

# modalisa

Création, analyse de questionnaires et d'entretiens pour  
Windows 10 et MacOS 10

## 9



### Analyses multivariées

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

Informations complémentaires : [www.modalisa.com](http://www.modalisa.com) - [modalisa@kynos.info](mailto:modalisa@kynos.info)

<b>La statistique multivariée .....</b>	<b>4</b>
<b>Ecart à l'indépendance .....</b>	<b>4</b>
Exemple .....	4
<b>Test du Khi<sup>2</sup> .....</b>	<b>4</b>
<b>Pourcentage de l'Écart Maximum (P.E.M.).....</b>	<b>5</b>
<b>Inventaire de variables discriminantes .....</b>	<b>6</b>
<b>Analyse factorielle des correspondances (A.F.C.).....</b>	<b>6</b>
<b>Typologies ou classifications automatiques.....</b>	<b>6</b>
<b>Profils de modalités .....</b>	<b>7</b>
<b>Création d'un profil de modalité .....</b>	<b>7</b>
<b>Définition des options du plan de création de profil .....</b>	<b>7</b>
<b>Lire un fichier de profil de modalités déjà créé.....</b>	<b>8</b>
<b>Plan de profils de modalités .....</b>	<b>8</b>
<b>Profil de variables .....</b>	<b>8</b>
<b>Principe et fonction .....</b>	<b>8</b>
<b>Interprétation des résultats.....</b>	<b>9</b>
<b>Précaution .....</b>	<b>9</b>
<b>Options de calcul.....</b>	<b>9</b>
<b>Présentation à l'écran des variables discriminantes.....</b>	<b>9</b>
Tris sur toutes les rubriques de calcul.....	10
Création de groupe de questions .....	10
Profil de modalités pour la question PCS obtenu à partir de la liste des variables discriminantes.....	10
<b>Création d'un groupe de questions sur l'ensemble des questions du profil ou un groupe choisi .....</b>	<b>10</b>
Lecture de la première ligne.....	10
Par variable.....	10
Analyse factorielle (ACP) sur la liste des liens par paires .....	11
<b>Exemple de carte d'analyse factorielle .....</b>	<b>11</b>
<b>Plan de variables discriminantes .....</b>	<b>11</b>
<b>Analyse Factorielle des Correspondances (A.F.C.) .....</b>	<b>12</b>
<b>Principe.....</b>	<b>12</b>
<b>Création .....</b>	<b>12</b>
Affichage des liens entre modalités et seuils pour les contributions .....	14
Affichage des valeurs des contributions.....	15
Marquer les modalités d'une question.....	15
Variables supplémentaires ou illustratives .....	15
Impression de l'AFC.....	15
<b>A.F.C à partir d'un tableau de nombres à importer.....</b>	<b>Erreur ! Signet non défini.</b>
Afficher les résultats d'un plan d'AFC déjà exécuté (extension AFC®).....	<b>Erreur ! Signet non défini.</b>
<b>Plan d'A.F.C.....</b>	<b>15</b>
<b>Présentation des questions numériques en classes .....</b>	<b>Erreur ! Signet non défini.</b>
Stats n Var / n Sous-pop .....	<b>Erreur ! Signet non défini.</b>
Tableau de moyennes pour plusieurs variables.....	<b>Erreur ! Signet non défini.</b>
Tableau de moyennes pour plusieurs sous-populations et plusieurs variables .....	<b>Erreur ! Signet non défini.</b>
<b>Analyse en Composantes Principales (A.C.P.).....</b>	<b>16</b>

<b>Principe</b> .....	Erreur ! Signet non défini.
<b>Création</b> .....	<b>16</b>
Création d'une nouvelle ACP à partir de variables de l'enquête .....	16
<b>Modification</b> .....	Erreur ! Signet non défini.
<b>A partir d'un tableau de nombre à importer</b> .....	Erreur ! Signet non défini.
<b>A partir d'une matrice de corrélation</b> .....	<b>17</b>
Afficher les résultats d'un plan d'ACP déjà exécuté (extension.Mo2.®).....	17
<b>Plan d'A.C.P.</b> .....	<b>17</b>
<b>Plan de typologies</b> .....	<b>18</b>

## La statistique multivariée

Le tri croisé est un des outils statistiques fondamentaux lorsqu'on recherche les liens entre variables qualitatives et à construire des typologies/segmentations.

Quand le nombre de variables est élevé, les méthodes d'analyse multivariée sont utiles, par exemple la construction de profils de modalités ou l'analyse factorielle des correspondances. Ces techniques de calcul s'articulent toutes autour d'un calcul complémentaire au tri croisé, celui de l'écart à l'indépendance.

### Ecart 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 observés sont les effectifs de chaque case au croisement de deux modalités. L'effectif théorique de chaque case est calculé de la manière suivante :

$(\text{Total ligne} * \text{total colonne}) / \text{Effectif total}$ .

En d'autres termes, on considère que chaque case du tableau représente la même proportion que le total de sa ligne 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 réellement observés. C'est cette relation qui est calculée dans l'écart à l'indépendance.

### Exemple

(adapté de Philippe Cibois)

Série du bac	Orientation			Total
	Université	Classes Préparatoires	Orientation Professionnelle	
A	30860	2938	16638	50436
BDD'	40926	2573	24157	67656
C	14914	11447	4035	30396
Technique	7822	97	42885	50804
<b>Total</b>	<b>94522</b>	<b>17055</b>	<b>87715</b>	<b>199292</b>

Série du bac	Orientation			Total
	Université	Classes Préparatoires	Orientation professionnelle	
A	23921,2	4316,2	22198,6	50436
BDD'	32088,5	5789,9	29777,6	67656
C	14416,5	2601,2	13378,3	30396
Technique	24095,8	4347,7	22360,5	50804
<b>Total</b>	<b>94522</b>	<b>17055</b>	<b>87715</b>	<b>199292</b>

Série du bac	Orientation		
	Université	Classes Préparatoires	Orientation professionnelle
A	6938,8	-1378,2	-5560,6
BDD'	8837,5	-3216,9	-5620,6
C	497,5	8845,8	-9343,3
Technique	-16273,8	-4250,7	20524,5

Khi <sup>2</sup>	Orientation
------------------	-------------

par case

Série du bac	Université	Classes Préparatoires	Orientation professionnelle	Total
A	2012,7	440,1	1392,9	3845,6
BDD'	2433,9	1787,3	1060,9	5282,1
C	17,2	30081,1	6525,3	36623,5
Technique	10991	4155,9	18839,2	33986
<b>Total</b>	<b>15454,8</b>	<b>36464,3</b>	<b>27818,3</b>	<b>79737,3</b>

Dans cet exemple, 30 860 élèves ayant passé le Baccalauréat de la série A vont à l'Université.

L'effectif théorique correspondant se calcule ainsi :  $(94522 * 50436) / 199292 = 23921,2$ .

L'écart à l'indépendance se calcule ainsi :  $30860 - 23921,2 = 6938,8$ .

Autrement dit, si l'effectif constaté de la case correspondait à la fois à la proportion des élèves de la série A dans l'ensemble de la population, et à la proportion de ceux qui vont à l'Université, alors cet effectif serait de 23 921,2 élèves.

Mais l'effectif réellement observé dans la population est nettement plus élevé (30860 élèves), et les écarts pour les autres orientations que prennent les élèves de la série A sont négatifs (-1378,2 et -5560,6), on peut en déduire que le Baccalauréat de série A conduit plus souvent à l'Université qu'en classes préparatoires ou en 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 du nombre d'individus dans l'échantillon, mais il n'y a pas de règle absolue. 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 à l'indépendance 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'analyses, un résultat complémentaire peut apporter une aide, le  $\text{Khi}^2$  par case.

### Test du $\text{Khi}^2$

Le  $\text{Khi}^2$  par case pondère l'écart à l'indépendance. Il ne faut pas le confondre avec le test du  $\text{Khi}^2$  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 de la série A à l'Université est assez proche de celui qui lie la série C et les classes préparatoires. Comparée à la différence entre l'écart minimum (-162173) et l'écart maximum (20524), une différence de 2000 est relativement faible.

Sur la base de ces valeurs absolues, il est tentant de conclure que le lien Série C-Classes préparatoires est équivalent au lien Série A-Université. Or le calcul du  $\text{Khi}^2$  par case montre que ce lien est considérablement plus élevé :  $2012 / 30081$  versus  $6938 / 8845$ .

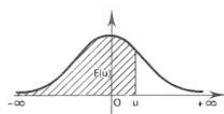
La somme des Khi2 calculés pour chaque case mesure le Khi2 global d'un tableau. Associé au degré de liberté et au niveau de signification, le Khi2 global permet le test du Khi2 sur un tableau de contingence.

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 : peut-on affirmer avec un degré de certitude donné que deux variables croisées ne sont pas indépendantes ?

Exemple d'hypothèse : « il y a 99% de chances ou plus que les variables « état civil » et « nombre d'enfants » ne soient pas indépendantes ». Pour évaluer le test, on se sert de la table suivante, qui figure dans les manuels de statistiques.

$F(u)$  : probabilité de trouver une valeur inférieure à  $u$  (probabilité cumulée), égale à la surface hachurée :

$$u = \frac{X - \bar{X}}{\sigma}$$



L'emploi de cette table exige par conséquent la standardisation préalable de la valeur de  $X$  dont on veut connaître la probabilité cumulée;  $u$  se lit dans la première colonne pour sa partie entière et sa première décimale, la deuxième décimale se trouvant dans la première ligne.

$u$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Exemples de lecture

Pour  $u = +0,92$ ,  $F(u) = 0,8212$ .  
 Pour  $u = -0,92$ ,  $F(u) = 1 - 0,8212 = 0,1788$ .  
 Pour  $F(u) = 0,6020$ ,  $u \approx +0,26$  (correspondant à  $0,26\sigma + \bar{x}$ ).  
 Pour  $F(u) = 0,3527$ ,  $1 - F(u) = 0,6473$  et  $u \approx -0,38$ .

L'hypothèse d'indépendance peut être rejetée si la somme des Khi2 par case dépasse la valeur inscrite au croisement du degré de liberté ((nombre de lignes - 1) x (le nombre de colonnes - 1)) et du seuil de probabilité choisi. Les tables donnent en général les valeurs pour 90%, 95% et 99%. Modalisa propose également ces 3 seuils.

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 ce 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 analysées 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.

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

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. A 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 tant mis en pourcentages, on parlera donc de Pourcentage de l'Écart Maximum ou P.E.M." Philippe 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 Philippe 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 2x2 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. Philippe 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.

### 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.

### Analyse factorielle des correspondances (A.F.C.)

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 donn.es. 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, L'analyse factorielle, PUF, Que sais-je, 1983, 126 p.

Philippe Cibois, L'analyse des données en 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 d'analyses 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.

### Typologies ou classifications automatiques

Dans Modalisa 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 permet de créer une ou plusieurs questions fermées 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 modalités

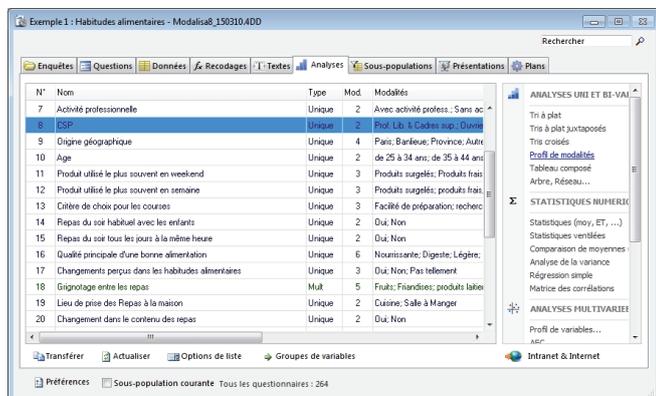
Lors de la réalisation d'un profil de modalités, Modalisa inventorie tous les tris croisés demandés et sélectionne les modalités croisées pour lesquelles l'écart et/ou le Khi2 par case sont au moins égal au seuil choisi. Le test global du Khi2 n'est pas effectué ici.

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

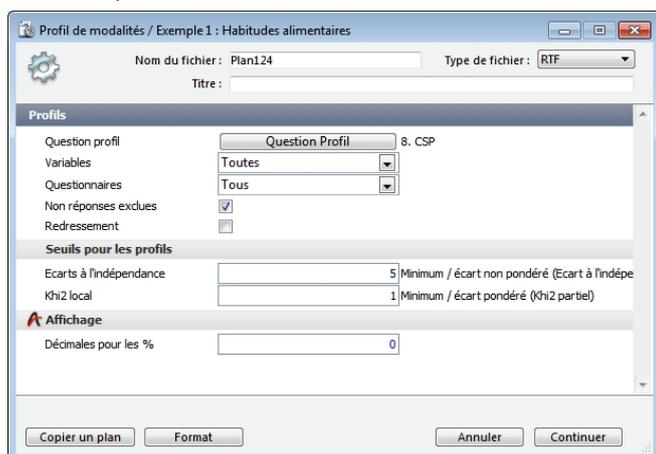
Pour plus amples précisions se reporter au Chapitre 7 et aux indications bibliographiques, en particulier à l'ouvrage de Philippe Cibois.

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).

## Création d'un profil de modalité



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



## 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>a</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.

Variable	Modalité	Effectifs	Ecarts	Khi2	PEM	Test Khi2 local
9. Origine géograp...	Paris	51	8	1,530	25 %	**
13. Critère de choix...	recherche d'équilibre	48	8	1,518	26 %	**
16. Qualité principa...	Equilibrée	94	20	5,190	35 %	***
21. Changement d...	Moins de plats	61	17	6,303	50 %	***
22. Changement d...	Moins copieux	61	9	1,436	21 %	**
25. TV regardée pe...	TV Non	121	25	6,266	46 %	***

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). Pondère 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.

Profil de modalités / Exemple 1 : Habitudes alimentaires

8. CSP  
Modalité 1. Prof. Lib. & Cadres sup. (L51 = 57,2%)

Variable	Modalité	Effectifs	Ecarts	Khi2	PEM	Test Khi2 local
21. Changement d...	Moins de plats	61	17	6,303	50 %	***
25. TV regardée pe...	TV Non	121	25	6,266	46 %	***
16. Qualité principa...	Equilibrée	94	20	5,190	35 %	***
13. Critère de choix...	recherche d'équilibre	48	8	1,518	26 %	**
9. Origine géograp...	Paris	51	8	1,530	25 %	**
22. Changement d...	Moins copieux	61	9	1,436	21 %	**

Groupes de questions Tous Transférer AFC << >> Fermer

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

### 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.

### Plan de profils de modalités

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

Plan - Profils / Exemple 1 : Habitudes alimentaires

Choix de la question sur laquelle les profils seront calculés

7	Activité professionnelle	Un	2	Avec activité profess
8	CSP	Un	2	Prof. Lib. & Cadres su
9	Origine géographique	Un	4	Paris; Banlieue; Provi
10	Age	Un	2	de 25 à 34 ans; de 3
11	Produit utilisé le plus souvent en weekend	Un	3	Produits surgelés; Prc
12	Produit utilisé le plus souvent en semaine	Un	3	Produits surgelés; prc
13	Critère de choix pour les courses	Un	3	Facilité de préparation
14	Repas du soir habituel avec les enfants	Un	2	Oui; Non
15	Repas du soir tous les jours à la même heure	Un	2	Oui; Non
16	Qualité principale d'une bonne alimentation	Un	6	Nourissante; Digeste
17	Changements perçus dans les habitudes alimentaires	Un	3	Oui; Non; Pas telleme
18	Grignotage entre les repas	Multi	5	Fruits; Friandises; pro
19	Lieu de prise des Repas à la maison	Un	2	Cuisine; Salle à Mang
20	Changement dans le contenu des repas	Un	2	Oui; Non
21	Changement du nombre de plats	Un	3	Plus de plats; Moins
22	Changement dans l'abondance des portions	Un	3	Plus copieux; Moins
22	Classes sur Si oui, dépense moyenne par personne	Un	4	de >=120 à <126; de

Chercher  
Par numéro 8. Annuler OK

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

Plan - Profils / Exemple 1 : Habitudes alimentaires

Nom du fichier : Plan125 Type de fichier : RTF

Titre :

Profils

Question profil Question Profil 8. CSP

Variables Toutes

Questionnaires Tous

Non réponses exclues

Redressement

Seuils pour les profils

Ecarts à l'indépendance 5 Minimum / écart non pondéré (Ecart à l'indépe

Khi2 local 1 Minimum / écart pondéré (Khi2 partiel)

Affichage

Décimales pour les % 0

Copier un plan Format Enregistrer et exécuter 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 sous-population 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 Philippe CIBOIS (cf. ref. Biblio en fin de documentation).

## Profil de variables

### Principe et fonction

Profil de variables

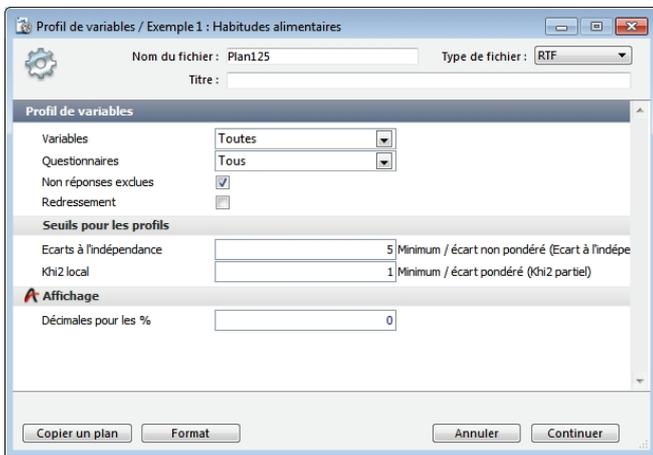
Annuler

1. Sélection de questions/variables
2. Lire un fichier de profils déjà créé (Extension : "\_DIR")

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 a priori les modalités à partir desquelles les groupes pertinents peuvent être constitués.



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. (exemple : 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 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 constituent 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.

## Précaution

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. paragraphe précédent).

## 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 Moy...	Ecart No...	Ecart So...	Mo...	Khi2 parti...	Khi2 parti...
1. Nombre d'enfants	11	7	68,6	9,8	26,8	3,1
2. Activité professionnelle	16	6	44,9	7,5	14,4	2,1
3. CSP	17	14	186,8	13,3	87,9	6,1
4. Origine géographique	16	8	75,3	9,4	32,9	4,1
5. Age	15	10	103,0	10,3	29,7	3,1
6. Produit utilisé le plus ...	24	5	30,7	6,1	17,3	3,1
7. Produit utilisé le plus ...	20	19	173,9	9,2	82,9	4,1
8. Critère de choix pour l...	17	18	161,5	9,0	59,6	3,1
9. Repas du soir habitue...	16	2	14,9	7,5	5,3	2,1
10. Repas du soir tous l...	11	5	40,7	8,1	8,6	1,1
11. Qualité principale d'u...	19	15	152,3	10,2	85,7	5,1
12. Changements perçu...	15	8	68,5	8,6	26,9	3,1
13. Grignotage entre les ...	13	15	102,1	6,8	47,8	3,1
14. Lieu de prise des Re...	11	8	68,3	8,5	16,5	2,1
15. Changement dans le ...	22	8	71,7	9,0	45,6	5,1

Lecture de la première ligne

La variable n<sub>3</sub> 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	11
Nombre d'écarts à l'indépendance supérieurs au seuil choisi de 5	7
Somme de ces écarts	68,6
Moyenne de ces écarts	9,8
Somme des Khi2 partiels supérieurs au seuil choisi de 1	26,8
Moyenne de ces écarts	3,8
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 d'un 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 pour l'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.

Variable	Ecart	Khi2	PEM Test	Khi2 loc
1. Nombre d'enfants	6	3,0	15,0	*
2. Activité professionnelle	34	15,0	41,0	***
1. Nombre d'enfants	7	3,0	19,0	**
2. Activité professionnelle	6	1,0	22,0	*
2. Activité professionnelle	10	4,0	37,0	***
2. Activité professionnelle	18	4,0	90,0	***
3. CSP	17	5,0	17,0	**
3. CSP	12	10,0	27,0	***
3. CSP	8	2,0	12,0	*
3. CSP	39	22,0	40,0	***
3. CSP	33	15,0	36,0	***
3. CSP	22	11,0	43,0	***
3. CSP	49	21,0	62,0	***
4. Origine géographique	7	2,0	17,0	**
4. Origine géographique	8	1,0	18,0	*

### 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 +

### Par variable

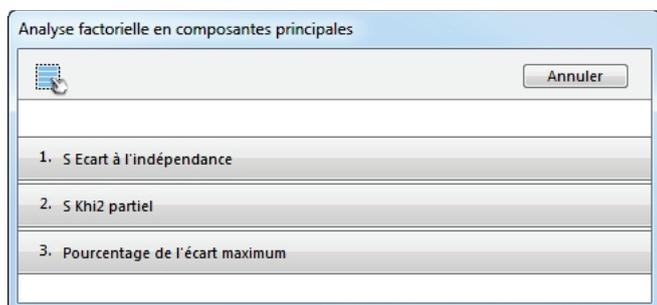
Le bouton "Par variable" permet d'afficher la liste des liens relatifs à une variable.

Variable	Ecart	Khi2	PEM Test	Khi2 local
1. Nombre d'enfants				
3 lien(s) (Test du khi2 >= 90%)				
2. Activité professionnelle	6	3,0	15,0	*
5. Age	34	15,0	41,0	***
9. Repas du soir habituel avec les...	7	3,0	19,0	**

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.



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.

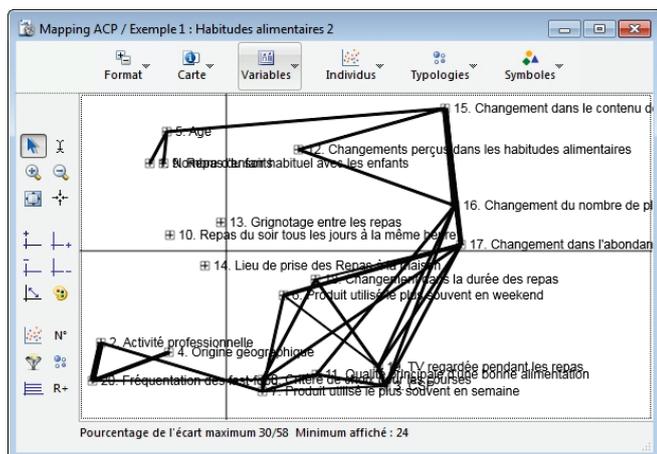
Quand les questions excentrées sont proches, elles ont des liens simultanés forts. Les liens entre questions sont marqués par des lignes dont l'épaisseur est proportionnelle à la force relative de l'attraction (1 à 5 points). Possibilité de visualiser les attractions question par question.

## Exemple de carte 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.



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

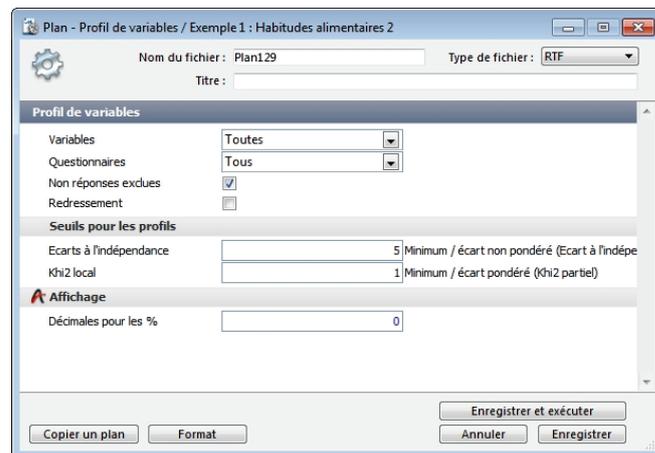
Sous-sélection des liens pour la question PCS :

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).

## Plan de variables discriminantes

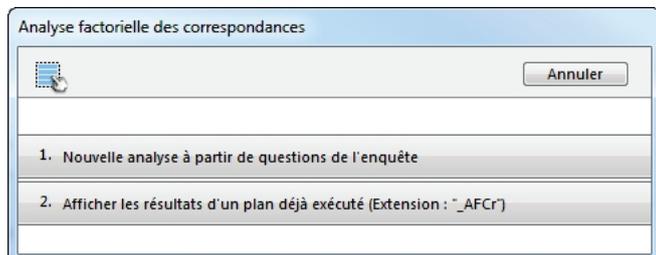


Le paramétrage est analogue à celui du plan de profils de modalités.

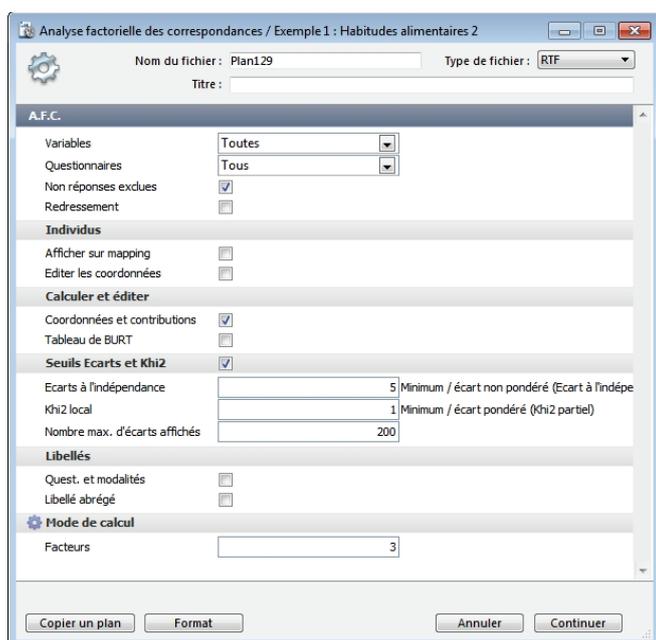
# Analyse Factorielle des Correspondances (A.F.C.)

## Principe

Modalisa permet de réaliser une AFC à partir des questions de l'enquête courante ou de relire une analyse demandée précédemment.



## Création



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 calculée 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 égal 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.

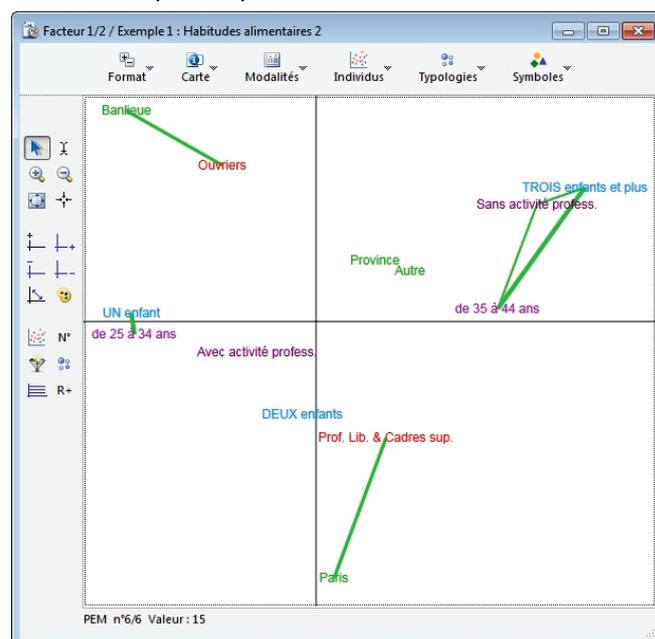
Cocher les cases des résultats souhaités : les cases "Coordonnées et Contributions" et "écarts et Khi2" sont cochées par défaut, possibilité d'y ajouter le "Tableau de BURT".

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 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.

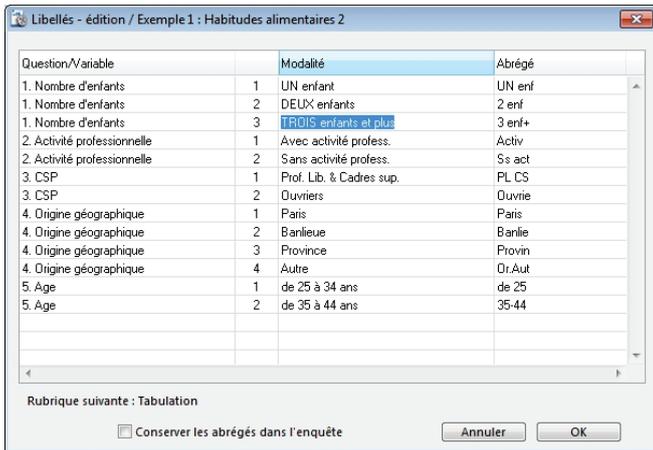
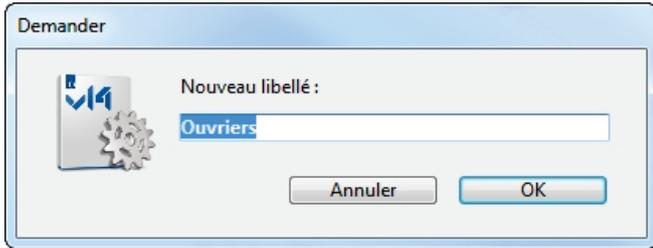
Gestion de l'affichage de la carte d'AFC (mapping) Les options du menu Format et Facteurs, ainsi que les icônes de la barre verticale située à gauche de la carte, permettent de jouer sur l'affichage de la carte.

Les intitulés de modalités 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.



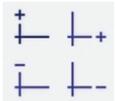
Pour changer la taille du mapping utiliser les loupes, choix Zoom + ou Zoom - du menu Format



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



permet de localiser les individus soit sous la forme de points, soit en reprenant les N° de questionnaires

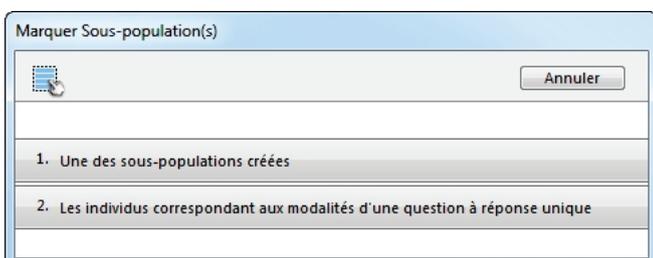


pour marquer une sous-population (Hommes/-25 ans/Paris)

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 sous-population est choisie. Les sous-populations sont gardées en mémoire.

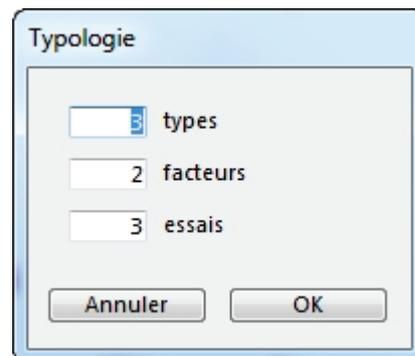
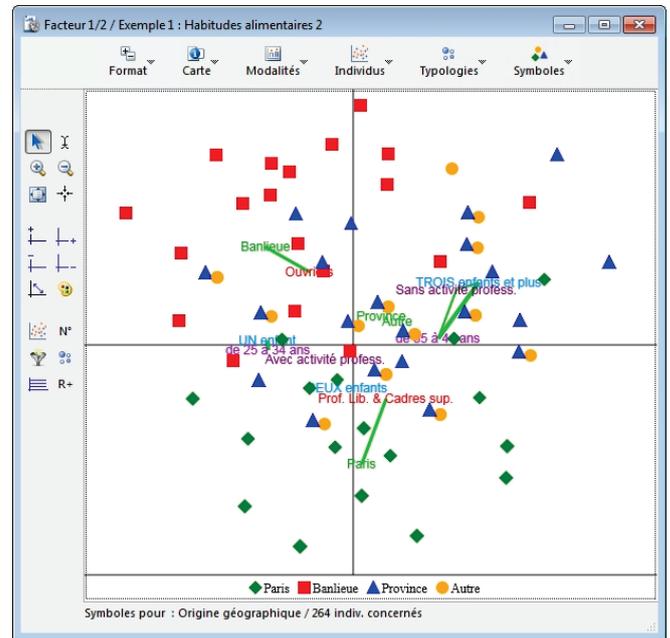
Cette fonction est surtout utilisable avec un écran couleur.



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 sous-populations ou les numéros de questionnaires. Il suffit de cliquer sur ce bouton pour réafficher les derniers types calculés.

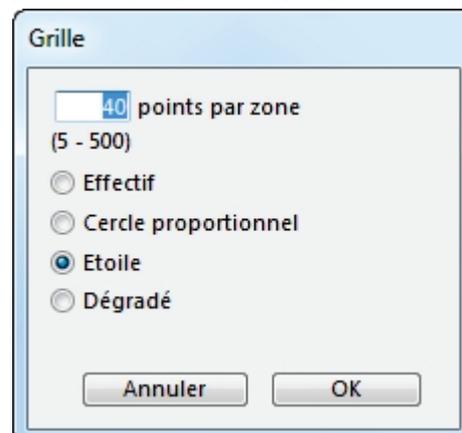


active l'écran de choix de typologie à réaliser



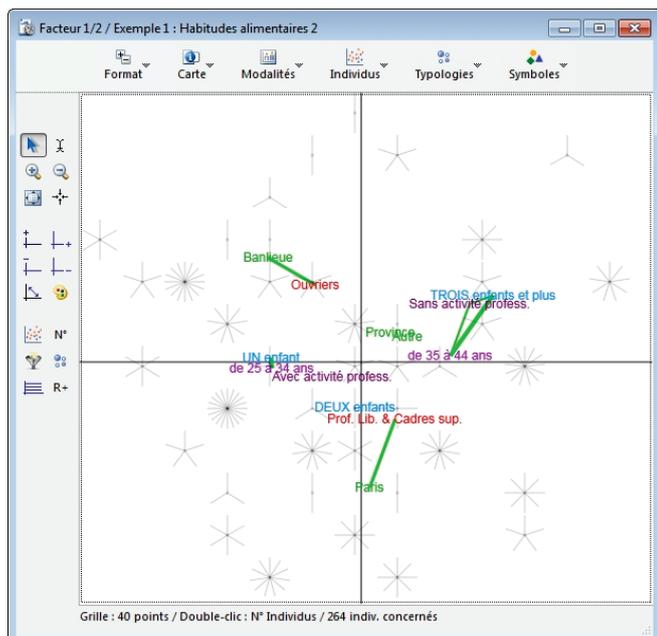
Les trois options de menu typologie à savoir Centres mobiles, Ascendante hiérarchique A-H et CM puis AH affichent l'écran ci-dessus.

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).

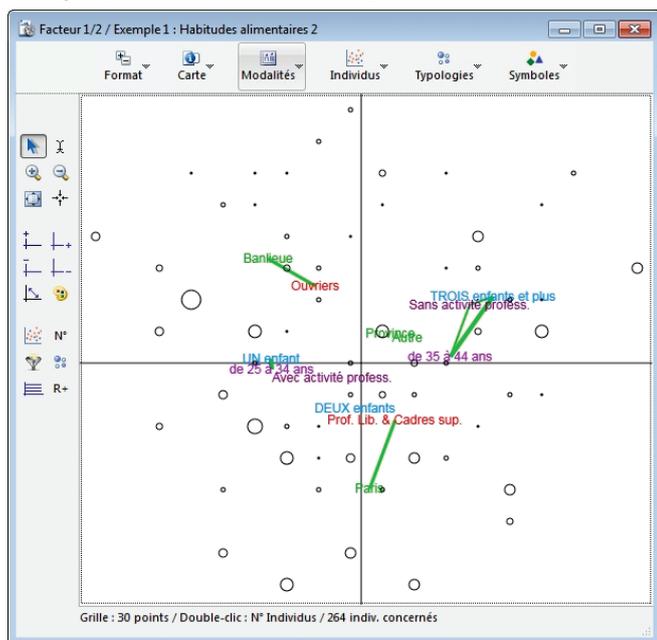


L'icône en forme de tableau quadrillé permet de créer une grille d'échelle variable permettant de repérer le nombre des

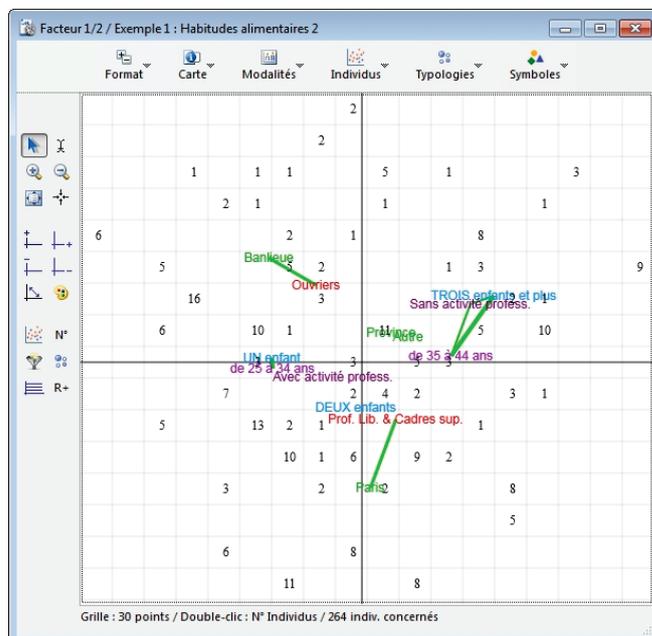
individus ainsi que la forme de leur représentation sur la carte.



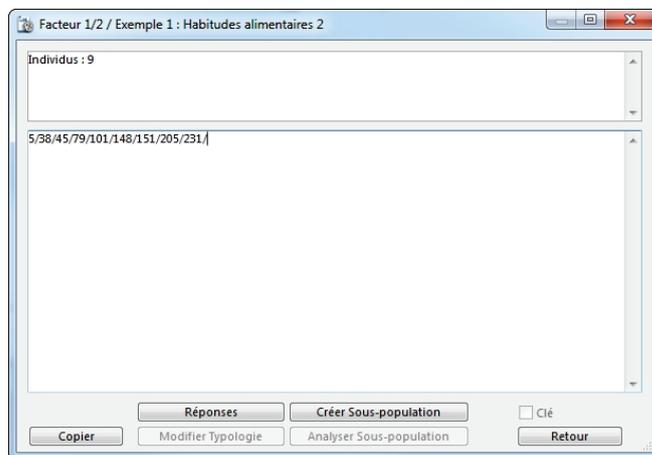
Sur la représentation en .toile le nombre de branche est égal au nombre des individus présents.



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. 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.

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".

Contributions		Absol. (1)	Relat. (1)	Absol. (2)	Relat. (2)
UN enfant	UN enf	12,49	0,36	0,03	0,00
DEUX enfants	2 enf	0,08	0,00	5,90	0,13
TROIS enfants et plus	3 enf+	19,89	0,51	7,68	0,12
Avec activité profess.	Activ	2,93	0,28	1,25	0,07
Sans activité profess.	Ss act	11,09	0,28	4,83	0,08
Prof. Lib. & Cadres s...	PL CS	3,03	0,14	12,96	0,40
Ouvriers	Ouvrie	4,08	0,14	17,47	0,40
Paris	Paris	0,09	0,00	31,40	0,58
Banlieue	Banlie	7,85	0,19	15,08	0,24
Province	Provin	1,67	0,06	2,77	0,06
Autre	Or.Aut	0,68	0,01	0,31	0,00
de 25 à 34 ans	de 25	18,04	0,66	0,13	0,00
de 35 à 44 ans	35-44	18,03	0,66	0,14	0,00

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 ».

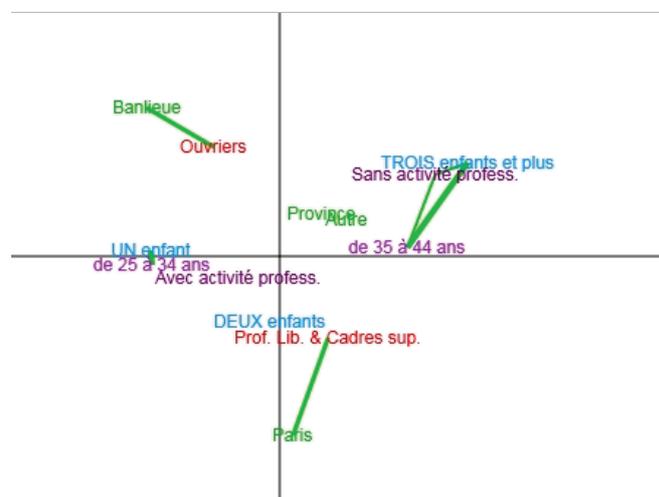
Modalisa<sup>a</sup> s'appuie sur les résultats de l'AFC existante pour positionner les modalités des nouvelles variables.

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.
- 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.

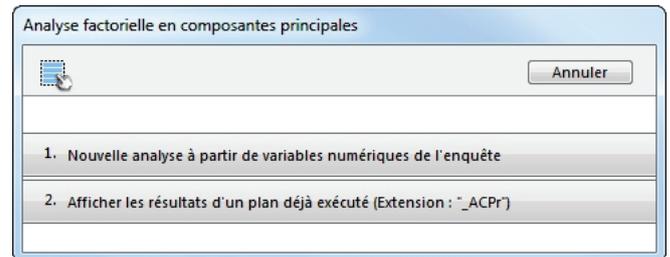


### Plan d'A.F.C.

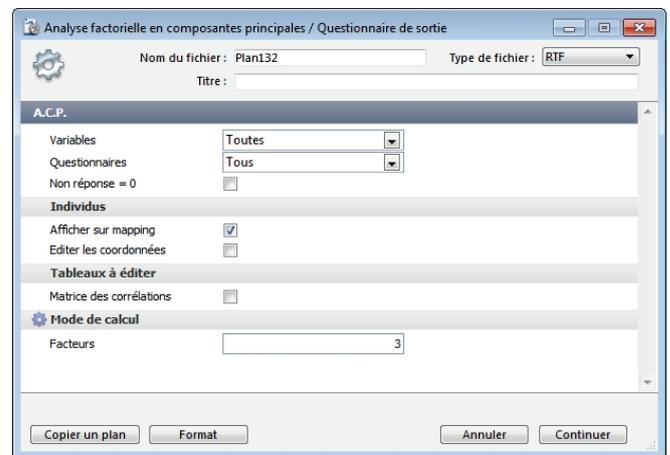
Cette option présente les mêmes choix qu'en création d'AFC par le Menu Analyse Saisir le titre du plan (27 caractères au maximum)

# Analyse en Composantes Principales (A.C.P.)

## Création



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.

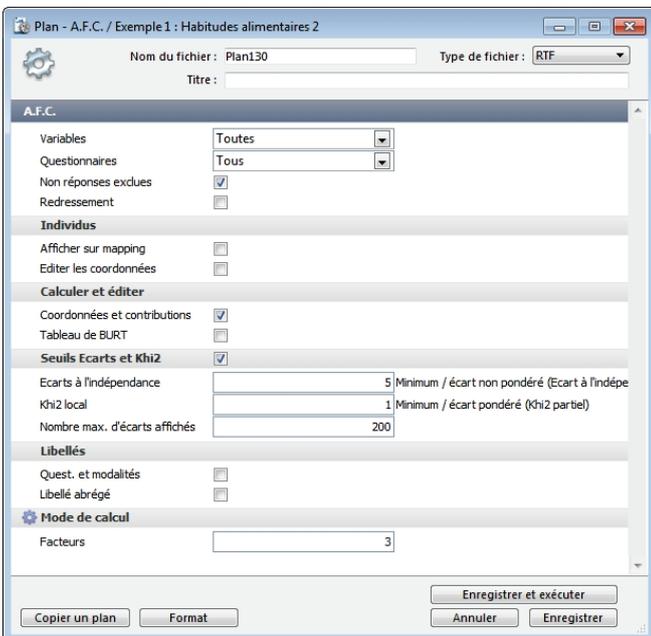


## Création d'une nouvelle ACP à partir de variables de l'enquête

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 égal à 3. Indiquer les résultats souhaités: "Matrices des corrélations" ou "Coordonnées des individus", les deux cases sont cochées par défaut.



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 Modalisa en sélectionne 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.

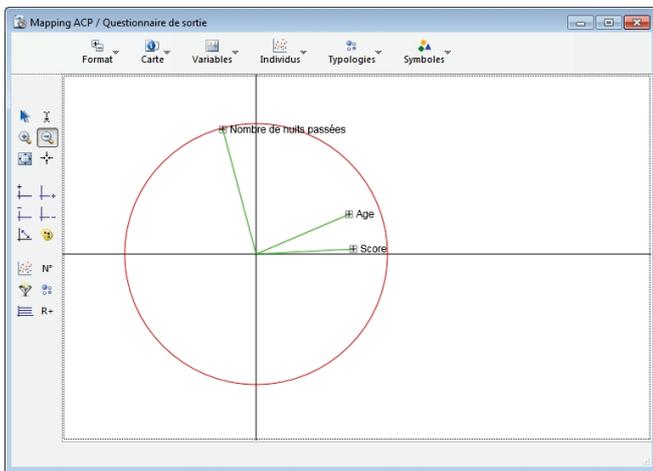
Ces choix pourront être modifiés après visualisation du mapping.

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.



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 ».

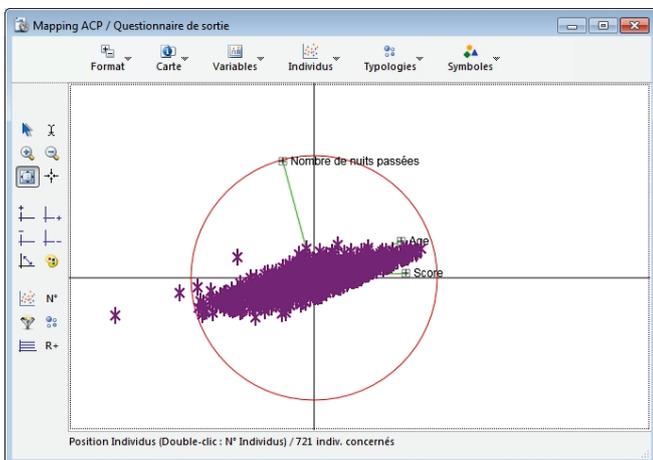
Ces choix effectués, cliquer sur Continuer. Modalisa<sup>a</sup> calcule l'ACP.

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.



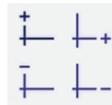
Pour changer la taille du mapping utiliser les loupes, choix Zoom + ou Zoom - du menu Format



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



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

### A partir d'une matrice de corrélation

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

### Plan d'A.C.P.

Cette option présente les mêmes choix qu'en menu Analyse Saisir le titre du plan (27 caractères au maximum)

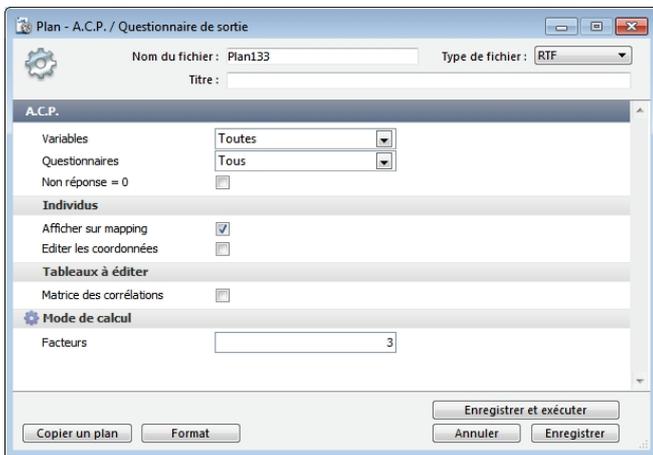
Indiquer les questions sur lesquelles sera calculée l'ACP: toutes les questions ou une sélection de questions, et création, dans ce cas, d'un groupe de questions ou choix sur liste d'un groupe déjà créé.

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.

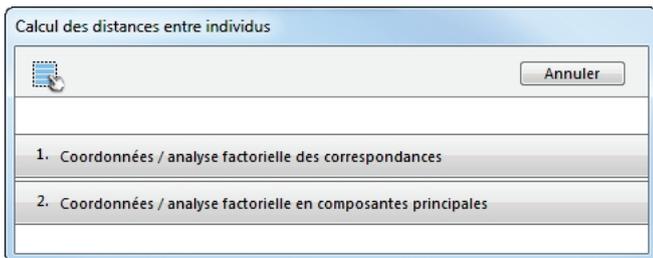


Il est possible de copier un plan déjà créé afin de n'avoir qu'à le modifier partiellement, cliquer dans ce cas sur le bouton « Copier un plan ».

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

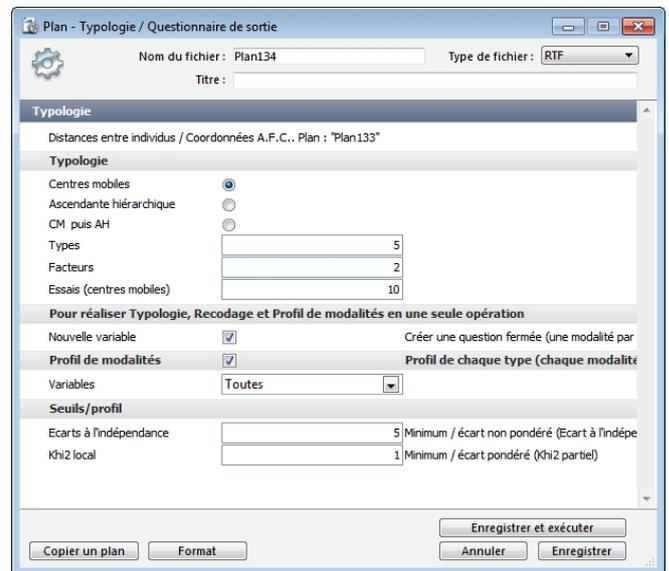
## Plan de typologies

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 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.