

Master TRIED,

**PROJET D'ANALYSE DES  
DONNEES QUALITATIVES**

Réalisé par:

**FOUTSE YUEHGOH**

Année universitaire : 2018/2019

## **INTRODUCTION:**

L'Analyse des Correspondances Multiples (ACM) est une analyse qui permet d'étudier une population d'individus décrits par des variables qualitatives. Ces données quantitatives fournissent un bon portrait global d'une population ou d'une région géographique. Elles permettent aussi d'observer des tendances sur une certaine durée. Elles sont très souvent utiles quand on cherche à expliquer le comment et le pourquoi des choses. Fournissent une information enrichie, approfondie et diversifiée, reposant sur quelques individus ou quelques cas (un ensemble fini). L'application la plus courante de l'ACM est le traitement de l'ensemble des réponses à une enquête. Chaque question constitue une variable dont les modalités sont les réponses proposées (parmi lesquelles chaque enquêté doit faire un choix unique).

Dans cette étude, nous allons utiliser cette méthode d'ACM sur un ensemble de jeu de données issue d'une enquête. La procédure d'enquête est utilisée quand, dans un domaine donné, on se trouve confronté à une situation d'incertitude quant aux causes d'un état de chose. De ce fait on est amené à poser des questions, souvent à des personnes, pour inventorier leurs opinions, leurs pratiques, leur situation, leur passé. De ce vaste coup de filet sans hypothèse préalable, on espère tirer des explications sur les phénomènes en cause. Dans notre cas, nous voulons comprendre l'opinion des gens concernant la pratique des organismes génétiquement modifié. L'ACM va étudier la variabilité des individus.

Cependant, avant de mettre en œuvre une ACM, il convient de se demander qu'elle analyse précisément devront nous étudier, et ceci dépendra des données à notre disposition et aussi en prenant compte de la problématique que nous essayons de résoudre.

Ainsi, pour présenter notre travail, nous allons commencer par définir notre problématique, l'objectif de cette étude et les démarches suivies avec l'interprétation des différentes analyses obtenues puis une conclusion.

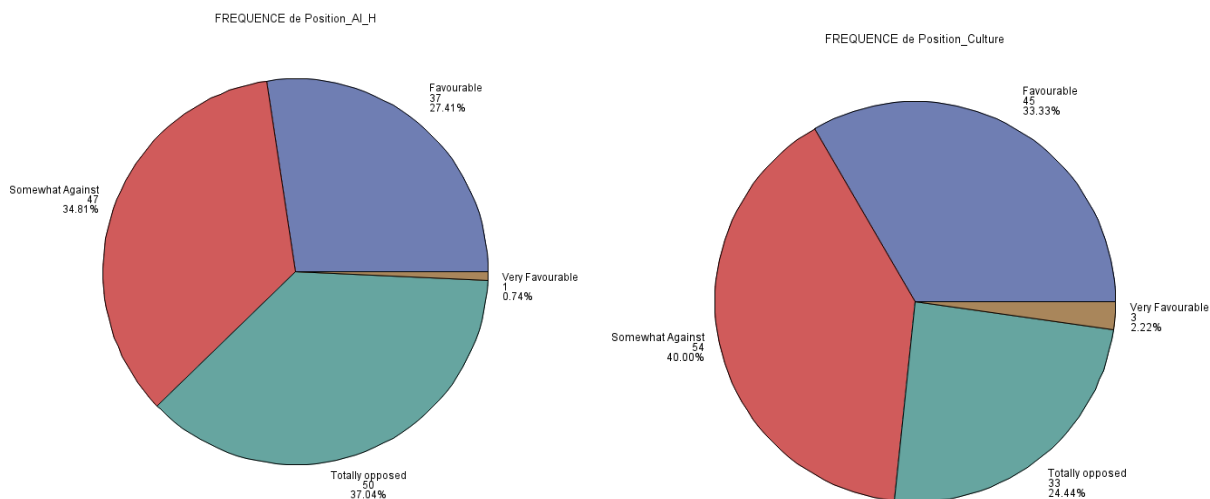
**PROBLEMATIQUE :** La France ainsi que plusieurs pays dans le monde sont inquiets en ce qui concerne les organismes génétiquement modifiés. Un organisme génétiquement modifié (OGM) est un organisme (animal, végétal, champignon, micro-organisme) dont on a modifié le matériel génétique d'une manière qui ne s'effectue pas naturellement, pour lui conférer une ou plusieurs caractéristiques recherchées. L'utilisation d'OGM la plus connue est celle qui est faite en agriculture pour conférer de nouveaux caractères aux plantes pour leur utilisation en alimentation animale et humaine. En France, la culture du maïs transgénique MON810 (seule plante autorisée à la culture en Europe) est interdite depuis 2008. Ainsi, deux étudiants d'Agro-campus ont mené une enquête avec 135 participants afin d'avoir une impression globale des personnes vis-à-vis de cette pratique de la modification génétique des organismes, appelée en anglais *genetically modified organisms* (GMOs). Pour obtenir des données pour cette étude, ils ont posé 21 questions qui ont été divisées en deux groupes, 16 questions directement liées à l'opinion des participants sur la pratique du GMO et 5 questions décrivant les participants. Les questions sont présentées dans la table en Annexe du document.

**OBJECTIF :** Utilisant ce questionnaire, nous voulons d'une part caractériser nos participants en termes de leurs relations avec la GMO et d'autre part essayer de voir si cette caractérisation a une relation avec

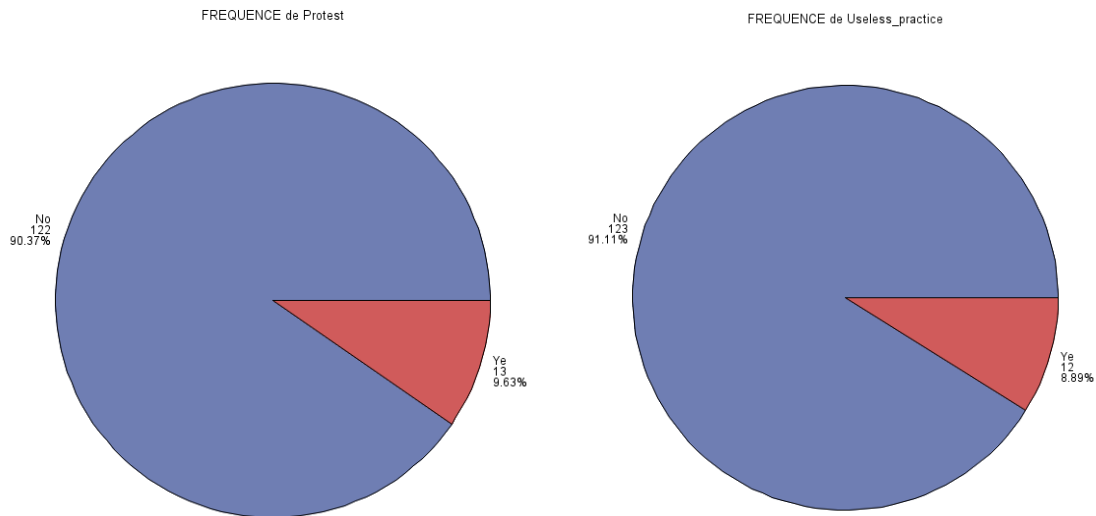
les variables descriptives. Ainsi, la question 4 des variables descriptives sera important l'or de l'interprétation des résultats car il se pourrait que ceux qui auront répondu oui à cette question pourrait avoir des connaissances scientifique sur la GMO comparé à d'autres participants. Alors l'or de cette étude on essayera de voir si c'est le cas.

**DEMARCHE** : Pour mener à bien cette étude, nous allons premièrement faire une étude uni dimensionnelle pour essayer de voir la répartition des modalités de chaque variables afin d'appliquer des traitements nécessaire pour mener à bien notre études. Ensuite nous allons faire une étude bidimensionnelle afin de voir si il y'a ou pas des dépendances entre les variables supplémentaire (a expliquée) et les variables actifs. Nous allons ensuit passer à l'ACM afin d'étudier les différences et les ressemblances des individus. Avec les résultats obtenus de l'ACM, nous allons effectuer une classification, et enfin faire une discrimination sur les composantes de l'ACM.

**Etude Unidimensionnelle** : Nous avons visualisée chaque variable afin de voir le pourcentage de chaque modalité dans chaque variable ainsi que le nombre de modalités de chaque variable. La variable avec le plus grand nombre de modalité est la profession, qui a 10 modalités. Et le plus petit nombre de modalité c'est 2.



Nous avons remarquée à travers les figures ci-déçu qu'en réponse à la question **"What do you think of the inclusion of GM raw materials in products for human consumption?" (position.AI.H)**, une seule personne a répondu very Favourable. Alors il est raisonnable de le mettre dans le groupe des favorable. Aussi, en réponse a la question **"What is your view of GMO cultivation in France?" (Position.culture)** nous avons fait pareil car juste 3 personnes on répondu very favorable. Et ainsi nous avons réduit le nombre de modalités de ces variables. Ceci est bien car l'ACM est gourmande par rapport aux nombres de dimensions (pour le calcul).



On remarque également des modalités rares dans d'autres variables comme ces deux figures mais on ne peut pas faire pareil car les modalités s'opposent. Par exemple pour la question **"Have you ever taken part in an anti-GMO protest?"**, sur 135 participants, juste 13 participants ont répondu oui. Aussi, pour la question **"Do you think that GMO are a useless scientific practice?"** sur 135 participants, juste 12 participants ont répondu oui.

On peut supposer que ces 12 personnes font partie des 13 qui ont répondu oui à la question sur les protest, nous allons décider de ce débarrasser de la variable qui représente cette question de useless\_practice et laisser la variable concernant les protest. Nous agissons ainsi car ces deux questions renvoient presque un même message, ainsi nous diminuons le nombre de variable et par conséquent le nombre de modalités.

Cette étude nous a permis de détecter les catégories rares, et les méthodes qui ont été appliquées à pour but aussi de s'assurer que ces catégories n'influencent pas trop notre analyse, c'est pourquoi nous avons regroupé certaines catégories similaires.

**Etude Bidimensionnelle** : Après le traitement effectué suite aux observations de l'étude unidimensionnelle, nous allons maintenant effectuer une étude bidimensionnelle afin de comparer nos deux groupes deux à deux. Ici nous voulons tester l'existence d'une relation entre ces groupes, voir si il y'a ou pas d'indépendance entre ces deux groupes et aussi voir la fréquence de chaque variable active dans les variables supplémentaires. Nous allons réaliser ces études à l'aide des tableaux de fréquence ou de contingence et tests du Chi-2, fournis par la fonction PROC FREQ. Cette procédure est à la fois : une procédure descriptive qui produit des tableaux de fréquences, de contingence ou des tris croisés multidimensionnels. Elle permet de calculer les effectifs et les fréquences de chaque valeur prise par une variable ; c'est une procédure statistique permettant de faire certains tests comme le test du chi2. Nous allons juste considérer une variable pour les présenter ici comment l'analyse se fait :

Fréquence  
Attendu  
Khi-2 de cellule  
Pourcentage  
Pct de ligne  
Pct de col.

Table de Profession par Implicated					
Profession	Implicated				Total
	A little	A lot	Certain extent	Not at all	
Accredited Prof	1	1	1	0	3
	0.6889	0.8	1.1778	0.3333	2.22
	0.1405	0.05	0.0268	0.3333	
	0.74	0.74	0.74	0.00	
	33.33	33.33	33.33	0.00	
Civil Servant	3	6	0	0	9
	2.0867	2.4	3.5333	1	6.67
	0.4215	5.4	3.5333	1	
	2.22	4.44	0.00	0.00	
	33.33	66.67	0.00	0.00	
Manual Labourer	0	1	0	0	1
	0.2296	0.2667	0.3926	0.1111	0.74
	0.2296	2.0167	0.3926	0.1111	
	0.00	0.74	0.00	0.00	
	0.00	100.00	0.00	0.00	
Other	2	2	5	0	9
	2.0867	2.4	3.5333	1	6.67
	0.0022	0.0867	0.6088	1	
	1.48	1.48	3.70	0.00	
	22.22	22.22	55.56	0.00	
Retailer	0	0	1	2	3
	0.6889	0.8	1.1778	0.3333	2.22
	0.6889	0.8	0.0268	8.3333	
	0.00	0.00	0.74	1.48	
	0.00	0.00	33.33	66.67	
Retired	1	6	5	2	14
	3.2148	3.7333	5.4963	1.5556	10.37
	1.6259	1.3762	0.0448	0.127	
	0.74	4.44	3.70	1.48	
	7.14	42.86	35.71	14.29	
Senior Management	4	4	6	3	17
	3.9037	4.5333	6.6741	1.8889	12.59
	0.0024	0.0627	0.0681	0.6536	
	2.96	2.96	4.44	2.22	
	23.53	23.53	35.29	17.65	

On remarque sur cette figure que tous les accredited professors sont impliquée dans ce débat, ainsi que la majeure partie des étudiants, juste 6 ne sont pas impliquées. Nous avons un seule Manual laborour qui est lui aussi tres bien impliquée, le majeur nombre de civil servant sont impliquée eux aussi et à fond. Parmi le peu de retired qui ont été interrogé un seul est moyennement impliquée et les deux autres pas du tout. Parmi les Senior Management, on voie 3 qui ne sont pas impliquée. Tous les techniciens sont impliqués. Les unemployed ne sont pas trop implique, un seule qui est à fond dans le débat, un autre qui est impliquée moyennement et deux autres pas du tout.

Student	18	11	34	6	69
	15.844	18.4	27.089	7.6667	51.11
	0.2933	2.9761	1.7632	0.3623	
	13.33	8.15	25.19	4.44	
	26.09	15.94	49.28	8.70	
Technician	2	4	0	0	6
	1.3778	1.6	2.3556	0.6667	4.44
	0.281	3.6	2.3556	0.6667	
	1.48	2.96	0.00	0.00	
	33.33	66.67	0.00	0.00	
Unemployed	0	1	1	2	4
	0.9185	1.0867	1.5704	0.4444	2.96
	0.9185	0.0042	0.2072	5.4444	
	0.00	0.74	0.74	1.48	
	0.00	25.00	25.00	50.00	
Total	31	36	53	15	135
	22.96	26.67	39.26	11.11	100.00

Statistiques pour la table de Profession par Implicated

Statistique	DDL	Valeur	Prob
Khi-2	27	47.9152	0.0078
Test du rapport de vraisemblance	27	49.8895	0.0047
Khi-2 de Mantel-Haenszel	1	2.2877	0.1304
Coefficient Phi		0.5958	
Coefficient de contingence		0.5118	
V de Cramer		0.3440	
WARNING: 85% des cellules ont des effectifs attendus inférieurs à 5. (Asymptotique) Le Khi-2 n'est peut-être pas un test valide.			

Test Khi-2 de Pearson	
Khi-2	47.9152
DDL	27
Pr > Khi-2 asymptotique	0.0078
Pr >= Khi-2 exact	.

Taille de l'échantillon = 135

On a effectué l'analyse sur un échantillon très petit, ainsi le test du  $\chi^2$  ne peut pas être très fiable. Cependant, le test du  $\chi^2$  montre une probabilité de rejeter à tort l'hypothèse d'indépendance entre les deux variables de 0,78% (0,0078), ce qui est assez petit : Cette conclusion est corroborée par les valeurs des trois derniers coefficients, phi, contingence et V de Cramer : pour ce dernier, par exemple, la liaison observée correspond à 34,40% de liaison entre les deux variables (ce qui est assez peu). On peut peut-être rejeter cette hypothèse ce qui voudrait dire qu'il y'a dépendance. Mais puis que l'effectif est petits, on va vérifier cette relation avec le teste de Fisher. Le teste nous donne une P-value de  $0,0078 < 0,05$  alors on rejette l'hypothèse d'Indépendance. Donc on voit bien une **dépendance** entre l'implication des enquêter aux débat sur la GMO avec leurs profession.

**Analyse factoriel (ACM)** : Avant de mettre en œuvre une ACM, il convient de se demander quelle analyse précisément devons-nous étudier étant donné que nous avons deux groupes, les variables représentant l'opinion des participants et les variables décrivant les participants. Nous allons faire une étude en passant les variables décrivant l'opinion des participants (groupe1) comme variables actifs, et les variables descriptives (groupe2) comme variables supplémentaires. Avec cette approche, nous pourrions faire apparaître les liaisons intéressantes entre les deux groupes plus rapidement. Aussi, cette pratique est bonne quand on a un grand questionnaire, on économise un temps de calcul considérable car l'analyse n'a pas besoin d'être effectuée sur la totalité des tableaux des réponses.

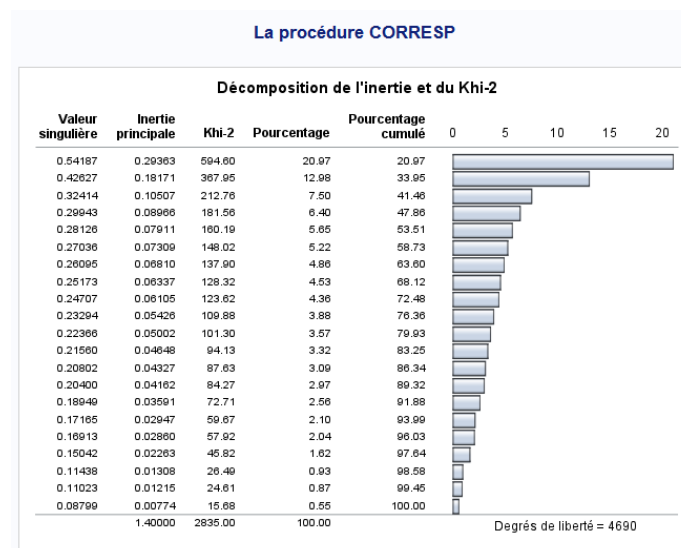
Tout le propos de l'analyse des correspondances multiples consiste à résumer un nombre élevé de variables ici les variables actifs, par un nombre faible de facteurs, le but étant de résumer les relations entre les variables. Une information portée par une variable peut être étudiée en fonction de ses modalités. En ACM, on se focalise premièrement à étudier ces modalités car elles représentent les variables ainsi que le groupe d'individus qui ont choisi ces modalités. L'ACM travaille à partir d'un tableau disjonctif complet. Avant d'aller dans l'analyse, nous rappelons quelques points importants pour la compréhension des nuages de points de l'analyse :

- Tous les individus ont le même poids car il n'y a aucune raison d'accorder plus d'importance aux réponses d'un enquêté plus que celle d'un autre.
- Si un individu possède une modalité rare, cela le caractérise beaucoup plus que s'il possédait une modalité fréquente. Si une modalité est possédée par tous les individus, elle ne caractérise donc aucun d'entre eux. Les modalités rares sont généralement très loin du barycentre.
- Si deux individus possèdent beaucoup de modalités en commun, alors la distance entre eux sera petite, cette distance sera grande ou 0 si ils possèdent très peu de modalités en commun ou si elles ont tout en commun respectivement.
- Les variables actifs servent à calculer la distance entre les individus et les variables supplémentaires sont utilisées pour l'interprétation des résultats.
- L'ACM va extraire les principales dimensions de variabilité avant de les décrire par les modalités, il s'intéressera aux associations entre modalités et fournira une visualisation d'ensemble de cette association entre modalités, une représentation simultanée des individus et des modalités.
- La proximité entre les modalités traduit une association forte entre ces modalités
- La proximité entre les individus sur les plans traduit une proximité entre les réponses.

- Les individus sont du côté des modalités qu'ils prennent et les modalités sont du côté des individus qui les prennent.
- Deux variables qualitatives sont liées si les modalités de l'une s'associent de façons particulières aux modalités de l'autre.

Ces points étant mentionnés, nous passons donc à l'application de l'ACM à l'aide du logiciel SAS en lui passant comme variables actives (ceux qui vont contribuer à la construction des axes), les variables représentant les 16 premières questions (groupe1) à l'exception de la 15<sup>ème</sup> question que nous avons décidée d'ôter de l'analyse unidimensionnelle de supprimer. Nous lui passons comme variables supplémentaires (celle qui ne contribue pas à la construction de l'axe), les 5 dernières questions, les variables descriptives.

Nous commençons premièrement par rechercher l'axe qui maximise l'inertie. Nous allons faire le choix grâce à la figure ci-dessous :

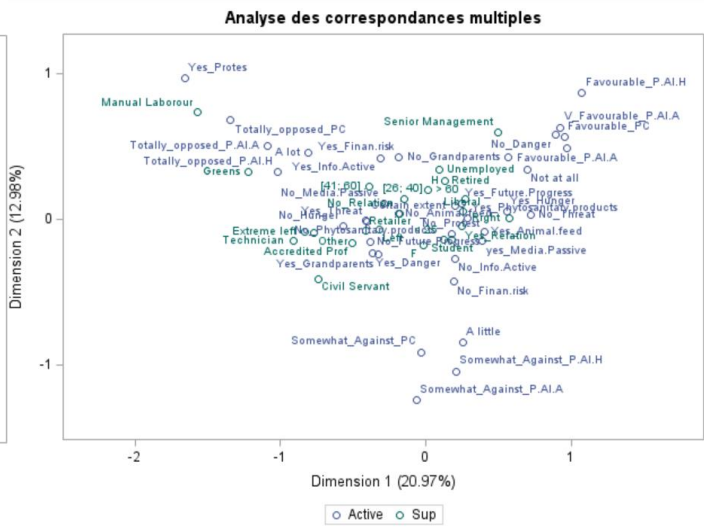
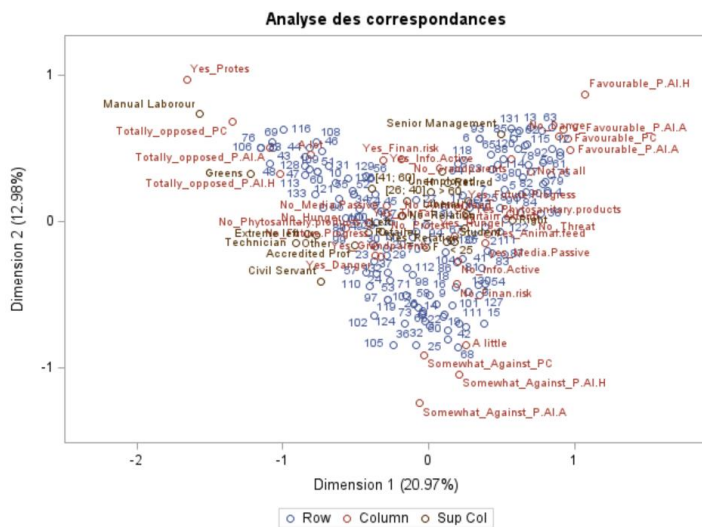


Utilisant l'histogramme des valeurs propres pour faire le choix des axes à retenir, nous allons utiliser le critère du « coude » : on observe deux décrochements (coude), suivi d'une décroissance régulière, entre le premier, deuxième, et le troisième facteur, ce qui pourrait inciter à n'examiner que les 2 premiers facteurs.

Nous voyons que le diagramme en barres met en évidence le premier axe a une inertie supérieure au deuxième axe, nous allons interpréter le premier plan factoriel en priorité. D'après la règle de coude, nous allons faire le choix des deux premiers axes. Le pourcentage d'inertie expliquée par ces deux axes est de **33,95%**.

Nous allons donc poursuivre par l'analyse du nuage des individus ainsi que les modalités des variables actives et supplémentaires. Pour plus de visibilité, nous allons afficher deux graphes, l'un sans les individus, juste les modalités et l'autre des individus et les modalités. Nous obtenons donc les nuages de points suivant:





L'allure du nuage des individus prennent la forme d'une fonction convexe. Les littératures appellent ce caractéristique **effet de Guffman**. Cet effet a tendance à répartir sur un axe les individus selon les modalités croissante (l'axe horizontal), dans notre cas nous avons totally\_opposed à gauche, somewhat\_against au milieu et very\_favourable à droite. Oppose sur l'autre axe (l'axe vertical), les individus extrêmes qui prennent de faibles ou fortes valeurs aux individus de valeurs moyen (effectif), dans notre cas ont la remarque avec Manual\_Laborour. Les modalités de faible effectif comme manual\_Laborour, yes\_protes sont loin de l'origine.

L'axe 1 oppose les totally\_opposed des Favourable et very\_Favourable. L'axe 2 oppose ceux qui sont impliquée A lot à ceux qui sont impliquée A little. Elle oppose aussi la modalité senior management unemploymement, retired, avec technician, manual laborour, civils servants entre les deux groupes.

Les variables supplémentaires présentes deux types d'informations :

- Sa révèle une structure forte entre la variable profession et l'identification a un parti politique. Vue par leurs proximité sur le plan des nuage de point .
- Ça n'identifie aucune structure particulière avec les variables age, sex, ou profession en relation à l'agriculture, nourriture industrielle et pharmacie industriel que l'on s'est posé au début, la question 4 du group2. Ils sont tous près de l'origine.

Puis que le poids de nous nos individus sont pareil, pour voir les qu'elles individus contribuent aux axes correspondant, nous allons comparer leurs contribués a 0,0074(leurs poids commun), nous obtenons les individus suivants pour ceux qui sont significativement **> 0,0074** :

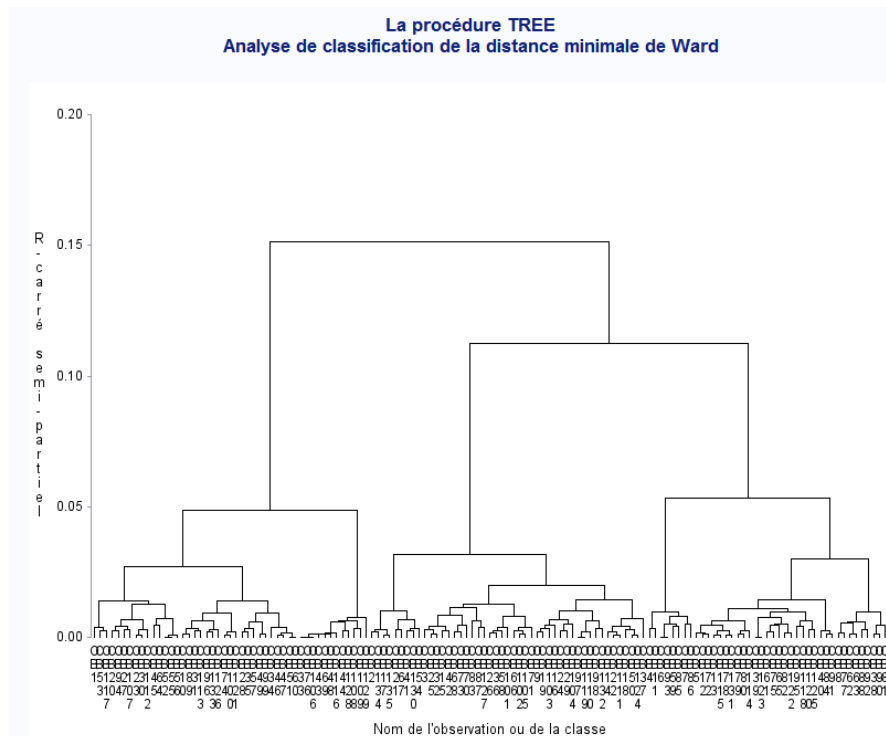
Dim1	Dim2
4 (0.0169)	25(0.0266), 22(0.0225)
33(0.0332)	26(0.0163)
48(0.0296)	32(0.0209)
59(0.0204)	36(0.0291)
69(0.0289)	69(0.0120), 105 (0.0293)
106 (0.0332), 116 (0.0248),	42(0.0213), 124 (0.0199)



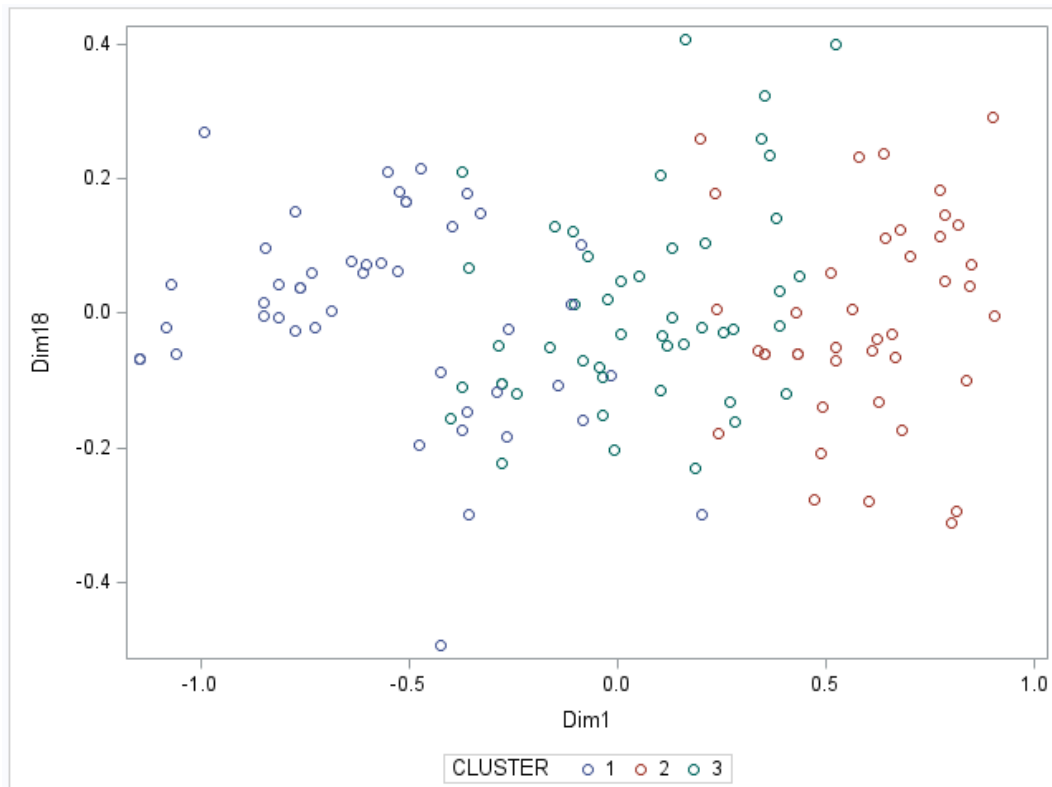
**Classification** : La classification d'individus décrits par des variables qualitatives est un vaste domaine d'étude et de nombreuses mesures de ressemblance existent. Cependant, ces méthodes sont difficile à mettre en œuvre et en pratique il est courant d'utiliser l'ACM comme prétraitement: les variables qualitatives sont transformées en composantes principales, c.-à-d. en variables quantitatives. La classification est construite à partir des (premières) dimensions factorielles. Pour donc réaliser la classification, nous allons faire à nouveau l'ACM cette fois avec nos 21 dimensions, puis extraire uniquement les observations de nos 21 dimensions et le stocker dans un fichier. Nous allons ensuite extraire nos variables supplémentaires et les stocker également dans un autre fichier puis joindre les deux nouveaux fichiers. Ainsi nous aurons nos variables supplémentaires ainsi que les observations de nos 21 dimensions. Avec ces données, nous allons tester quelques méthodes de classification :

- **Classification hiérarchique** : une hiérarchie est extrêmement utile pour déterminer le nombre de classes. Ce choix peut être fait à partir de l'allure générale de l'arbre, du niveau des nœuds (ces niveaux peuvent être représentées par un diagramme en barres), du nombre de classes (qui doit être ni trop faible ni trop grand) et de l'interopérabilité des classes. Nous allons donc nous servir de cette méthode pour déterminer le nombre de classes pour notre classification.

A l'aide de la fonction **proc tree** nous obtenons la figure suivante :



Nous remarquons les trois classes distinctes. A l'aide de la fonction **step disc** de SAS, nous avons obtenu les dimensions les plus importantes pour la variable représentant la question 4 (Relation). Nous allons donc présenter notre nuage de point sur l'une de ces dimensions importantes pour obtenir la figure ci-dessous :



La discrimination n'est pas parfaite mais on peut clairement voir nos 3 classes.

- **Analyse factoriel discriminante:** Cette analyse est une méthode qui permet de maitre en évidence les liaisons entre un caractère qualitatif à expliquer et un ensemble de caractères quantitative explicatifs. Ici le caractère qualitatif à expliquer sera la variable Relation qui représente la question **“Is your profession or education in any way linked to agriculture, the food industry or the pharmaceutical industry?”** l'ensemble quantitative sont les composantes de l'ACM que nous avons extraire. C'elle que nous alons applique fonctionne comme suit :
  - L'analyse discriminante permet de générer des règles d'affectation et de calculer les probabilités d'appartenance à chaque classe pour chaque observation a posteriori
  - Possibilité de validation croisée, en séparant l'échantillon en deux blocs :
    - avec un échantillon d'apprentissage pour créer le modèle et élaborer les règles de décisions ou d'affectation
    - un second échantillon, l'échantillon-test, pour estimer les performances du modèle
  - Sont algorithmes a pour but de maximiser la distance inter-groupe à partir des barycentres et minimiser la distance entre individus intra-groupe.
  - Les directions factorielles discriminantes successives sont alors déterminées, tandis que des graphiques (plans factoriels) permettent de visualiser les individus ou les variables expliquant la construction des axes

Nous allons l'appliquée à l'aide de la fonction SAS PROC DISCRIM avec l'option all crossvalidate canonical, en utilisant pour variables, les dimensions importante obtenu par la fonction step disc (dim18 dim1 dim5 dim3 dim6 dim20 dim14). Ce sont ces dimensions (plan factoriels) qui vont permettre de visualiser les individus ou les variables expliquant la construction des axes. Le résultat de notre analyse est le suivant :

**La procédure DISCRIM**  
**Résultats de classification pour données de calibration : WORK.SORTIEACM1**  
**Résultats de validation croisée utilisant Fonction discriminante linéaire**

Nombre d'observations et probabilités a posteriori de la moyenne classifiées dans Relation		
De Relation	No_Relation	Yes_Relation
No_Relation	54 0.7426	25 0.6852
Yes_Relation	18 0.6480	38 0.7137
Total	72 0.7190	63 0.7024
A priori	0.5	0.5

Estimations du compte des erreurs de la probabilité a posteriori pour Relation			
Estimation	No_Relation	Yes_Relation	Total
Stratifié	0.2841	0.2989	0.2915
Non stratifi	0.2331	0.3445	0.2888
A priori	0.5000	0.5000	

Notre classification n'est pas bonne, nous obtenons juste un pourcentage de **68,15%** de bien classé. Avec **28,41%** de ceux qui ont répondu **No** a la question ont été mal classe et **29,89%** de ceux qui ont répondu **yes** a la question ont été mal classe.

**Conclusion** : La relation entre les trois nuages de point suggère 3 différentes positions concernant la pratique des organismes génétiquement modifier. Ces position doivent être comparée en terme de profession ainsi que le partie politique des individus enquêtée, car nous avons remarquée l'or de l'étude que ces deux variables supplémentaires semble être très lie. Contrairement aux trois autre variables supplémentaires, le sexe, l'âge et le faite que leurs profession soit liée ou pas avec les industries ce qui pourrais apporter plus de notions scientifique sur cette pratique des organismes génétiquement modifier.

## Annexe

**Lien vers les données** : <http://factominer.free.fr/bookV2/gmo.csv>

### **Group1: (Opinions des participants)**

<b>Variables</b>	<b>Modalités</b>
1. Do you feel implicated in the debate about GMO? (Implicated)	a lot, to a certain extent, a little, not at all
2. What is your view of GMO cultivation in France? (Position Culture)	very favourable, favourable, somewhat against, totally opposed
3. What do you think of the inclusion of GM raw materials in products for human consumption? (Position AI H)	very favourable, favourable, somewhat against, totally opposed
4. What do you think of the inclusion of GM raw materials in products to be fed to animals? (Position AI A)	very favourable, favourable, somewhat against, totally opposed
5. Have you ever taken part in an anti-GMO protest? (Protest)	Yes/No
6. Do you think the media communicate enough information about GMO? (Media Passive)	Yes/No
7. Do you take it upon yourself to find out more information about GMO? (Info Active)	Yes/No
8. Do you think that GMO might enable us to reduce the use of fungicides? (Phytosanitary products)	Yes/No
9. Do you think that GMO might enable us to reduce the problems of hunger in the world?(Hunger)	Yes/No
10. Do you think that the use of GMO might help to improve farmers' lives? (Animal feed)	Yes/No
11. Do you think that GMO might lead to future scientific advances? (Future Progress)	Yes/No
12. Do you think that GMO represent a danger to our health? (Danger)	Yes/No
13. Do you think that GMO represent a possible danger to the environment? (Threat)	Yes/No
14. Do you think that GMO represent a financial risk for farmers? (Finan risk)	Yes/No

15. Do you think that GMO are a useless scientific practice? (Useless practice)	Yes/No
16. Do you think our grandparents' generation had a healthier diet than us? (Grandparents)	Yes/No

**Group2: (Variables descriptives)**

<b>Variables</b>	<b>Modalités</b>
1. Sex	male, female
2. Professional status	farmer, student, manual labourer, senior management, civil servant, accredited professional, technician, retailer, other profession, unemployed, retired
3. Age	-25 years, 25-40 years, 40-60 years, +60 years
4. Is your profession or education in any way linked to agriculture, the food industry or the pharmaceutical industry? (Relation)	Yes/No
5. Which political movement do you most adhere to? (Political Party)	extreme left, green, left, liberal, right, extreme right