

1 Objectif

Analyse des correspondances multiples – Comparaison de logiciels.

L'analyse des correspondances multiples (ACM) est une technique factorielle qui s'applique aux tableaux « individus - variables », ces dernières étant exclusivement catégorielles. On peut la voir comme une variante de l'analyse en composantes principales (*il y a de nombreuses manières de la voir en réalité*). D'ailleurs, lorsque les descripteurs sont tous binaires, les deux approches se rejoignent. A la sortie, nous obtenons une description synthétique des données qui préserve les proximités entre les individus, met en évidence les liaisons entre les variables et, éventuellement, permettre de situer des groupes d'observations partageant les mêmes caractéristiques. Je n'en dirai pas plus, il existe de très nombreux supports de qualité sur internet qui permettent de situer relativement simplement les tenants et aboutissants de cette méthode¹.

Nous avons décrit la mise en œuvre de l'ACM dans plusieurs tutoriels, notamment sous Tanagra² et sous R avec le package FactoMiner³. Les sorties du composant « MULTIPLE CORRESPONDENCE ANALYSIS » ayant été complétées dans la version 1.4.48 de Tanagra, je me suis dit que la meilleure manière de situer les avancées est de comparer les tableaux de résultats fournis par Tanagra avec ceux des autres logiciels. Ce travail permet aux étudiants de se dégager de l'autocratie des logiciels : comprendre la méthode est le plus important, qu'importent les outils, tous – *normalement, si ce n'est pas le cas, il faut comprendre pourquoi* – fourniront en définitive les mêmes résultats.

Dans un premier temps, nous décrivons la mise en œuvre de Tanagra sur une base de données comportant 8403 individus décrits par 23 variables. Nous mettrons en parallèle les sorties de SAS 9.3 (PROC CORRESP). Par la suite, nous mènerons la même étude avec R via le package « ade4 », puis avec STATISTICA. Nous constaterons que si les solutions numériques sont identiques, la manière de les présenter peut différer d'un logiciel à l'autre.

2 Données

Le fichier « **loisirs.xls** » décrit les loisirs de (**n = 8403**) individus. Etant relativement nombreuses, il paraît illusoire de vouloir analyser finement les observations dans l'espoir d'identifier des groupes particuliers. Graphiquement, nous aurons un amas de points où il serait très difficile de distinguer quoique ce soit. Nous nous concentrerons donc avant tout sur l'analyse des relations entre les descripteurs dans ce tutoriel.

Les variables actives (**p = 18**) de l'étude recensent les occupations (ex. lecture oui/non, cinéma oui/non, etc.). Nous chercherons à identifier les concomitances et les incompatibilités entre ces activités, que nous essaierons de caractériser à l'aide de (**q = 5**) variables illustratives relatives à la signalétique des individus (ex. âge, profession, etc.).

¹ Ex. Labat, « Analyse de données – Ordinations et regroupements », « [Chapitre 4. L'analyse des Correspondances Multiples \(AFCM\)](#) ». Abdi & Valentin, « [Multiple Correspondence Analysis](#) », in Encyclopedia of Measurement and Statistics, 2007. Sans compter les excellents ouvrages en français qui font référence.

² Tutoriel Tanagra, « [AFCM – Races canines](#) ».

³ Tutoriel Tanagra, « [Analyse de Correspondances Multiples avec R](#) ».

Voici une description sommaire de la base :

Reading n:2757 y:5646	Listening_music n:2456 y:5947	Cinema n:5044 y:3359	Show n:5978 y:2425	Exhibition n:5808 y:2595	Computer n:5245 y:3158	Sport n:5308 y:3095
Walking n:4228 y:4175	Travelling n:5040 y:3363	Playing_music n:6943 y:1460	Collecting n:7541 y: 862	Volunteering n:7118 y:1285	Mechanic n:4864 y:3539	
Gardening n:5047 y:3356	Knitting n:6990 y:1413	Cooking n:4717 y:3686	Fishing n:7458 y: 945	TV n0:1017 n1:1223 n2:2156 n3:1775 n4:2232	Sex F:4616 M:3787	Age (45,55]:1837 (35,45]:1646 (25,35]:1302 (55,65]:1257 (65,75]: 937 [15,25]: 857 (Other): 567
Marital_status Divorcee : 792 Married :4333 Remarried: 404 Single :2140 Widower : 734	Profession Employee :2552 NotAvailable :1498 Manual_labourer :1161 Management :1052 Unskilled_worker: 792 Foreman : 735 (Other) : 613	nb_activitees Min. : 0.000 1st Qu.: 4.000 Median : 7.000 Mean : 6.866 3rd Qu.: 9.000 Max. :16.000				

La situation de la variable quantitative « nb_activitees » (nombre d'activités) est assez particulière. Bien qu'elle décrive un comportement, nous l'utiliserons quand même comme variable supplémentaire. Nous constaterons qu'elle caractérise à merveille l'attitude des individus par rapport aux loisirs qu'ils affectionnent.

Enfin, les $p = 18$ descripteurs actifs (Reading... TV) totalisent $M = 39$ modalités. Au maximum, nous pouvons extraire $M - p = 39 - 18 = F_{\max} = 21$ facteurs avec l'ACM.

3 ACM avec Tanagra (1.4.48)

3.1 ACM avec SAS

Nous commençons avec la **PROC CORRESP** de **SAS 9.3** dans un premier temps puisque ses sorties serviront de repères. Après avoir importé les données⁴, nous introduisons la commande suivante :

```
proc corresp mca data = mesdata.loisirs dimens = 2;
  tables Reading Listening_music Cinema Show
  Exhibition Computer Sport Walking Travelling
  Playing_music Collecting Volunteering Mechanic
  Gardening Knitting Cooking Fishing TV Sex
  Age Marital_status Profession;
  supplementary Sex Age Marital_status
  Profession;
run;
```

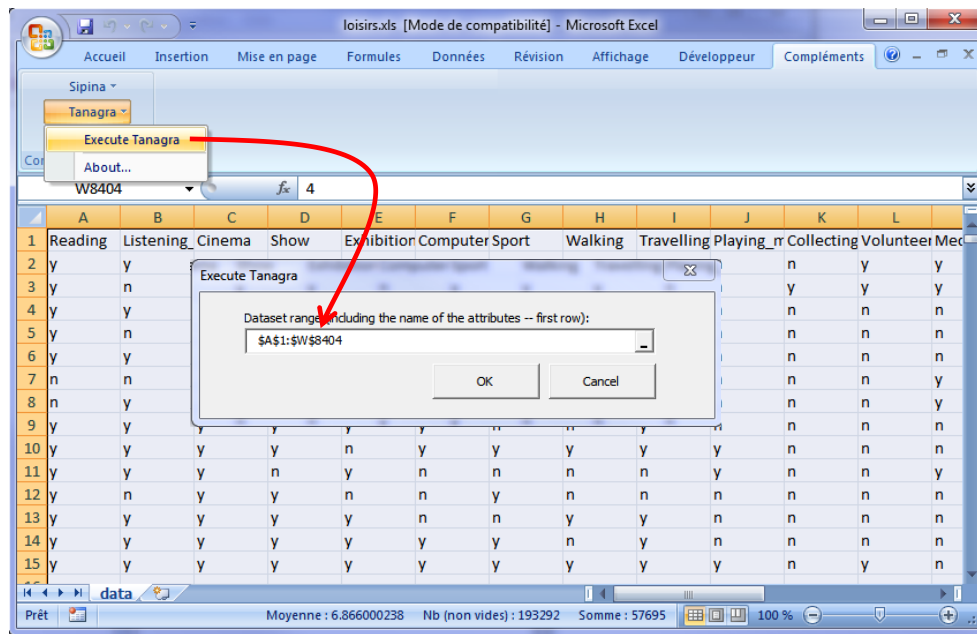
Nous n'avons pas utilisé « nb_activitees » car SAS ne semble pas savoir gérer les variables supplémentaires quantitatives dans l'ACM (*je n'ai pas su faire en tous les cas, si un lecteur charitable*

⁴ Cf. <http://tutoriels-data-mining.blogspot.fr/2012/04/la-proc-logistic-de-sas-93.html>

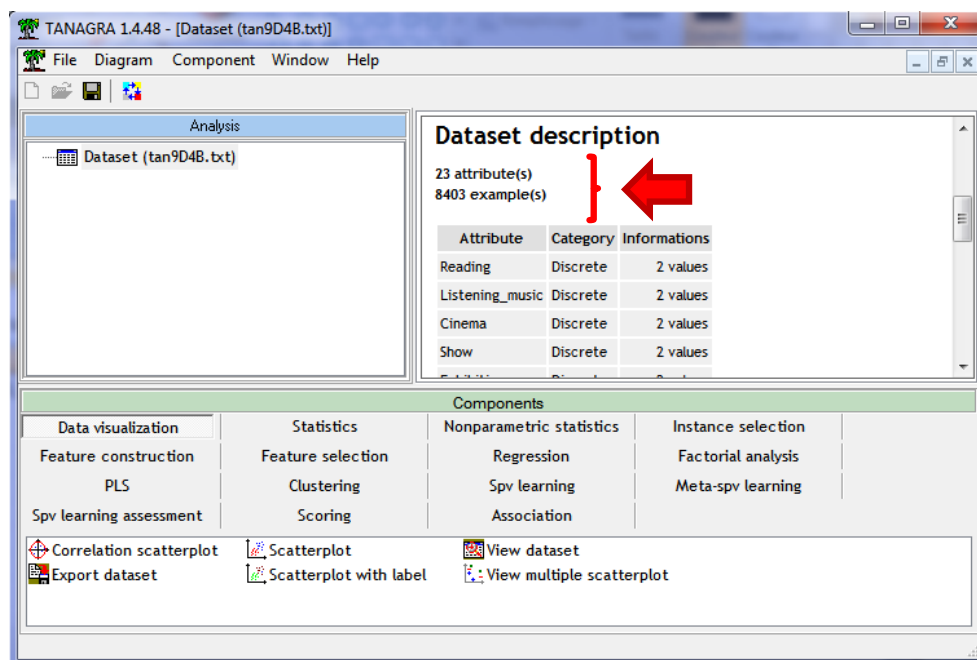
a une idée...). Cela ne porte pas à conséquence dans la comparaison des résultats puisqu'elle ne pèse pas dans la construction des axes factoriels.

3.2 Importation de données dans Tanagra

Le plus simple toujours est de charger le fichier « **loisirs.xls** » dans Excel, puis de l'envoyer à Tanagra via la macro complémentaire **Tanagra.xla**⁵. Une boîte de dialogue permet de vérifier les coordonnées des cellules (**\$A\$1:\$W\$8404**). Nous validons en cliquant sur le bouton OK.



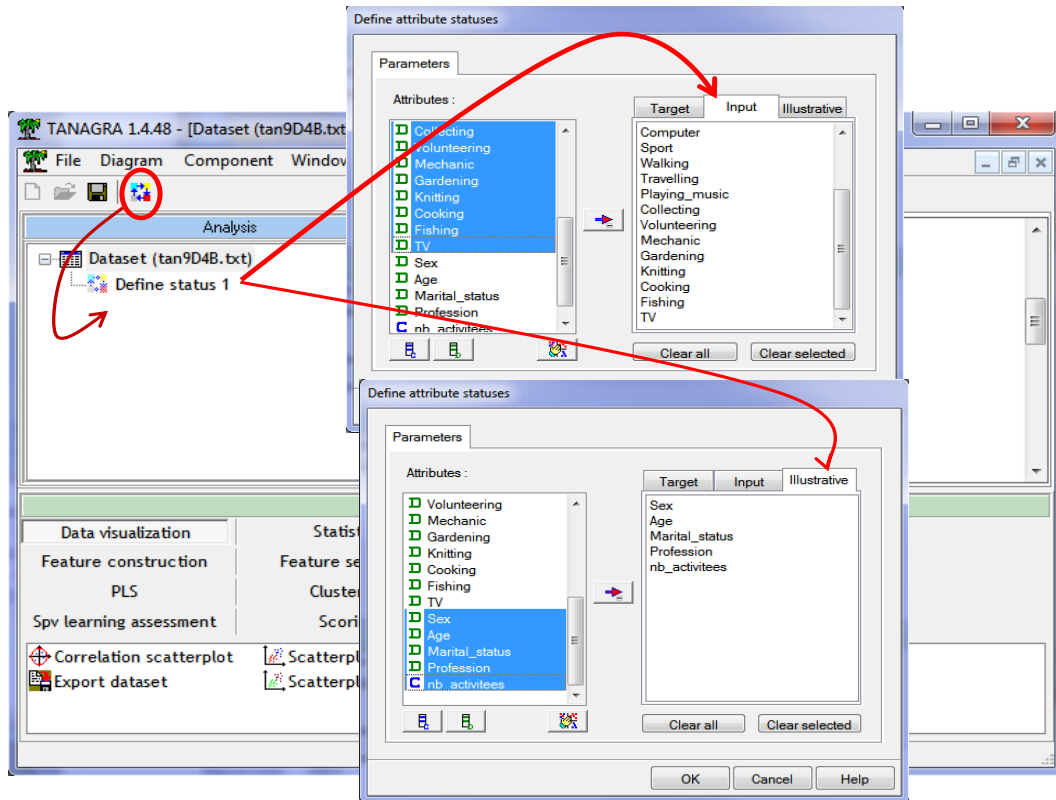
Tanagra est automatiquement démarré et les données chargées. Nous avons 23 variables et 8403 observations dans la base.



⁵ Voir <http://tutoriels-data-mining.blogspot.fr/2010/08/ladd-in-tanagra-pour-excel-2007-et-2010.html> pour Excel ; <http://tutoriels-data-mining.blogspot.fr/2011/07/tanagra-addon-pour-openoffice-33.html> pour Open et Libre Office.

3.3 Paramétrage de l'ACM

Nous devons spécifier les variables actives (INPUT : Reading ... TV) et supplémentaires (ILLUSTRATIVE) de l'étude. Nous utilisons le composant DEFINE STATUS pour ce faire. Les actives sont forcément toutes qualitatives (D pour discrète), les illustratives peuvent être discrètes ou continues (C).



Puis, nous plaçons le composant MULTIPLE CORRESPONDENCE ANALYSIS. Nous actionnons le menu contextuel PARAMETERS pour spécifier les paramètres de l'algorithme.

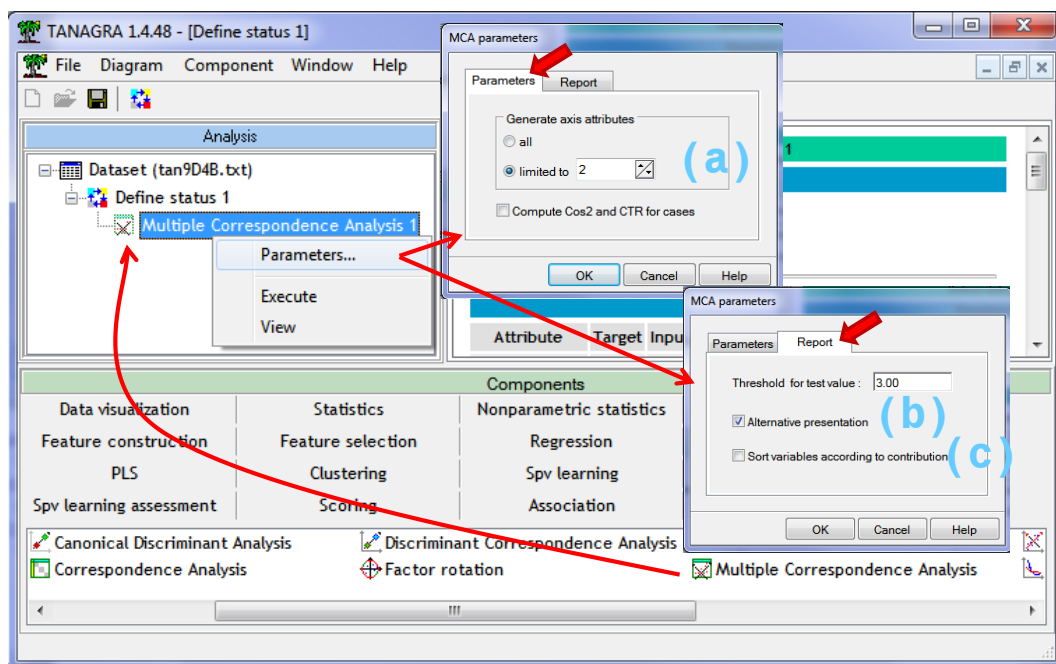


Figure 1 - Paramétrage du composant "Multiple Correspondence Analysis" - Tanagra

Nous limitons les calculs à (**F = 2**) facteurs (a) (nous verrons pourquoi par la suite) ; nous choisissons la présentation alternative des rapports (b) (toutes les informations sont réunies dans un seul tableau, cet agencement est avantageux si le nombre de facteurs est limité) ; nous ne trions pas les variables selon leurs contributions (c) (pour l’instant).

Nous validons ces spécifications et nous cliquons sur le menu VIEW pour accéder aux résultats.

3.4 Lecture des résultats

Les sorties sont subdivisées en sections. Voyons en le détail et comparons-les avec celles de SAS.

3.4.1 Description du problème

Tanagra regroupe dans un premier temps les informations relatives aux données traitées.

Problem statement	
# of instances	8403
# of variables	18
# of variable values	39
Max # of factors	21
# of factors extracted	2
Total inertia	1.166667

$I_{tot} = \frac{M}{p} - 1 = \frac{39}{18} - 1$

L’inertie totale à décomposer est égale à $I_{tot} = 1.166667$.

3.4.2 Valeurs propres – Parts d’inertie expliquée

Le tableau des valeurs propres λ_f indique la décomposition de l’inertie sur les facteurs de l’ACM.

Eigen values

Axis	Eigen value	% explained	Histogram	% cumulated
1	0.197712	16.95%		16.95%
2	0.080649	6.91%		23.86%
3	0.072022	6.17%		30.03%
4	0.062872	5.39%		35.42%
5	0.058460	5.01%		40.43%
6	0.055812	4.78%		45.22%
7	0.055523	4.76%		49.98%
8	0.053308	4.57%		54.55%
9	0.053044	4.55%		59.09%
10	0.049130	4.21%		63.30%
11	0.046493	3.99%		67.29%
12	0.045074	3.86%		71.15%
13	0.043513	3.73%		74.88%
14	0.043361	3.72%		78.60%
15	0.040800	3.50%		82.09%
16	0.037982	3.26%		85.35%
17	0.037335	3.20%		88.55%
18	0.036219	3.10%		91.66%
19	0.034968	3.00%		94.65%
20	0.032342	2.77%		97.42%
21	0.030046	2.58%		100.00%

Tanagra

The CORRESP Procedure						
Inertia and Chi-Square Decomposition						
Singular Value	Principal Inertia	Chi-Square	Percent	Cumulative Percent	4	8 12 16 20
0.44465	0.19771	41234	16.95	16.95	*****	
0.28399	0.08065	16820	6.91	23.86	*****	
0.26837	0.07202	15021	6.17	30.03	*****	
0.25074	0.06287	13113	5.39	35.42	*****	
0.24179	0.05846	12192	5.01	40.43	*****	
0.23625	0.05581	11640	4.78	45.22	*****	
0.23563	0.05552	11580	4.76	49.98	*****	
0.23089	0.05331	11118	4.57	54.55	*****	
0.23031	0.05304	11063	4.55	59.09	*****	
0.22165	0.04913	10246	4.21	63.30	*****	
0.21562	0.04649	9697	3.99	67.29	*****	
0.21231	0.04507	9401	3.86	71.15	*****	
0.20860	0.04351	9075	3.73	74.88	*****	
0.20823	0.04336	9043	3.72	78.60	*****	
0.20199	0.04080	8509	3.50	82.09	****	
0.19489	0.03798	7921	3.26	85.35	****	
0.19322	0.03733	7786	3.20	88.55	****	
0.19031	0.03622	7554	3.10	91.66	****	
0.18700	0.03497	7293	3.00	94.65	****	
0.17984	0.03234	6745	2.77	97.42	***	
0.17334	0.03005	6266	2.58	100.00	***	
Total	1.16667	243317	100.00			

Degrees of Freedom = 1444

SAS

Comme $\sum_{f=1}^{F_{max}} \lambda_f = I_{tot}$, nous pouvons afficher l'information en proportion d'inertie expliquée. SAS indique la statistique du $\chi^2 = 243317$ du tableau de Burt [*degré de liberté* = $(M - 1) \times (M - 1) = 38 \times 38 = 1444$]. Ce χ^2 est aussi réparti sur les facteurs.

Nous avons reproduit le calcul du χ^2 du tableau de Burt sous R avec les outils du package « ade4 ».

```
#chargement des données
donnees <- read.table(file="loisirs.txt",header=T,sep="\t")
#package ade4
library(ade4)
#sélectionner les variables actives
donnees.active <- subset(donnees,select=1:18)
#construire le tableau de burt
tableau.burt <- acm.burt(donnees.active,donnees.active)
#calculer le khi-2
print(chisq.test(tableau.burt))
```

Nous obtenons :

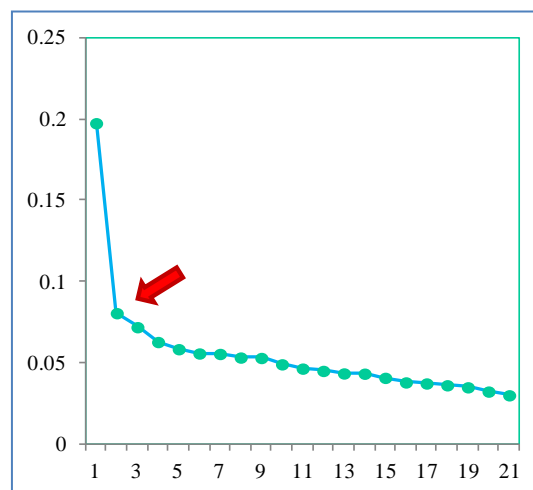
```
Pearson's Chi-squared test

data: tableau.burt
X-squared = 243317, df = 1444, p-value < 2.2e-16
```

Le choix du nombre de facteurs F à retenir est toujours délicat en analyse factorielle. Elle l'est d'autant plus en ACM que les parts d'inerties sont plus dispersées sur les axes.

La règle de Kaiser revient à sélectionner les facteurs pour lesquels les valeurs propres sont supérieures à leur moyenne c.-à-d. $\lambda_f \geq \frac{I_{tot}}{M-p} = \frac{\frac{M-1}{p} - 1}{M-p} = \frac{1}{p} = \frac{1}{18} = 0.0556$. Dans notre cas, nous retiendrons les 6 premiers. C'est ce que met en évidence Tanagra dans le tableau des valeurs propres (lignes avec fond bleu ciel). Malheureusement, cette procédure est souvent trop permissive. Nous serions emmenés à sélectionner trop d'axes⁶.

Finalement, un bon « éboulis des valeurs propres » avec détection du « coude » reste une alternative simple et viable. Ici, le choix de **F = 2** axes n'est pas dénué de bon sens (*4 axes aussi ? Cf. section 3.5*).



⁶ Cf. une règle plus stricte dans : G. Saporta, « [Inférence sur les valeurs propres et autres indices en ACP, AFC et ACM](#) ».

3.4.3 Caractérisation des facteurs – Variables actives

Values	Overall			Factor 1				Factor 2			
	Attribute = Value	Mass	Sq.Dist	Inertia	coord	v.test	cos2	ctr (%)	coord	v.test	cos2
Reading = y	0.0373	0.4883	0.0182	-0.34128	-44.766	0.2385	2.20	0.02481	3.255	0.0013	0.03
Reading = n	0.0182	2.0479	0.0373	0.6989	44.766	0.2385	4.50	-0.05081	-3.255	0.0013	0.06
-	-	-	-	-	-	Tot.ctr.	6.70	-	-	Tot.ctr.	0.09
Listening_music = y	0.0393	0.413	0.0162	-0.3373	-48.111	0.2755	2.26	-0.09957	-14.202	0.024	0.48
Listening_music = n	0.0162	2.4214	0.0393	0.81674	48.111	0.2755	5.48	0.2411	14.202	0.024	1.17
-	-	-	-	-	-	Tot.ctr.	7.74	-	-	Tot.ctr.	1.65
Cinema = y	0.0222	1.5016	0.0333	-0.76429	-57.17	0.389	6.56	-0.43047	-32.2	0.1234	5.10
Cinema = n	0.0333	0.6659	0.0222	0.50897	57.17	0.389	4.37	0.28667	32.2	0.1234	3.40
-	-	-	-	-	-	Tot.ctr.	10.93	-	-	Tot.ctr.	8.50
Show = y	0.016	2.4652	0.0395	-0.97212	-56.753	0.3834	7.66	-0.26952	-15.735	0.0295	1.44
Show = n	0.0395	0.4057	0.016	0.39435	56.753	0.3834	3.11	0.10933	15.735	0.0295	0.59
-	-	-	-	-	-	Tot.ctr.	10.77	-	-	Tot.ctr.	2.03
Exhibition = y	0.0172	2.2382	0.0384	-0.94475	-57.885	0.3988	7.75	0.01221	0.748	0.0001	0.00
Exhibition = n	0.0384	0.4468	0.0172	0.42211	57.885	0.3988	3.46	-0.00546	-0.748	0.0001	0.00
-	-	-	-	-	-	Tot.ctr.	11.21	-	-	Tot.ctr.	0.00
Computer = n	0.0347	0.6021	0.0209	0.44399	52.448	0.3274	3.46	0.1872	22.114	0.0582	1.51
Computer = y	0.0209	1.6609	0.0347	-0.7374	-52.448	0.3274	5.74	-0.31091	-22.114	0.0582	2.50
-	-	-	-	-	-	Tot.ctr.	9.20	-	-	Tot.ctr.	4.01
Sport = y	0.0205	1.715	0.0351	-0.70138	-49.092	0.2868	5.09	-0.30274	-21.19	0.0534	2.33
Sport = n	0.0351	0.5831	0.0205	0.40896	49.092	0.2868	2.97	0.17652	21.19	0.0534	1.36
-	-	-	-	-	-	Tot.ctr.	8.06	-	-	Tot.ctr.	3.68
Walking = y	0.0276	1.0127	0.028	-0.4175	-38.028	0.1721	2.43	0.3288	29.949	0.1068	3.70
Walking = n	0.028	0.9875	0.0276	0.41227	38.028	0.1721	2.40	-0.32468	-29.949	0.1068	3.65
-	-	-	-	-	-	Tot.ctr.	4.84	-	-	Tot.ctr.	7.35
Travelling = y	0.0222	1.4987	0.0333	-0.72931	-54.608	0.3549	5.98	-0.01233	-0.923	0.0001	0.00
Travelling = n	0.0333	0.6673	0.0222	0.48664	54.608	0.3549	3.99	0.00823	0.923	0.0001	0.00
-	-	-	-	-	-	Tot.ctr.	9.97	-	-	Tot.ctr.	0.01
Playing_music = n	0.0459	0.2103	0.0097	0.20976	41.928	0.2092	1.02	0.03376	6.749	0.0054	0.06
Playing_music = y	0.0097	4.7555	0.0459	-0.99749	-41.928	0.2092	4.86	-0.16055	-6.749	0.0054	0.31
-	-	-	-	-	-	Tot.ctr.	5.88	-	-	Tot.ctr.	0.37
Collecting = n	0.0499	0.1143	0.0057	0.07057	19.132	0.0436	0.13	-0.05443	-14.756	0.0259	0.18
Collecting = y	0.0057	8.7483	0.0499	-0.61733	-19.132	0.0436	1.10	0.47614	14.756	0.0259	1.60
-	-	-	-	-	-	Tot.ctr.	1.22	-	-	Tot.ctr.	1.79
Volunteering = y	0.0085	5.5393	0.0471	-0.77622	-30.231	0.1088	2.59	0.22851	8.9	0.0094	0.55
Volunteering = n	0.0471	0.1805	0.0085	0.14013	30.231	0.1088	0.47	-0.04125	-8.9	0.0094	0.10
-	-	-	-	-	-	Tot.ctr.	3.06	-	-	Tot.ctr.	0.65
Mechanic = y	0.0234	1.3744	0.0322	-0.43065	-33.671	0.1349	2.19	0.43928	34.346	0.1404	5.60
Mechanic = n	0.0322	0.7276	0.0234	0.31333	33.671	0.1349	1.60	-0.31962	-34.346	0.1404	4.07
-	-	-	-	-	-	Tot.ctr.	3.79	-	-	Tot.ctr.	9.67
Gardening = n	0.0334	0.6649	0.0222	0.17671	19.864	0.047	0.53	-0.54887	-61.697	0.453	12.46
Gardening = y	0.0222	1.5039	0.0334	-0.26576	-19.864	0.047	0.79	0.82542	61.697	0.453	18.74
-	-	-	-	-	-	Tot.ctr.	1.32	-	-	Tot.ctr.	31.21
Knitting = n	0.0462	0.2021	0.0093	0.04808	9.802	0.0114	0.05	-0.18328	-37.365	0.1662	1.92
Knitting = y	0.0093	4.9469	0.0462	-0.23784	-9.802	0.0114	0.27	0.90666	37.365	0.1662	9.52
-	-	-	-	-	-	Tot.ctr.	0.32	-	-	Tot.ctr.	11.45
Cooking = n	0.0312	0.7814	0.0244	0.31303	32.459	0.1254	1.55	-0.3253	-33.731	0.1354	4.09
Cooking = y	0.0244	1.2797	0.0312	-0.40058	-32.459	0.1254	1.98	0.41629	33.731	0.1354	5.24
-	-	-	-	-	-	Tot.ctr.	3.52	-	-	Tot.ctr.	9.33
Fishing = n	0.0493	0.1267	0.0062	0.00476	1.225	0.0002	0.00	-0.10363	-26.685	0.0848	0.66
Fishing = y	0.0062	7.8921	0.0493	-0.03754	-1.225	0.0002	0.00	0.81784	26.685	0.0848	5.18
-	-	-	-	-	-	Tot.ctr.	0.01	-	-	Tot.ctr.	5.84
TV = n2	0.0143	2.8975	0.0413	-0.18988	-10.225	0.0124	0.26	0.1004	5.406	0.0035	0.18
TV = n4	0.0148	2.7648	0.0408	0.15043	8.293	0.0082	0.17	-0.06669	-3.676	0.0016	0.08
TV = n1	0.0081	5.8708	0.0475	-0.27331	-10.339	0.0127	0.31	-0.13431	-5.081	0.0031	0.18
TV = n3	0.0117	3.7341	0.0438	-0.03349	-1.588	0.0003	0.01	0.25292	11.997	0.0171	0.93
TV = n0	0.0067	7.2625	0.0488	0.45949	15.629	0.0291	0.72	-0.34638	-11.782	0.0165	1.00
-	-	-	-	-	-	Tot.ctr.	1.46	-	-	Tot.ctr.	2.37

Figure 2 - Caractérisation des facteurs - Variables actives - Tanagra

La version 1.4.48 de Tanagra intègre un affichage optionnel qui regroupe en un seul emplacement toutes les informations relatives aux variables (modalités) actives (Figure 2). Il évite d'avoir à jongler avec plusieurs tableaux disséminés à différents endroits. Cette nouvelle fonctionnalité est particulièrement pratique lorsque le nombre de facteurs est relativement réduit. Dans le cas contraire, lorsqu'il est élevé, la profusion des chiffres peut rebuter l'utilisateur. Nous avons toujours la possibilité de revenir à un affichage en tableaux séparés.

La partie « **Overall** » indique le rôle de chaque modalité dans l'analyse : « mass » correspond à son poids c.-à-d. pour une modalité « k » d'effectif n_k d'une variable quelconque, elle est égale à $mass(k) = \frac{n_k}{n \times p}$; puis nous avons le carré de la distance de la modalité avec le barycentre **G** du nuage de points modalités $d^2(k, G) = \frac{n}{n_k} - 1$; enfin, nous avons l'inertie de la modalité, $inertie(k) = mass(k) \times d^2(k, G) = \frac{1}{p} \times \left(1 - \frac{n_k}{n}\right)$. La somme des inerties de l'ensemble des modalités est égale à l'inertie totale $I_{tot} = 1.66667$.

Ainsi pour la modalité « Reading = y » d'effectif $n_k = 5646$, nous obtenons $mass(k) = \frac{5646}{8403 \times 18} = 0.0373$; $d^2(k, G) = \frac{8403}{5646} - 1 = 0.4883$; $inertie(k) = \frac{1}{18} \times \left(1 - \frac{5646}{8403}\right) = 0.0182$.

Ensuite, pour les facteurs successifs, nous disposons pour chaque modalité : de sa coordonnée (coord) ; de sa valeur test⁷ ; de son cosinus carré indiquant la qualité de représentation (\cos^2) ; de sa contribution indiquant son impact (ctr).

Le déchiffrement d'un tel déluge d'informations est toujours compliqué. Tanagra surligne les coordonnées des modalités (avec des couleurs différentes selon que la valeur est positive ou négative) qui paraissent importantes pour chaque facteur. Il s'appuie sur la règle suivante : (1) le \cos^2 de la modalité est supérieur à $1/F_{max}$; (2) sa contribution supérieure à son poids ; (3) sa valeur test supérieure à la valeur définie lors du paramétrage de la méthode (par défaut **3**, Figure 1).

Le **premier facteur** caractérise les personnes qui apprécient les occupations qui « bougent » (cinéma, spectacle, voyages, sport, etc.) ou « modernes » (informatique) en opposition avec ceux qui, de toute manière, ne s'intéressent pas à grand-chose (pas de lecture, pas de musique). Dans la section suivante, en étudiant les variables supplémentaires, nous verrons que ces traits révèlent un antagonisme entre d'une part, les jeunes, célibataires, occupant des professions intermédiaires ou de management, et d'autre part, les personnes âgées, veufs, ayant occupé des emplois non qualifiés.

Sur le **second facteur**, les tenants des loisirs tranquilles se démarquent (pêche, cuisine, tricot, mécanique, marche, etc.). Il s'agit surtout de personnes en couple (mariés) dans la force de l'âge (45 – 65 ans). Cette préférence ne semble pas avoir de lien particulier avec le type de profession.

Nous nous en tiendrons à cette lecture très simplifiée dans ce tutoriel, notre principal objectif étant de comparer les outils.

A ce propos, SAS fournit les mêmes informations mais en plusieurs tableaux. Nous les avons réunies dans une seule copie d'écran en n'affichant que les premières modalités (Figure 3).

⁷ Cf. Tutoriel Tanagra, « [Interpréter la 'valeur test'](#) ».

Summary Statistics for the Column Points			
	Quality	Mass	Inertia
Reading_n	0.2398	0.0182	0.0320
Reading_y	0.2398	0.0373	0.0156
Listening_music_n	0.2995	0.0162	0.0337
Listening_music_y	0.2995	0.0393	0.0139

Column Coordinates		
	Dim1	Dim2
Reading_n	-0.6989	-0.0508
Reading_y	0.3413	0.0248
Listening_music_n	-0.8167	0.2411
Listening_music_y	0.3373	-0.0996

Squared Cosines for the Column Points		
	Dim1	Dim2
Reading_n	0.2385	0.0013
Reading_y	0.2385	0.0013
Listening_music_n	0.2755	0.0240
Listening_music_y	0.2755	0.0240

Partial Contributions to Inertia for the Column Points		
	Dim1	Dim2
Reading_n	0.0450	0.0006
Reading_y	0.0220	0.0003
Listening_music_n	0.0548	0.0117
Listening_music_y	0.0226	0.0048

Figure 3 - Caractérisation des facteurs - Variables actives - SAS

Notons que SAS fournit l'inertie relative (*Inertia*) des modalités c.-à-d. les valeurs sont normalisées de manière à ce que leur somme fasse 1. Ainsi, nous avons $inertie_{rel}(k) = \frac{inertie(k)}{I_{tot}} = \frac{1 - \frac{n_k}{M-p}}{1 - \frac{n_k}{M-p}}$. Pour la modalité « reading = y », $inertie_{rel}(k) = \frac{1 - \frac{5646}{8403}}{1 - \frac{5646}{8403}} = 0.0156$.

'Quality' correspond à la somme des COS² sur les facteurs de l'étude.

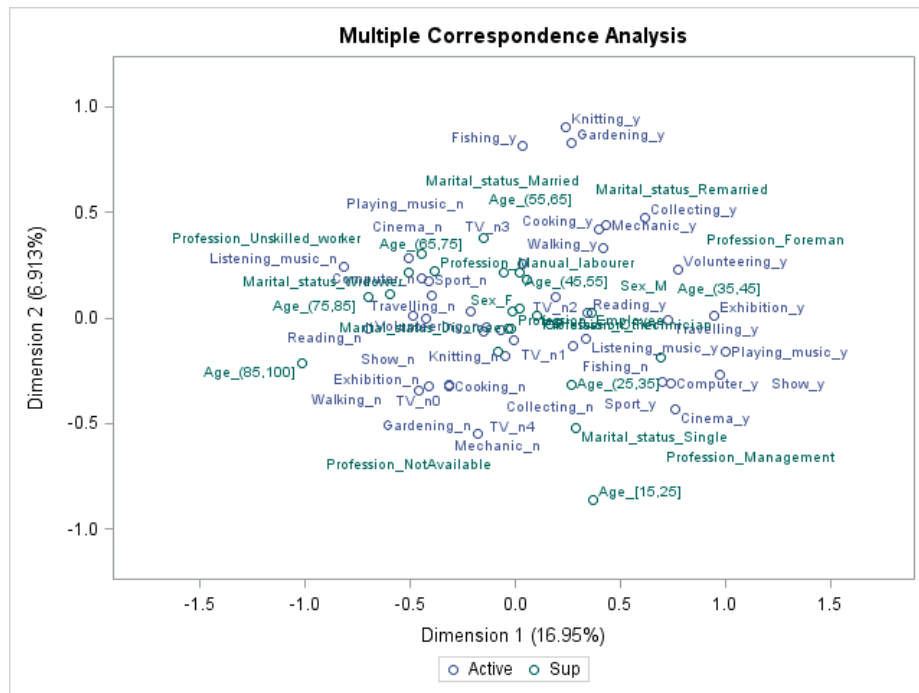
3.4.4 Caractérisation des facteurs – Variables supplémentaires qualitatives

Supplementary variables (discrete)						
Values	Factor 1			Factor 2		
	coord	v.test	cos2	coord	v.test	cos2
Sex = F	-0.01758	-1.779	0.0004	0.04203	4.253	0.0022
Sex = M	0.02143	1.779	0.0004	-0.05123	-4.253	0.0022
Age = (55,65]	0.15308	5.885	0.0041	0.37974	14.599	0.0254
Age = (45,55]	-0.02194	-1.064	0.0001	0.21262	10.309	0.0126
Age = (25,35]	-0.26738	-10.495	0.0131	-0.31486	-12.358	0.0182
Age = (75,85]	0.70149	15.862	0.0299	0.10066	2.276	0.0006
Age = (35,45]	-0.20096	-9.092	0.0098	-0.02025	-0.916	0.0001
Age = [15,25]	-0.36956	-11.416	0.0155	-0.86045	-26.580	0.0841
Age = (65,75]	0.44733	14.526	0.0251	0.30146	9.789	0.0114
Age = (85,100]	1.01451	9.400	0.0105	-0.21437	-1.986	0.0005
Marital_status = Married	0.05502	5.204	0.0032	0.21344	20.187	0.0485
Marital_status = Remarried	-0.05352	-1.103	0.0001	0.18011	3.710	0.0016
Marital_status = Single	-0.28849	-15.457	0.0284	-0.52252	-27.997	0.0933
Marital_status = Divorcee	0.03375	0.998	0.0001	-0.04854	-1.435	0.0002
Marital_status = Widower	0.50933	14.443	0.0248	0.21664	6.143	0.0045
Profession = Management	-0.69283	-24.024	0.0687	-0.18662	-6.471	0.0050
Profession = NotAvailable	0.08448	3.607	0.0015	-0.15983	-6.824	0.0055
Profession = Employee	0.01683	1.019	0.0001	0.03153	1.909	0.0004
Profession = Manual_labourer	0.38242	14.035	0.0234	0.22312	8.189	0.0080
Profession = Foreman	-0.36580	-10.381	0.0128	0.02161	0.613	0.0000
Profession = Other	-0.10025	-1.478	0.0003	0.01317	0.194	0.0000
Profession = Unskilled_worker	0.59571	17.614	0.0369	0.11339	3.353	0.0013
Profession = Technician	-0.16535	-3.393	0.0014	-0.03053	-0.626	0.0000

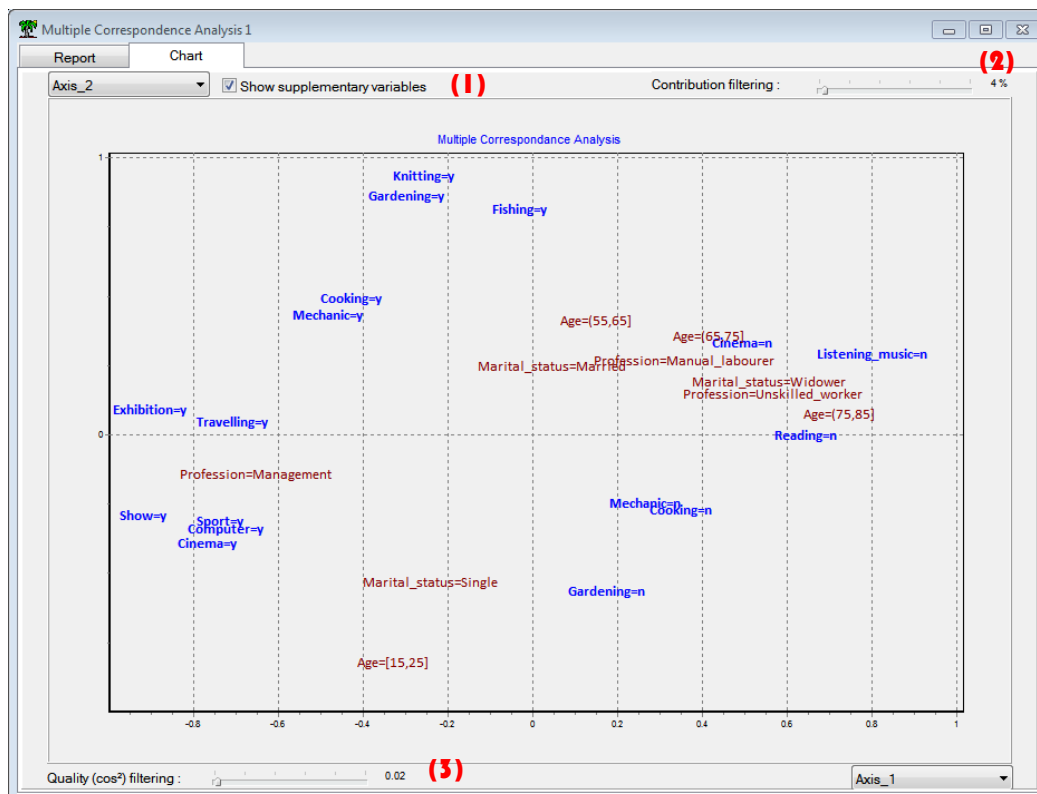
Figure 4 - Caractérisation des facteurs - Variables supplémentaires qualitatives – Tanagra

Les représentations graphiques permettent d'obtenir une vision synthétique des résultats de l'analyse factorielle (à condition de ne pas leur faire raconter des choses qui n'existent pas). Dans Tanagra, le **nuage des modalités actives et illustratives** sont accessibles dans l'onglet CHART. Bien évidemment, nous y retrouvons les informations décrites dans les tableaux précédents.

SAS intègre pareillement ce type de graphique dans son rapport.



Tanagra propose plusieurs options pour clarifier l'affichage qui peut être excessivement confus lorsque le nombre de modalités est élevé.



Il est possible : (1) d'intégrer ou non les variables supplémentaires dans le graphique ; (2) de filtrer les modalités selon leurs contributions aux axes (ici, seules les modalités actives contribuant plus de 4% sur au moins un des deux axes sont affichées) ; (3) de filtrer les modalités selon la qualité de leur représentation (ici, seules passent les modalités actives ou illustratives avec un \cos^2 supérieur à 2% sur au moins un des deux axes) ; et enfin, de perturber (très légèrement) aléatoirement la position des points pour éviter leurs superpositions (jittering)⁸.

3.4.7 Fonctions de projection

Tout comme en analyse en composantes principales, il doit être possible de projeter des individus supplémentaires dans l'espace factoriel. Tanagra sait le faire automatiquement. Mais il fournit également les informations idoines – les coefficients des fonctions scores – pour que l'on puisse le réaliser en dehors du logiciel (dans le tableur Excel par exemple).

Factor Score Coefficients
Applied to indicator matrix i.e. columns are dummy variables

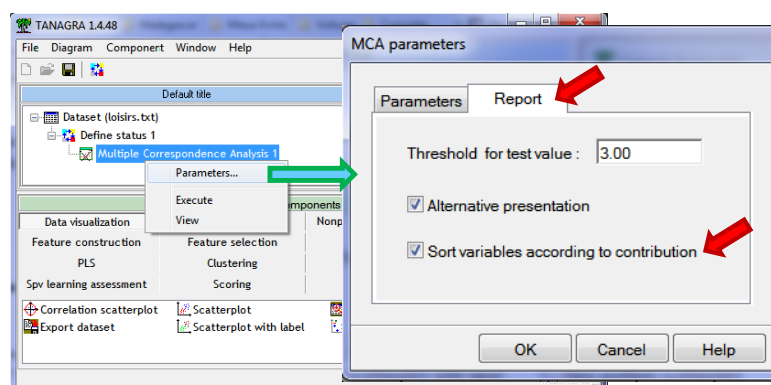
Attribute = Value	Axis_1	Axis_2
Reading = y	-0.0426403	0.0048539
Reading = n	0.0873221	-0.0099403
Listening_music = y	-0.0421431	-0.0194785
Listening_music = n	0.1020459	0.0471656
Cinema = y	-0.0954925	-0.0842114
Cinema = n	0.0635922	0.0560797
Show = y	-0.1214597	-0.0527260
Show = n	0.0492706	0.0213885

Figure 6 - Coefficients des fonctions de projections - 8 premières modalités

Nous n'affichons que les 8 premiers coefficients dans notre copie d'écran (Figure 6). On notera qu'ils s'appliquent sur les indicatrices des modalités des individus c.-à-d. après codage disjonctif complet de la description des individus.

3.4.8 Tri des variables selon leurs contributions

Mettre des couleurs facilite la lecture du tableau de caractérisation des facteurs (Figure 2). Il est possible d'aller plus loin en triant les variables de manière décroissante selon leurs contributions aux axes. Tanagra procède par blocs : il trie sur le premier axe les variables ayant une contribution supérieure à $(1/p)$; puis il fait de même pour les variables restantes sur le second axe, etc. Nous modifions le paramétrage du composant MULTIPLE CORRESPONDANCE ANALYSIS.



⁸ <http://blogs.sas.com/content/iml/2011/07/05/jittering-to-prevent-overplotting-in-statistical-graphics/>

Nous obtenons :

Values Attribute = Value	Overall				Factor 1			Factor 2				
	Mass	Sq.Dist	Inertia	coord	v.test	cos2	ctr (%)	coord	v.test	cos2	ctr (%)	
Exhibition = y	0.0172	2.2382	0.0384	-0.94475	-57.885	0.3988	7.75	0.01221	0.748	0.0001	0	
Exhibition = n	0.0384	0.4468	0.0172	0.42211	57.885	0.3988	3.46	-0.00546	-0.748	0.0001	0	
-	-	-	-	-	-	-	Tot.ctr.	11.21	-	-	Tot.ctr.	0
Cinema = y	0.0222	1.5016	0.0333	-0.76429	-57.17	0.389	6.56	-0.43047	-32.2	0.1234	5.1	
Cinema = n	0.0333	0.6659	0.0222	0.50897	57.17	0.389	4.37	0.28667	32.2	0.1234	3.4	
-	-	-	-	-	-	-	Tot.ctr.	10.93	-	-	Tot.ctr.	8.5
Show = y	0.016	2.4652	0.0395	-0.97212	-56.753	0.3834	7.66	-0.26952	-15.735	0.0295	1.44	
Show = n	0.0395	0.4057	0.016	0.39435	56.753	0.3834	3.11	0.10933	15.735	0.0295	0.59	
-	-	-	-	-	-	-	Tot.ctr.	10.77	-	-	Tot.ctr.	2.03
Travelling = y	0.0222	1.4987	0.0333	-0.72931	-4.608	0.3549	5.98	-0.01233	-0.923	0.0001	0	
Travelling = n	0.0333	0.6673	0.0222	0.48664	4.608	0.3549	3.99	0.00823	0.923	0.0001	0	
-	-	-	-	-	-	-	Tot.ctr.	9.97	-	-	Tot.ctr.	0.01
Computer = n	0.0347	0.6021	0.0209	0.44399	-52.448	0.3274	3.46	0.1872	22.114	0.0582	1.51	
Computer = y	0.0209	1.6609	0.0347	-0.7374	52.448	0.3274	5.74	-0.31091	-22.114	0.0582	2.5	
-	-	-	-	-	-	-	Tot.ctr.	9.2	-	-	Tot.ctr.	4.01
Sport = y	0.0205	1.715	0.0351	-0.70138	-49.092	0.2868	5.09	-0.30274	-21.19	0.0534	2.33	
Sport = n	0.0351	0.5831	0.0205	0.40896	49.092	0.2868	2.97	0.17652	21.19	0.0534	1.36	
-	-	-	-	-	-	-	Tot.ctr.	8.06	-	-	Tot.ctr.	3.68
Listening_music = y	0.0393	0.413	0.0162	-0.3373	-48.111	0.2755	2.26	-0.09957	-14.202	0.024	0.48	
Listening_music = n	0.0162	2.4214	0.0393	0.81674	48.111	0.2755	5.48	0.2411	14.202	0.024	1.17	
-	-	-	-	-	-	-	Tot.ctr.	7.74	-	-	Tot.ctr.	1.65
Reading = y	0.0373	0.4883	0.0182	-0.34128	-44.766	0.2385	2.2	0.02481	3.255	0.0013	0.03	
Reading = n	0.0182	2.0479	0.0373	0.6989	44.766	0.2385	4.5	-0.05081	-3.255	0.0013	0.06	
-	-	-	-	-	-	-	Tot.ctr.	6.7	-	-	Tot.ctr.	0.09
Playing_music = n	0.0459	0.2103	0.0097	0.20976	41.928	0.2092	1.02	0.03376	6.749	0.0054	0.06	
Playing_music = y	0.0097	4.7555	0.0459	-0.99749	-41.928	0.2092	4.86	-0.16855	-6.749	0.0054	0.31	
-	-	-	-	-	-	-	Tot.ctr.	5.88	-	-	Tot.ctr.	0.37
Gardening = n	0.0334	0.6649	0.0222	0.17671	19.864	0.047	0.53	-0.54887	-61.697	0.453	12.46	
Gardening = y	0.0222	1.5039	0.0334	-0.26576	-19.864	0.047	0.79	0.82542	61.697	0.453	18.74	
-	-	-	-	-	-	-	Tot.ctr.	1.82	-	-	Tot.ctr.	31.21
Knitting = n	0.0462	0.2021	0.0093	0.04808	9.802	0.0114	0.05	-0.18328	-37.365	0.1662	1.92	
Knitting = y	0.0093	4.9469	0.0462	-0.23784	-9.802	0.0114	0.27	0.90666	37.365	0.1662	9.52	
-	-	-	-	-	-	-	Tot.ctr.	0.32	-	-	Tot.ctr.	11.45
Mechanic = y	0.0234	1.3744	0.0322	-0.43065	-33.671	0.1349	2.19	0.43928	34.346	0.1404	5.6	
Mechanic = n	0.0322	0.7276	0.0234	0.31333	33.671	0.1349	1.6	-0.31962	-34.346	0.1404	4.07	
-	-	-	-	-	-	-	Tot.ctr.	1.79	-	-	Tot.ctr.	9.67
Cooking = n	0.0312	0.7814	0.0244	0.31303	32.459	0.1254	1.55	-0.3253	33.731	0.1354	4.09	
Cooking = y	0.0244	1.2797	0.0312	-0.40058	-32.459	0.1254	1.98	0.41629	-33.731	0.1354	5.24	
-	-	-	-	-	-	-	Tot.ctr.	3.52	-	-	Tot.ctr.	9.33
Walking = y	0.0276	1.0127	0.028	-0.4175	-38.028	0.1721	2.43	0.3288	29.949	0.1068	3.7	
Walking = n	0.028	0.9875	0.0276	0.41227	38.028	0.1721	1.4	-0.32468	-29.949	0.1068	3.65	
-	-	-	-	-	-	-	Tot.ctr.	4.84	-	-	Tot.ctr.	7.35
Fishing = n	0.0493	0.1267	0.0062	0.00476	1.225	0.0002	0	-0.10363	-26.685	0.0848	0.66	
Fishing = y	0.0062	7.8921	0.0493	-0.03754	-1.225	0.0002	0	0.81784	26.685	0.0848	5.18	
-	-	-	-	-	-	-	Tot.ctr.	0.01	-	-	Tot.ctr.	5.84
TV = n2	0.0143	2.8975	0.0413	-0.18988	-10.225	0.0124	0.26	0.1004	5.406	0.0035	0.18	
TV = n4	0.0148	2.7648	0.0408	0.15043	8.293	0.0082	0.17	-0.06669	-3.676	0.0016	0.08	
TV = n1	0.0081	5.8708	0.0475	-0.27331	-10.339	0.0127	0.31	-0.13431	-5.081	0.0031	0.18	
TV = n3	0.0117	3.7341	0.0438	-0.03349	-1.588	0.0003	0.01	0.25292	11.997	0.0171	0.93	
TV = n0	0.0067	7.2625	0.0488	0.45949	15.629	0.0291	0.72	-0.34638	-11.782	0.0165	1	
-	-	-	-	-	-	-	Tot.ctr.	1.46	-	-	Tot.ctr.	2.37
Collecting = n	0.0499	0.1143	0.0057	0.07057	19.132	0.0436	0.13	-0.05443	-14.756	0.0259	0.18	
Collecting = y	0.0057	8.7483	0.0499	-0.61733	-19.132	0.0436	1.1	0.47614	14.756	0.0259	1.6	
-	-	-	-	-	-	-	Tot.ctr.	1.22	-	-	Tot.ctr.	1.79
Volunteering = y	0.0085	5.5393	0.0471	-0.77622	-30.231	0.1088	2.59	0.22851	8.9	0.0094	0.55	
Volunteering = n	0.0471	0.1805	0.0085	0.14013	30.231	0.1088	0.47	-0.04125	-8.9	0.0094	0.1	
-	-	-	-	-	-	-	Tot.ctr.	3.06	-	-	Tot.ctr.	0.65

Figure 7 - Caractérisation des facteurs - Tri selon les contributions

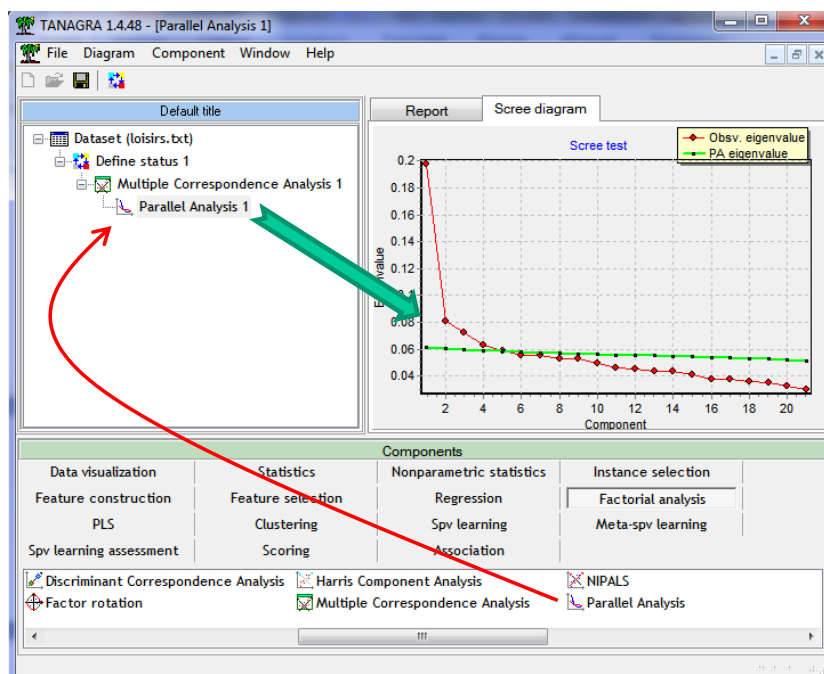
Nous visualisons mieux le clivage entre les variables sur les deux premiers axes. Nous constatons également que « cinéma » joue un rôle ambivalent.

3.5 Détection du nombre adéquat de facteurs – Analyse parallèle

La détermination du nombre « optimal » de facteurs est un problème délicat soulevons-nous plus haut. A la lumière du « scree plot », nous avons proposé une modélisation en 2 axes, tout en soulignant qu'un second coude semblait perceptible pour une solution en 4 axes.

Dans cette section nous utilisons l'analyse parallèle⁹ pour obtenir une autre estimation. L'idée repose sur le calcul des valeurs propres de chaque facteur sous l'hypothèse nulle (H0) d'absence de liaison. Pour ce faire, nous nous appuyons sur une méthode de Monte-Carlo. Elle consiste à enchaîner les actions suivantes : (1) modifier aléatoirement la position des valeurs dans les colonnes, ces dernières étant traitées individuellement, indépendamment les unes des autres, on casse ainsi l'éventuel lien pouvant exister entre les variables ; (2) lancer l'ACM sur ces données modifiées ; (3) collecter les valeurs propres associées à chaque facteur. En répétant un certain nombre de fois (R = 200 fois par défaut dans Tanagra) cette procédure, nous obtenons la distribution empirique des valeurs propres sous H0 [$\lambda_k(H0)$] pour chaque facteur. Nous considérons alors qu'un axe est intéressant si sa valeur propre observée λ_k est supérieure au quantile d'ordre 0.95 (pour un test unilatéral à 5%) de $\lambda_k(H0)$.

Nous insérons le composant PARALLEL ANALYSIS à la suite de l'analyse des correspondances multiples dans le diagramme. Nous actionnons directement le menu VIEW.



Avec cette approche, une modélisation en 4 facteurs semble crédible en définitive. En creusant un peu, je me suis rendu compte que le 4^{ème} facteur retraçait une relation assez incongrue entre l'appétence à la Télévision et le bénévolat, expliquée par la classe d'âge des personnes.

⁹ Franklin, Gibson, Robertson, Pohlmann, Fralish, « [Parallel Analysis : A Method for Determining Significant Components](#) », Journal of Vegetation Science, 6 :99-106, 1995. L'idée est développée pour l'analyse en composantes principales mais elle est très facilement transposable à l'analyse des correspondances multiples.


```
#coordonnées
print(ade4.acm$co)
#evaluation des colonnes
eval.acm <- inertia.dudi(ade4.acm,col.inertia=T)
#contributions
print(eval.acm$col.abs/100)
#cos2 (signé) + inertie relative
print(eval.acm$col.rel/100)
```

Nous affichons uniquement les 8 premières modalités.

Coordinates			Contributions			Cos ²			
	Comp1	Comp2		Comp1	Comp2		Comp1	Comp2	con.tra
Reading.n	0.698896310	-0.050812554	Reading.n	4.50	0.06	Reading.n	23.85	-0.13	3.20
Reading.y	-0.341278273	0.024812294	Reading.y	2.20	0.03	Reading.y	-23.85	0.13	1.56
Listening_music.n	0.816740533	0.241100351	Listening_music.n	5.48	1.17	Listening_music.n	27.55	2.40	3.37
Listening_music.y	-0.337298596	-0.099569945	Listening_music.y	2.26	0.48	Listening_music.y	-27.55	-2.40	1.39
Cinema.n	0.508970535	0.286667333	Cinema.n	4.37	3.40	Cinema.n	38.90	12.34	1.90
Cinema.y	-0.764289187	-0.430470386	Cinema.y	6.56	5.10	Cinema.y	-38.90	-12.34	2.86
Show.n	0.394345359	0.109333343	Show.n	3.11	0.59	Show.n	38.34	2.95	1.37
Show.y	-0.972122292	-0.269523597	Show.y	7.66	1.44	Show.y	-38.34	-2.95	3.39

4.2 Caractérisation à l'aide des variables supplémentaires

Des instructions additionnelles sont nécessaires pour la prise en compte des variables qualitatives supplémentaires, lesquelles doivent être également recodées en indicatrices.

```
#colonnes supplémentaires - var. catégorielles
indicators.sup <- acm.disjonctif(subset(donnees,select=19:22))
colsup <- supcol(ade4.acm,indicators.sup)
print(colsup$cosup)
```

Voici les coordonnées des modalités sur les 2 premiers facteurs :

	Comp1	Comp2
Sex.F	-0.01758075	0.04202668
Sex.M	0.02142930	-0.05122661
Age..25.35.	-0.26738296	-0.31485847
Age..35.45.	-0.20095930	-0.02025289
Age..45.55.	-0.02193895	0.21262313
Age..55.65.	0.15307920	0.37973843
Age..65.75.	0.44733456	0.30146093
Age..75.85.	0.70149394	0.10066300
Age..85.100.	1.01450515	-0.21436885
Age..15.25.	-0.36955960	-0.86044927

Pour la variable supplémentaire quantitative, nous calculons les corrélations avec les facteurs.

```
#colonne supplémentaire - var. quantitative
correl <- sapply(ade4.acm$li,function(x){cor(x,donnees$nb_activitees)})
print(correl)
```

Nous obtenons :

```

> #colonne supplémentaire - var. quantitativie
> correl <- sapply(ade4.acm$li,function(x){cor(x,donnees$nb_activitees)})
> print(correl)
      Axis1      Axis2
-0.9753459  0.1980007

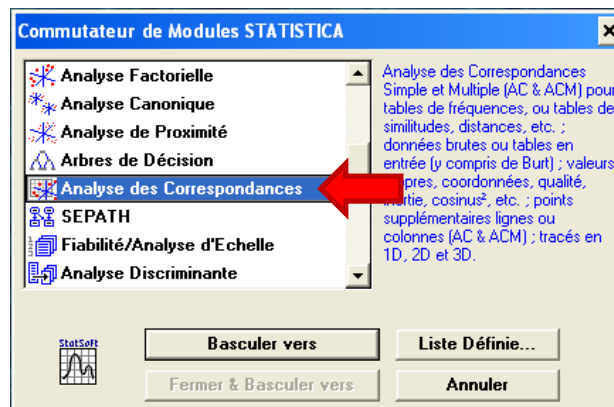
```

5 ACM avec STATISTICA 5.5

Ma version de STATISTICA est ancienne mais convient parfaitement pour ce que nous souhaitons réaliser. De toute manière, l'ACM n'a pas connu d'évolutions significatives ces dernières années.

5.1 Démarrage, importation des données et paramétrage de la méthode

Au démarrage de STATISTICA, nous choisissons le module 'Analyse des Correspondances'.



Nous importons le fichier « loisirs.txt ». Nous définissons ensuite les paramètres de la méthode.

The screenshot shows the 'Analyse des Correspondances Multiple (ACM)' dialog box with the following settings and annotations:

- (1)** The 'Méthode' section has 'Analyse des Correspondances Multiple (ACM)' selected.
- (2)** The 'Entrée' section has 'Données brutes (nécessite tabulation)' selected.
- (3)** The 'Sélectionner les Variables de Classement Spécifiant la Tab' dialog is open, showing a list of variables (1-READING to 23-NB_ACTIV) with variables 1-12 selected.
- (4)** The 'Sélectionner les Variables Supplémentaires' dialog is open, showing a list of variables (1-READING to 22-PROFESSI) with variables 19-22 selected.

Nous choisissons une ACM (1), sur des données brutes c.-à-d. un tableau individus x variables (2), nous sélectionnons les variables de l'étude (toutes sauf NB_ACTIVITEES, comme pour SAS je n'ai pas su paramétrer la gestion des variables illustratives quantitatives) (3), et nous spécifions les variables supplémentaires (4).

5.2 Accès aux résultats

Un panneau résume les principaux résultats de l'analyse. Nous observons, entre autres, les valeurs propres associées aux facteurs.

Différents boutons donnent accès aux sorties détaillées. En cliquant sur le bouton 'Coordonnées colonne', nous obtenons les informations relatives aux modalités pour les deux premiers axes.

Dans la partie haute, nous avons les variables actives...

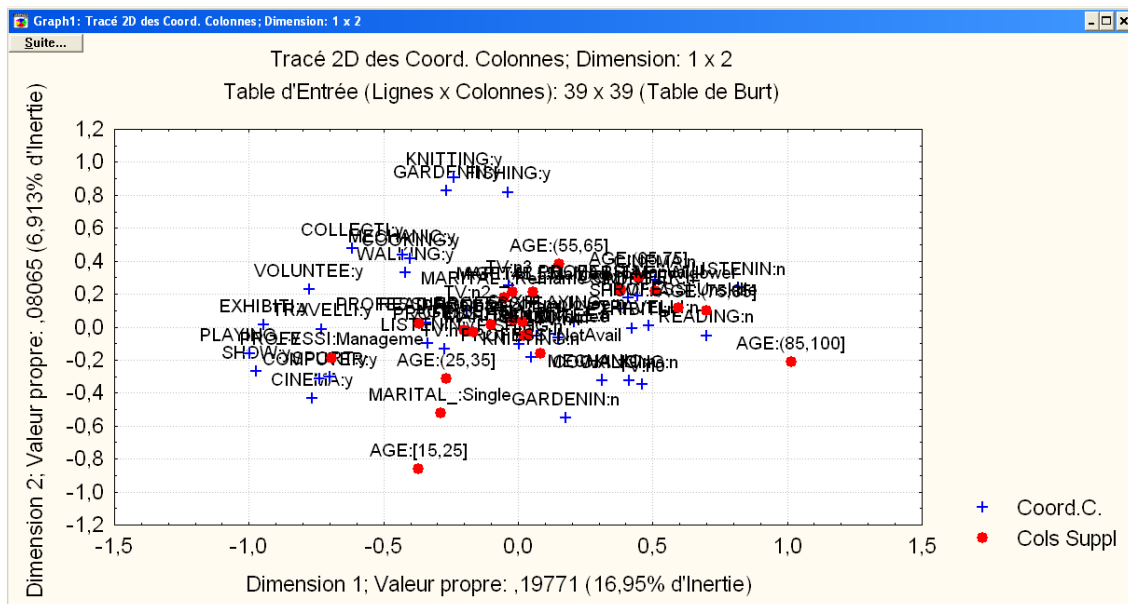
NonLigne	Ligne Numéro	Coord. Dim. 1	Coord. Dim. 2	Masse	Qualité	Inertie Relative	Inertie Dim. 1	Cosinus² Dim. 1	Inertie Dim. 2	Cosinus² Dim. 2
READING :y	1	-.3413	.0248	.0373	.2398	.0156	.0220	.2385	.0003	.0013
READING :n	2	.6989	-.0508	.0182	.2398	.0320	.0450	.2385	.0006	.0013
LISTENIN :y	3	-.3373	-.0996	.0393	.2995	.0139	.0226	.2755	.0048	.0240
LISTENIN :n	4	.8167	.2411	.0162	.2995	.0337	.0548	.2755	.0117	.0240
CINEMA :y	5	-.7643	-.4305	.0222	.5124	.0286	.0656	.3890	.0510	.1234
CINEMA :n	6	.5090	.2867	.0333	.5124	.0190	.0437	.3890	.0340	.1234
SHOW :y	7	-.9721	-.2695	.0160	.4128	.0339	.0766	.3834	.0144	.0295
SHOW :n	8	.3943	.1093	.0395	.4128	.0137	.0311	.3834	.0059	.0295
EXHIBITI :y	9	-.9447	.0122	.0172	.3989	.0329	.0775	.3988	.0000	.0001
EXHIBITI :n	10	.4221	-.0055	.0384	.3989	.0147	.0346	.3988	.0000	.0001
COMPUTER :n	11	.4440	.1872	.0347	.3856	.0179	.0346	.3274	.0151	.0582
COMPUTER :y	12	-.7374	-.3109	.0209	.3856	.0297	.0574	.3274	.0250	.0582
SPORT :y	13	-.7014	-.3027	.0205	.3403	.0301	.0509	.2868	.0233	.0534
SPORT :n	14	.4090	.1765	.0351	.3403	.0175	.0297	.2868	.0136	.0534
WALKING :y	15	-.4175	.3288	.0276	.2789	.0240	.0243	.1721	.0370	.1068
WALKING :n	16	.4123	-.3247	.0280	.2789	.0237	.0240	.1721	.0365	.1068
TRAVELLI :y	17	-.7293	-.0123	.0222	.3550	.0286	.0598	.3549	.0000	.0001
TRAVELLI :n	18	.4866	.0082	.0333	.3550	.0191	.0399	.3549	.0000	.0001
PLAYING :n	19	.2098	.0338	.0459	.2146	.0083	.0102	.2092	.0006	.0054
PLAYING :y	20	-.9975	-.1606	.0097	.2146	.0393	.0486	.2092	.0031	.0054
COLLECTI :n	21	.0706	-.0544	.0499	.0695	.0049	.0013	.0436	.0018	.0259
COLLECTI :y	22	-.6173	.4761	.0057	.0695	.0427	.0110	.0436	.0160	.0259
VOLUNTEE :y	23	-.7762	.2285	.0085	.1182	.0403	.0259	.1088	.0055	.0094
VOLUNTEE :n	24	.1401	-.0413	.0471	.1182	.0073	.0047	.1088	.0010	.0094
MECHANIC :y	25	-.4306	.4393	.0234	.2753	.0276	.0219	.1349	.0560	.1404
MECHANIC :n	26	.3133	-.3196	.0322	.2753	.0201	.0160	.1349	.0407	.1404

Dans la partie basse, nous disposons des informations sur les variables supplémentaires.

Coordonnées Colones et Contributions à l'Inertie										
Table d'Entrée (Lignes x Colones): 39 x 39 (Table de Burt)										
Inertie Totale=1.1667										
NonLigne	Ligne Numéro	Coord. Dim. 1	Coord. Dim. 2	Masse	Qualité	Inertie Relative	Inertie Dim. 1	Cosinus ² Dim. 1	Inertie Dim. 2	Cosinus ² Dim. 2
SEX: F		-.0176	.0420		.0025			.0004		.0022
SEX: M		.0214	-.0512		.0025			.0004		.0022
AGE: (55,65]		.1531	.3797		.0295			.0041		.0254
AGE: (45,55]		-.0219	.2126		.0128			.0001		.0126
AGE: (25,35]		-.2674	-.3149		.0313			.0131		.0182
AGE: (75,85]		.7015	.1007		.0306			.0299		.0006
AGE: (35,45]		-.2010	-.0203		.0099			.0098		.0001
AGE: [15,25]		-.3696	-.8604		.0996			.0155		.0841
AGE: (65,75]		.4473	.3015		.0365			.0251		.0114
AGE: (85,100]		1.0145	-.2144		.0110			.0105		.0005
MARITAL: Married		.0550	.2134		.0517			.0032		.0485
MARITAL: Remarrie		-.0535	.1801		.0018			.0001		.0016
MARITAL: Single		-.2885	-.5225		.1217			.0284		.0933
MARITAL: Divorcee		.0337	-.0485		.0004			.0001		.0002
MARITAL: Widower		.5093	.2166		.0293			.0248		.0045
PROFESSI: Manageme		-.6928	-.1866		.0737			.0687		.0050
PROFESSI: NotAvail		.0845	-.1598		.0071			.0015		.0055
PROFESSI: Employee		.0168	.0315		.0006			.0001		.0004
PROFESSI: Manual_l		.3824	.2231		.0314			.0234		.0080
PROFESSI: Foreman		-.3658	.0216		.0129			.0128		.0000
PROFESSI: Other		-.1003	.0132		.0003			.0003		.0000
PROFESSI: Unskille		.5957	.1134		.0383			.0369		.0013
PROFESSI: Technici		-.1654	-.0305		.0014			.0014		.0000

« Qualité » correspond à la somme des \cos^2 sur les facteurs générés.

Nous pouvons obtenir la représentation des modalités actives et supplémentaires dans le plan factoriel en actionnant le bouton 'Tracé des coordonnées - 2D'.



6 Conclusion

Concernant l'analyse des correspondances multiples, et nous l'avons constaté dans la majorité de nos comparatifs, nous obtenons exactement les mêmes résultats numériques d'un logiciel à l'autre. Et pour cause. Les algorithmes sont parfaitement identifiés. Seul le mode de présentation peut différer parfois. Mais ce n'est en rien rédhibitoire. L'étudiant qui se présente à un entretien ne doit donc pas avoir d'inquiétude s'il n'a pas vu tel ou tel outil à l'Université. Seule la maîtrise des techniques compte véritablement.