

Université : UADB

Domaine : Sciences et Technologies

Mention : Mathématiques appliquées



Etablissement : UFR/SATIC

Niveau : Master I

Spécialité : Statistique et Informatique
Décisionnelle

PROJET D'ANALYSE DES DONNÉES



RÉALISÉ PAR :

DALYÓ SID OUSMANE OURBA

PROFESSEUR :

DR CHEIKH TIDIANE SECK

Table des matières

Liste des tableaux et graphiques.....	3
Liste des tableaux	3
Liste des graphiques	3
Objectifs	4
A) Exploration des liaisons deux à deux entre les variables du tableau de données.	4
B) Analyse en composante principale	6
4) Interprétation du premier plan principal du nuage des individus par la mise en évidence d'une typologie des éleveurs.	7
C) Classification	12
Annexes : liste des codes R utilisés dans le projet.	19

Liste des tableaux et graphiques

Liste des tableaux

Tableau 1: résultat des croisements du test de Fisher.....	5
Tableau 2: contribution et cosinus carré des individus extrêmes	7
Tableau 3: individus dont les contributions sont supérieurs à la moyenne des contribution sur la dimension 1	8
Tableau 4: individus dont les contributions sont supérieurs à la moyenne des contribution sur la dimension 2	10
Tableau 5: Modalités significatives sur la dimension 1.....	11
Tableau 6: Modalités significatives sur la dimension 2.....	12
Tableau 7: variables caractérisant la partition en 10 classes	13
Tableau 8: classe 1.....	14
Tableau 9: classe 2.....	14
Tableau 10: classe 3.....	15
Tableau 11: classe 4.....	15
Tableau 12: description par les composantes principales.....	16
Tableau 13: description des individus de la classe 1	16
Tableau 14: description des individus de la classe 2	17
Tableau 15: description des individus de la classe 3	17
Tableau 16: description des individus de la classe 4	17

Liste des graphiques

Graphique 1: Diagramme des valeurs propres.....	6
Graphique 2: nuage des individus	7
Graphique 3: nuages des individus après suppressions des individus extrêmes	8
Graphique 4: nuages des modalités après suppression des individus extrêmes	11
Graphique 5: Visualisation des classes et le dendrogramme sur le premier plan principal de l'ACM.....	18

Objectifs

L'objectif de ce projet est d'analyser des données sur les éleveurs de "ladoum" de la région de Thiès. Spécifiquement on cherchera à faire une typologie de ces éleveurs et à les caractériser pour mieux comprendre la situation de ce secteur d'élevage. Les données se trouvent dans le fichier "ladoum.xls" donné lors du cours.

A) Exploration des liaisons deux à deux entre les variables du tableau de données.

Le test de khi-deux permet de calculer la liaison entre deux variables qualitatives. Cependant, la condition d'application du khi-deux stipule que toutes les valeurs du tableau de contingence obtenu lors du croisement des variables doivent être supérieures ou égales à 5. Nos variables ne respectant pas cette règle alors nous ne pouvons l'utiliser. Nous allons alors utiliser un test équivalent non paramétrique : le test exact de Fisher. On pose les hypothèses suivantes :

H_0 : les variables sont indépendantes

H_1 : les variables ne sont pas indépendantes.

Le tableau 1 ci-dessous nous montre l'ensemble des croisements des tests effectués.

Les valeurs dont les p-value sont colorés en bleues sont des p-value < 5% donc significatif.

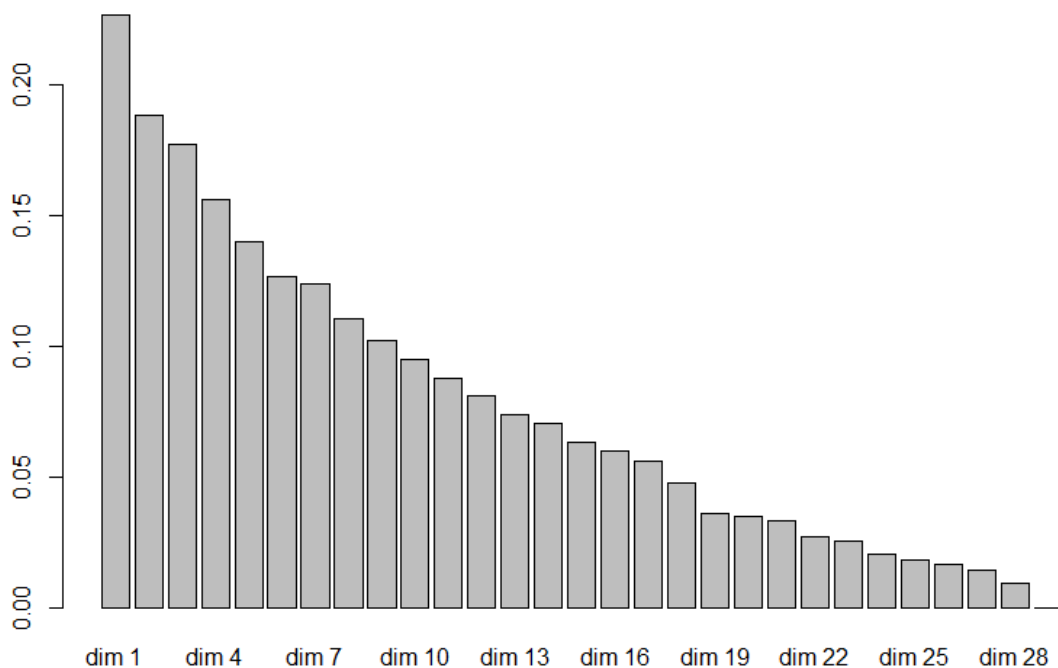
Tableau 1: résultat des croisements du test de Fisher.

Variables	age	ethnie	statut. matrimonial	niveau_ etude	type_ habitat	type_ propriete	act_ princ	duree_ elevage	raison_ elev	localisatio_ bergerie	nbre_ brebis	total_ ovin
sexe	0,221	0,3171	0,0261	0,2099	0,3413	0,3386	0,03638	0,7244	0,5259	0,1458	0,2883	0,1464
age		0,001912	0,02015	0,01625	0,06683	0,00534	0,1438	0,04813	0,333	0,9019	0,3832	0,499
ethnie			0,2945	0,08632	0,01261	0,03094	0,706	0,3245	0,4932	0,1317	0,9568	0,4118
statut.matrimonial				0,2383	1	0,1446	0,3298	0,571	1	1	1	0,2371
niveau_etude					0,6059	0,01466	0,01674	0,8562	0,6427	0,05613	0,5308	0,1914
type_habitat						1	0,02561	0,6072	0,43	0,5132	1	0,5555
type_propriete							0,1532	0,1721	0,8065	0,2363	0,5159	0,5382
act_princ								0,7751	0,796	0,007893	0,0592	0,189
duree_elevage									0,0293	0,1058	0,4691	0,1593
raison_elev										0,07549	0,1318	0,557
localisatio_ bergerie											0,06485	0,01198
nbre_brebis												1,89E-07

Conclusion : les résultats du tableau 1 nous montrent que chaque variable est au moins liée à une autre variable. On peut donc appliquer une Analyse en composante multiple pour aller au-delà de la liaison entre variables et chercher les modalités qui pourraient être associées.

B) Analyse en composante multiple

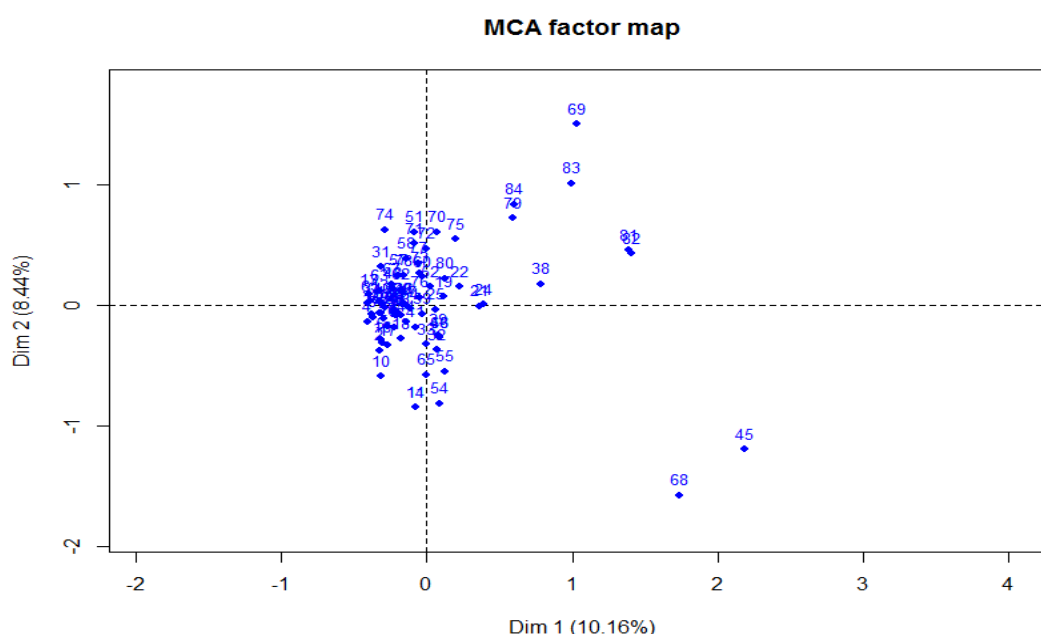
- 1) Pour trouver le nombre d'axe à retenir, on utilise le critère du coude sur le diagramme de décroissance des valeurs propres.
- 2) L'inertie d'un axe mesure la liaison entre l'axe et les variables. Cette inertie est donc égale à la moyenne des rapports de corrélations des variables avec l'axe.
- 3) Après l'exécution de l'ACM avec le package FactoMineR, on visualise le diagramme des valeurs propres. On remarque ainsi, un coude (creux) notable entre la 3^{ième} et la 4^{ième} dimension, on retiendra donc alors 3 axes dans la suite de notre étude.



Graphique 1: *Diagramme des valeurs propres*

4) Interprétation du premier plan principal du nuage des individus par la mise en évidence d'une typologie des éleveurs.

L'analyse de la représentation du graphique 1 des individus détecte des individus rares qui influencent très fortement les résultats observés. On notera dans un premier temps le caractère extrême de ces individus avant de les retirer pour procéder à la suite de l'analyse. En observant le graphique, on peut noter que 7 individus sont très particuliers et contribuent énormément à la construction du premier plan principal. Ce sont les individus 45, 68, 69, 74, 81, 82, et 83.



Graphique 2: nuage des individus

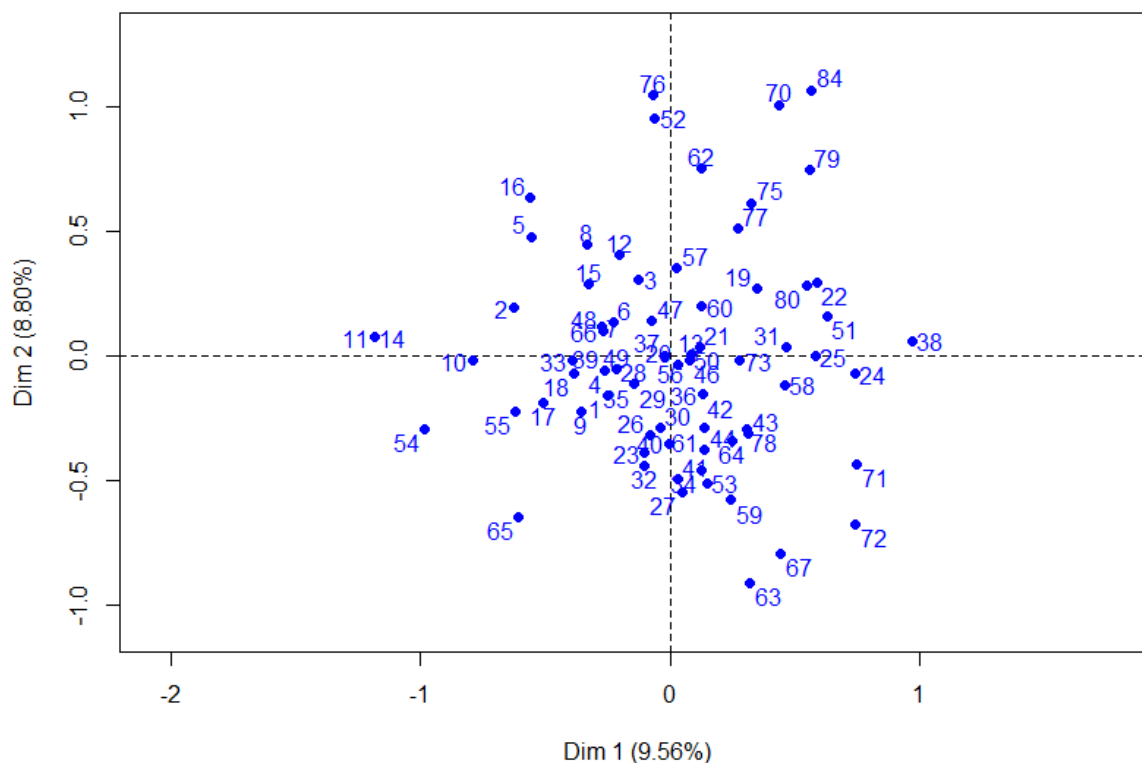
L'observation des contributions et des cosinus carrés de ces individus dans le tableau 3 nous confirme notre assertion. Le cumul de leur contribution à la construction du premier plan principal atteint alors 62,6%.

Tableau 2: contribution et cosinus carré des individus extrêmes

individus	dim 1	ctr	cos2	dim 2	ctr	cos2
45	2,185	25,068	0,564	-1,191	8,969	0,168
68	1,741	15,919	0,355	-1,575	15,675	0,29
69	1,033	5,602	0,105	1,504	14,3	0,223
74	-0,291	0,446	0,01	0,628	2,491	0,048
81	1,384	10,064	0,474	0,459	1,333	0,052
82	1,409	10,418	0,497	0,432	1,179	0,047
83	0,995	5,202	0,269	1,01	6,445	0,277

Après la suppression de ses individus, notre base de données comporte alors 77 individus et nous obtenons le graphique des individus suivants :

Graphes des individus après suppression des individus extrêmes



Graphique 3: nuages des individus après suppressions des individus extrêmes

Pour interpréter ce graphique en fonction des dimensions, on se basera sur les individus éloignés du centre du graphique. Les individus sont dites éloignés du centre de gravité si leur contributions est grande c'est-à-dire nettement supérieure à la contribution moyenne qui est

$\frac{1}{\text{nombre d'individus}} * 100$. Dans notre cas on a une contribution moyenne de

$\frac{1}{77} * 100 = 1,30\%$ environ. Cependant, On n'interprète que les individus bien représentés sur un plan.

Tableau 3: individus dont les contributions sont supérieurs à la moyenne des contribution sur la dimension 1

individus	dim 1	ctr	cos2
10	-0,793	4,444	0,350
11	-1,185	9,927	0,640
14	-1,185	9,927	0,640
16	-0,562	2,229	0,197
38	0,970	6,649	0,170

54	-0.984	6,837	0,413
67	0.445	1,397	0,102
70	0.435	1,337	0,061
71	0.749	3,961	0,138
72	0.742	3,890	0,123
84	0.567	2,270	0,418

La dimension 1 semble opposer les individus 38, 67, 70, 71, 72, 84 situé à droite de l'axe du graphique et caractérisé par des coordonnées positive sur l'axe à des individus comme 10, 11, 14, 16, 54 situé à gauche du graphique caractérisé par une coordonnée fortement négative sur l'axe.

On remarque que le groupe auquel les individus 67, 71 et 72 appartiennent (caractérisés par une coordonnée positive sur l'axe) possède :

- une forte fréquence des modalités sexe=F, total_ovin=d:10-30, ethnie=Peul, type_propriete=Collectif, nbre_brebis=b:10-20 et statut.matrimonial=Célibataire .
- une faible fréquence des modalités sexe=H, total_ovin=o:0-10, duree_elevage=d:20-40, statut.matrimonial=Marié, nbre_brebis=b:0-10 et type_propriete=Individuel.

Le groupe auquel les individus 70 et 84 appartiennent (caractérisés par une coordonnée positive sur l'axe) possède :

- une forte fréquence des modalités type_habitat=Ferme, age=a:25-45, ethnie=Serer et duree_elevage=d:0-10.
- une faible fréquence des modalités type_habitat=Maison, ethnie=Peul, age=a:60 et + et act_princ=AP8.

Le groupe auquel les individus 14, 11, 54, 16 et 10 appartiennent (caractérisés par une coordonnée négative sur l'axe) possède :

- une forte fréquence des modalités age=a:60 et +, raison_elev=Heritage, sexe=H, localisatio_bergerie=Peri-urbain, duree_elevage=d:20-40 et type_habitat=Maison.
- une faible fréquence des modalités duree_elevage=d:0-10, age=a:25-45, raison_elev=Economique, sexe=F, localisatio_bergerie=Urbain et type_habitat=Ferme (du plus rare au plus commun).

Tableau 4: individus dont les contributions sont supérieures à la moyenne des contribution sur la dimension 2

individus	Dim 2	Ctr	Cos2
52	0,952	6,949453	0,002
62	0,751	4,325617	0,025
63	-0,908	6,326378	0,035
67	-0,791	4,804797	0,044
70	1,007	7,774396	0,024
71	0,435	1,452891	0,305
72	-0,435	3,484414	0,324
75	0,612	2,875943	0,192
76	1,044	8,354685	0,014
79	0,744	4,245536	0,015
84	1,065	8,698113	0,147

La dimension 2 semble opposer les individus 52, 62, 70, 75, 76, 79 et 84 qui se situe en haut de l'axe et caractérisé par des coordonnées fortement positive au individus 63, 67, 71 et 72 qui eux se situe en bas de l'axe et se caractérisant par des coordonnées fortement négatives sur l'axe.

Le groupe auquel les individus 52, 62, 70, 75, 76, 79 et 84 appartiennent (caractérisés par des coordonnées positives sur l'axe) possède :

- une forte fréquence des modalités type_habitat=Ferme, age=a:25-45, ethnie=Serer et duree_elevage=d:0-10.
- une faible fréquence des modalités type_habitat=Maison, ethnie=Peul, age=a:60 et + et act_princ=AP8.

Le groupe auquel les individus 63, 67, 71 et 72 appartiennent (caractérisés par des coordonnées négatives sur l'axe) possède :

- une forte fréquence des modalités sexe=F, total_ovin=d:10-30, ethnie=Peul, type_propriete=Collectif, nbre_brebis=b:10-20 et statut.matrimonial=Célibataire

AP8	-0.27883291	2.605428e-02
Urbain	-0.04332883	1.656881e-02
Marié	-0.38258129	1.234459e-02
b:0-10	-0.21706043	1.020945e-02
H	-0.16694692	1.003541e-02
o:0-10	-0.17927219	1.827540e-04
Peri-urbain	-0.62179964	3.959629e-06
a:60 et +	-0.45817585	5.147950e-07
d:20-40	-0.46015940	2.894990e-08
Heritage	-0.68215237	1.160726e-08

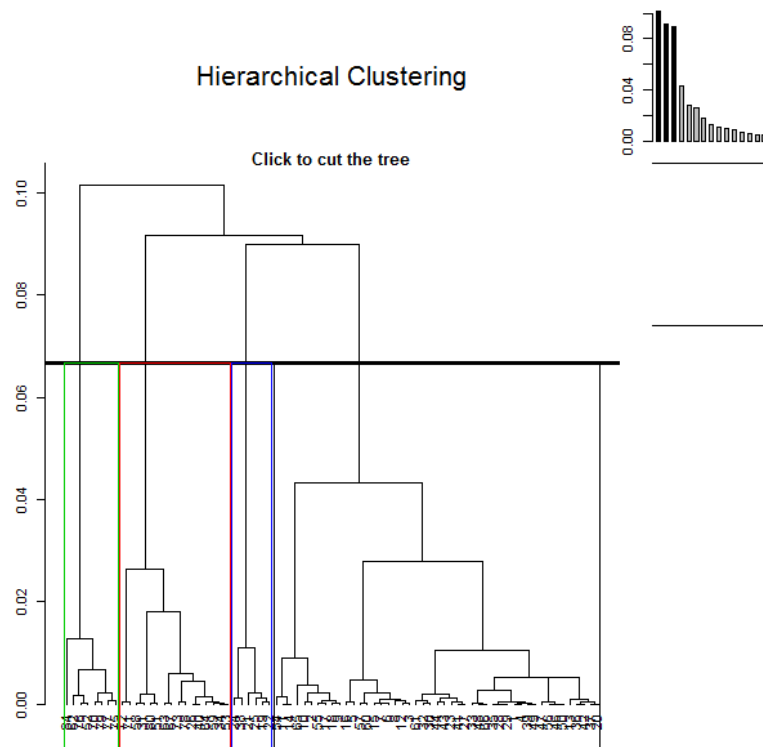
La dimension 2 est caractérisée par les modalités significatives suivantes :

Tableau 6: Modalités significatives sur la dimension 2

\$`Dim 2`\$category	Estimate	p.value
Ferme	0.3558532	8.108421e-07
wolof	0.2459378	5.074788e-06
H	0.2623501	1.115269e-05
Serer	0.5217149	3.588742e-05
o:0-10	0.1732936	1.623242e-04
a:25-45	0.2215336	2.030273e-03
AP2	0.1946669	5.405356e-03
Autrui	0.8390259	8.748998e-03
b:0-10	0.1750107	3.206205e-02
d:20-40	0.1964494	4.046757e-02
b:10-20	-0.1750107	3.206205e-02
AP3	-0.3797451	2.931070e-02
d:10-20	-0.2266351	2.195471e-03
Individuel	-0.1568087	8.695513e-04
d:10-30	-0.1732936	1.623242e-04
a:45-60	-0.2562600	1.133026e-04
Collectif	-0.6822172	1.172614e-05
F	-0.2623501	1.115269e-05
Maison	-0.3558532	8.108421e-07
Peul	-0.3427039	6.268143e-08

C) Classification

- 1) L'utilisation d'une CAH sur les facteurs de l'ACM permet la transformation des variables qualitatives en composantes principales qui sont quantitatives ce qui stabilisera les classes obtenues par l'élimination du bruit.
- 2) Pour réaliser une classification, on peut utiliser le critère d'agrégation de Ward dont le but est d'agréger les individus en minimisant l'inertie intra classe et en maximisant l'inertie interclasse.
- 3) Après la réalisation de la classification avec FactoMineR, on observe le diagramme des inerties et on remarque un coude entre le 3^{ième} et le 4^{ième} bâton. En coupant, l'arbre au niveau du 3^{ième} bâton, on obtient alors 4 classes.



graphique 1: Dendrogramme

Les classes peuvent être décrites par :

- Les variables et des modalités
- Les axes factoriels
- Les individus

✓ Descriptions par les variables et/ou des modalités

Tableau 7: variables caractérisant la partition en 10 classes

variables	p.value	df
type_habitat	5.939776e-09	3
age	7.454587e-09	6
act_princ	8.186607e-09	12
localisatio_bergerie	1.520087e-08	6
sexe	1.523927e-07	3
ethnie	9.725197e-05	9
total_ovin	2.021832e-04	3
nbre_brebis	7.053493e-04	3
niveau_etude	1.350151e-03	9
type_propriete	4.172149e-02	6

Les variables type_habitat et age sont celles qui caractérisent le plus la partition en 4 classes. Ainsi, on remarque dans toutes les classes la présence de ces deux variables exemptés dans la 4^{ième} classe.

Tableau 8: classe 1

\$`classe 1`						
	cla/Mod	Mod/Cla	global	p.value	v.test	
sexe=H	59.37500	100.000000	83.116883	4.420454e-05	4.084328	
age=a:60 et +	93.75000	39.473684	20.779221	4.739030e-05	4.068138	
type_habitat=Maison	55.07246	100.000000	89.610390	2.923844e-03	2.975633	
localisatio_bergerie=Peri-urbain	90.90909	26.315789	14.285714	3.119461e-03	2.955714	
total_ovin=o:0-10	63.63636	73.684211	57.142857	4.408410e-03	2.847356	
raison_elev=Heritage	100.00000	18.421053	9.090909	5.247926e-03	2.791408	
nbre_brebis=b:0-10	54.28571	100.000000	90.909091	6.395910e-03	2.726762	
niveau_etude=Secondaire	90.00000	23.684211	12.987013	6.657215e-03	2.713522	
niveau_etude=Moyen	43.54839	71.052632	80.519481	4.457905e-02	-2.008605	
act_princ=AP3	12.50000	2.631579	10.389610	3.362421e-02	-2.124550	
ethnie=Serer	0.00000	0.000000	6.493506	2.914072e-02	-2.181578	
nbre_brebis=b:10-20	0.00000	0.000000	9.090909	6.395910e-03	-2.726762	
total_ovin=d:10-30	30.30303	26.315789	42.857143	4.408410e-03	-2.847356	
type_habitat=Ferme	0.00000	0.000000	10.389610	2.923844e-03	-2.975633	
age=a:25-45	0.00000	0.000000	14.285714	2.508432e-04	-3.661397	
sexe=F	0.00000	0.000000	16.883117	4.420454e-05	-4.084328	

Tableau 9: classe 2

\$`classe 2`						
	cla/Mod	Mod/Cla	global	p.value	v.test	
sexe=F	100.000000	54.166667	16.883117	1.358472e-08	5.678543	
nbre_brebis=b:10-20	100.000000	29.166667	9.090909	1.439217e-04	3.801330	
total_ovin=d:10-30	51.515152	70.833333	42.857143	1.119871e-03	3.258540	
localisatio_bergerie=Urbain	38.095238	100.000000	81.818182	2.861937e-03	2.982191	
niveau_etude=Moyen	37.096774	95.833333	80.519481	1.989728e-02	2.328279	
type_propriete=Collectif	58.333333	29.166667	15.584416	4.056875e-02	2.047911	
type_habitat=Maison	34.782609	100.000000	89.610390	4.212145e-02	2.032318	
type_habitat=Ferme	0.000000	0.000000	10.389610	4.212145e-02	-2.032318	
act_princ=AP2	7.142857	4.166667	18.181818	2.976414e-02	-2.173215	
age=a:60 et +	6.250000	4.166667	20.779221	1.314469e-02	-2.479825	
localisatio_bergerie=Peri-urbain	0.000000	0.000000	14.285714	1.140786e-02	-2.529951	
total_ovin=o:0-10	15.909091	29.166667	57.142857	1.119871e-03	-3.258540	
nbre_brebis=b:0-10	24.285714	70.833333	90.909091	1.439217e-04	-3.801330	
sexe=H	17.187500	45.833333	83.116883	1.358472e-08	-5.678543	

Tableau 10: classe 3

\$`classe 3`						
	cla/Mod	Mod/cla	global	p.value	v.test	
age=a:25-45	63.636364	77.77778	14.285714	4.523484e-06	4.585755	
type_habitat=Ferme	75.000000	66.66667	10.389610	9.327313e-06	4.432208	
ethnie=Serer	80.000000	44.44444	6.493506	4.464056e-04	3.511011	
total_ovin=o:0-10	18.181818	88.88889	57.142857	4.504353e-02	2.004248	
total_ovin=d:10-30	3.030303	11.11111	42.857143	4.504353e-02	-2.004248	
age=a:45-60	4.000000	22.22222	64.935065	8.177517e-03	-2.644651	
ethnie=Peul	3.773585	22.22222	68.831169	3.455848e-03	-2.923982	
type_habitat=Maison	4.347826	33.33333	89.610390	9.327313e-06	-4.432208	

Tableau 11: classe 4

\$`classe 4`						
	cla/Mod	Mod/cla	global	p.value	v.test	
act_princ=AP4	100.000000	66.66667	5.194805	1.108422e-05	4.394864	
localisatio_bergerie=Campagne	100.000000	50.00000	3.896104	2.734108e-04	3.639271	
niveau_etude=Primaire	66.666667	33.33333	3.896104	1.510595e-02	2.429829	
localisatio_bergerie=Urbain	3.174603	33.33333	81.818182	9.334749e-03	-2.599545	
act_princ=AP8	0.000000	0.00000	61.038961	2.504389e-03	-3.022811	

✓ Description par les composantes principales

Le tableau 12 nous montre d'abord que les individus que la classe 1 possèdent de faibles coordonnées sur le premier axe. Ensuite, il nous montre que ceux de la classe 2 possèdent de faibles coordonnées sur le deuxième et troisième axe. Aussi, ceux de la classe 3 ont des coordonnées sur le premier et le deuxième axe. Enfin, les individus de la classe 4 ont des coordonnées élevées sur le premier et le troisième axe.

Tableau 12: description par les composantes principales

Link between the cluster variable and the quantitative variables						
=====						
	Eta2	P-value				
Dim.2	0.6399118	3.574632e-16				
Dim.1	0.5629670	3.948786e-13				
Dim.3	0.5131301	1.945831e-11				
Description of each cluster by quantitative variables						
=====						
\$`classe 1`						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Dim.1	-6.351505	-0.3164176	-1.451986e-17	0.322806	0.4286968	2.132181e-10
\$`classe 2`						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Dim.1	3.409637	0.2491632	-1.451986e-17	0.2298130	0.4286968	6.504928e-04
Dim.3	-2.845875	-0.1953473	1.414138e-16	0.3657830	0.4026856	4.428960e-03
Dim.2	-4.940098	-0.3464691	-1.137201e-16	0.2544065	0.4114371	7.808333e-07
\$`classe 3`						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Dim.2	5.964389	0.7737414	-1.137201e-16	0.2546070	0.4114371	2.455507e-09
Dim.1	2.221984	0.3003430	-1.451986e-17	0.2390431	0.4286968	2.628440e-02
Dim.3	-2.428062	-0.3082849	1.414138e-16	0.3637997	0.4026856	1.517976e-02
\$`classe 4`						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Dim.3	5.450600	0.8660784	1.414138e-16	0.3350015	0.4026856	5.020016e-08
Dim.1	3.291626	0.5568106	-1.451986e-17	0.2717113	0.4286968	9.961003e-04

✓ Description par les individus

Il existe deux types d'individus spécifiques pour décrire les classes :

- Les individus les plus proches du centre de classe
- Les individus les plus éloignés des centres des autres classes

Tableau 13: description des individus de la classe 1

\$`para`					\$dist				
Cluster: 1					Cluster: 1				
39	7	49	18	1	14	11	54	10	2
0.1004045	0.1164195	0.1292939	0.1645506	0.1975518	1.499790	1.499790	1.239990	1.095091	1.040597

L'individu 39 appartient à la classe 1 et est le plus proche du centre de cette classe.

L'individu 14 appartient à la classe 1 et est le plus éloigné des centres des classes 2, 3 et 4.

Tableau 14: description des individus de la classe 2

\$`para`						\$dist					
cluster: 2						cluster: 2					
64	34	53	59	43		72	71	63	67	78	
0.02167352	0.17076490	0.19429188	0.24211780	0.24804847		1.758699	1.520202	1.172140	1.167129	0.906857	

L'individu 64 appartient à la classe 2 et est le plus proche du centre de cette classe.

L'individu 72 appartient à la classe 2 et est le plus éloigné des centres des classes 1, 3 et 4.

Tableau 15: description des individus de la classe 3

\$`para`						\$dist					
cluster: 3						cluster: 3					
79	77	70	75	52		84	70	79	76	75	
0.2689266	0.2712165	0.2713319	0.3310341	0.4677755		1.747048	1.283749	1.134125	1.080960	1.042025	

L'individu 79 appartient à la classe 3 et est le plus proche du centre de cette classe.

L'individu 84 appartient à la classe 3 et est le plus éloigné des centres des classes 1, 2 et 4.

Tableau 16: description des individus de la classe 4

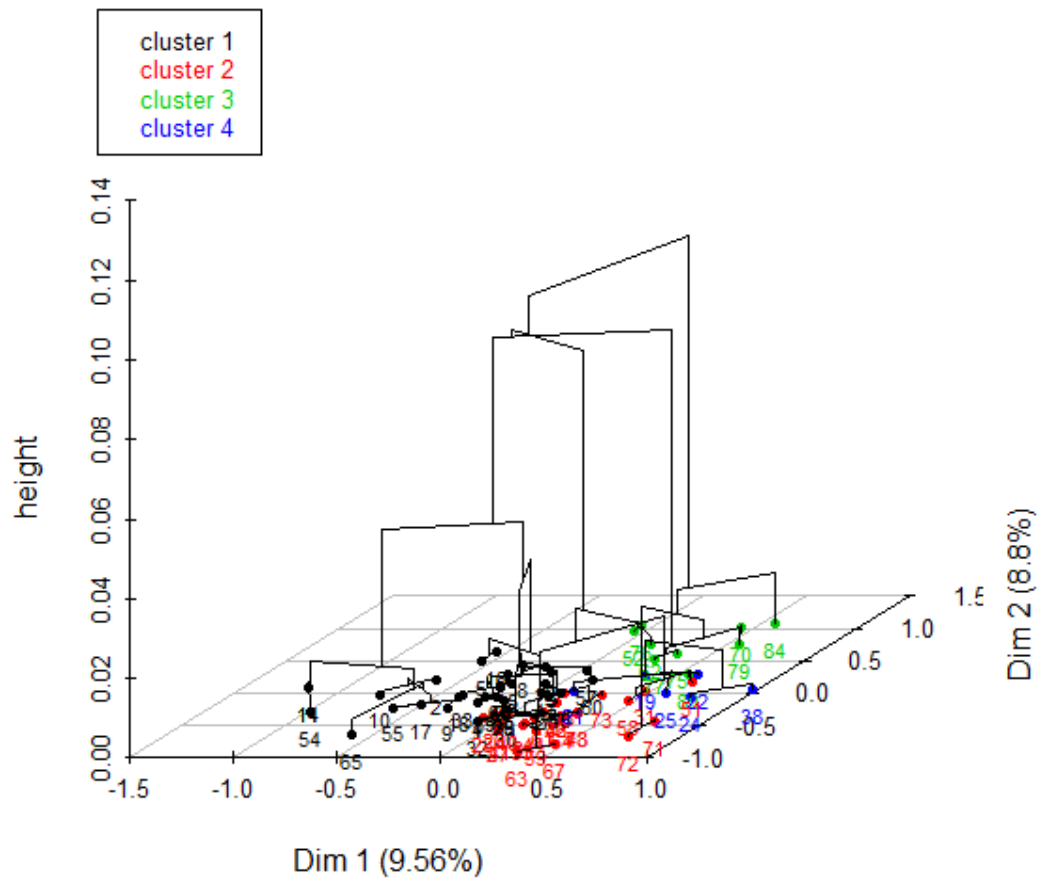
\$`para`						\$dist					
cluster: 4						cluster: 4					
25	24	19	22	21		38	24	25	22	21	
0.1753479	0.3505578	0.3991166	0.4198196	0.4460022		1.8610803	1.4245232	1.0359809	0.9799912	0.8744584	

L'individu 25 appartient à la classe 4 et est le plus proche du centre de cette classe.

L'individu 14 appartient à la classe 4 et est le plus éloigné des centres des classes 1,2, 4.

- 4) Visualisation les classes et le dendrogramme sur le premier plan principal de l'ACM.

Hierarchical clustering on the factor map



Graphique 5: Visualisation des classes et le dendrogramme sur le premier plan principal de l'ACM

- 5) Effectuation de la méthode des k-means sur ce tableau et comparaison des résultats avec ceux de la CAH.

Annexes : liste des codes R utilisés dans le projet.

```
##### Importation de la base de données
## Choix du dossier de travail
setwd("C:/Users/Sid Ousmane/Desktop/ADD/projet")
ladoum <- read.csv("ladoum.csv", sep = ";", header = T)

## Suppression de la variable religion car representant une seule modalité
ladoum <- ladoum[,-6]

## Transformation des variables qualitatives en factor
ladoum$sexe<- factor(ladoum$sexe, labels = c("H","F"))
ladoum$ethnie <- factor(ladoum$ethnie,labels = c("Peul", "Wolof","Serer", "Soce") )
ladoum$statut.matrimonial <- factor(ladoum$statut.matrimonial, labels = c("Marié",
"Célibataire", "Veuf"))
ladoum$niveau_etude <- factor(ladoum$niveau_etude, labels = c("Primaire",
"Moyen", "Secondaire", "Supérieur"))
ladoum$type_habitat <- factor(ladoum$type_habitat, labels = c("Maison", "Ferme"))
ladoum$type_propriete <- factor(ladoum$type_propriete, labels = c("Individuel",
"Collectif", "Autrui"))
ladoum$sact_princ <- factor(ladoum$sact_princ, labels = c("AP1", "AP2",
"AP3","AP4","AP7", "AP8"))
ladoum$raison_elev <- factor(ladoum$raison_elev, labels = c("Economique",
"Heritage", "Affection"))
ladoum$localisatio_bergerie<- factor(ladoum$localisatio_bergerie, labels = c("Peri-
urbain", "Urbain", "Campagne"))

## Transformation des variables quantitatives en classes d'intervalles
ladoum$age <- cut(ladoum$age, c(25,45,60, 71))
levels(ladoum$age) <- c("a:25-45", "a:45-60", "a:60 et +")
ladoum$duree_elevage <- cut(ladoum$duree_elevage, c(0,10,20,40))
levels(ladoum$duree_elevage) <- c("d:0-10", "d:10-20", "d:20-40")
ladoum$nbre_brebis <- cut(ladoum$nbre_brebis, c(0,10,20,30))
levels(ladoum$nbre_brebis) <- c("b:0-10", "b:10-20", "b:20-30")
```

```
ladoum$total_ovin <- cut(ladoum$total_ovin, c(0,10,30,61))
levels(ladoum$total_ovin) <- c("o:0-10", "d:10-30", "d:30 et +")
```

```
### Exploration des liaisons des variables de la base de données deux a deux
chisq.test()
fisher.test()
```

```
## tableau de contingence
table(ladoum$sexe, ladoum$age)
table(ladoum$sexe, ladoum$ethnie)
table(ladoum$sexe, ladoum$statut.matrimonial)
table(ladoum$sexe, ladoum$niveau_etude)
table(ladoum$sexe, ladoum$type_habitat)
table(ladoum$sexe, ladoum$type_propriete)
table(ladoum$sexe, ladoum$sact_princ)
table(ladoum$sexe, ladoum$duree_elevage)
table(ladoum$sexe, ladoum$raison_elev)
table(ladoum$sexe, ladoum$localisatio_bergerie)
table(ladoum$sexe, ladoum$nbre_brebis)
table(ladoum$sexe, ladoum$total_ovin)
```

```
table(ladoum$age, ladoum$ethnie)
table(ladoum$age, ladoum$statut.matrimonial)
table(ladoum$age, ladoum$niveau_etude)
table(ladoum$age, ladoum$type_habitat)
table(ladoum$age, ladoum$type_propriete)
table(ladoum$age, ladoum$sact_princ)
table(ladoum$age, ladoum$duree_elevage)
table(ladoum$age, ladoum$raison_elev)
table(ladoum$age, ladoum$localisatio_bergerie)
table(ladoum$age, ladoum$nbre_brebis)
table(ladoum$age, ladoum$total_ovin)
```

```

table(ladoun$ethnie, ladoun$statut.matrimonial)
table(ladoun$ethnie, ladoun$niveau_etude)
table(ladoun$ethnie, ladoun$type_habitat)
table(ladoun$ethnie, ladoun$type_propriete)
table(ladoun$ethnie, ladoun$sact_princ)
table(ladoun$ethnie, ladoun$ duree_elevage)
table(ladoun$ethnie, ladoun$raison_elev)
table(ladoun$ethnie, ladoun$localisatio_bergerie)
table(ladoun$ethnie, ladoun$nbre_brebis)
table(ladoun$ethnie, ladoun$total_ovin)

```

```

table(ladoun$statut.matrimonial, ladoun$niveau_etude)
table(ladoun$statut.matrimonial, ladoun$type_habitat)
table(ladoun$statut.matrimonial, ladoun$type_propriete)
table(ladoun$statut.matrimonial, ladoun$sact_princ)
table(ladoun$statut.matrimonial, ladoun$ duree_elevage)
table(ladoun$statut.matrimonial, ladoun$raison_elev)
table(ladoun$statut.matrimonial, ladoun$localisatio_bergerie)
table(ladoun$statut.matrimonial, ladoun$nbre_brebis)
table(ladoun$statut.matrimonial, ladoun$total_ovin)

```

```

table(ladoun$niveau_etude, ladoun$type_habitat)
table(ladoun$niveau_etude, ladoun$type_propriete)
table(ladoun$niveau_etude, ladoun$sact_princ)
table(ladoun$niveau_etude, ladoun$ duree_elevage)
table(ladoun$niveau_etude, ladoun$raison_elev)
table(ladoun$niveau_etude, ladoun$localisatio_bergerie)
table(ladoun$niveau_etude, ladoun$nbre_brebis)
table(ladoun$niveau_etude, ladoun$total_ovin)

```

```
table(ladoum$type_habitat, ladoum$type_propriete)
table(ladoum$type_habitat, ladoum$sact_princ)
table(ladoum$type_habitat, ladoum$duree_elevage)
table(ladoum$type_habitat, ladoum$raison_elev)
table(ladoum$type_habitat, ladoum$localisatio_bergerie)
table(ladoum$type_habitat, ladoum$nbre_brebis)
table(ladoum$type_habitat, ladoum$total_ovin)
```

```
table(ladoum$type_propriete, ladoum$sact_princ)
table(ladoum$type_propriete, ladoum$duree_elevage)
table(ladoum$type_propriete, ladoum$raison_elev)
table(ladoum$type_propriete, ladoum$localisatio_bergerie)
table(ladoum$type_propriete, ladoum$nbre_brebis)
table(ladoum$type_propriete, ladoum$total_ovin)
```

```
table(ladoum$sact_princ, ladoum$duree_elevage)
table(ladoum$sact_princ, ladoum$raison_elev)
table(ladoum$sact_princ, ladoum$localisatio_bergerie)
table(ladoum$sact_princ, ladoum$nbre_brebis)
table(ladoum$sact_princ, ladoum$total_ovin)
```

```
table(ladoum$duree_elevage, ladoum$raison_elev)
table(ladoum$duree_elevage, ladoum$localisatio_bergerie)
table(ladoum$duree_elevage, ladoum$nbre_brebis)
table(ladoum$duree_elevage, ladoum$total_ovin)
```

```
table(ladoum$raison_elev, ladoum$localisatio_bergerie)
table(ladoum$raison_elev, ladoum$nbre_brebis)
table(ladoum$raison_elev, ladoum$total_ovin)
```

```
table(ladoun$localisatio_bergerie, ladoun$nbre_brebis)
```

```
table(ladoun$localisatio_bergerie, ladoun$total_ovin)
```

```
table(ladoun$nbre_brebis, ladoun$total_ovin)
```

```
## fisher.test entre les deux variable qualitatives deux à deux
```

```
fisher.test(ladoun$sexe, ladoun$age)
```

```
fisher.test(ladoun$sexe, ladoun$ethnie)
```

```
fisher.test(ladoun$sexe, ladoun$statut.matrimonial)
```

```
fisher.test(ladoun$sexe, ladoun$niveau_etude)
```

```
fisher.test(ladoun$sexe, ladoun$type_habitat)
```

```
fisher.test(ladoun$sexe, ladoun$type_propriete)
```

```
fisher.test(ladoun$sexe, ladoun$sact_princ)
```

```
fisher.test(ladoun$sexe, ladoun$duree_elevage)
```

```
fisher.test(ladoun$sexe, ladoun$raison_elev)
```

```
fisher.test(ladoun$sexe, ladoun$localisatio_bergerie)
```

```
fisher.test(ladoun$sexe, ladoun$nbre_brebis)
```

```
fisher.test(ladoun$sexe, ladoun$total_ovin)
```

```
fisher.test(ladoun$age, ladoun$ethnie)
```

```
fisher.test(ladoun$age, ladoun$statut.matrimonial)
```

```
fisher.test(ladoun$age, ladoun$niveau_etude)
```

```
fisher.test(ladoun$age, ladoun$type_habitat)
```

```
fisher.test(ladoun$age, ladoun$type_propriete)
```

```
fisher.test(ladoun$age, ladoun$sact_princ)
```

```
fisher.test(ladoun$age, ladoun$duree_elevage)
```

```
fisher.test(ladoun$age, ladoun$raison_elev)
```

```
fisher.test(ladoun$age, ladoun$localisatio_bergerie)
```

```
fisher.test(ladoun$age, ladoun$nbre_brebis)
```

```
fisher.test(ladoun$age, ladoun$total_ovin)
```

```
fisher.test(ladoun$ethnie, ladoun$statut.matrimonial)
fisher.test(ladoun$ethnie, ladoun$niveau_etude)
fisher.test(ladoun$ethnie, ladoun$type_habitat)
fisher.test(ladoun$ethnie, ladoun$type_propriete)
fisher.test(ladoun$ethnie, ladoun$sact_princ)
fisher.test(ladoun$ethnie, ladoun$ duree_elevage)
fisher.test(ladoun$ethnie, ladoun$raison_elev)
fisher.test(ladoun$ethnie, ladoun$localisatio_bergerie)
fisher.test(ladoun$ethnie, ladoun$nbre_brebis)
fisher.test(ladoun$ethnie, ladoun$total_ovin)
```

```
fisher.test(ladoun$statut.matrimonial, ladoun$niveau_etude)
fisher.test(ladoun$statut.matrimonial, ladoun$type_habitat)
fisher.test(ladoun$statut.matrimonial, ladoun$type_propriete)
fisher.test(ladoun$statut.matrimonial, ladoun$sact_princ)
fisher.test(ladoun$statut.matrimonial, ladoun$ duree_elevage)
fisher.test(ladoun$statut.matrimonial, ladoun$raison_elev)
fisher.test(ladoun$statut.matrimonial, ladoun$localisatio_bergerie)
fisher.test(ladoun$statut.matrimonial, ladoun$nbre_brebis)
fisher.test(ladoun$statut.matrimonial, ladoun$total_ovin)
```

```
fisher.test(ladoun$niveau_etude, ladoun$type_habitat)
fisher.test(ladoun$niveau_etude, ladoun$type_propriete)
fisher.test(ladoun$niveau_etude, ladoun$sact_princ)
fisher.test(ladoun$niveau_etude, ladoun$ duree_elevage)
fisher.test(ladoun$niveau_etude, ladoun$raison_elev)
fisher.test(ladoun$niveau_etude, ladoun$localisatio_bergerie)
fisher.test(ladoun$niveau_etude, ladoun$nbre_brebis)
fisher.test(ladoun$niveau_etude, ladoun$total_ovin)
```

```
fisher.test(ladoun$type_habitat, ladoun$type_propriete)
```



```
fisher.test(ladoun$type_habitat, ladoun$sact_princ)
fisher.test(ladoun$type_habitat, ladoun$duree_elevage)
fisher.test(ladoun$type_habitat, ladoun$raison_elev)
fisher.test(ladoun$type_habitat, ladoun$localisatio_bergerie)
fisher.test(ladoun$type_habitat, ladoun$nbre_brebis)
fisher.test(ladoun$type_habitat, ladoun$total_ovin)
```

```
fisher.test(ladoun$type_propriete, ladoun$sact_princ)
fisher.test(ladoun$type_propriete, ladoun$duree_elevage)
fisher.test(ladoun$type_propriete, ladoun$raison_elev)
fisher.test(ladoun$type_propriete, ladoun$localisatio_bergerie)
fisher.test(ladoun$type_propriete, ladoun$nbre_brebis)
fisher.test(ladoun$type_propriete, ladoun$total_ovin)
```

```
fisher.test(ladoun$sact_princ, ladoun$duree_elevage)
fisher.test(ladoun$sact_princ, ladoun$raison_elev)
fisher.test(ladoun$sact_princ, ladoun$localisatio_bergerie)
fisher.test(ladoun$sact_princ, ladoun$nbre_brebis)
fisher.test(ladoun$sact_princ, ladoun$total_ovin)
```

```
fisher.test(ladoun$duree_elevage, ladoun$raison_elev)
fisher.test(ladoun$duree_elevage, ladoun$localisatio_bergerie)
fisher.test(ladoun$duree_elevage, ladoun$nbre_brebis)
fisher.test(ladoun$duree_elevage, ladoun$total_ovin)
```

```
fisher.test(ladoun$raison_elev, ladoun$localisatio_bergerie)
fisher.test(ladoun$raison_elev, ladoun$nbre_brebis)
fisher.test(ladoun$raison_elev, ladoun$total_ovin)
```

```
fisher.test(ladoun$localisatio_bergerie, ladoun$nbre_brebis)
```

```

fisher.test(ladoum$localisatio_bergerie, ladoum$total_ovin)

fisher.test(ladoum$nbre_brebis, ladoum$total_ovin)

## Utilisation de l'analyse en composante Multiple
## chargement du packages FactoMineR
library("FactoMineR")
res.mca <- MCA(ladoum)
## Affichons le diagramme des valeurs propres pour determiner le nombre d'axes.
barplot(res.mca$eig[,1])

##### typologie des individus
summary(res.mca, ncp=2, nbelements = Inf)

## ACM après retrait des individus extrême
ext.mca <- MCA(ladoum[-c(45,68,69,74,81,82,83),], ncp =3)

## Typologie des individus
plot(ext.mca, title="Graphes des individus après suppression des individus extrêmes",
invisible = c("var"), autoLab = "yes")
summary(ext.mca, ncp = 2, nbelements = Inf)
#affichage des contributions

##### Typologie des modalités
dimdesc(ext.mca)

##### Classification
res.hcpc <- HCPC(ext.mca)

## description des classes

# Description par les variables et des modalités
res.hcpc$desc.var$test.chi2
res.hcpc$desc.var$category

```

```
# Description par les composantes principales  
res.hcpc$desc.axes
```

#4) Visualisation les classes et le dendrogramme sur le premier plan principal de l'ACM.

```
plot(res.hcpc)
```

5) Effectuation de la méthode des k-means sur ce tableau et comparaison des résultats avec ceux de la CAH.