CLA BY MODALITE

CLA	MODALITE		
	!AGR1	!AGR2	! Total
Frequency!			
Percent !			
Row Pct !			
Col Pct	!AGR1	!AGR2	! Total
1	3	2	5
	11.11	7.41	18.52
	60.00	40.00	
	21.43	15.38	
2	5	2	7
	18.52	7.41	25.93
	71.43	28.57	
	35.71	15.38	
3	5	5	10
	18.52	18.52	37.04
	50.00	50.00	
	35.71	38.46	
4	1	4	5
	3.70	14.81	18.52
	20.00	80.00	
	7.14	30.77	
Total	14	13	27
	51.85	48.15	100.00

CHI2 des tableaux de contingence croisant la variable de classe et les variables analysées
(triés par niveau de significativité)

Variable	Nbre de modalités	Effectif pondéré du tableau	Nbre de cases d'effectif <5	Nbre de cases d'effectif <5 (en %)	CHI2	Degrés de liberté	PROBA
Taille	3	27	8	67	54.00	6	0.0000
Poids	3	27	9	75	49.76	6	0.0000
Vélocité	3	27	10	83	25.70	6	0.0003
Affection	2	27	4	50	15.16	3	0.0017
Agressivité	2	27	5	63	3.25	3	0.3542
Intelligence	3	27	12	100	3.87	6	0.6942

CAH sur un tableau disjonctif complet : les races de chiens
La macro PARTQUAL : partition en 4 classes
Caractérisation de chaque classe de la partition par les variables analysées
(triées par niveau de significativité)

Cl4	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
1	Taille	5	22.00	2	0.0000
	Affection	5	4.64	1	0.0312
	Vélocité	5	6.34	2	0.0420
	Intelligence	5	2.55	2	0.2800
	Poids	5	1.85	2	0.3972
	Agressivité	5	0.13	1	0.7154
2	Taille	7	20.00	2	0.0000
	Poids	7	16.62	2	0.0002
	Affection	7	3.22	1	0.0730
	Vélocité	7	4.57	2	0.1017
	Agressivité	7	1.07	1	0.2999
	Intelligence	7	0.65	2	0.7216
3	Vélocité	10	10.55	2	0.0051
	Poids	10	9.29	2	0.0096
	Taille	10	8.00	2	0.0183
	Affection	10	1.91	1	0.1667
	Intelligence	10	0.41	2	0.8144
	Agressivité	10	0.01	1	0.9067
4	Poids	5	22.00	2	0.0000
	Affection	5	5.38	1	0.0203
	Vélocité	5	4.24	2	0.1200
	Taille	5	4.00	2	0.1353
	Agressivité	5	2.03	1	0.1540
	Intelligence	5	0.26	2	0.8774

Caractérisation de la partition par les modalités des variables analysées
(triées par niveau de significativité)

Modalité	Label	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
POI3	POI3	Poids	5	22.00	3	0.0001
TAI2	TAI2	Taille	5	22.00	3	0.0001
TAI1	TAI1	Taille	7	20.00	3	0.0002
POI1	POI1	Poids	8	16.30	3	0.0010
TAI3	TAI3	Taille	15	12.00	3	0.0074
POI2	POI2	Poids	14	11.46	3	0.0095
VEL3	VEL3	Vélocité	9	10.80	3	0.0129
VEL1	VEL1	Vélocité	10	8.82	3	0.0317
AFF1	AFF1	Affection	13	7.86	3	0.0490
AFF2	AFF2	Affection	14	7.30	3	0.0630
VEL2	VEL2	Vélocité	8	6.08	3	0.1078
INT1	INT1	Intelligence	8	2.08	3	0.5566
AGR2	AGR2	Agressivité	13	1.69	3	0.6399
AGR1	AGR1	Agressivité	14	1.57	3	0.6670
INT2	INT2	Intelligence	13	1.30	3	0.7289
INT3	INT3	Intelligence	6	0.49	3	0.9205

CAH sur un tableau disjonctif complet : les races de chiens
 La macro PARTQUAL : partition en 4 classes
 Caractérisation de chaque classe de la partition par les modalités des variables analysées
 (triées par niveau de significativité)

Cl4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
1	sur-représentées	TAI2	TAI2	Taille	5	100.0	18.5	0.0000	4.21685
		VEL2	VEL2	Vélocité	4	80.0	29.6	0.0172	2.11610
		AFF2	AFF2	Affection	5	100.0	51.9	0.0248	1.96342
		INT2	INT2	Intelligence	4	80.0	48.1	0.1399	1.08061
		POI2	POI2	Poids	4	80.0	51.9	0.1860	0.89277
		AGR1	AGR1	Agressivité	3	60.0	51.9	0.5377	-0.09459
1	sous-représentées	INT3	INT3	Intelligence	1	20.0	22.2	0.6969	0.51544
		AGR2	AGR2	Agressivité	2	40.0	48.1	0.5377	0.09459
		POI1	POI1	Poids	1	20.0	29.6	0.5281	0.07057
		VEL1	VEL1	Vélocité	1	20.0	37.0	0.3715	-0.32799
		POI3	POI3	Poids	0	0.0	18.5	0.3262	-0.45043
		TAI1	TAI1	Taille	0	0.0	25.9	0.1920	-0.87038
		INT1	INT1	Intelligence	0	0.0	29.6	0.1440	-1.06236
		VEL3	VEL3	Vélocité	0	0.0	33.3	0.1061	-1.24737
		AFF1	AFF1	Affection	0	0.0	48.1	0.0248	-1.96342
		TAI3	TAI3	Taille	0	0.0	55.6	0.0098	-2.33352
Cl4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
2	sur-représentées	TAI1	TAI1	Taille	7	100.0	25.9	0.0000	4.72937
		POI1	POI1	Poids	7	100.0	29.6	0.0000	4.28814
		VEL1	VEL1	Vélocité	5	71.4	37.0	0.0427	1.71964
		AFF2	AFF2	Affection	6	85.7	51.9	0.0478	1.66631
		AGR1	AGR1	Agressivité	5	71.4	51.9	0.2237	0.75985
		INT1	INT1	Intelligence	3	42.9	29.6	0.3322	0.43385
2	sous-représentées	VEL2	VEL2	Vélocité	2	28.6	29.6	0.6678	0.43385
		INT2	INT2	Intelligence	3	42.9	48.1	0.5461	0.11570
		INT3	INT3	Intelligence	1	14.3	22.2	0.4976	-0.00607
		AGR2	AGR2	Agressivité	2	28.6	48.1	0.2237	-0.75985
		TAI2	TAI2	Taille	0	0.0	18.5	0.1920	-0.87038
		POI3	POI3	Poids	0	0.0	18.5	0.1920	-0.87038
		AFF1	AFF1	Affection	1	14.3	48.1	0.0478	-1.66631
		VEL3	VEL3	Vélocité	0	0.0	33.3	0.0358	-1.80119
		POI2	POI2	Poids	0	0.0	51.9	0.0019	-2.88900
		TAI3	TAI3	Taille	0	0.0	55.6	0.0009	-3.12406
Cl4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
3	sur-représentées	POI2	POI2	Poids	10	100.0	51.9	0.0001	3.67558
		VEL3	VEL3	Vélocité	8	80.0	33.3	0.0002	3.58997
		TAI3	TAI3	Taille	10	100.0	55.6	0.0004	3.38495
		AFF1	AFF1	Affection	7	70.0	48.1	0.0891	1.34605
		INT3	INT3	Intelligence	3	30.0	22.2	0.3873	0.28628
		AGR2	AGR2	Agressivité	5	50.0	48.1	0.5982	-0.24861
		INT1	INT1	Intelligence	3	30.0	29.6	0.6506	-0.38692
3	sous-représentées	AGR1	AGR1	Agressivité	5	50.0	51.9	0.5982	0.24861
		INT2	INT2	Intelligence	4	40.0	48.1	0.4018	-0.24861
		VEL2	VEL2	Vélocité	2	20.0	29.6	0.3494	-0.38692
		AFF2	AFF2	Affection	3	30.0	51.9	0.0891	-1.34605
		TAI2	TAI2	Taille	0	0.0	18.5	0.0767	-1.42797
		POI3	POI3	Poids	0	0.0	18.5	0.0767	-1.42797
		TAI1	TAI1	Taille	0	0.0	25.9	0.0219	-2.01600
		POI1	POI1	Poids	0	0.0	29.6	0.0110	-2.29210
VEL1	VEL1	Vélocité	0	0.0	37.0	0.0023	-2.83305		

CAH sur un tableau disjonctif complet : les races de chiens
 La macro DESQUAL : description de la partition en 4 classes avec des variables supplémentaires
 Tableau(x) de contingence croisant la variable de classe et les variables explicatives

----- Variable=Fonction -----

TABLE OF CLA BY MODALITE

CLA	MODALITE				Total
Frequency!					
Percent !					
Row Pct !					
Col Pct	!CHAS	!COMP	!UTIL	!	Total
1	2	3	0		5
	7.41	11.11	0.00		18.52
	40.00	60.00	0.00		
	22.22	30.00	0.00		
2	1	6	0		7
	3.70	22.22	0.00		25.93
	14.29	85.71	0.00		
	11.11	60.00	0.00		
3	6	1	3		10
	22.22	3.70	11.11		37.04
	60.00	10.00	30.00		
	66.67	10.00	37.50		
4	0	0	5		5
	0.00	0.00	18.52		18.52
	0.00	0.00	100.00		
	0.00	0.00	62.50		
Total	9	10	8		27
	33.33	37.04	29.63		100.00

CHI2 des tableaux de contingence croisant la variable de classe et les variables explicatives
 (triés par niveau de significativité)

Variable	Nbre de modalités	Effectif pondéré du tableau	Nbre de cases d'effectif <5	Nbre de cases d'effectif <5 (en %)	CHI2	Degrés de liberté	PROBA
Fonction	3	27	9	75	25.56	6	0.0003

Caractérisation de chaque classe de la partition par les variables explicatives
 (triées par niveau de significativité)

CLA	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
1	Fonction	5	2.26	2	0.3230
2	Fonction	7	7.31	2	0.0258
3	Fonction	10	4.11	2	0.1283
4	Fonction	5	11.88	2	0.0026

Caractérisation de la partition par les modalités des variables explicatives
 (triées par niveau de significativité)

Modalité	Label	Variable	Effectif pondéré	CHI2	Degrés de liberté	PROBA
UTIL	UTIL	Fonction	8	11.91	3	0.0077
COMP	COMP	Fonction	10	9.02	3	0.0291
CHAS	CHAS	Fonction	9	4.63	3	0.2011

CAH sur un tableau disjonctif complet : les races de chiens

La macro DESQUAL : description de la partition en 4 classes avec des variables supplémentaires
 Caractérisation de chaque modalité de la variable CL4 par les modalités des variables explicatives
 (triées par niveau de significativité)

CL4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
1	sur-représentées	COMP	COMP	Fonction	3	60.0	37.0	0.2495	0.67607
		CHAS	CHAS	Fonction	2	40.0	33.3	0.5527	-0.13257
1	sous-représentées	UTIL	UTIL	Fonction	0	0.0	29.6	0.1440	-1.06236
CL4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
2	sur-représentées	COMP	COMP	Fonction	6	85.7	37.0	0.0042	2.63919
		CHAS	CHAS	Fonction	1	14.3	33.3	0.2240	-0.75882
2	sous-représentées	UTIL	UTIL	Fonction	0	0.0	29.6	0.0567	-1.58273
CL4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
3	sur-représentées	CHAS	CHAS	Fonction	6	60.0	33.3	0.0341	1.82347
		UTIL	UTIL	Fonction	3	30.0	29.6	0.6506	-0.38692
3	sous-représentées	COMP	COMP	Fonction	1	10.0	37.0	0.0311	-1.86456
CL4	Modalités	Modalité	Label	Variable	Effectif pondéré	Fréquence dans la classe(%)	Fréquence dans la pop. (%)	PROBA	Val.Test
4	sur-représentées	UTIL	UTIL	Fonction	5	100.0	29.6	0.0007	3.19727
4	sous-représentées	CHAS	CHAS	Fonction	0	0.0	33.3	0.1061	-1.24737
		COMP	COMP	Fonction	0	0.0	37.0	0.0767	-1.42797

Annexe: liste des paramètres des macros

La macro ACP

DATAACT = nom de table SAS
DATASUP z nom de table SAS
VARACT = liste de variables
VARSUP = liste de variables
ID = variable
IDSUP = liste de variables
POIDS = variable
IMPRESS = O ou N
CORR = O ou N
VECP= n ou MAX
IOA = n
IOS = n
IVA = n
IVS = n
PARTIEL = n
LSIZE = longueur
NOTES = O ou N
CLASSES = liste de variables
FMTCLAS = liste de formats
AXECLAS = n
OUT = nom de table SAS
NAXER = n
FILL = VAR et/ou IND et/ou BARY et/ou ALL

La macro PLOTACP

DATA = nom de table SAS
AXEH = n
AXEV = n
POINTS = OBSACT et/ V ou OBSSUP et/ou VARSUP et/ou VARACT et/ V ou AXEUNI
IDSUP = variable
CERCLE = O ou N
NOTES = O ou N

La macro AIDEAFC

DATA = nom de table SAS
ID = variable
IOA = n
IOS = n
IVA = n
IVS = n
PARTIEL = n
LSIZE = longueur
NOTES = O ou N

La macro AIDEACM

DATA = nom de table SAS
DATAINIT = nom de table SAS
ID = variable
IOA = n
IOS = n
IVA = n
IVS = n
WEIGHT = nom de variable
PARTIEL = n
MISSING = OUI ou NON
LSIZE = longueur
NOTES = O ou N
ANALYSE = TABLES ou VAR
VARACT = liste de noms de variables
VARSUP = liste de noms de variables
FORMAT = instruction format
NBMODACT = nombres des modalités des variables actives
NBMODSUP = nombres des modalités des variables supplémentaires

La macro PLOTCOR

DATA = nom de table SAS
AXEH = n
AXEV = n
POINTS = OBSACT et/ou OBSSUP et/ou VARSUP et/ou VARACT
ID = variable
NOTES = O ou N

La macro CAHNUM

DATA = nom de table SAS
VAR = liste de variables
ID = variable
TABN = nom de table SAS
POIDS = variable
ARBRE = OUI ou NON
CCC = OUI ou NON
NOEUDS = n ou MAX
LSIZE = longueur
NOTES = O ou N
METHODE = méthode utilisée
REDUC = OUI ou NON

La macro CAHQVAL

DATA = nom de table SAS
VAR = liste de variables
ID = variable
TABN = nom de table SAS
POIDS = variable
ARBRE = OUI ou NON
CCC = OUI ou NON
NOEUDS = n ou MAX
LSIZE = longueur
NOTES = O ou N
METHODE = méthode utilisée

La macro PARTNUM

TABN = nom de table SAS
DATA = nom de table SAS
VAR = liste des variables
ID = variable
POIDS = nom de variable
NCL= nombre de classes
LISTE = OUI ou NON
AFFECT = OUI ou NON
OUTPART = nom de table SAS
VARCLASS = nom de variable
DESC = OUI ou NON
LSIZE = longueur
NOTES = O ou N
CONSOLID = OUI ou NON
OBSSUP = OUI ou NON
REDUC = OUI ou NON

La macro PARTQUAL

TABN = nom de table SAS
DATA = nom de table SAS
VAR = liste des variables
ID = variable
POIDS = nom de variable
NCL= nombre de classes
LISTE = OUI ou NON
AFFECT = OUI ou NON
OUTPART = nom de table SAS
VARCLASS = nom de variable
DESC = OUI ou NON
LSIZE = longueur
NOTES = O ou N
CONSOLID = OUI ou NON
OBSSUP = OUI ou NON
LIBVAR = liste de libellés
NBMODA = liste des nombres de modalités
PROBA = probabilité

La macro DESNUM

DATA = nom de table SAS
VAR = liste de variables
POIDS = variable
VARCLASS = nom de variable
MISSING = OUI ou NON
NCL = n
NOTES = O ou N

La macro DESQUAL

DATA = nom de table SAS
VAR = liste de variables
POIDS = variable
VARCLASS = nom de variable
MISSING = OUI ou NON
NCL = n
LIBVAR = liste de libellés
NBMODA = liste des nombres de modalités
PROBA = probabilité
NOTES = O ou N