

Création, analyse de questionnaires et d'entretiens pour MAC & PC



# Analyses multivariées

Procédures de réalisation d'Analyses Factorielles de Correspondances, d'Analyses en Composante Principale, de typologies, de Profils de Modalités et de Variables...

#### **STATISTIQUE MULTIVARIEE**

Profils de modalités, AFC et représentation graphique des écarts à l'indépendance et du khi2.

Le **tri croisé** est un des outils fondamentaux dont l'analyste dispose lorsqu'il cherche à mettre en valeur les liens entre variables de type qualitatif et à construire des typologies.

Quand le nombre de variables est élevé, des méthodes plus systématiques sont utiles, par exemple, la construction de *profil de modalités* et l'*analyse factorielle des correspondances*.

Ces techniques de calcul s'articulent toutes autour d'un calcul complémentaire au tri croisé, le calcul de l'écart à l'indépendance.

L'écart à l'indépendance représente la différence entre l'effectif observé et l'effectif théorique.

Dans un tableau de contingence, les effectifs de chaque case au croisement de deux modalités sont les effectifs observés.

L'effectif théorique de chaque case se calcule de la manière suivante : le total ligne de la case est multiplié par le total colonne correspondant, puis le résultat est divisé par l'effectif total.

En d'autres termes, on considère que chaque case du tableau représente la même proportion que le total de sa ligne  $\underline{et}$  de sa colonne par rapport à l'ensemble des effectifs.

Ce résultat n'a pas d'intérêt en lui-même, mais relativement aux effectifs réels. C'est cette relation qui est calculée dans l'écart à l'indépendance.

### Exemple de tableau de contingences (adapté de P. Cibois)

Ef. observé		]							
Série du bac.	Université	Classes prép.	Orient. prof.	Total					
A	30860	2938	16638	50436					
BDD'	40926	2573	24157	67656					
С	14914	11447	4035	30396					
Technique	7822	97	42885	50804					
Total	94522	17055	87715	199292					
Ef.théorique		Destination							
Série du bac.	Université	Classes prép.	Orient, prof.	Total					
4	07004 0	474.6.0	00400.4	50476					

wai ie aa bae.		widewere et apt	with territy of with	1.0.101
A	23921,2	4316,2	22198,6	50436
BDD'	32088,5	5789,9	29777,6	67656
С	14416,5	2601,2	13378,3	30396
Technique	24095,8	4347,7	22360,5	50804
Total	94522	17055	87715	199292

Ecart		Destination						
Série du bac.	Université	Classes prép.	Orientations prof.					
A	6938,8	-1378,2	-5560,6					
BDD'	8837,5	-3216,9	-5620,6					
С	497,5	8845,8	-9343,3					
Technique	-16273.8	-4250.7	20524.5					

Khi2 par case				
Série du bac.	Université	Classes prép.	Orient. prof.	Total
A	2012,7	440,1	1392,9	3845,6
BDD'	2433,9	1787,3	1060,9	5282,1
С	17,2	30081,1	6525,3	36623,5
Technique	10991,0	4155,9	18839,2	33986
Total	15454,8	36464,3	27818,3	79737,3

Dans cet exemple, il y a 30 860 élèves de la série A qui vont à l'Université.

Effectif théorique correspondant :  $94 \quad 522*50$  $436/199 \quad 292 = 23 \quad 921,2$ 

Écart à l'indépendance : 30 860-23 921,2 = 6938,8.

Autrement dit:

Si l'effectif constaté de la case correspondait

À la fois à la proportion de la Série A dans l'ensemble de la population

Et à la proportion de ce qui vont à l'Université

Alors cet effectif réel serait de 23 921,2.

Puisqu'il est nettement plus élevé dans notre population, et que les écarts pour les autres destinations de la série A sont négatifs, on en déduit que cette série conduit plus à l'Université qu'aux classes préparatoires ou à l'orientation professionnelle.

La somme des écarts à l'indépendance d'un tableau est toujours égale à 0.

Une des difficultés consiste à déterminer un seuil minimal à partir duquel un écart à l'indépendance est significatif. Ceci dépend, bien entendu, du nombre d'individus dans l'échantillon, mais il n'y a pas de règles absolues. Aucun manuel sérieux n'en propose d'ailleurs.

Le chiffre 5 peut être un repère. Si l'effectif réel, l'effectif théorique et l'écart sont respectivement inférieurs à 5, il est risqué d'affirmer que le résultat ne relève pas du hasard.

Malgré cette relative indétermination dans les règles d'analyse, un résultat complémentaire peut apporter une aide, le Khi2 par case.

### Note sur le Khi2 par case, le test du Khi2 et le profil des modalités

Le **Khi2 par case** pondère l'écart à l'indépendance. Il ne faut pas le confondre avec le test du Khi2 global sur l'ensemble d'un tableau de tri croisé.

Chaque écart est rapporté à ses marges (ligne et colonne). Ce calcul est nécessaire lorsqu'on veut comparer deux modalités d'une variable quant à leur force respective d'attraction avec une autre variable.

Dans notre exemple, l'écart entre la Série A et l'Université (6938) est assez proche de celui qui lie la série C et les classes préparatoires (8845). Une différence de 2000 est relativement faible si on la compare avec la différence entre l'écart minimum (-16273) et l'écart maximum (20524). Sur la base de ces valeurs absolues, on peut être tenté de conclure que le lien C-Classes préparatoires n'est pas sensiblement plus fort que le lien A-Université. En fait ce lien est considérablement plus élevé. C'est ce que montre le Khi2 par case : 2012/30081 au lieu de 6938/8845. La somme des Khi2 calculés pour chaque case mesure le Khi2 global d'un tableau. Associé au degré de liberté et au niveau de signification, le Khi2 global permet le **test du Khi2** sur un tableau de contingence.

Ce test est un des moyens de savoir si l'on peut extrapoler un résultat sur la population dont est issu un échantillon. La question est la suivante : Peuton affirmer avec un degré de certitude donné que deux variables croisées ne sont pas ind pendantesÈ?

Exemple d'hypothèse : "Il y a 99 % de chances ou plus pour que les variables "Etat civil" et "Nombre d'enfants" ne soient pas indépendantes. Pour évaluer le test on se sert d'une table fournie dans les manuels de statistiques.

L'hypothèse d'indépendance peut être rejetée si la somme des Khi2 par case dépasse la valeur inscrite au croisement du **degré de liberté** du tableau ((nombre de lignes -1) multiplié par (nombre de colonnes -1)) et du **seuil de probabilité** choisi (les tables donnent en général les valeurs pour 99%, 95%, 90%).

Le test porte sur le rejet de l'hypothèse d'indépendance. Il ne mesure pas la force de l'attraction entre les variables. L'analyse précise du tableau croisé est nécessaire.

Les types de calcul mentionnés ci-dessus peuvent être effectués pour chaque tri croisé. L'analyse successive de chaque tri croisé est d'autant plus fastidieuse que le nombre de variables est élevé. De plus, le constat d'un lien entre deux variables n'exclut pas que ce lien dépende d'une troisième variable à détecter.

Plusieurs techniques peuvent aider à résoudre œ genre de problèmes sous Modalisa:

1) Filtrer les tris croisés selon le résultat du **test du Khi2**. Ce test porte sur l'ensemble des attractions ou rejets entre modalités analysés dans un tableau de contingence. Cf.: les tris croisés à l'écran ou les plans de tris croisés.

2) Faire établir par le programme le **profil des modalités** d'une question. Modalisa inventorie tous les tris croisés demandés et sélectionne les modalités croisées pour lesquelles l'écart et/ou le Khi2 par case sont au moins égal au seuil choisi.

Le test global du Khi2 n'est pas effectué ici. Un écart entre deux modalités peut être retenu sans que le global soit élevé, et ceci d'autant plus que le nombre des modalités est élevé.

Le profil de modalités permet de comparer entre elles les modalités d'une question. Pour une typologie prenant en compte simultanément un nombre élevé de variables, il est difficile de dégager des tendances globales dans le fouillis des listings d'écarts et de Khi2.

L'AFC peut aider à y mettre un peu d'ordre.

# Note sur le P.E.M. (Pourcentage de l'Écart Maximum).

Le P.E.M. local mesure le lien entre deux modalités. Le P.E.M. global généralise le calcul de ce lien pour deux questions fermées.

À la différence des indices cités ci-dessus (écart à l'indépendance, Khi2 partiel) le P.E.M. local fait référence au lien le plus fort possible entre deux modalités.

Pour une case d'un tableau croisé, le calcul de la liaison maximum consiste à y mettre "l'effectif le plus fort possible qui soit compatible avec les marges du tableau et à calculer l'écart à l'indépendance maximum qu'il y aurait dans ce cas. La force de la liaison sera simplement calculée en examinant le rapport entre l'écart à l'indépendance observé et l'écart maximum à l'indépendance calculé précédemment. Ce rapport étant mis en pourcentages, on parlera donc de Pourcentage de l'Ecart Maximum ou P.E.M." Ph. Cibois, 1993, p.45.

En généralisant le P.E.M. local on obtient le P.E.M. global. Au lieu d'une case d'un tableau, c'est le tableau complet qui est pris en compte. Pour le mode de calcul du P.E.M. global, voir Ph. Cibois, 1993, p.59-62.

Un des avantages principaux de ces indices tient à ce qu'ils sont peu sensibles à une augmentation des effectifs.

Dans Modalisa, ces deux indices sont disponibles pour les tris croisés, les profils de modalités et l'inventaire des variables discriminantes. Une présentation graphique des liens et de leur force relative peut être affichée sur une carte d'analyse factorielle des correspondances - P.E.M. local ainsi que sur le mapping résumant les liens entre variables - P.E.M. global -.

Le P.E.M. local permet de filtrer les tris croisés pertinents selon qu'ils possèdent au moins une case avec un P.E.M. local significatif. Cette possibilité de filtre des tris est active pour les tris croisés à l'écran et pour les plans de tris croisés.

Pour qu'un P.E.M. local soit retenu il faut que l'effectif soit suffisant et que l'écart à l'indépendance soit significatif. Pour évaluer si un lien est significatif, Modalisa effectue le test du khi2 sur un tableau 2\*2 incluant l'effectif de la case en position 1/1 et le solde des effectifs dans les cases adjacentes. Le P.E.M. n'est pas retenu si la signification du test est inférieure à 90 %.

NB: Pour calculer la liaison maximum d'un tableau de contingences il faut qu'il y ait un ordre sur les lignes et sur les colonnes (Cf. Ph. Cibois, 1993, pp. 59-62). Par défaut Modalisa considère qu'il n'existe pas d'ordre a priori sur les modalités. Le programme permute le tableau selon l'ordre du premier facteur d'une analyse factorielle. Si vous souhaitez éviter cette permutation et figer l'ordre des modalités de telle ou telle question, choisissez dans le Menu Classement, l'article Modifier Questions & Recodages, puis l'option 4.

### Note sur l'inventaire de variables discriminantes.

L'analyse factorielle des correspondances est un outil puissant mais difficile à interpréter quand le nombre de questions et de modalités est très élevé.

De plus elle a tendance à mettre en valeur les écarts sur les modalités aux effectifs faibles.

Le profil de modalité présente les résultats croisés à partir d'une seule question.

Entre les tris à plat, les tris croisés et les profils de modalités d'une part et les techniques plus sophistiquées des analyses factorielles d'autre part, il s'agit ici de faire émerger les questions à partir desquelles il y a le plus de chances de discriminer les individus étudiés en prenant en compte l'ensemble des attractions simultanées entre questions et, si de telles questions existent dans l'enquête, de les sélectionner pour réaliser des cartes d'analyse factorielle pertinentes et lisibles.

Le programme permet de classer les questions/variables selon leur caractère discriminant par un inventaire systématique des écarts positifs à l'indépendance, du Khi2 partiel, du P.E.M. global et du test du Khi2 entre toutes les questions choisies.

Interprétation des résultats:

1) Plus la moyenne du P.E.M. global, la somme des écarts inventoriés, la somme des khi2 partiels ou le nombre de tests du khi2 significatifs pour une variable sont élevés, plus la variable a de chances d'être discriminante relativement au groupe des variables choisies pour l'analyse.

2) Le détail des liens variable par variable permet de détecter des groupes de variables qui pourraient être traités dans une analyse factorielle au niveau des modalités.

### Note sur l'analyse factorielle des correspondances

L'AFC a pour objectif une présentation graphique synthétique et optimale des proximités et distances entre les modalités, proximités et distances calculées systématiquement sur la base des Khi2 par case.

L'AFC est une des techniques d'analyse des données. Elle s'est particulièrement développée en France sous l'impulsion de J.P. Benzecri.

Les a priori de Modalisa s'inspirent des contributions de Philippe Cibois à la didactique de l'analyse factorielle. Cf. notamment:

Philippe Cibois, <u>L'analyse factorielle</u>, PUF, Que sais-je, 1983, 126 p..

Philippe Cibois, <u>L'analyse des données en</u> sociologie, PUF, Le sociologue, 1984, 218 p.

L'analyse factorielle sert d'abord à simplifier la lecture des grands tableaux de nombres.

Elle permet de remplacer un tableau de nombres par une bonne approximation de celui-ci, permettant une représentation graphique des liens entre les modalités de plusieurs variables.

L'analyse peut être utile sur un tri croisé entre 2 variables seulement, mais on travaille en général sur des tris croisés multiples.

L'interprétation d'un mapping (représentation graphique des modalités sur 2 axes) est assez délicate.

Le fait que deux modalités soient proches ne signifie pas nécessairement qu'elles sont en attraction. Il est rare en effet que tous les liens entre modalités puissent être représentés en deux dimensions.

Pour appréhender la position relative d'un point, il faudrait pouvoir le représenter simultanément en plusieurs dimensions, autant qu'il y a de facteurs dans l'analyse.

Pour aider à l'interprétation, les programmes proposent généralement des informations chiffrées indiquant le poids relatif de chaque point pour chaque facteur.

Dans Modalisa l'approche est un peu différente. On part du principe qu'un résumé du lien entre deux modalités peut être exprimé par l'écart à l'indépendance ou le Khi2 par case (écart pondéré par ses marges).

Pourquoi ne pas montrer graphiquement ce lien? N'est-ce pas une manière de résoudre le problème de l'interprétation de la proximité entre points sur tel ou tel axe?

Si l'on accepte ce point de vue, l'analyse factorielle perd son caractère ésotérique et se réduit à une technique permettant de représenter de manière suggestive le maximum d'informations sur les attractions entre modalités.

L'AFC devient une technique exploratoire qui permet un premier débroussaillage des données et qui peut suggérer des pistes fructueuses.

Il faut toujours revenir aux tris croisés ou aux profils de modalités pour vérifier les suggestions de l'approche graphique.

Le fait que l'AFC révèle peu d'attractions entre variables peut être un résultat intéressant en tant que tel.

### Note sur les typologies ou classifications automatiques

Dans Modalisa<sup>™</sup> les classifications automatiques sont construites sur la base des distances entre individus sur les cartes d'AFC (ou d'ACP). Voir en fin de documentation les précisions sur les caractéristiques de chaque forme de typologie.

Modalisa<sup>™</sup> permet de créer une ou plusieurs questions fermées à partir des typologies: chaque modalité de la question créée représente un type.

Il devient donc possible de tester la pertinence d'une typologie sur la population étudiée en créant des profils de modalités à partir des nouvelles questions créées.

Les modalités d'une question-type ainsi créée peuvent être représentées sur une carte d'AFC en variable supplémentaire.

L'alternance entre les différentes techniques présentées ci-dessus permet d'affiner progressivement l'analyse et de justifier pas à pas les résultats synthétiques.

### **PROFILS DE MODALITES**

Lors de la réalisation d'un profil de modalités Modalisa<sup>TM</sup> inventorie tous les tris croisés demandés et sélectionne les modalités croisées pour lesquelles l'écart et/ou le Khi2 par case sont au moins égal au seuil choisi. Le test global du Khi2 n'est pas effectué ici.

Un écart entre deux modalités peut être retenu sans que le Khi2 global soit élevé, et ceci d'autant plus que le nombre de modalités est élevé.

Pour plus amples précisions se reporter au Chapitre 7 et aux indications bibliographiques, en particulier l'ouvrage de Ph. Cibois.

Afin de préparer sélectivement des plans d'AFC il est possible d'effectuer un calcul systématique des écarts à l'indépendance (pondérés et non-pondérés), à partir d'une question de départ (à réponse unique ou à réponses multiples) croisée avec toutes les autres questions (ou un groupe choisi de questions).

Procéder comme suit:

# Profil de modalités Annuler O Créer les profils des modalités pour une question choisie 1 O Lire un fichier de profils déjà créé (Extension: "\_PROr.TXT") 2

#### Création d'un profil de modalité

🖬 Profil de modalités / Habitudes alimentaires : démo 📃 🔲 🗙								
Choix de la question sur laquelle les profils seront calcul	és							
Variable	Туре	Mod.	Code	Classement	N*			
Nombre d'enfants	Un	3		1	<b></b>			
Activité professionnelle	Un	2		2				
CSP	Un	2		3				
Origine géographique	Un	4		4				
Age	Un	2		5				
Produit utilisé le plus souvent en weekend	Un	3		6				
Produit utilisé le plus souvent en semaine	Un	3		7				
Critère de choix pour les courses	Un	3		8				
Repas du soir habituel avec les enfants	Un	2		9				
Repas du soir tous les jours à la même heure	Un	2		10				
Qualité principale d'une bonne alimentation	Un	6		11	-			
Chercher-								
Par numéro 💌 4.			Annule	er C	IK			

Choisir d'abord la question sur laquelle les profils doivent être calculés pour chaque modalité de la question, y compris les Sans réponses.

Profil de modalité	s / Habitude	s alimentaires : démo	1	_ 🗆 🗵
Titre du plan	Plan051		Ext.: PRO	
Question profil	4. Origine géog	graphique		
Questions à croiser	Toutes	O Sélection		
Questionnaires	Tous	O Sous-population		
	Redressem	ent		
Sans réponses	O Incluses	• Exclues	Décimales po	our les %
Seuils pour les profils				
Minimum / écart no Minimum / écart po	n pondéré (Eca ndéré (Khi2 part	t à l'indépendance) iel)		5
Copier un plan	.   Q	uestion Profil	Annuler	Continuer

# Définition des options du plan de création de profil

- questions à croiser
- tous les questionnaires ou une sous-population
- prise en compte ou non d'un redressement
- inclusion ou non des sans réponses

- seuils pour les écarts à prendre en compte (Ecart à l'indépendance et Khi2 par case sont expliqués en fin de documentation).

Modalisa<sup>™</sup> donne, après avoir effectué les calculs et avant de les afficher à l'écran, la possibilité de stocker les résultats et libellés dans un fichier externe.

Un écran présente pour la première modalité de la question de départ la liste des effectifs, écarts et Khi2 pour chaque modalité des questions sélectionnées pour lesquelles des écarts ont été trouvés.

Question	Modalité	Effectifs	Ecarts	Khi2	PEM	Test Khi2	
2SP	Prof. Lib. & Cadres sup.	51	8	1,530	25 %	"	ī.
Critère de choix pour les courses	Facilité de préparation	30	7	1,872	13 %	1	
Repas du soir habituel avec les enfants	Non	20	5	1,803	14 %		
réquentation des fast-food	Oui	30	15	14,758	39 %		
réquentation des fast-food / Grignotage entre les repas	Oui	30	15	13,923	39 %	ш	1
		ब				D	đ

Lecture de la première ligne du tableau:

Dans la catégorie Professions libérales & Cadres supérieurs, 51 individus sont originaires de Paris.

La valeur théorique pour ce croisement de modalités est de 42,8. L'écart à l'indépendance est donc de 8,2 (51-42,8).

Pondérée par les marges du tableau (total ligne et total colonne) cette valeur devient 1,5. C'est la colonne Khi2 (par case).

La colonne P.E.M. affiche la valeur du P.E.M. local. Cette notion est expliquée en fin de documentation.

La colonne de droite présente les résultats du test du Khi2 pour chaque tableau réduit à 2 lignes et 2 colonnes. De 1 à 3 "**|**" selon la signification du test, respectivement 90, 95 et 99 %.

Pour chaque lien affiché, Modalisa calcule le test du Khi2 sur un tableau 2\*2 incluant l'effectif de la case en position 1/1 et le solde des effectifs dans les cases adjacentes.

Cette information permet notamment d'exclure des liens relevant de fluctuations d'échantillonnage.

L'ascenseur permet de faire défiler les résultats pour chaque modalité de la question.

"Groupe de questions" permet de créer un nouveau groupe de questions au vu de l'inventaire effectué.

Le bouton "Tous" affiche la liste complète des écarts inventoriés.

Modalité	Question	Modalité	Effectifs	Ecarts	Khi2	PEM	Test Khi2	
Paris	Repas du soir habituel avec les er	Non	20	5	1,803	14 %		
Paris	Fréquentation des fast-food	Oui	30	15	14,758	39 %	III	
Paris	Fréquentation des fast-food / Grig	Oui	30	15	13,923	39 %	III	
Banlieue	CSP	Ouvriers	32	9	3,824	31 %	III	
Banlieue	Age	de 25 à 34 ans	32	6	1,142	21 %		
Banlieue	Fréquentation des fast-food	Oui	17	6	3,290	15 %	1	
Banlieue	Fréquentation des fast-food / Grig	Oui	17	6	3,443	15 %		
Province	Critère de choix pour les courses	recherche d'équilibre	37	6	1,102	15 %		
Province	Qualité principale d'une bonne alir	Equilibrée	66	8	1,221	14 %	l	
Province	Fréquentation des fast-food	Non	100	17	3,409	74 %	III	
Province	Fréquentation des fast-food / Grig	Non	100	17	3,541	16 %	III	
Autre	Produit utilisé le plus souvent en s	produits frais	14	5	2,987	51 %		
	<b>x</b>							Þ

Un simple clic sur l'intitulé de la colonne effectue un tri par ordre croissant.

Le bouton Chercher affiche ce dialogue de sélection des profils à afficher



# Lire un fichier de profil de modalités déjà créé

Cette option permet de relire un fichier de profil de modalité (extension.Mo2.®), créé à partir d'un plan de tri.

Choisir le fichier à relire; les résultats s'affichent sur le même principe que précédemment.

#### **PROFIL DE VARIABLES**

#### **Principe et fonction**



Lors du dépouillement des résultats d'une enquête par questionnaires incluant de nombreuses questions fermées (variables qualitatives) avant ou après recodage, une des tâches les plus délicates de l'analyste consiste à repérer les questions qui permettent de constituer des groupes d'individus qui se ressemblent suffisamment à l'intérieur de chaque groupe tout en se distinguant de manière significative des individus des autres groupes.

Les analyses factorielles et autres profils de modalités ont été élaborés pour faciliter ce genre d'exercice.

Ces techniques se révèlent souvent décevantes quand le nombre de questions posées est élevé et quand la connaissance de la population étudiée n'est pas suffisante pour déceler à priori les modalités à partir desquelles les groupes pertinents peuvent être constitués.

🚊 Pr	ofil de variable:	: / Habitudes alimentaires : démo	)	_ 🗆 ×				
Tit	re du plan	Plan053	Ext.: DIS					
Qu	estions à croiser	⊙ Toutes O Sélection						
Qu	uestionnaires	Tous     O Sous-population						
		Redressement						
Sa	ans réponses	O Incluses 💿 Exclues	Décimales pour les %					
Se	uils pour les profils							
M M	Minimum / écart non pondéré (Ecart à l'indépendance)         5           Minimum / écart pondéré (Khi2 partiel)         1							
	Copier un plan	Question Profil	Annuler Conti	inuer				

Les analyses les plus "faciles" et "gratifiantes" de ce point de vue sont celles qui s'appuient sur des théories sociales bien établies à partir desquelles un corpus d'hypothèses peut être constitué a priori. (ex. PCS / parcours éducatif ou opinion politique).

Les investigations sur ce type d'enquêtes permettent souvent de confirmer et d'affiner les présupposés théoriques qui ont présidé à l'élaboration du questionnaire.

Sans négliger les avantages d'une étude s'appuyant sur des hypothèses bien établies, sans négliger non plus l'importance décisive de la qualité du questionnaire et de l'échantillon.

Ce module se propose de faciliter des investigations permettant de remettre au premier plan l'approche empirique au moment de l'analyse. Plus les questions et les modalités sont nombreuses, plus leur combinatoire potentielle est complexe.

Moins les a priori sont nombreux sur les questions et les modalités discriminantes, moins il est aisé de repérer les associations pertinentes.

L'analyse factorielle des correspondances est un outil puissant mais difficile à interpréter quand le nombre de questions et de modalités est très élevé. De plus elle a tendance à mettre en valeur les écarts sur les modalités aux effectifs faibles.

Le profil de modalité présente les résultats croisés à partir d'une seule question.

Entre les tris à plat, les tris croisés et les profils de modalités d'une part et les techniques plus sophistiquées des analyses factorielles d'autre part, il s'agit ici de faire émerger les questions à partir desquelles il y a le plus de chances de discriminer les individus étudiés en prenant en compte l'ensemble des attractions simultanées entre questions et, si de telles questions existent dans l'enquête, de les sélectionner pour réaliser une carte d'analyse factorielle pertinente et lisible.

Le programme permet de classer les questions (variables) selon leur caractère discriminant d'après le critère de l'écart à l'indépendance par un inventaire systématique des écarts positifs à l'indépendance, du Khi2 partiel, du P.E.M. global et du test du Khi2 entre toutes les questions choisies.

#### Interprétation des résultats:

1) Plus la moyenne du P.E.M. global, la somme des écarts inventoriés, la somme des Khi2 partiels ou le nombre de tests du Khi2 significatifs pour une variable sont élevés, plus la variable a de chances d'être discriminante relativement au groupe des variables choisies pour l'analyse.

2) Le détail des liens variable par variable permet de détecter des groupes de variables qui pourraient être traités dans une analyse factorielle au niveau des modalités.

Une première liste permet d'appréhender globalement le résultat variable par variable.

Pour analyser chaque somme des écarts, deux options sont proposées:

1) Par variables deux à deux.

La somme des écarts de chaque variable est détaillée selon le ou les sous-ensembles qui la constitue au niveau des variables liées, c'est-à-dire la ou les sommes des écarts inventoriés pour chaque variable croisée.

Cette option permet de détecter les variables les plus fortement liées simultanément entre elles.

Les liens et les distances de telle modalité à telle modalité mis en évidence sur une AFC classique sont ici masqués.

2) Par modalités à partir d'une variable choisie (équivalent de "Profil de modalités").

La somme des écarts d'une variable est détaillée selon les sous-ensembles qui la constitue au niveau de chacune de ses modalités, c'est-à-dire le ou les écarts inventoriés pour chaque modalité de la variable choisie avec les modalités des variables croisées.

Les sélections, les tris et les graphiques d'analyse factorielle permettent d'affiner progressivement l'analyse.

#### Précautions

Prévoir un temps de traitement correspondant au nombre de questions à analyser, en particulier si le groupe de questions comporte une sélection importante de questions à réponses multiples.

Nombre de tris croisés à réaliser:  $(n^{*}(n-1))/2$ , pour n = nombre de questions choisies.

Le plan: "Variables discriminantes" peut être utile pour différer le traitement.

### **Options de calcul**

Le paramétrage est analogue à celui du profil de modalités (cf. § précédant).

Profil de variable	s / Habitudes alimentaires : démo	
Titre du plan	Plan053	Ext.: DIS
Questions à croiser	• Toutes • Sélection	
Questionnaires	Tous O Sous-population	
	Redressement	
Sans réponses	O Incluses O Exclues	Décimales pour les %
Seuils pour les profils		
Minimum / écart no Minimum / écart po	n pondéré (Ecart à l'indépendance) ndéré (Khi2 partiel)	5
Copier un plan	Question Profil	Annuler Continuer

### Présentation à l'écran des variables discriminantes

Les listes de résultats et les graphiques peuvent à tout moment être imprimés, exportés et copiés vers le presse-papiers.

Le format présente les données calculées sous forme de liste.

Variable	PEM Moyenne	Ecarts Nombre	Ecarts Somme	Ecarts Moyenne	Khi2 partiel Somme	Khi2 partiel Moyenne	Test du khi2 99%	Test du khi2 95%	Test du khi2 90%	T k
1. Nombre d'enfants	13%	7	68,6	9,8	26,7	3,8	1	1	1	
2. Activité professionnelle	24%	10	73,8	7,4	23,9	2,4	3	0	2	
3. CSP	17%	17	205,7	12,1	93,5	5,5	5	1	1	
4. Origine géographique	16%	14	129,7	9,3	57,8	4,1	2	2	1	
5. Age	14%	15	134,3	9,0	37,7	2,5	3	1	1	
6. Produit utilisé le plus souvent en weekend	24%	6	35,9	6,0	22,0	3,7	0	1	1	
<ol> <li>Produit utilisé le plus souvent en semaine</li> </ol>	22%	25	225,1	9,0	97,6	3,9	4	3	6	
8. Critère de choix pour les courses	17%	23	193,9	8,4	70,5	3,1	5	2	3	
<ol><li>Repas du soir habituel avec les enfants</li></ol>	17%	3	20,1	6,7	7,0	2,4	0	2	0	
10. Repas du soir tous les jours à la même heu	14%	7	52,8	7.5	12,3	1,8	1	0	2	
11. Qualité principale d'une bonne alimentation	20%	17	163,5	9,6	92,4	5,4	4	1	1	
12. Changements perçus dans les habitudes a	17%	13	108,7	8,4	43,6	3,4	1	3	1	
<ol> <li>Grignotage entre les repas</li> </ol>	20%	26	532,6	20,5	1442,3	55,5	2	2	1	ĺ
	4									١

#### Lecture de la première ligne

La variable n°3 comprend 3 modalités et s'appelle "Nombre d'enfants".

Le croisement de cette variable avec toutes les autres variables du groupe donne les résultats suivants:

Moyenne des P.E.M. globaux	14
Nombre d'écarts à l'indépendance supérieurs au seuil choisi de 5	7
Somme de ces écarts	63
Moyenne de ces écarts	9
Somme des Khi2 partiels supérieurs au seuil choisi de 1	26
Moyenne de ces écarts	3,7
Nombre de tests du Khi2 dont la signification est supérieure ou égale à 99%	1
Nombre de tests du Khi2 dont la signification est supérieure ou égale à 95% et inférieure à 99%	1

Dans cette analyse la variable la plus discriminante est la PCS, sauf pour le nombre d'écarts retenus.

#### Tris sur toutes les rubriques de calcul

Pour le test du Khi2, la dernière colonne affiche le nombre de tris croisés pour lequel le test n'est pas fiable parce qu'au moins une case du tableau des valeurs théoriques est inférieure à 5.

#### Création de groupe de questions

Sur l'ensemble des questions ou un groupe choisi

Profil de modalité pour une des questions

Sélectionner la question et cliquer sur le bouton PROFIL.

#### Profil de modalités pour la question PCS obtenu à partir de la liste des variables discriminantes.

Voir le chapitre "Profil des modalités" pour un commentaire sur les valeurs affichées sur cette liste.

L'ascenseur permet de trier la colonne correspondante par ordre croissant.

#### Création de groupe de questions, sur l'ensemble des questions du profil ou un groupe choisi.

Cliquez sur le bouton "Groupe de questions" pour afficher la liste des questions concernées par le profil. Sélectionnez les questions du groupe à créer.

#### AFC

Sur l'ensemble des questions ou un groupe choisi sur la liste.

Si vous souhaitez que le plan d'analyse factorielle affiche automatiquement le groupe de questions courant, enregistrez ce groupe sur le premier dialogue présenté.

Inventaire des attractions par paires de questions

Bouton "Par paire" sur la liste des variables discriminantes.

📄 Profil de variables / Ha	abitudes alimentaires : démo				. 🗆 🗙
Nombre de paires: 71 (Test du l	<hi2>=90%)</hi2>	Ecarts	Khi2	PEM	Test 🔺 Khi2
1. Nombre d'enfants	2. Activité professionnelle	6	3	15 %	1
1. Nombre d'enfants	5. Age	34	15	41 %	l ii l
1. Nombre d'enfants	9. Repas du soir habituel avec les e	7	3	19%	i ii
2. Activité professionnelle	5. Age	6	1	22 %	1 i
2. Activité professionnelle	7. Produit utilisé le plus souvent en s	10	4	37 %	i ii ii
2. Activité professionnelle	21. Fréquentation des fast-food	18	4	90 %	
2. Activité professionnelle	21R1. Fréquentation des fast-food /	24	7	18 %	
3. CSP	<ol> <li>Origine géographique</li> </ol>	17	5	17 %	
3. CSP	<ol> <li>Produit utilisé le plus souvent en s</li> </ol>	12	10	27 %	
3. CSP	<ol> <li>Critère de choix pour les courses</li> </ol>	8	2	12 %	1
3. CSP	<ol> <li>11. Qualité principale d'une bonne a</li> </ol>	39	22	40 %	
3. CSP	<ol> <li>16. Changement du nombre de plats</li> </ol>	33	15	36 %	
3. CSP	17. Changement dans l'abondance	22	11	43 %	
3. CSP	20. TV regardée pendant les repas	49	21	62 %	
4. Origine géographique	8. Critère de choix pour les courses	12	3	20 %	1
1		t	t	Ļ	t
Copier C		Retour			
					Ē

Lecture de la première ligne

Le tri croisé entre la variable 1-Nombre d'enfants et la variable 5-Classes d'âges donne les résultats suivants:

Somme des écarts à l'indépendance supérieurs au seuil choisi de 5	34
Somme des Khi2 partiels supérieurs au seuil choisi de 1	15
P.E.M. global	41%
Signification du test du khi réalisé entre les deux variables	99% ou plus

#### Par variable

Le bouton "Par variable" permet d'afficher la liste des liens relatifs à une variable.



Seuls les tris croisés dont le test du Khi2 est supérieur ou égal à 90 % sont affichés.

### Analyse factorielle (ACP) sur la liste des liens par paires

Les valeurs pour le calcul de l'ACP peuvent être sélectionnées.

Analyse factorielle en composantes principales	
	Annuler
S Ecart à l'indépendance	1
🔿 S Khi2 partiel	2
O Pourcentage de l'écart maximum	3

La carte factorielle situe graphiquement les questions les unes par rapport aux autres en prenant en compte l'ensemble des attractions simultanées en fonction de leur force relative.

Les questions les plus excentrées par rapport à l'origine des axes sont les plus discriminantes.

Quand les questions excentrées sont proches, elles ont des liens simultanés forts. Les liens entre questions sont marqués par des lignes dont l'épaisseur est proportionnelle à la force relative de l'attraction (1 à 5 points). Possibilité de visualiser les attractions question par question.

# Exemples de cartes d'analyse factorielle

Lors du premier affichage, le programme ne dessine que les liens les plus importants. Dans cet exemple 11 P.E.M. sur 19.

L'article "Lien suivant" du menu "Carte" permet d'afficher par ordre décroissant les liens restants.

Avec l'article "Seuil pour Lien" du menu "Carte" vous pouvez déterminer le lien minimum à afficher.



Dans cet exemple la variable Activité professionnelle à un lien avec la variable Age. On constate une forte liaison entre Age et Nombre d'enfants.

Le menu "Variables" permet de dessiner les liens variable par variable.

Sous-sélection des liens pour la question PCS:

Toutes les options de visualisation des cartes factorielles sont disponibles (zoom, déplacement, choix des facteurs, inversion des facteurs, modification des libellés, copier-coller,...)

NB1. Comme pour toute analyse factorielle, une carte ne présente que deux facteurs. Les effets de perspective peuvent donc exister.

NB2. A la différence d'une ACP classique sur variable quantitative, les oppositions sur la carte ne marquent pas une corrélation globale négative (coefficient de corrélation proche de -1).

Cette distance marque ici les liens relativement moins forts (relativement aux liens les plus forts inventoriés).

# A.F.C.: ANALYSES FACTORIELLES DES CORRESPONDANCES

Modalisa permet de réaliser une AFC à partir des questions de l'enquête courante ou d'un tableau de chiffres, au format ASCII extérieur, soit de relire une analyse demandée précédemment.

Analyse factorielle des correspondances	
	Annuler
O Nouvelle analyse à partir de questions de l'enquête	1
O Nouvelle analyse à partir d'un tableau de nombres à importer	2
<ul> <li>Afficher les résultats d'un plan déjà exécuté (Extension: "_AFCr.TXT")</li> </ul>	3

Paramètrer l'AFC à partir de l'écran suivant...

Analyse factoriel	le des correspondances / Habitude	es alimenta
Titre du plan	Plan058	Ext.: AFC
Questions	Toutes     O Sélection	
Questionnaires	Tous     O Sous-population	
	Redressement	
Sans réponses	O Incluses O Exclues	Facteurs 3
Libellés	Modalités     O Quest. et modalités	🖸 Libellé long 🔿 Libellé abrégé
Individus	Afficher sur mapping	Editer les coordonnées
Calculer et éditer	Coordonnées et contributions Seuils Ecarts et Khi2	Tableau de BURT
	Minimum / écart non pondéré (Ecart à Minimum / écart pondéré (Khi2 partiel)	l'indépendance) [5 [1
Copier un plar	1	Annuler Continuer

# Nouvelle analyse à partir de questions de l'enquête

La création du plan d'analyse factorielle des correspondances permet, comme pour tous les autres plans, de définir les éléments devant être pris en compte pour les calculs demandés ainsi que la présentation souhaitée pour les résultats.

Saisir le titre du plan (27 caractères au maximum)

Choisir les questions sur lesquelles sera calculée l'AFC:

Toutes les questions (option utilisable dans le cadre d'enquêtes comprenant peu de questions).

Une sélection de questions, dans ce cas créer un groupe de questions ou choisir un groupe déjà créé.

Important : la taille du tableau principal (tableau de BURT à deux dimensions) pour le calcul de l'AFC en mémoire vive peut être calculé en prenant le carré de la somme des modalités des questions choisies.

La limite théorique de la taille d'un tableau à 2 dimensions est de 100 000 lignes sur 100 000 colonnes, ce qui représente 10 milliards de cellules!

Indiquer si l'AFC doit porter sur tous les questionnaires ou sur une sous-population, et dans ce cas, la sélectionner, puis sur les données brutes

ou redressées, et dans ce cas, sélectionner de même un redressement. Indiquer si les sans réponses doivent être prises ou non en compte.

Préciser le nombre de facteurs à extraire, par défaut ce nombre est égale à 3.

Définir la présentation des modalités sur la carte: soit les modalités seules, en entier ou en abrégés, soit les modalités avec les titres des questions.

Ces choix pourront être modifiés après visualisation de la carte.

Cocher les cases des résultats souhaités : les cases "Coordonnées et Contributions" et "Écarts et Khi2" sont cochées par défaut, possibilité d'y ajouter le "Tableau de BURT".

Redéfinir éventuellement les seuils minimaux pour les écarts: par défaut 5 pour les écarts non pondérés et 1 pour les écarts pondérés.

Il est possible de copier un plan déjà créé afin de n'avoir qu'à le modifier partiellement, cliquer dans ce cas sur le bouton Copier un plan.

Ces choix effectués, cliquer sur Continuer. Modalisa<sup>TM</sup> calcule l'AFC, opération qui peut prendre du temps si les questions et/ou les questionnaires à prendre en compte sont en nombre élevé; il est préférable dans ce cas de lancer les plans à partir du Menu PLAN.



NB: Les lignes et colonnes dont le total est égal à 0 ne sont pas prises en compte dans les calculs.

Gestion de l'affichage de la carte d'AFC (mapping)

Les options du menu Format et Facteurs, ainsi que les icônes de la barre verticale située à gauche de la carte, permettent de jouer sur l'affichage de la carte. Les intitulés de modalité peuvent être librement modifiés en cliquant sur le nom de la modalité, manipulation qui ouvre une fenêtre permettant d'effectuer la modification. Il est également possible de modifier le choix de présentation des modalités défini dans le plan, grâce aux différents choix "Libellés" du menu Modalités≤.

sélectionner ce pointeur dans la barre des icônes pour modifier les libellés de modalités

Demand	er
<b>?</b>	Nouveau libellé :
	Profession libérale & Cadre sup.
	Annuler OK



pour changer la taille du mapping utiliser les loupes, choix Zoom + ou Zoom - du menu Format pour déplacer la carte dans la fenêtre,

cliquer sur la main

permet de remettre la carte dans la position d'origine

à utiliser pour passer au facteur suivant



à utiliser pour permuter les axes.



123

permet de localiser les individus soit sous la forme de points, soit en reprenant les N° de questionnaires pour marquer une sous-population (Hommes/-25 ans/Paris).

Permet aussi de marquer les individus correspondant aux modalités d'une question à réponse unique (Hommes/Femmes).

Ce bouton change de fonction dès qu'une souspopulation est choisie. Les sous-populations sont gardées en mémoire.

Cette fonction est surtout utilisable avec un écran couleur.

Marquer Sous-population(s)	
·	Annuler
O Une des sous-populations créées	1
<ul> <li>C Les individus correspondant aux modalités d'une question à réponse unique</li> </ul>	2

NB: Dès qu'une typologie est calculée ce bouton change de fonction. Les types sont gardés en mémoire si l'on décide par exemple d'afficher les souspopulations ou les numéros de questionnaires. Il suffit de cliquer sur ce bouton pour réafficher les derniers types calculés.



active l'écran de choix de typologie à réaliser





Les trois options de menu typologie à savoir Centres mobiles, Ascendante hiérarchique A-H et CM puis AH affichent l'écran ci-dessus.

Pour plus ample information sur les typologies se reporter aux paragraphes qui leurs sont consacrés (Typologie dans la partie Analyse, plan de typologie dans la partie Plans...).



L'icône en forme de tableau cadrillé permet de créer une grille d'échelle variable permettant de repérer le nombre des individus ainsi que la forme de leur représentation sur la carte



Sur la représentation en étoile le nombre de branche est égal au nombre des individus présents.



La taille des cercles est ici proportionnelle au nombre des individus.





ce pointeur permet d'afficher les numéros des individus présents dans une cellule de la grille de la carte (cf. grille ci-dessus).



Lorsque la grille n'est pas activée ce pointeur permet d'afficher les références des individus présents sur un point.



cet icône n'est actif qu'avec un écran couleurs. Il permet de faire varier la couleur des éléments constitutifs des cartes



### Affichage des liens entre modalités et seuils pour les contributions

Pendant une AFC, Modalisa enregistre les écarts à l'indépendance pondérés et non-pondérés et les P.E.M. locaux.

Le menu "Carte" permet:

- de gérer l'affichage des liens à l'écran

#### Article "Lien suivant"

Dessine à l'écran le lien suivant par ordre décroissant.

#### Article "Masquer tous les liens"

Efface de la carte tous les liens dessinés.

#### Article "Afficher les liens"

Permet de sélectionner le type de liens et le seuil minimal pour le lien courant. Cf. ci-après l'article "Seuil pour liens et contributions".

#### Article "Liste des liens"

Permet de consulter, imprimer ou exporter la liste des liens calculés.

Article "Seuil pour liens et contributions".

Dans la première zone vous sélectionnez le lien courant et le minimum affiché pour ce type de lien.

NB: Si un seuil a été choisi pour les contributions, seules seront affichées les modalités (et donc les liens) pour lesquelles le seuil est atteint.

Seuils pour Liens et contributions		
Seuils pour les liens affichés		
Pourcentage de l'écart maximum     Ecart à l'indépendance     Khi2 par case     Aucun lien		
Minimum demandé pour le lien courant (	0-100%)	14
Minimum / Contributions		
Contribution absolue (0-100%) Moyenne Contribution relative (0-100%)	: 5,5	0,0
	Annuler	OK

La zone du bas permet de fixer un seuil pour les contributions absolues et relatives sur les deux facteurs courants.

Il suffit que la valeur soit atteinte sur l'un des deux facteurs pour que la modalité soit affichée. La moyenne des contributions absolues est proposée.

Dans la plupart des cas cette valeur moyenne permet d'exclure les modalités dont la contribution absolue au facteur n'est pas significative.

### Affichage des valeurs des contributions

Sur le mapping, les contributions absolues sont affichées par défaut sous forme de rectangles. Elles peuvent être également affichées en ellipses, ou masquées.

Pour obtenir la liste des valeurs des contributions absolues et relatives, sélectionnez l'article "Contributions-Liste" du menu "Carte".

	s					
Eacteure:			1		2	
* Information:			- 28 au	nev	32.26	5%
26 Information:			Bool B	JAS Dolot A	Bool E	J/6 Dolot
Lonaldans.		LODDA			DSDI. I	
JURDANIE		JURDAN	0,00	0,00	0,00	0,00
LIBAN		LIBAN	0,01	0,00	0,00	0,00
LAUS		LAUS	0,01	0,00	0,00	0,00
JAPUN		JAPUN	0,00	0,00	0,00	0,00
TALLE		TALIE	0,01	0,00	0,00	0,00
ISRAEL		ISRAEL	0,00	0,00	0,01	0,00
JAMAIQUE		JAMAIU	0,00	0,00	0,00	0,00
IVUIRE (COTE D')	ļ	IVUIRE	0,00	0,00	0,09	0,00
MALAISIE		MALAIS	0,00	0,00	0,00	0,00
MADAGASCAR		MADAG.	0,00	0,00	0,00	0,00
MALI		MALI	0,32	0,03	0,03	0,00
MALAWI		MALAW	0,00	0,00	0,00	0,00
MACEDUINE		MACEDI	0,00	0,00	0,00	0,00
LIBYE		LIBAE	0,00	0,00	0,00	0,00
LIBERIA		LIBERI	0,00	0,00	0,00	0,00
MACAU		MACAU	0,00	0,00	0,00	0,00
LUXEMBOORG		LUXEME	0,00	0,00	0,00	0,00
THAILANDE		THAILA	0,00	0,00	0,00	0,00
TCHECUSLOVAU	UIE	TCHECU	0,00	0,00	0,00	0,00
TRINITE ET TUB	AGU	TRINU	0,00	0,00	0,00	0,00
TOGO		TOGU	0,00	0,00	0,00	0,00
TCHAD		TCHAD	0,00	0,00	0,00	0,00
SULTANAT D'UM	AN .	SULIAN	0,00	0,00	0,01	0,00
SUISSE		SUISSE	0,00	0,00	0,00	0,00 🖵
1		1	Ŧ	+	Ŧ	+
Tri/départ				Déta	il : Doul	ble-clic
Détail	E	xporter			Reto	ur 📘

Aucune option d'affichage du mapping n'est irréversible, ces options étant destinées à permettre à l'utilisateur de trouver, par essais successifs, la présentation la plus satisfaisante par rapport à ses besoins.

#### Marquer les modalités d'une question

Dans le menu "Modalités" l'article "Marquer une question" permet de faire apparaître les libellés d'une question sélectionnée en gras sur le mapping, afin d'améliorer la lisibilité.

Pour annuler ce marquage, sélectionnez l'article "Annuler marques".

### Variables supplémentaires ou illustratives

Pour afficher sur la carte une ou plusieurs variables supplémentaires choisir l'article Supplémentaire dans le Menu Modalités.

Modalisa<sup>™</sup> s'appuie sur les résultats de l'AFC existante pour positionner les modalités des nouvelles variables.

Les libellés de modalités de variables illustratives sont affichés en italique.

#### Impression de l'AFC

Il est possible d'imprimer l'ensemble des résultats ainsi que le mapping

• impression des contributions: menu Fichier, choix Imprimer Résultats.

	Facteur nº1 : 62,6%			Facteur nº2 : 12,4%			Facteur nº3 : 12,0%		
	Position	C. Abs.	C. Rel.	Position	C. Abs.	C. Re1.	Position	C. Abs.	C. Rel.
Homme	-0,019	0,02	0,00	-0,412	44,71	0,94	-0,044	0,54	0,01
Femme	-0,055	0,14	0,01	0,468	50,60	0,94	-0,010	0,02	0,00
001	-0,557	18,43	0,96	-0,003	0,00	0,00	-0,042	0,56	0,00
NON	0,757	22,32	0,96	0,005	0,00	0,00	-0,006	0,00	0,00
Conditionnement individu	0,833	24,78	0,96	0,037	0,25	0,00	-0,002	0,00	0,00
Conditionnement famillia	-0,533	17,78	0,95	-0,022	0,15	0,00	-0,043	0,62	0,00
Cinema	0,608	12,05	0,56	-0,002	0,00	0,00	0,030	0,16	0,00
Plage	0,122	0,76	0,08	-0,028	0,20	0,00	-0,317	26,75	0,53
Rue	-0,684	1,19	0,45	0,523	3,50	0,26	-0,387	1,98	0,14
Détaillant	-0,290	1,55	0,36	0,071	0,47	0,02	-0,258	6,41	0,28
Grande Surface	-0,114	0,93	0,06	-0,014	0,07	0,00	0,411	62,90	0,80

- impression de la valeur des écarts: menu Écarts, choix Valeur des écarts, puis imprimer ou exporter
- tableau de BURT: Modalisa exporte automatiquement un fichier (extension BRT) qui peut être ouvert à partir d'un tableur ou traitement de textes.
- -impression du mapping:

Pour plus de souplesse, le mapping peut également être copié en format PICT, lu, en tant que tel, par des programmes tels que MacDraw, afin de le retoucher (texte, couleurs, taille et épaisseur des traits...).



Par ailleurs, cette copie est lisible au format PICT et/ou bitmap dans un tableur ou un traitement de texte.

#### Suggestion

Le traitement de texte Word permet depuis la version 5.0 d'éditer un graphique PICT.

Pour imprimer une carte contenant un grand nombre de modalités il peut être utile de changer la police et/ou de réduire la taille de la police.

Faites un double-clic sur le graphique collé dans Word. Demandez Tout sélectionner et choisissez par exemple la police Times avec une taille inférieure à 9 points.

# **AFC** A PARTIR D'UN TABLEAU DE NOMBRE A IMPORTER

Analyse factorielle des correspondances	
	Annuler
O Nouvelle analyse à partir de questions de l'enquête	1
O Nouvelle analyse à partir d'un tableau de nombres à importer	2
<ul> <li>Afficher les résultats d'un plan déjà exécuté (Extension: "_AFCr.TXT")</li> </ul>	3

Choisir et ouvrir le fichier à importer, qui doit obligatoirement être en format ASCII (Excel format texte, par exemple).

Analyse factorielle des correspondances / Habitu	des alimenta 📃 🗖 🗙
Fichier :	
C:\MODALISA 4\Plan017_PR0i.TXT	
Libellés :	
<ul> <li>Présents dans le fichier</li> <li>(Entêtes de colonnes et de lignes)</li> </ul>	9 Aucun
Tableau des nombres :	
2 colonnes	
2 lignes	
Code ASCII / fin de rubr: 9 Code ASCII / fin de fiche: 13	
Annuler	OK

Il s'agit en général d'un tableau de tris croisés.

Si les libellés des lignes et des colonnes sont présents dans le fichier à importer, cliquer sur le bouton radio correspondant.

Dans ce cas, le nombre de lignes et de colonnes que vous saisissez ne doit pas comptabiliser la ligne et la colonne des libellés.

### Afficher les résultats d'un plan d'AFC déjà exécuté (extension AFC.®).

Dans la boite de dialogue que vous propose Modalisa<sup>™</sup>, choisir et ouvrir le plan d'AFC dont les résultats déjà calculés sont à importer.

# PRESENTATION DES QUESTIONS NUMERIQUES EN CLASSES

Variable	Туре	Mod.	Code	Classement	N° de recoda
Ancienneté dans la fonction	Num		4	4	3
Rang de Ancienneté dans la fonction	Num			4	
Ancienneté dans l'administration territoriale	Num		5	5	1
Nombre d'agents permanents	Num		12	12	1
• I					F

Cette option ne permet pas de créer un recodage, mais elle offre des possibilités qui ne sont pas actives sur le format de recodage en classes :

Les boutons variable suivante, variable précédente, atteindre, graphique, copier, dupliquer l'écran sont actives.

Le menu local sous-opulation permet de sélectionner une sous population parmi ceux existantes dans la base.

Mettre à jour Nombre de classes		Amplitude auto Une o Amplitude choisie Eff		ne classe par valeur Effectifs ± égaux		OK
4. Ancienneté dans la fonction						
Edition borne inf.	N*	Bornes	Amplitude	Effectifs	%	Cumul
0	1	Moins de 6	6	1 441	57,4%	57,4%
6	2	de 6 à moins de 12	6	720	28,7%	86,1%
12	3	de 12 à moins de 18	6	255	10,1%	96,3%
18	4	de 18 à moins de 24	6	77	3,0%	99,4%
24	5	de 24 à moins de 30	6	12	,4%	99,8%
30	6	30 et plus	6	3	,1%	100,0%
						Ī
<u> </u>						-

L'utilisateur peut choisir le nombre de classes, l'amplitude, une classe par valeur,...

Dans l'encadré en bas de la fenêtre ci-dessus, Modalisa donne les valeurs de la moyenne, le maximum, le minimum, l'écart-Type, l'étendue, le nombre de classes,...

L'activation de l'icône graphique fait apparaître une fenêtre de dialogue permettant à l'utilisateur de choisir le type de graphique désiré.

Graphique	
	Annuler
O Moyenne et ± Ecart-type	1
O Boîte de dispersion	2
O Effectifs pour chaque valeur de la variable	3
O Effectifs pour chaque classe	4

Le choix 4 (Effectifs pour chaque classe) donne le graphique ci-dessous.



### Stats n Var / n Sous-pop



Ce menu offre deux options, la première concerne les statistiques comparées pour toutes les modalités d'une question fermée. La seconde portera sur une selection de sous-population.

### Tableau de moyennes pour plusieurs variables

Choisissez d'abord la question fermée dans la fenêtre ci-dessous.

Variable	Туре	Mod.	Code	Classement	N*
Critère de choix pour les courses	Un	3		8	
Repas du soir habituel avec les enfants	Un	2		9	
Repas du soir tous les jours à la même heure		2		10	
Qualité principale d'une bonne alimentation		6		11	
Changements perçus dans les habitudes alimentaires	Un	3		12	
Grignotage entre les repas	Mult	5		13	
Grignotage entre les repassr	Mult	5		13	
Lieu de prise des Repas à la maison	Un	2		14	
Changement dans le contenu des repas	Un	2		15	
Changement du nombre de plats	Un	3		16	
Changement dans l'abondance des portions	Un	3		17	ľ
र					

Puis choisissez une ou plusieurs variables de type numérique.

ariable	Туре	Mod.	Code	Classement	N° de recoda
i oui, dépense moyenne par personne	Num		ĺ –	22	4
SSAI	Num			22	
ombre de repas pris en fast-food par mois	Num			23	1
d 1					Þ
ilection continue : Maiuscula alia / Sálastia	n discon	inue : Co	ntrôle-cli	-	

L'algorithme calcule cinq résumés statistiques (moyenne, écart-type, somme, nombre de réponses et nombre de valeurs manquantes), évalue les moyennes pour lesquelles l'écart est significatif au risque de 5 % et effectue le test de Fisher sur chaque variable numérique.

	Statistiques co	mparées / Hat	oitudes alimentai	ires : démo	
	Comparaison de I	moyennes			Fermer
				Ecart-type	Ecart-type 💌
<b>.</b>		22. Sioui, dépense	23. Nombre de repas pris en	Ecart-type Somme	
12	Non réponse			Nombre	۸.
	Fruits	5,27	3,67	Sans rép	
	Friandises	4,62	2,38		
	produits laitiers	6,01	2,33		
	Autres	5,82	2,64		
	Rien	6,22	2,38		
	GLOBAL	6,08	2,7		
					~
		4			
	Test de la moyer significativement Test de Fisher:	nne : Une case es différente de la m	t soulignée lorsque oyenne de la varial	la moyenne correspondan ble (risque de 5%)	te est
	22. Si oui, dépen	ise moyenne par p	personne: F=2,43	p=0,045 Discriminant	

Le menu local permet de passer d'un résumé statistique à un autre (moyenne, écart-type, nombre de réponses et nombre de valeurs manquantes).

Le bouton "Transposer" permet d'afficher les variables en lignes ou inversement.

NB : Les formules de calcul des écarts significatifs sont présentées en Fin de documentation.

Le bouton présentant deux petits fichiers permet de copier le tableau dans le presse-papiers, ainsi il possible de le coller dans un traitement de texte ou dans un tableur. Pour visualiser le tableau sous forme d'un graphique cliquer sur le bouton correspondant.



Ce graphique résume le tablau de moyennes pour plusieurs variables. Les modalités de la question fermée de type multiple "grignotage entre les repas" sont en abscisse, par contre la moyenne est sur l'axe des oordonnées.

Les batonnets représentent les deux variables numériques.

#### Tableau de moyennes pour plusieurs souspopulations et plusieurs variables

Choisissez une ou plusieurs sous populations dans la fenêtre ci-dessous.

🔛 Statistiques co	mparées / Habitudes alir	nentaires : dém	D		- 🗆 ×
	Effectifs lors de	la création ou de l	a mise à jour :	Sous-pop	Ensemble 🔺
CADRES 1 ENFANT	PARIS BANLIEUE			23	264
Dépenses moyenne	264	264			
dépenses moyenne	52	264			
Origine géographique	19	264			
Origine géographique	/ Autre			19	264
Origine géographique	/ Banlieue			53	264
Origine géographique / Banlieue					264
Origine géographique / Paris					264
Origine géographique / Paris					264
Origine géographique / Province					264
Origine géographique / Province					264
Origine géographique / Sans réponses					264
Origine géographique / Sans réponses					264
Sous-population 013				52	264
Information : Double	e-clic				
Tout	Chercher par nom				
Copier la liste	Chercher par variables	Détail	Annuler	Sél	ection 🚽
•					Þ

Choisissez ensuite les variables numériques à évaluer.

🧮 Statistiques comparées / Habitudes alimentaires : démo 📃 🗖 🔀						
Sélection de variable						
Variable	Туре	Mod.	Code	Classement	N° de recoda	
Si oui, dépense moyenne par personne	Num			22	4 🔶	
ESSAI	Num			22		
Nombre de repas pris en fast-food par mois	Num			23	1	
					$\overline{\mathbf{v}}$	
					Þ	
Sélection continue : Majuscule-clic / Sélectio	n discor	itinue : Co	ntrôle-clia	-		
				Annuler	ОК	

Les résumés statistiques seront ventilés pour chaque modalité de la question fermée ou pour chaque sous-population.

	🔜 Statistiques comparées / Habitudes alimentaires : démo							
	Comparaison de moyennes					Fermer		
				Ecart-type	Ec.	art-type 💌		
<b>11</b>		22. Sioui, dépense	22R. I	Moyenne Ecart-type Somme	re de en			
12	CADRES 1 ENFANT PARIS BAN	5,59		Nombre	2,3	1 🔺		
	Origine géographique / Autre			Sans rép				
	Origine géographique / Banlieue	5,23		104	2,6	9		
	Origine géographique / Paris	6,45		86	2,8	2		
	Origine géographique / Province	6,89		44,77	1,7	5		
	Sous-population 013	6,08		90,44	2,	7		
	GLOBAL	6,08		90,44	2,	7		
						~		
		4						
	Test de la moyenne : Une case est significativement différente de la mo Test de Fisher:	soulignée lorsqu oyenne de la varia	e la mo able (ri:	oyenne corres sque de 5%)	pondante es			
	22. Si oui, dépense moyenne par p	ersonne: F=0,55	p=0,3	3 Non discrim	iinant			

Le graphique en lignes qui suit résume les écarttypes des différents sous-populations.



### **REALISATION D'UNE ACP**



On peut calculer des ACP à partir des questions numériques de l'enquête courante, d'un tableau de chiffres extérieur au format ASCII, d'une matrice de corrélation précédemment créée et exportée ou relire une analyse demandée précédemment.

Titre du plan	Plan080	Ext.: ACP
Questions	• Toutes • Sélection	
Questionnaires	O Tous O Sous-population	Social
Facteurs	3	
ndividus	Afficher sur mapping	Editer les coordonnées
Tableaux à éditer	Matrice des corrélations	
Conier un pla	» 1	Annular Continuer

# Création d'une nouvelle ACP à partir de variables de l'enquête

Saisir le titre du plan (27 caractères au maximum)

Choisir les questions sur lesquelles sera calculée l'ACP : toutes les questions ou une sélection de questions, dans ce cas créer un groupe de questions ou choisir un groupe déjà créé.

Indiquer si l'ACP doit porter sur tous les questionnaires ou sur une sous-population, et dans ce cas, la sélectionner. Préciser le nombre de facteurs à extraire, par défaut ce nombre est égale à 3. Indiquer les résultats souhaités : "Matrices des corrélations" ou "Coordonnées des individus", les deux cases sont cochées par défaut.



Il est possible de copier un plan déjà créé afin de n'avoir qu'à le modifier partiellement, cliquer dans ce cas sur le bouton Copier un plan.

Ces choix effectués, cliquer sur Continuer. Modalisa<sup>™</sup> calcule l'ACP.

Le carré positionné devant les libellés représente les coordonnées de chaque variable par rapport aux facteurs choisis (minimum -1 maximum 1).

Le rayon du cercle est de 1.

La position des individus n'est pas calculée sur la même échelle que celle des variables.

Sont significatives : la direction et la distance relative des individus par rapport à l'origine des axes.

La concentration des individus peut être modifiée à l'aide de l'article Coefficient de dispersion du menu Individus.



Il est également possible de modifier le choix de présentation des modalités défini dans le plan, grâce aux différents choix "Libellés" du menu Données.



Sélectionner ce pointeur dans la barre des icônes pour modifier les libellés de modalités





pour changer la taille du mapping utiliser les icônes en forme de loupe, choix Zoom + ou Zoom du menu Format



pour déplacer la carte dans la fenêtre, cliquer sur la main



permet de remettre la carte dans la position d'origine



à utiliser pour passer au facteur suivant



à utiliser pour permuter les axes.



permet de localiser les individus soit sous la forme de points, soit en lettres (A,B,C...), soit en reprenant les numéros de questionnaires.



pour marquer une sous-population (Hommes/-25 ans/Paris).

permet de créer une question fermée à réponse unique à partir de la typologie.

NB: Dès qu'une typologie est calculée ce bouton change de fonction. Les types sont gardés en mémoire si l'on décide par exemple d'afficher les souspopulations ou les numéros de questionnaires.

Il suffit de cliquer sur ce bouton pour réafficher les derniers types calculés.



Pour plus ample information sur les typologies se reporter aux paragraphes qui leurs sont consacrés (Typologie dans la partie Analyse, plan de typologie dans la partie Plans...).

# Nouvelle ACP à partir d'un tableau de nombre à importer

Choisir et ouvrir le fichier à importer, qui doit obligatoirement être en format ASCII (Excel format texte, par exemple).

Les variables sont en colonnes et les individus en lignes.



#### Création d'une nouvelle analyse à partir d'une matrice de corrélation créée

Choisir dans la base le fichier à importer.

Cliquer sur Continuer pour obtenir le tableau des Valeurs Propres puis le mapping.

### Afficher les résultats d'un plan d'ACP déjà exécuté (extension.Mo2.®).

Dans la boite de dialogue que vous propose Modalisa<sup>™</sup>, choisir et ouvrir le plan d'ACP Mo2.® à importer.

#### **PLAN DE PROFILS DE MODALITES**

Choisir la question à réponse(s) unique ou multiples à traiter.

Variable	Type	Mod.	Code	Classement	N°	
Nombre d'enfants	Un	3		1		
Activité professionnelle	Un	2		2		
CSP	Un	2		3		
Origine géographique	Un	4		4		
Age	Un	2		5		
Produit utilisé le plus souvent en weekend	Un	3		6		
Produit utilisé le plus souvent en semaine	Un	3		7		
Critère de choix pour les courses	Un	3		8		
Repas du soir habituel avec les enfants	Un	2		9		
Repas du soir tous les jours à la même heure	Un	2		10		
Qualité principale d'une bonne alimentation	Un	6		11	-	

Pour lancer le plan de profils de modalités, sélectionner OK.

Plan - Profils / H	abitudes alimentaires : démo					
Titre du plan	Plan064	Ext.: PRO				
Question profil	3. CSP					
Questions à croiser	Toutes C Sélection					
Questionnaires	Tous C Sous-population					
	Redressement					
Sans réponses	O Incluses O Exclues	Décimales pour les %				
Seuils pour les profils						
Minimum / écart no Minimum / écart po	Minimum / écat non pondéré (Ecat à l'indépendance) 5 Minimum / écat pondéré (Khi2 patiel) 1					
Copier un plan	Question Profil	Annuler Enregistrer				

Après la saisie du titre du plan, définir les options de ce plan de profils et les seuils désirés.

Questions à croiser pour le calcul des profils : toutes les questions ou sur une sélection de questions (création d'un groupe de questions ou choix sur liste d'un groupe déjà créé).

Tri sur tous les questionnaires ou sur une souspopulation et, le cas échéant, choix sur liste d'une sous-population.

Traitement sur les données brutes ou sur les données redressées (Choix sur liste d'un redressement).

Inclusion ou non des sans réponses.

Nombre de décimales pour l'affichage des écarts: 0, 1 ou 2.

Seuils minimaux pour les écarts:

Options par défaut: 5 pour les écarts non pondérés; 1 pour les écarts pondérés.

Il est possible de changer la question sur laquelle les profils seront calculés en cliquant sur le bouton Question profil. Une fois ces choix effectués, cliquer sur "Enregistrer" afin de créer ce fichier plan de profils. Il sera possible de l'exécuter ultérieurement.

Les profils de modalités permettent de détecter rapidement les variables dont les modalités sont susceptibles de présenter des attractions significatives pour un seuil donné d'écarts à l'indépendance (pondérés ou non).

Cette méthode permet de détecter un groupe de variable à même de créer une AFC pertinente.

La méthode s'inspire de P. CIBOIS (cf. réf. Biblio. en fin de documentation).

#### **PLAN DE VARIABLES DISCRIMINANTES**



Le paramétrage est analogue à celui du plan de profils de modalités.

### PLAN D'AFC

Cette option présente les mêmes choix qu'en création d'AFC par le Menu Analyse

Saisir le titre du plan (27 caractères au maximum)

Plan - A.F.C. / Ha	bitudes alimentaires : démo		×
Titre du plan	Plan065	Ext.: AFC	
Questions	Toutes     O Sélection		
Questionnaires	Tous O Sous-population		
	Redressement		
Sans réponses	O Incluses O Exclues	Facteurs 3	
Libellés	● Modalités  ● Quest. et modalités	C Libellé long C Libellé abrégé	
Individus	C Afficher sur mapping	Editer les coordonnées	
Calculer et éditer	Coordonnées et contributions Seuils Ecarts et Khi2	🖵 Tableau de BURT	
	Minimum / écart non pondéré (Ecart à Minimum / écart pondéré (Khi2 partiel)	'indépendance) 5 1	
Copier un plan		Annuler Enregistrer	J

Choisir les questions sur lesquelles sera calculée l'AFC : toutes les questions ou une sélection de questions, dans ce cas créer un groupe de questions ou choisir un groupe déjà créé.

Indiquer si l'AFC doit porter sur tous les questionnaires ou sur une sous-population, et dans

ce cas, la sélectionner, puis sur les données brutes ou redressées, et dans ce cas, sélectionner de même un redressement.

Indiquer si les sans réponses doivent être prises ou non en compte.

Préciser le nombre de facteurs à extraire, par défaut trois.

Définir la présentation des modalités sur le mapping: soit les modalités seules, en entier ou en abrégés, soit les modalités avec les titres des questions.

Ces choix pourront être modifiés après visualisation du mapping.

Cocher les cases des résultats souhaités: les cases "Coordonnées et Contributions" et "Écarts et Khi2" sont cochées par défaut, possibilité d'y ajouter le "Tableau de Burt"

Redéfinir éventuellement les seuils minimaux pour les écarts: par défaut 5 pour les écarts non pondérés et 1 pour les écarts pondérés.

Il est possible de copier un plan déjà créé afin de n'avoir qu'à le modifier partiellement, cliquer dans ce cas sur le bouton Copier un plan.

Une fois ces choix effectués, cliquer sur Enregistrer afin de créer ce paramétrage de plan.

### PLAN D'ACP

Cette option présente les mêmes choix qu'en menu Analyse

Saisir le titre du plan (27 caractères au maximum)

Indiquer les questions sur lesquelles sera calculée l'ACP: toutes les questions ou une sélection de questions, et création, dans ce cas, d'un groupe de questions ou choix sur liste d'un groupe déjà créé.

Traitement sur tous les questionnaires ou sur une sous-population, et, le cas échéant, choix sur liste d'une sous-population.

Nombre de facteurs à extraire, par défaut trois.

Individus devant apparaître ou non sur le mapping.

Indiquer les résultats souhaités: "Matrices des corrélations" ou "Coordonnées des individus", les deux cases sont cochées par défaut.

Titre du plan	Plan066	Ext: ACP
Questions	Toutes     Sélection	
Questionnaires		
Facteurs	3	
ndividus	Afficher sur mapping	Editer les coordonnées
Tableaux à éditer	Matrice des corrélations	

Il est possible de copier un plan déjà créé afin de n'avoir qu'à le modifier partiellement, cliquer dans ce cas sur le bouton Copier un plan.

Une fois ces choix effectués, cliquer sur Enregistrer afin de créer ce paramétrage de plan.

### PLANS DE TYPOLOGIE



La constitution d'un plan de typologie suppose la création préalable d'un plan d'AFC ou d'ACP.



Nombre de type: minimum 2 et maximum 100.

Nombre de facteurs : la distance entre les individus est évaluée à partir de leur position sur deux, trois ou n dimensions (nombre de facteurs). Par défaut, Modalisa<sup>™</sup> propose deux facteurs (suffisants dans la plupart des cas).

Nombre d'essais (pour les typologies sur centres mobiles): en l'absence d'optimum le nombre d'essais permet à l'utilisateur de choisir un seuil qu'il juge adapté.

Les informations concernant les typologies ascendantes hiérarchiques fournissent: les numéros de nœuds, les deux individus (ou nœuds) associés pour créer le nouveau nœud, l'effectif puis la distance avec le nœud précédent.

Créer une question fermée: les modalités de la nouvelle variable fermée à réponse unique seront constituées par les types calculés.

La variable comportera donc autant de modalités que vous avez demandé de types.

Il n'est pas possible de tronquer le dendrogramme par plan. Cette option est disponible sur le graphique affichant le dendrogramme.

Il est possible de créer un profil dès lors que l'on a créé une question fermée.

Plan - Typologie / Habitudes alimentaires : démo				
Titre du plan	Plan068 Ext.: TYP	]		
Distances entre individus / Coordonnées A.C.P., Plan : "Plan066"				
Typologie	⊙ Centres mobiles O Ascendante hiérarchique ⊙ CM puis AH			
	5 types 2 facteurs			
	10 essais (centres mobiles)			
	Editer les informations sur la typologie AH (noeuds)			
V Créer une question fermée (une modelité per tune)				
<ul> <li>✓ Profil de chaque type (chaque modalité de la question fermée)</li> </ul>				
Questions/profil				
Seuils/profil	Seuils/profil Minimum / écart non pondéré (Ecart à l'indépendance) 5 Minimum / écart pondéré (Khí2 partiel) 1			
Copier un plan Annuler Enregistrer				