

## UNIWIN VERSION 9.7.0

# ANALYSE FACTORIELLE DES CORRESPONDANCES

Révision : 02/09/2023

Définition.....	1
Entrée des données .....	2
Données manquantes .....	4
Calculs.....	4
Entrée des données supplémentaires (colonnes).....	4
Entrée des données supplémentaires (lignes).....	5
Exemple : Fichier JOURNAUX .....	5
L'option Rapports .....	13
L'option Graphiques .....	14
Une rapide interprétation des résultats.....	20
Les variables internes créées par la procédure .....	20

### Définition

La méthode d'Analyse Factorielle des Correspondances Simples ou Binaires (AFC) permet de représenter graphiquement un tableau de contingence créé par le ou les croisements (tris croisés) de deux ou plusieurs variables qualitatives.

Voir la procédure 'Tri croisé' pour plus d'informations sur la façon de construire ce tableau.

La méthode vise à rassembler sur un ou plusieurs graphiques (plan factoriel) la plus grande partie possible de l'information contenue dans le tableau en s'attachant non pas aux valeurs absolues mais aux correspondances entre les caractéristiques, c'est-à-dire aux valeurs relatives.

La possibilité d'analyse de points lignes ou colonnes supplémentaires est offerte.

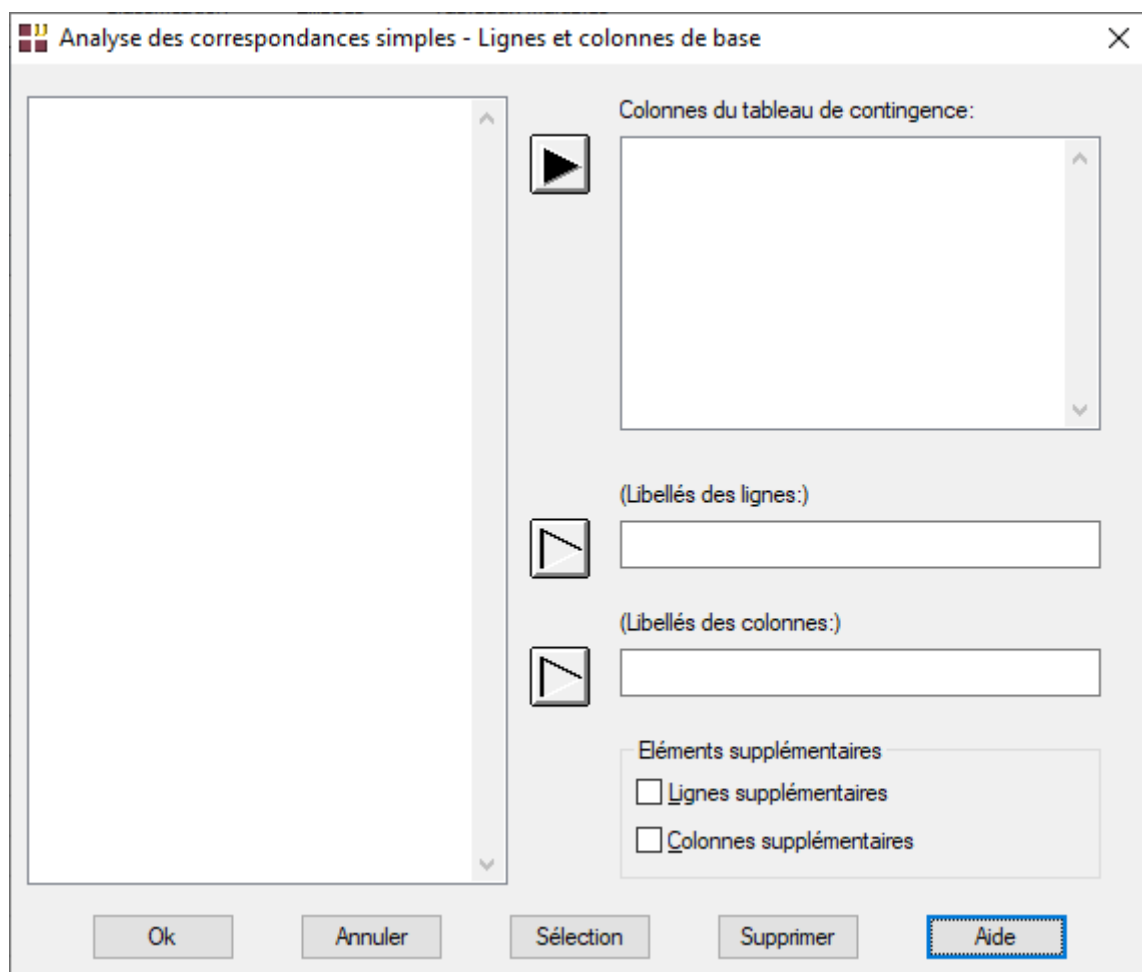
Après affichage du tableau et de l'histogramme des inerties, vous pouvez choisir le nombre d'axes factoriels à extraire.

Un rapport général de synthèse est proposé ainsi que l'ensemble des graphiques des plans factoriels lignes, colonnes, simultanés, barycentriques et de ceux relatifs aux lignes et colonnes supplémentaires.

De nombreux outils d'aide à l'interprétation sont fournis, notamment contributions et cosinus carrés.

## Entrée des données

Cliquons sur l'icône AFC dans le ruban Décrire. La boîte de dialogue montrée ci-dessous s'affiche :



Cette boîte de dialogue permet de choisir les variables de base qui définissent les lignes et colonnes actives du tableau de contingence.

- Lignes et colonnes de base

Les variables sélectionnées définissent les colonnes de base (actives) du tableau de contingence. Les éléments de ces variables constituent les lignes de base (actives).

Vous pouvez définir une variable contenant les libellés des lignes de base (actives) et une variable contenant les libellés des colonnes de base (actives).

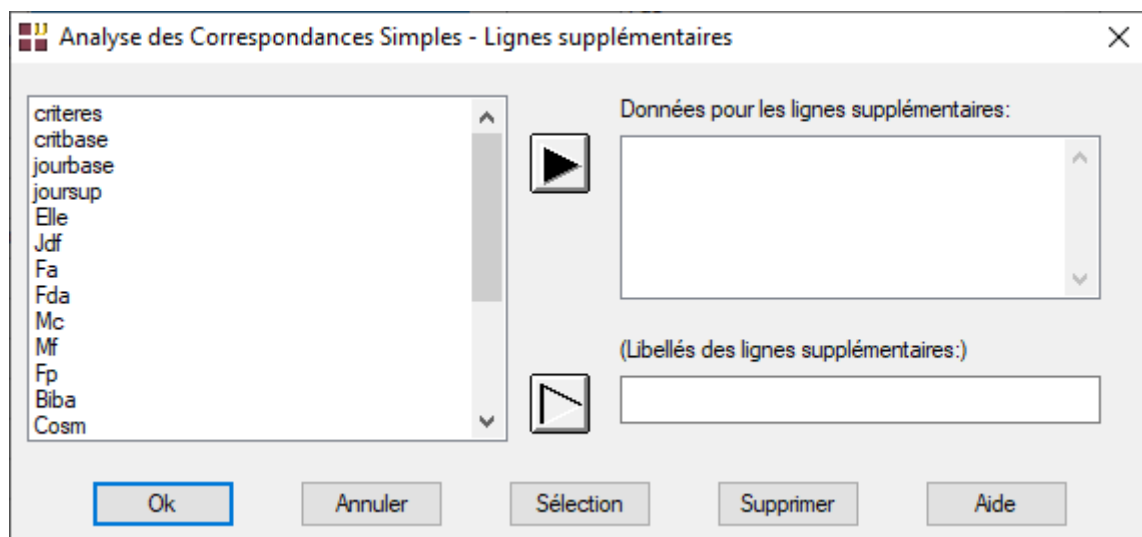
- Lignes et colonnes supplémentaires

Deux cases à cocher permettent d'indiquer si l'on désire analyser des lignes supplémentaires (passives) et/ou des colonnes supplémentaires (passives).

Si ces deux cases sont cochées, deux boîtes de dialogue complémentaires s'affichent. Elles sont montrées ci-dessous (avec le fichier JOURNAUX pour exemple).

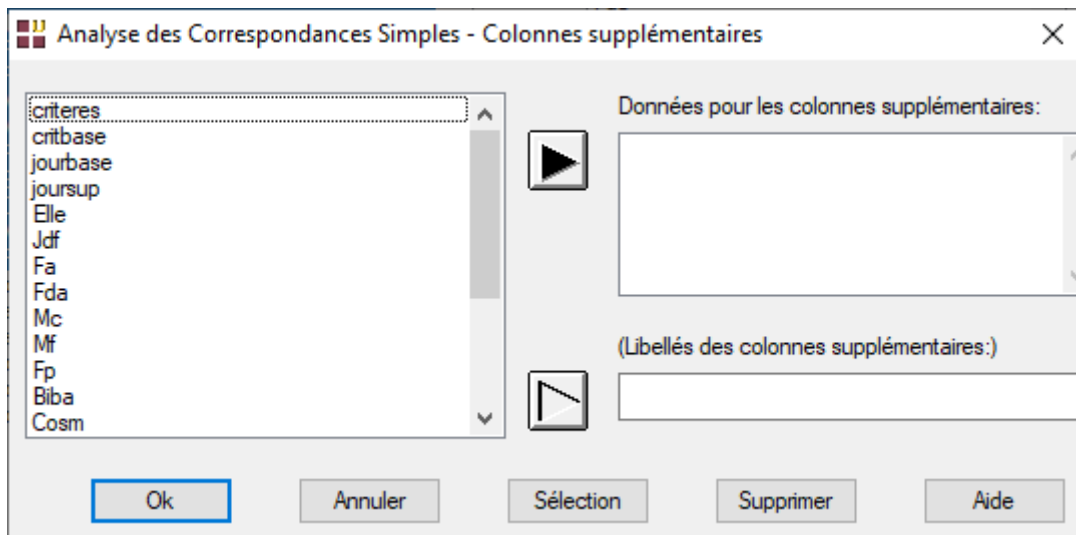
- Lignes supplémentaires

Chaque variable sélectionnée définit les données sur la variable de base associée pour les lignes supplémentaires (passives) du tableau de contingence. Une variable contenant les données des lignes supplémentaires (passives) a donc autant d'éléments qu'il y a de lignes supplémentaires. Il y a autant de variables définissant les données des lignes supplémentaires (passives) qu'il y a de variables de base. Vous pouvez définir une variable contenant les libellés des lignes supplémentaires.



- Colonnes supplémentaires

Chaque variable sélectionnée définit une colonne supplémentaire (passive) du tableau de contingence. Une colonne supplémentaire (passive) a donc autant d'éléments qu'il y a de lignes de base (actives). Vous pouvez définir une variable contenant les libellés des colonnes supplémentaires.



Dans toutes ces boîtes de dialogue, les zones de libellés sont optionnelles : UNIWIN génère automatiquement des libellés en cas de besoin.

### Données manquantes

Dans cette procédure les données manquantes ne sont pas autorisées, ni pour les lignes et colonnes de base, ni pour les lignes et colonnes supplémentaires.

### Calculs

UNIWIN diagonalise la matrice la matrice des covariances entre les colonnes. Il faut donc, de préférence, organiser vos variables de sorte qu'il y ait plus de lignes que de colonnes.

### Entrée des données supplémentaires (colonnes)

L'entrée des données pour les colonnes supplémentaires peut se faire de deux façons différentes :

- Les données supplémentaires sont dans d'autres variables du même fichier de données. Dans la boîte de dialogue Colonnes supplémentaires, il faut sélectionner ces variables.
- Les données supplémentaires sont dans un deuxième fichier. Il faut alors avant de débiter l'AFC, ouvrir les deux fichiers simultanément (le fichier des données de base et le fichier des données supplémentaires). Dans la boîte de dialogue Colonnes supplémentaires, il faut sélectionner les variables définissant les données supplémentaires dans ce deuxième fichier.

## Entrée des données supplémentaires (lignes)

L'entrée des données pour les lignes supplémentaires peut se faire de trois façons différentes :

- Les données supplémentaires sont dans d'autres variables du même fichier de données. Dans la boîte de dialogue Lignes supplémentaires, il faut sélectionner les variables définissant les données supplémentaires en respectant l'ordre défini lors de la sélection des données de base : la première variable sélectionnée définit les données supplémentaires pour la première variable de base, la deuxième variable sélectionnée définit les données supplémentaires pour la deuxième variable de base...
- Les données supplémentaires sont dans les mêmes variables que les variables contenant les données de base. Il faut alors utiliser le bouton Sélection dans la boîte de dialogue. Dans la boîte de dialogue Lignes et colonnes de base, sélectionner les variables puis cliquer sur le bouton Sélection et définir la sélection logique désirée. Dans la boîte de dialogue Lignes supplémentaires, sélectionner ces mêmes variables dans le même ordre puis cliquer sur le bouton Sélection et définir la sélection logique.
- Les données supplémentaires sont dans un deuxième fichier. Il faut alors avant de débiter l'AFC, ouvrir les deux fichiers simultanément (le fichier des données de base et le fichier des données supplémentaires). Dans la boîte de dialogue Lignes supplémentaires, il faut sélectionner les variables définissant les données supplémentaires dans ce deuxième fichier en respectant l'ordre défini lors de la sélection des variables de base : la première variable sélectionnée dans le deuxième fichier définit les données supplémentaires pour la première variable sélectionnée dans le premier fichier, la deuxième variable sélectionnée dans le deuxième fichier définit les données supplémentaires pour la deuxième variable sélectionnée dans le premier fichier...

### Exemple : Fichier JOURNAUX

Nous utiliserons le fichier JOURNAUX pour illustrer cette procédure.

Ce fichier contient les résultats, sous la forme d'un tableau de contingence (éclaté en variables colonnes), d'une enquête, comportant 18 questions (Oui, Non) définissant les lignes, effectuée auprès de lectrices de 17 magazines, définissant les colonnes.

Ainsi chaque nombre pour une ligne  $i$  et une colonne  $j$  du tableau de contingence, est le comptage du nombre de lectrices ayant données la réponse Oui à la question  $i$  et lisant le journal  $j$ .

Les colonnes définissent les 17 magazines (11 colonnes de base, 6 colonnes supplémentaires).

Les magazines de base (actifs) sont :

<i>Magazine</i>	<i>Nom de la variable</i>	<i>Libellé court</i>
ELLE	Elle	ELLE
JOURS DE FRANCE	Jdf	JDF
FEMME ACTUELLE	Fa	FA
FEMME D'AUJOURD'HUI	Fda	FDA
MARIE CLAIRE	Mc	MC
MARIE FRANCE	Mf	MF
FEMME PRATIQUE	Fp	FP
BIBA	Biba	BIBA
COSMOPOLITAN	Cosm	COSM
MODES ET TRAVAUX	Met	MET
PRIMA	Prim	PRIM

Les libellés courts sont contenus dans la variable jourbase.

Les magazines supplémentaires (passifs) sont :

<i>Magazine</i>	<i>Nom de la variable</i>	<i>Libellé court</i>
PARENTS	pare	PARE
ENFANTS	enfa	ENFA
L'EXPRESS	express	EXPR
LE POINT	point	POIN
LE NOUVEL OBSERVATEUR	obse	OBSE
PARIS MATCH	pm	PM

Les libellés courts sont contenus dans la variable joursup.

Pour chaque magazine, les lectrices ont répondu aux questions suivantes par Oui ou par Non :

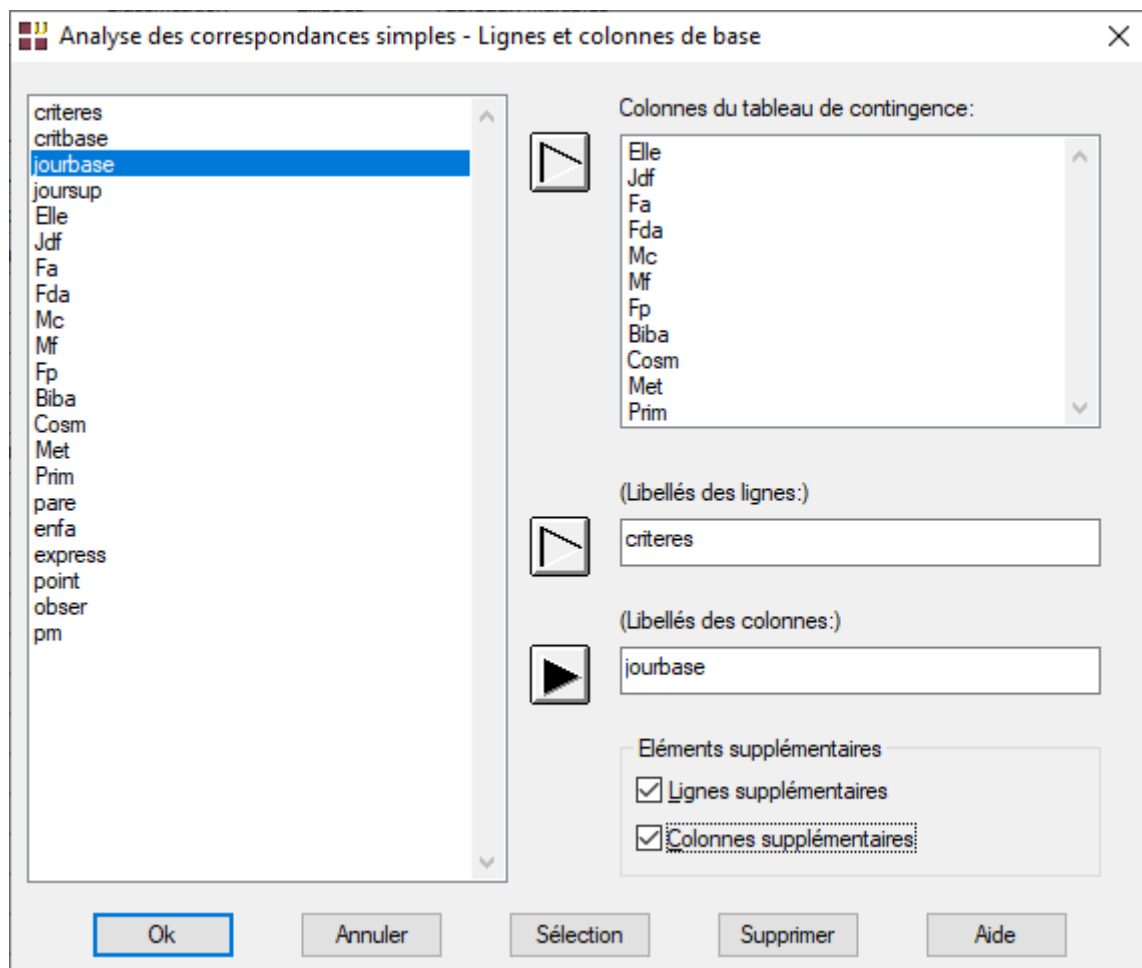
<i>Question</i>	<i>Libellé court</i>	<i>Base</i>
Avez-vous un réfrigérateur ?	Refri	Oui
Avez-vous une chaîne Hifi ?	Hifi	Oui
Avez-vous une machine à laver ?	Laver	Oui
Avez-vous une caméra ?	Camera	Oui
Avez-vous une voiture neuve ?	Voiture	Oui
Avez-vous un magnétoscope ?	Magneto	Oui
Avez-vous une résidence secondaire ?	Maison2	Non
Possédez-vous votre habitation principale ?	Maison1	Non
Avez-vous un compte bancaire ?	Banque	Non
Avez-vous une carte de crédit ?	Carte	Non
Avez-vous un appareil photo ?	Photo	Oui
Vos revenus sont-ils entre 120 et 240KF ?	Revenu1	Non
Vos revenus sont ils supérieurs à 240KF ?	Revenu2	Non

Avez-vous un chien ?	Chien	Non
Niveau de vie (plus de 5 biens)	Vie	Non
Poste d'encadrement ?	Cadre	Non
Enfants de moins de 15 ans ?	Enfant	Non
Avez-vous un téléviseur ?	Tv	Oui

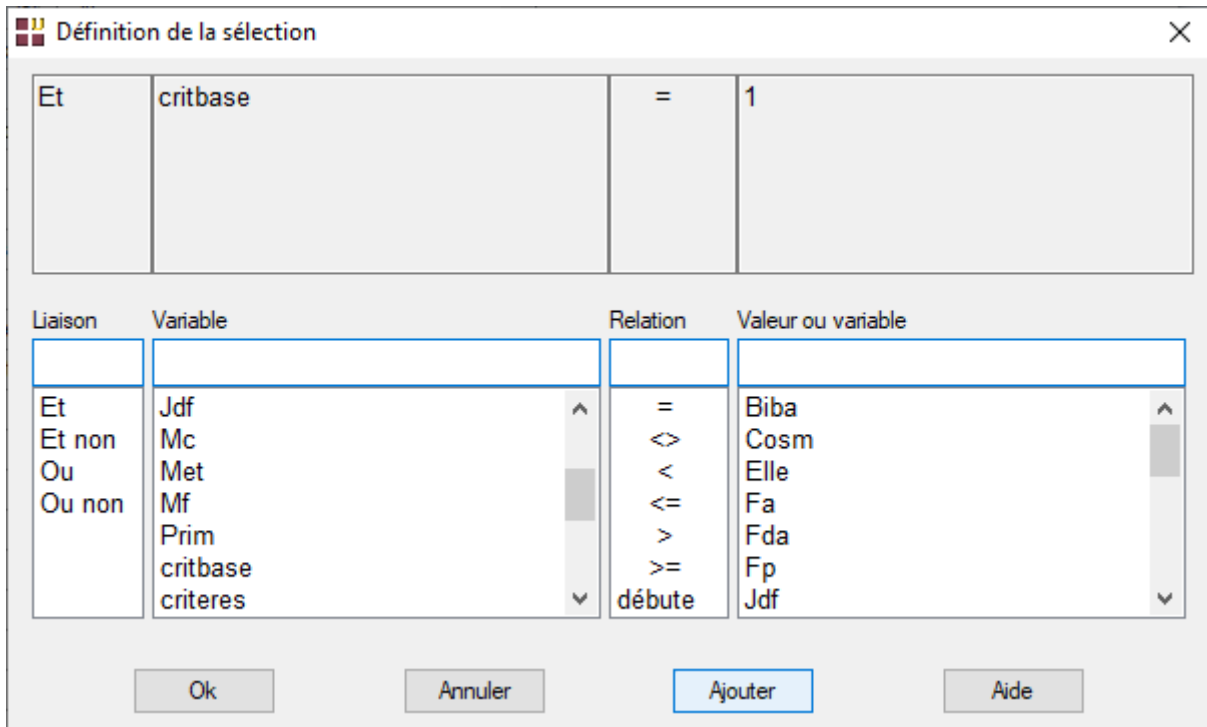
Les libellés courts des questions sont contenus dans la variable *criteres*. La variable *critbase* (codée 0 ou 1) indique si la question appartient aux questions de base (1) ou aux questions supplémentaires (0).

Cliquons sur l'icône AFC dans le ruban Décrire.

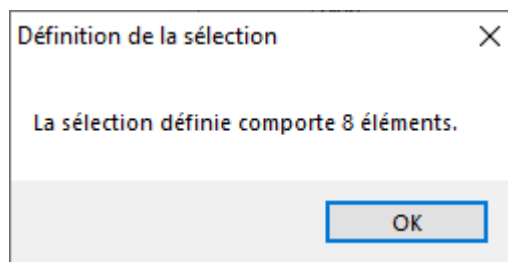
Nous choisissons les variables *Elle* à *Prim* comme variables de base, *jourbase* et *criteres* comme variables contenant les libellés respectifs des colonnes (journaux) et des lignes (questions).



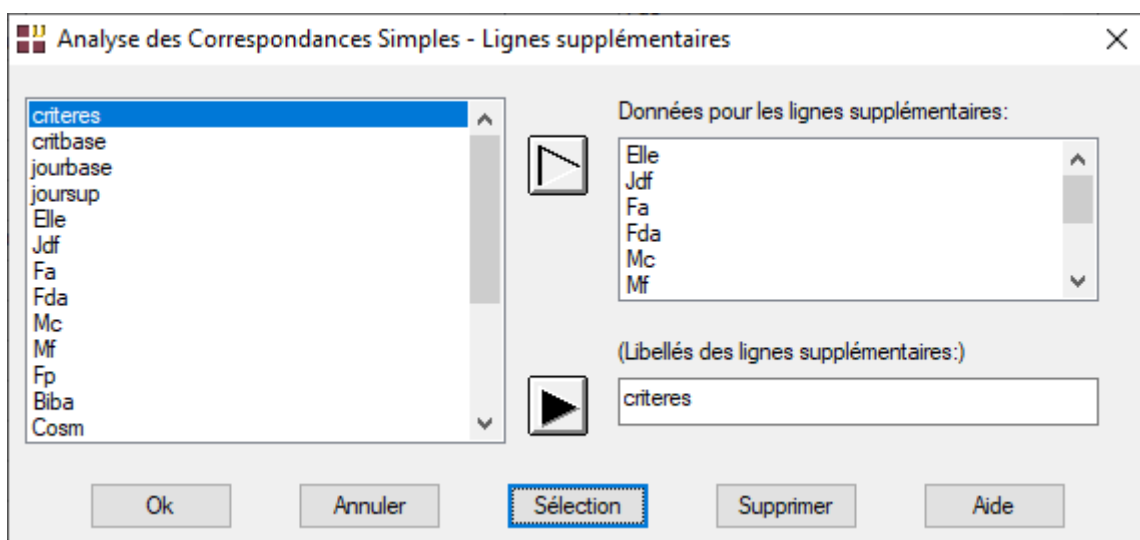
Cochons les deux cases Lignes supplémentaires et Colonnes supplémentaires, puis cliquons sur le bouton Sélection et sélectionnons les lignes qui correspondent aux critères de base.



Cliquons sur Ok. UNIWIN nous informe que 8 questions (lignes) sont sélectionnées puis affiche la boîte de dialogue Lignes supplémentaires.



Sélectionnons les variables *Elle* à *Prim* comme pour les Lignes supplémentaires, la variable *criteres* pour les libellés et cliquons sur le bouton Sélection.





Définissons la sélection montrée ci-dessous, qui sélectionne les questions (lignes) supplémentaires. UNIWIN nous informe qu'il y a 10 individus (questions) supplémentaires. Cliquons sur le bouton Ok.

Liaison	Variable	Relation	Valeur ou variable
Et	Jdf	=	Biba
Et non	Mc	<>	Cosm
Ou	Met	<	Elle
Ou non	Mf	<=	Fa
	Prim	>	Fda
	critbase	>=	Fp
	criteres	début	Jdf

Définition de la sélection

La sélection définie comporte 10 éléments.

OK

La boîte de dialogue Colonnes supplémentaires, montrée ci-dessous, s'affiche alors.

Analyse des Correspondances Simples - Colonnes supplémentaires

criteres  
critbase  
jourbase  
joursup  
Elle  
Jdf  
Fa  
Fda  
Mc  
Mf  
Fp  
Biba  
Cosm

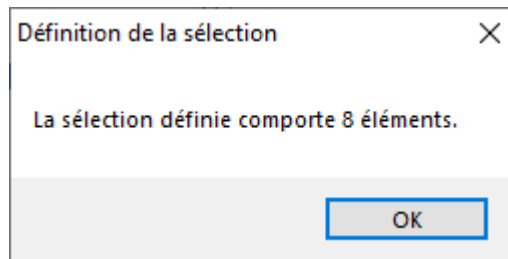
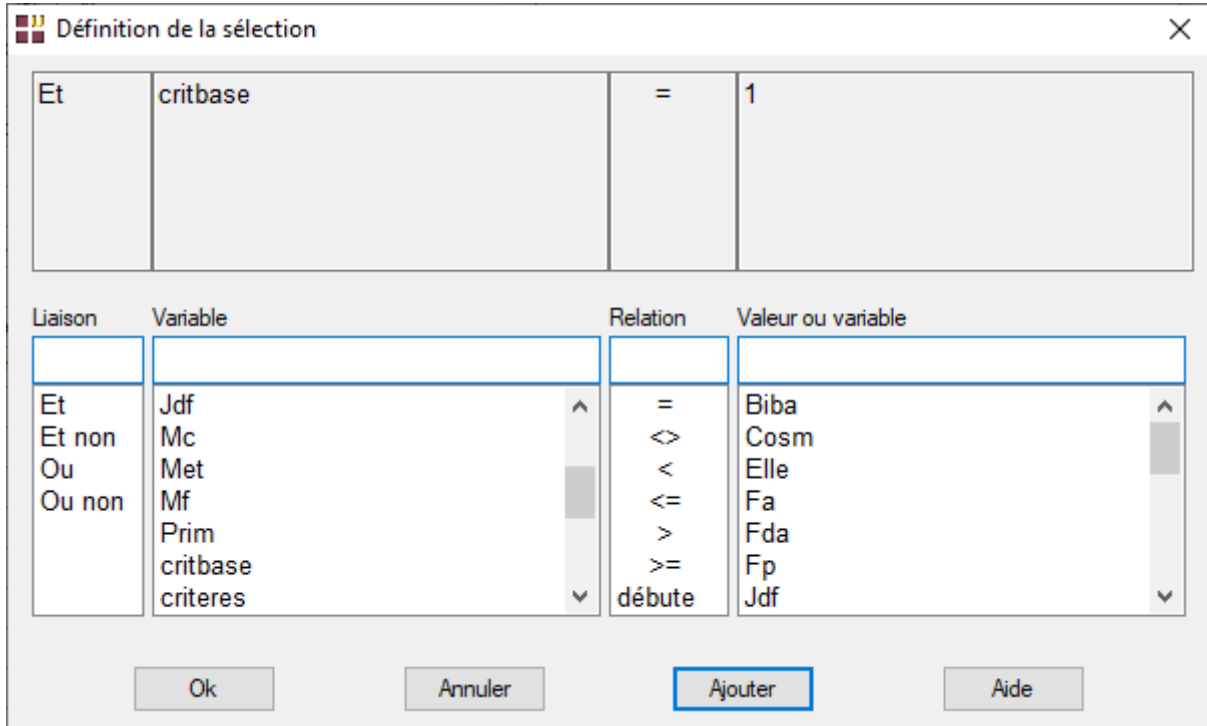
Données pour les colonnes supplémentaires:  
pare  
enfa  
express  
point  
obser  
pm

(Libellés des colonnes supplémentaires:)  
joursup

Ok Annuler Sélection Supprimer

Sélectionnons les variables *pare* à *pm* comme colonnes (journaux) supplémentaires et la variable *joursup* comme variable contenant les libellés de ces journaux supplémentaires.

Cliquons enfin sur le bouton Sélection et définissons la sélection des lignes (questions) comme précédemment pour les données de base.



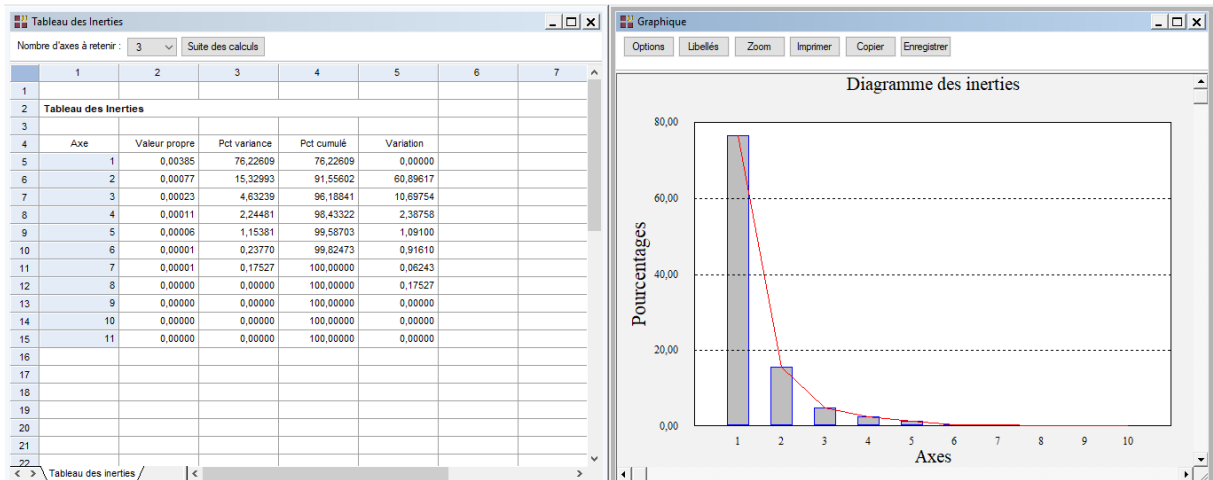
Le filtre défini sélectionne à nouveau nos 8 individus (questions de base).

Cliquons enfin sur le bouton Ok.

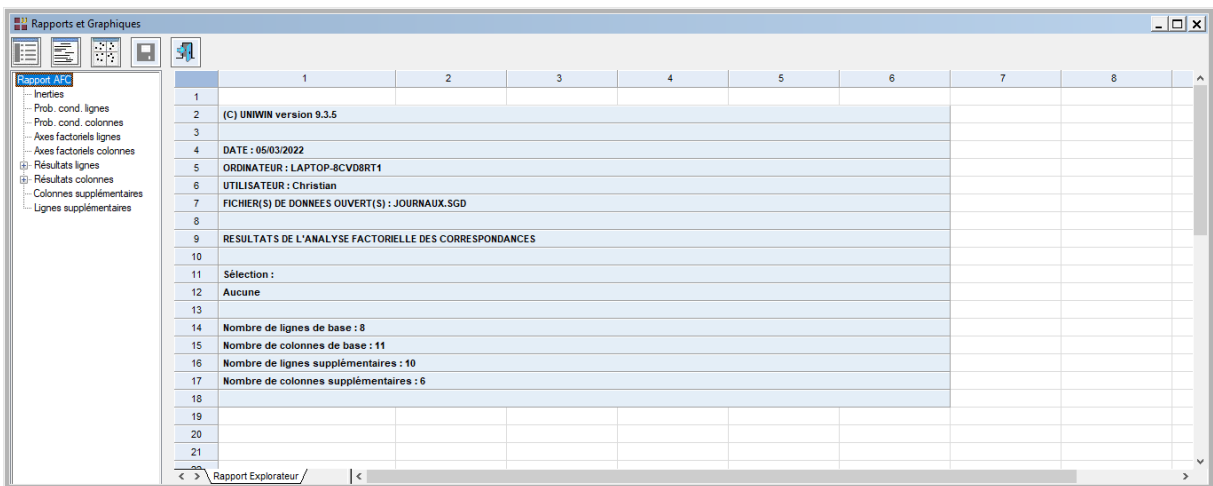
Après avoir renseigné cette boîte de dialogue, UNIWIN débute les calculs de l'Analyse des Correspondances Simples.


Après quelques instants, un tableau précisant l'inertie expliquée par les différents vecteurs propres issus de l'analyse apparaît ainsi qu'un diagramme des pourcentages d'inertie expliquée par chacun des axes.


L'option 'Nombre d'axes à retenir' permet de préciser le nombre d'axes à retenir pour l'analyse.

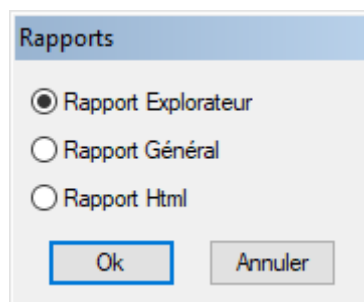



Cliquons sur le bouton 'Suite des calculs'. Après quelques instants, l'écran suivant s'affiche :

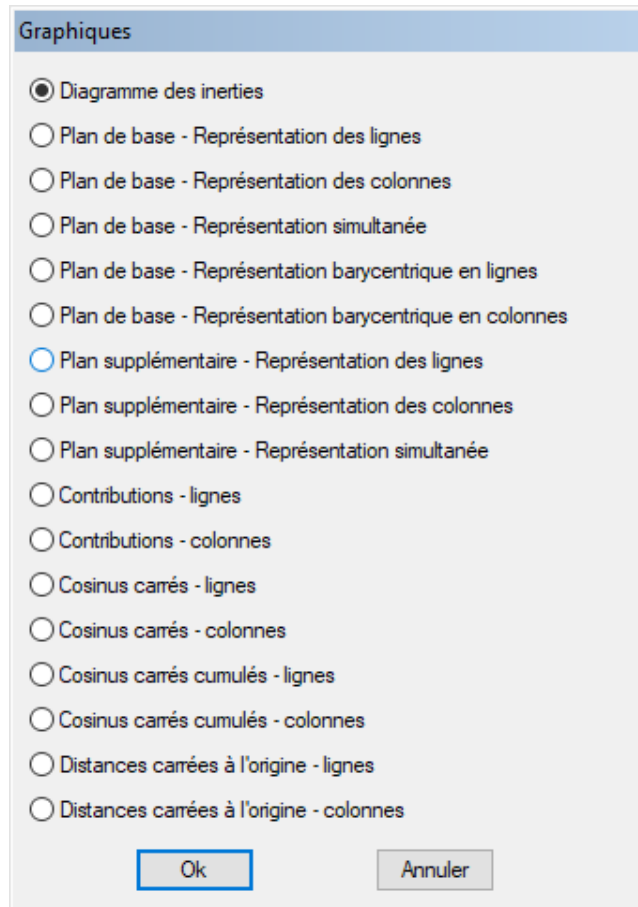



La barre d'outils 'Rapports et Graphiques' permet par l'icône 'Données'  de rappeler la boîte de dialogue d'entrée des données.

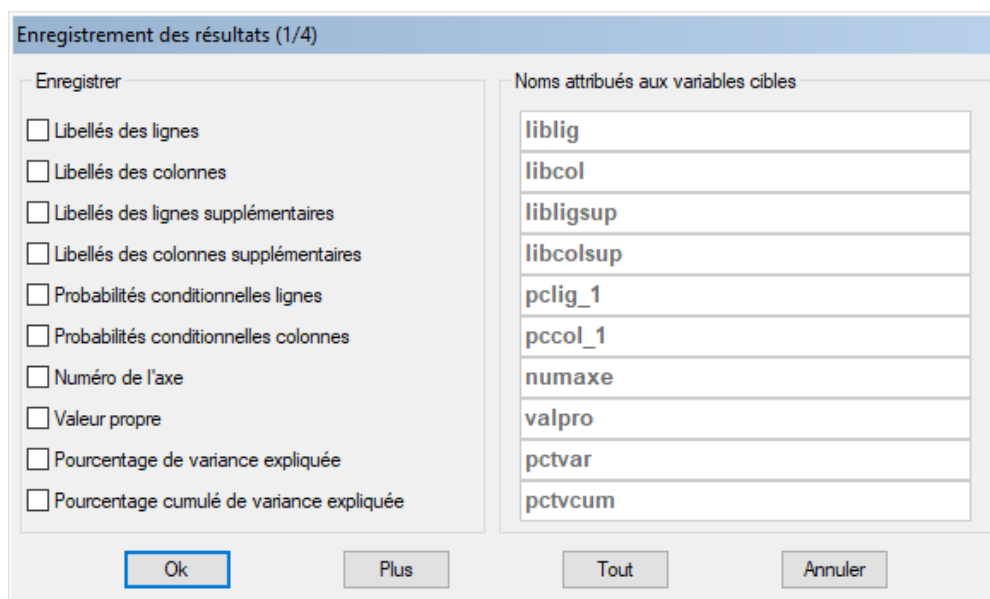
L'icône 'Rapports'  affiche la boîte de dialogue des options pour les rapports :



et l'icône 'Graphiques'  affiche la boîte de dialogue des options pour les graphiques :



L'icône 'Enregistrer'  permet de sélectionner les résultats de l'analyse à enregistrer dans un fichier.



Note : le bouton 'Plus' permet d'afficher la suite de la liste des variables.

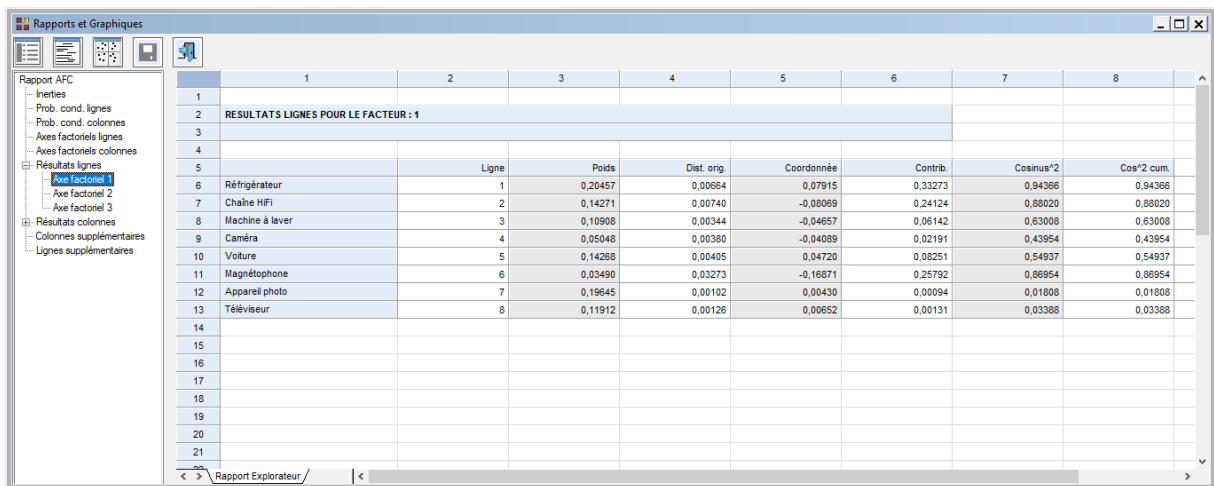
L'icône 'Quitter'  permet de quitter l'analyse.

## L'option Rapports

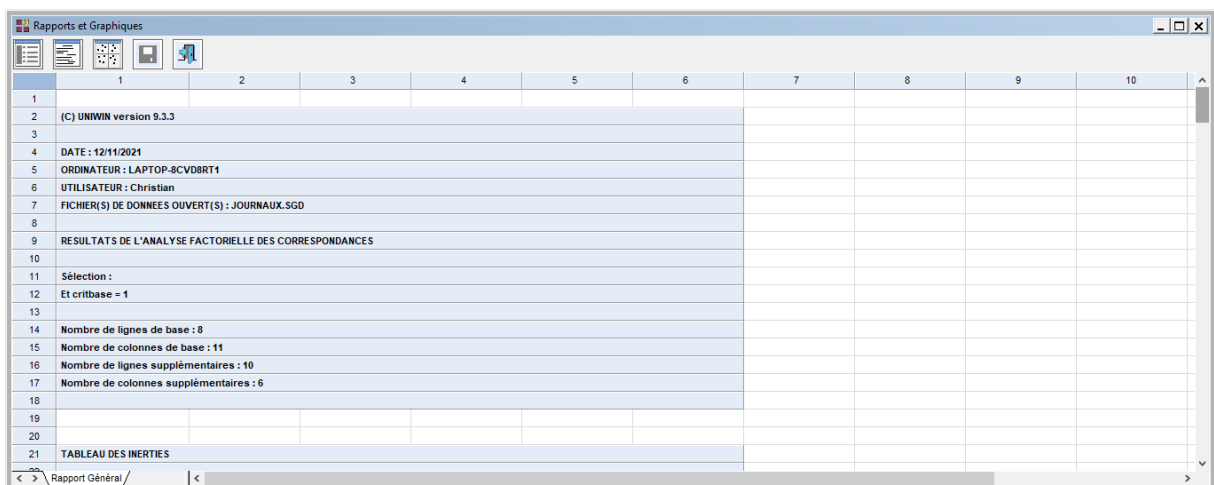
Cette option permet d'obtenir le rapport à l'écran sous la forme d'un explorateur, d'un tableur ou au format HTML.

L'impression des rapports fait appel à la procédure 'Aperçu avant impression'. Pour des informations sur cette procédure, voir le 'Manuel de l'Utilisateur'.

Voici trois exemples du rapport pour notre AFC : Explorateur, Général, HTML.



	1	2	3	4	5	6	7	8
1								
2	RESULTATS LIGNES POUR LE FACTEUR : 1							
3								
4								
5								
		Ligne	Poids	Dist. orig.	Coordonnée	Contrib.	Cosinus*2	Cos*2 cum.
6		Réfrigérateur	1	0,20457	0,00664	0,07915	0,33273	0,94366
7		Chaîne HIFI	2	0,14271	0,00740	-0,08069	0,24124	0,88020
8		Machine à laver	3	0,10908	0,00344	-0,04657	0,06142	0,63008
9		Caméra	4	0,05048	0,00380	-0,04089	0,02191	0,43954
10		Voiture	5	0,14268	0,00405	0,04720	0,08251	0,54937
11		Magnétophone	6	0,03490	0,03273	-0,16871	0,25792	0,86954
12		Appareil photo	7	0,19645	0,00102	0,00430	0,00094	0,01808
13		Téléviseur	8	0,11912	0,00126	0,00652	0,00131	0,03388
14								
15								
16								
17								
18								
19								
20								
21								



	1	2	3	4	5	6	7	8	9	10
1										
2	(C) UNIWIN version 9.3.3									
3										
4	DATE : 12/11/2021									
5	ORDINATEUR : LAPTOP-8CVDBRT1									
6	UTILISATEUR : Christian									
7	FICHIER(S) DE DONNEES OUVERT(S) : JOURNAUX.SGD									
8										
9	RESULTATS DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES									
10										
11	Sélection :									
12	Et critbase = 1									
13										
14	Nombre de lignes de base : 8									
15	Nombre de colonnes de base : 11									
16	Nombre de lignes supplémentaires : 10									
17	Nombre de colonnes supplémentaires : 6									
18										
19										
20										
21	TABLEAU DES INERTIES									

Rapports et Graphiques

(C) UNIWIN version 9.3.3

DATE : 12/11/2021  
 ORDINATEUR : LAPTOP-8CVD8RT1  
 UTILISATEUR : Christian  
 FICHER(S) DE DONNEES OUVERT(S) : JOURNAUX.SGD

RESULTATS DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES

Sélection :  
 Et critbase = 1

Nombre de lignes de base : 8  
 Nombre de colonnes de base : 11  
 Nombre de lignes supplémentaires : 10  
 Nombre de colonnes supplémentaires : 6

TABLEAU DES INERTIES

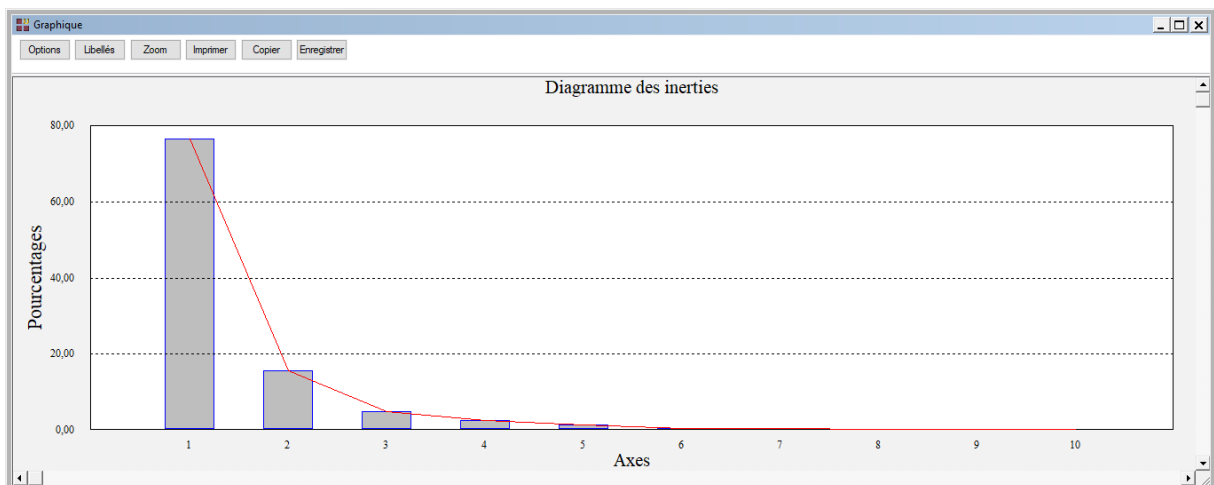
	Valeur propre	Pct variance	Pct cumulé	Variation
1	0,00385	76,22609	76,22609	0,00000
2	0,00077	15,32993	91,55602	60,89617
3	0,00023	4,63239	96,18841	10,69754
4	0,00011	2,24481	98,43322	2,38758

## L'option Graphiques

Cette option permet d'obtenir divers graphiques pour notre AFC.

- Diagramme des inerties

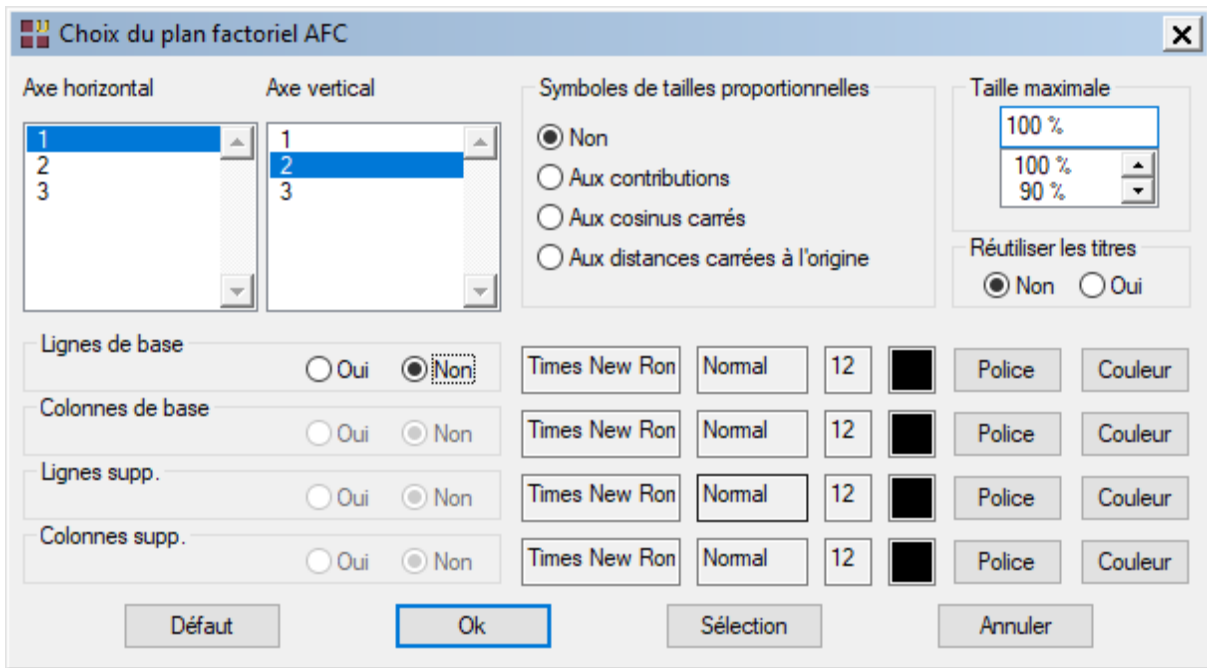
Ce graphique affiche les pourcentages d'inertie pour chacun des axes factoriels.



- Les options Plan de base

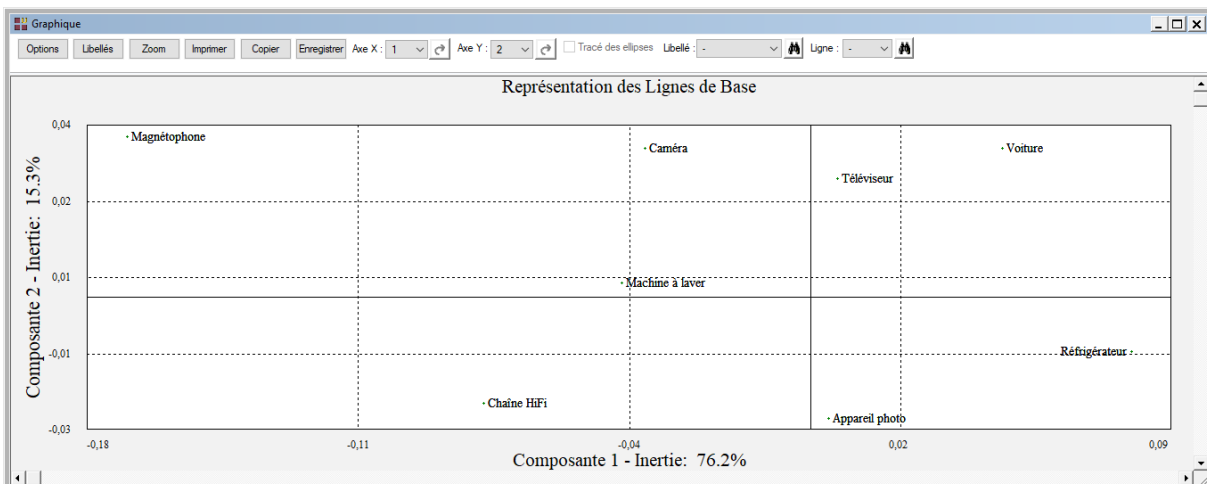
Ces options permettent de représenter les plans factoriels de base : lignes, colonnes, simultané (lignes et colonnes), barycentrique en lignes ou barycentrique en colonnes.

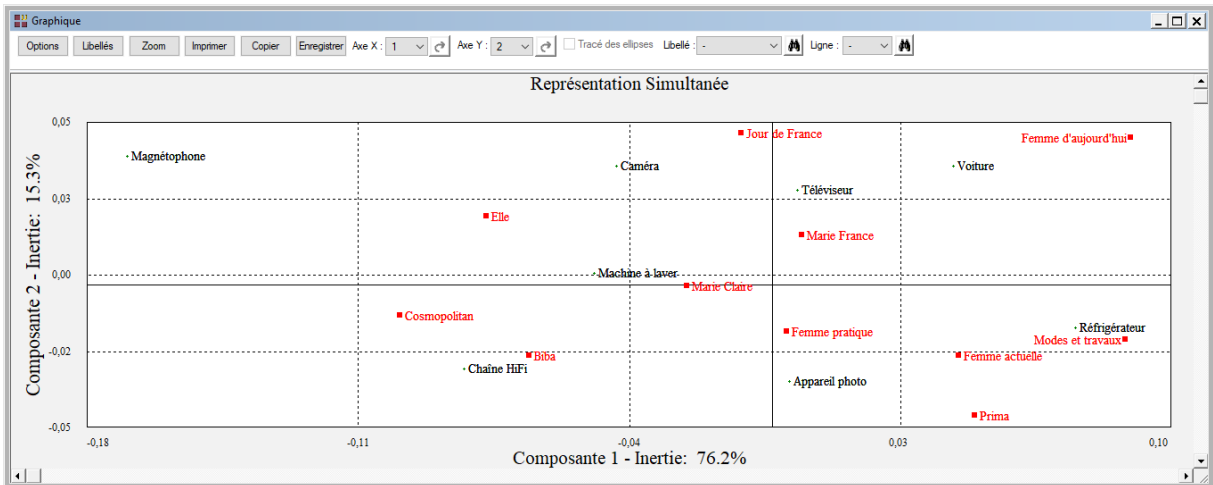
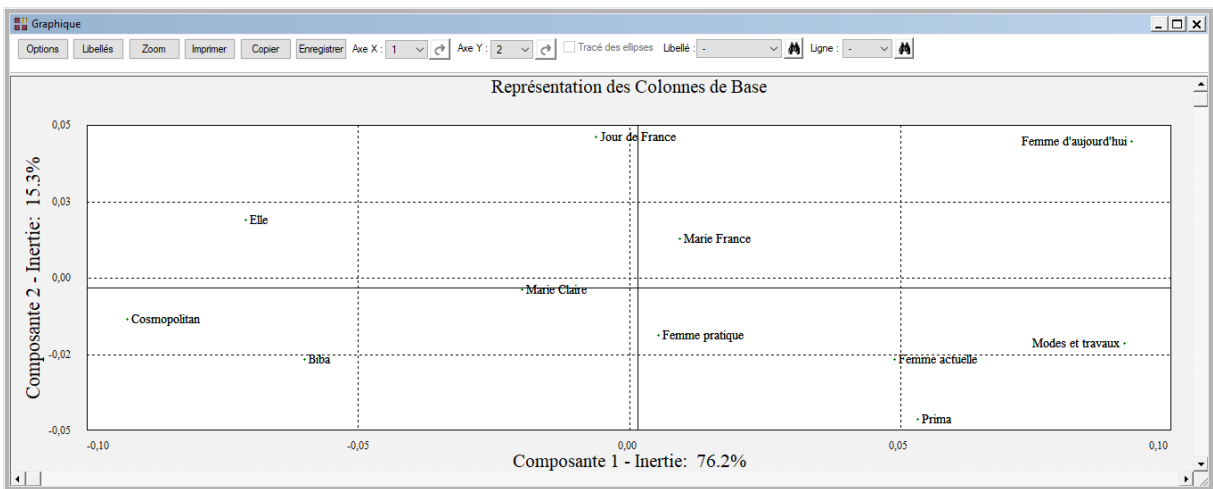
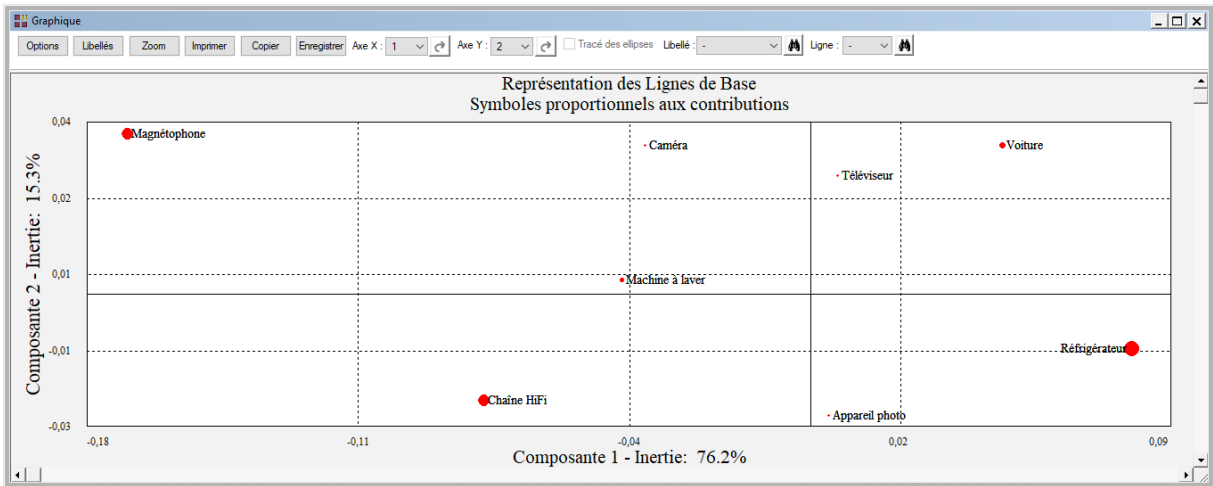
Choisissons la première option. Une boîte de dialogue permettant de choisir le plan factoriel s'affiche alors.



Elle permet également de préciser si l'on désire afficher les libellés des lignes, de choisir la couleur et la police, d'indiquer si les titres du graphique (titre 1, titre 2), doivent être conservés pour être réutilisés ultérieurement dans d'autres graphiques créés lors de cette même session de travail et de préciser si les symboles des points doivent être proportionnels aux contributions, aux cosinus carrés, aux distances carrées à l'origine ou tous de la même taille.

A noter que les options d'exploration graphique décrites dans le 'Manuel de Utilisateur' sont toutes accessibles, notamment l'interrogation d'un point et le zoom.



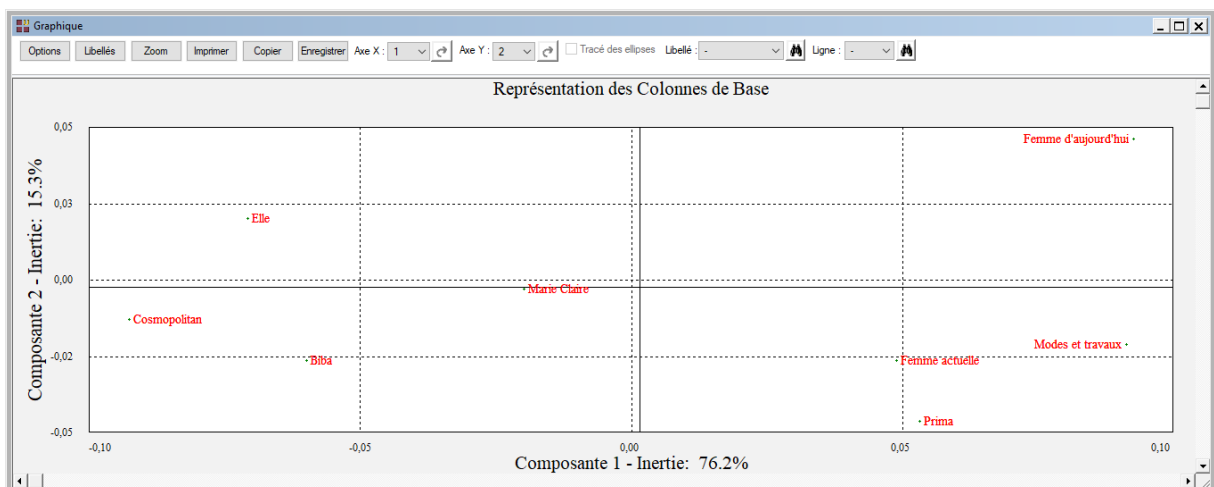
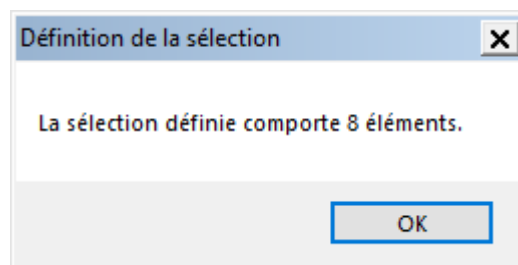
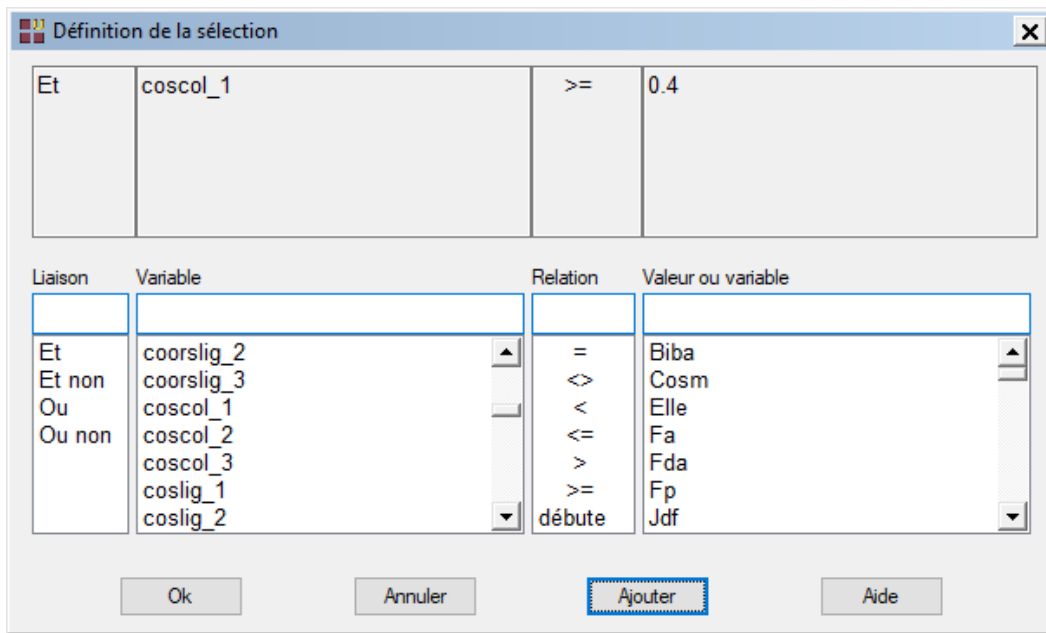


A noter également le bouton 'Sélection' qui permet de n'afficher que les points qui satisfont une condition logique, par exemple une condition basée sur les contributions ou les cosinus carrés. Ceci est très pratique dans le cas d'un nuage comportant beaucoup d'individus : seuls les points significatifs d'un point de vue statistique sont affichés.



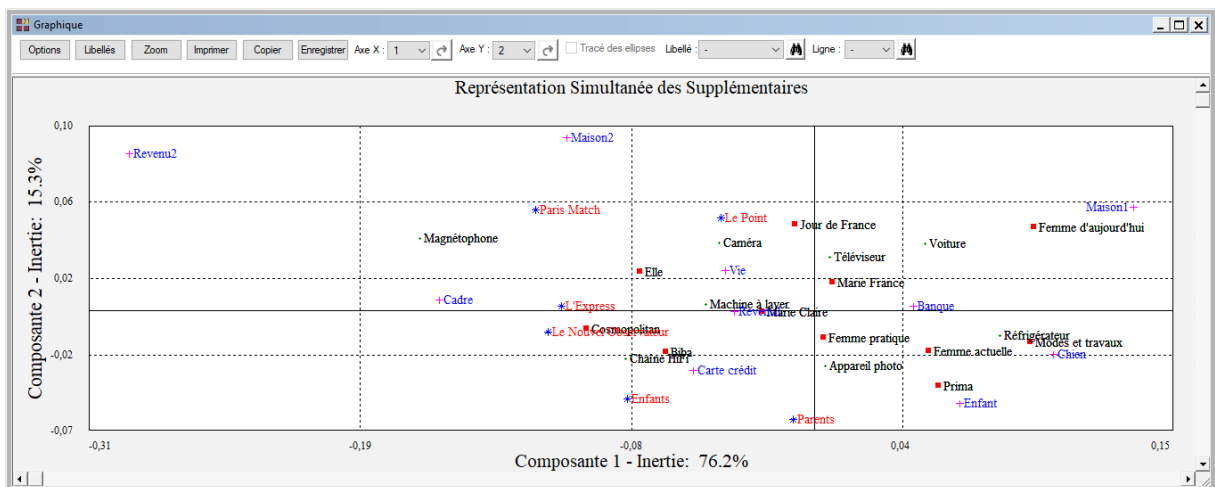
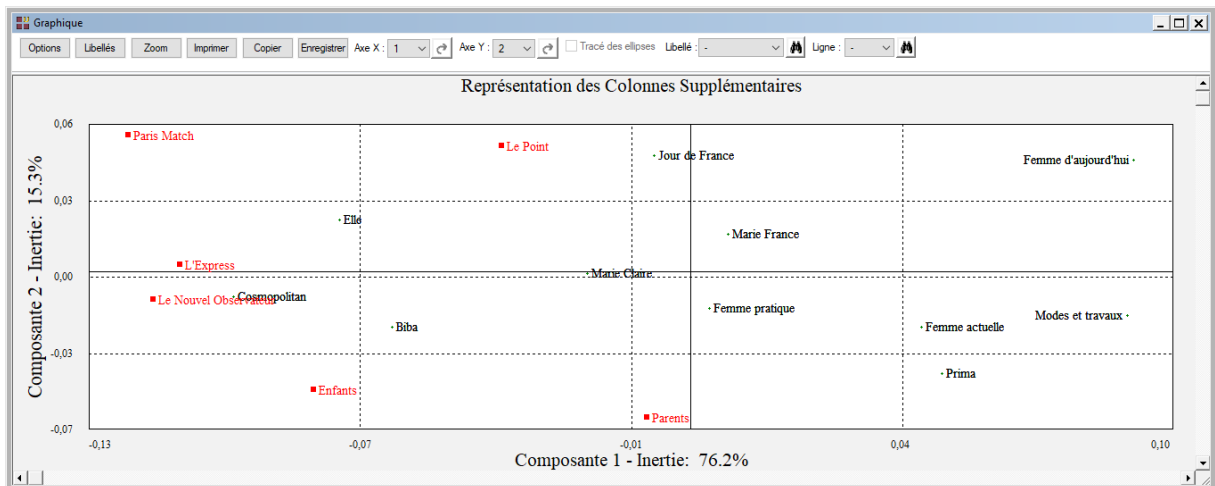
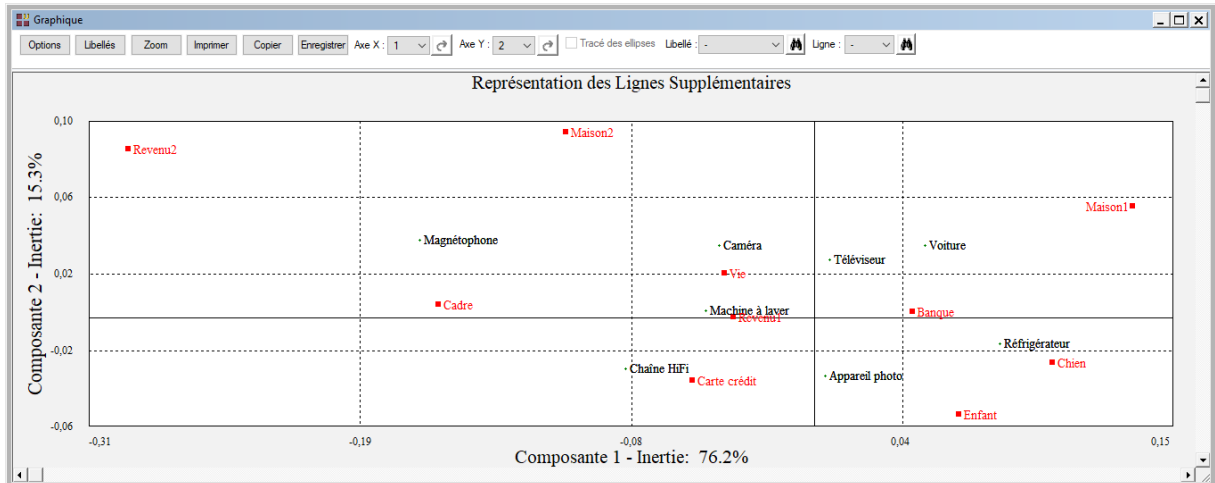
Les principaux résultats de l'analyse sont accessibles dans la boîte de dialogue 'Sélection' en plus des données du fichier JOURNAUX.

Par exemple, sélectionnons les journaux (colonnes) qui ont un cosinus carré sur le premier axe factoriel supérieur ou égal à 0,4. La variable coscol\_1 du fichier des résultats contient les cosinus carrés. Dans la boîte de dialogue Sélection entrons la sélection. Ce filtre sélectionne 8 journaux parmi les 11. Le graphique résultant de cette sélection est montré ci-après.



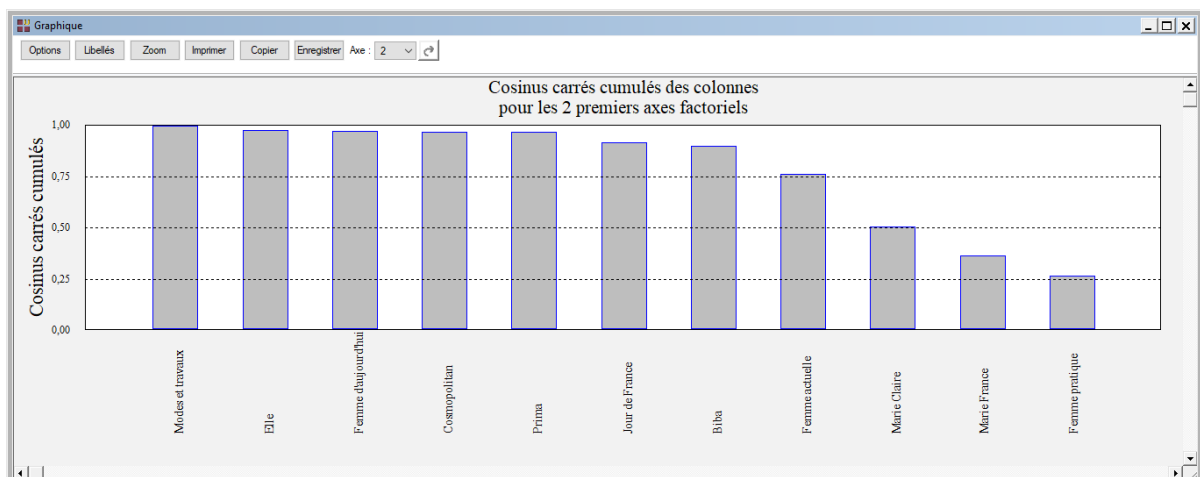
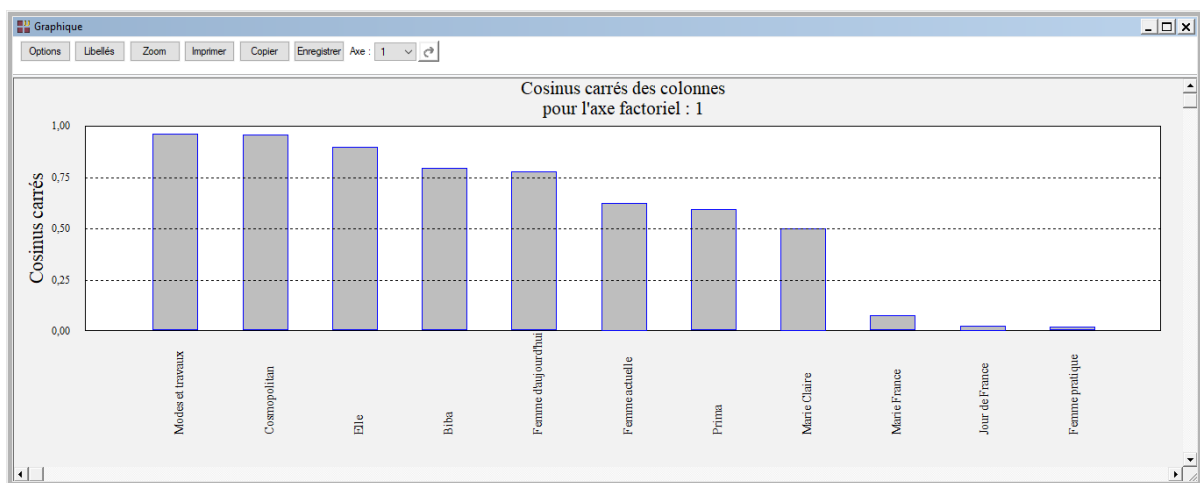
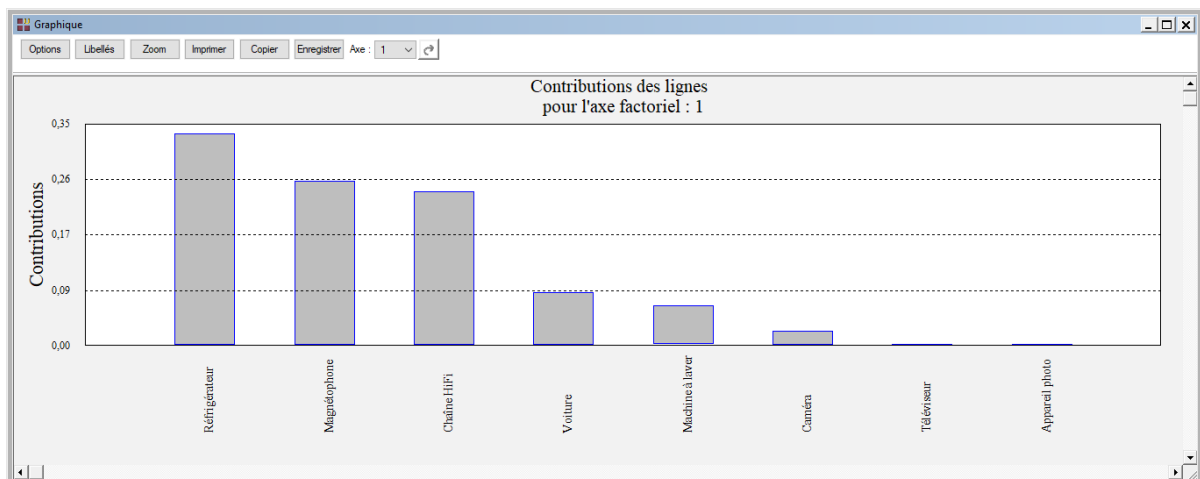
- Les options Plan supplémentaire

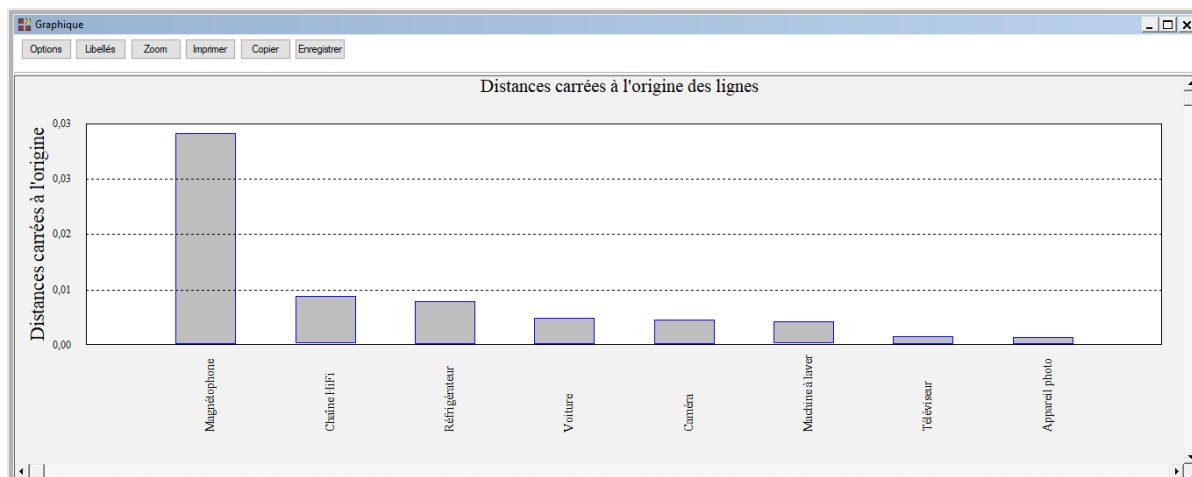
Ces options permettent de représenter les plans factoriels des lignes et des colonnes supplémentaires : lignes supplémentaires, colonnes supplémentaires ou lignes et colonnes supplémentaires. Comme pour les lignes et colonnes de base, le choix du plan factoriel est proposé et l'option de Sélection est active.



- Les options Contributions, Cosinus carrés, Cosinus carrés cumulés et Distances carrées à l'origine

Ces quatre options permettent de visualiser des diagrammes en bâtons ordonnés de diverses caractéristiques statistiques calculées pour les lignes ou les colonnes lors de l'analyse. Ils mettent en évidence les individus qui contribuent fortement à la définition des axes factoriels, qui sont bien projetés dans les plans factoriels et qui sont éloignés de l'origine.





## Une rapide interprétation des résultats

Cet exemple montre que le premier axe factoriel caractérise les lectrices habitant en ville, à hauts revenus, lisant Elle, Biba et Cosmopolitan et les lectrices n'habitant pas en ville, possédant leur habitation principale, un chien et lisant Modes et Travaux.

Le deuxième axe factoriel est relatif à l'âge des lectrices. Les plus âgées, possédant une voiture neuve et deux résidences, lisent Jour de France et Femme d'Aujourd'hui. Les plus jeunes lisent Femme Pratique et Prima.

## Les variables internes créées par la procédure

Voici la liste des variables internes créées par la procédure. Ces variables peuvent notamment être utilisées avec l'option 'Sélection'. A noter que certaines des variables mentionnées ci-dessous peuvent ne pas apparaître, en fonction des options choisies.

<i>Variable</i>	<i>Contenu</i>
liblig	Libellés des lignes
libcol	Libellés des colonnes
libligsup	Libellés des lignes supplémentaires
libcolsup	Libellés des colonnes supplémentaires
pclig	Probabilités conditionnelles lignes
pccol	Probabilités conditionnelles colonnes
numaxe	Numéros des axes
valpro	Valeurs propres
pctvar	Pourcentages de variance expliquée
pctvcum	Pourcentages cumulés de variance expliquée
variation	Variations de la variance expliquée
covcol	Matrices des covariances colonnes

marglig	Loi marginale lignes
margcol	Loi marginale colonnes
dolig	Distances carrées à l'origine lignes
docol	Distances carrées à l'origine colonnes
coorscol	Coordonnées des colonnes supplémentaires
cplig	Composantes principales lignes
cpcol	Composantes principales colonnes
conlig	Contributions lignes
concol	Contributions colonnes
coslig	Cosinus carrés lignes
colcol	Cosinus carrés colonnes
coorslig	Coordonnées des lignes supplémentaires
coorlig	Coordonnées des lignes de base
coorcol	Coordonnées des colonnes de base
simul	Représentation simultanée des lignes et des colonnes de base
barlig	Représentation barycentrique lignes de base
barcol	Représentation barycentrique colonnes de base
suplig	Représentation lignes supplémentaires
supcol	Représentation colonnes supplémentaires
simsup	Représentation simultanée des lignes et des colonnes supplémentaires