

Analyse Typologique avec SPSS

Cas :

On souhaite classifier le comportement de consommateurs en fonction de leurs attitudes vis à vis du shopping.

Dans le cadre d'un pré-test, une étude quantitative a été réalisée sur un échantillon de 20 personnes.

Les consommateurs ont exprimé leur degré d'accord avec les affirmations suivantes sur une échelle en 7 points (1 = n'est pas du tout d'accord – 7 = est d'accord) :

V1 = le shopping est amusant

V2 = le shopping est mauvais pour le budget

V3 = je profite du shopping pour manger à l'extérieur

V4 = j'essaie de trouver les meilleures affaires quand je fais du shopping

V5 = le shopping ne m'intéresse pas

V6 = vous pouvez économiser beaucoup d'argent en comparant les prix

Une fois les données recueillies, il vous est demandé de réaliser une analyse typologique afin de distinguer des groupes d'individus puis de décrire les caractéristiques des groupes obtenus.

Pour cela, les données recueillies seront traitées à partir du logiciel SPSS.

Le fichier Shopping contient les réponses des individus aux questions posées et peut être téléchargé sur le site <http://www.marketeurexpert.com>.

Traitement des données avec SPSS

Une fois ouvert, le logiciel SPSS propose deux types d’affichage :

- Affichage des données : vous pouvez lire les données recueillies. Les colonnes représentent les variables mesurées (V1 à V6). Les lignes présentent les réponses de chaque individu aux questions posées.
- Affichage des variables (voir copie d’écran suivante) : vous avez accès aux propriétés des variables utilisées pour l’analyse.

	caseno	v1	v2	v3	v4	v5	v6	var	var	var	var	var
1	1	6	4	7	3	2	3					
2	2	2	3	1	4	5	4					
3	3	7	2	6	4	1	3					
4	4	4	6	4	5	3	6					
5	5	1	3	2	2	6	4					
6	6	6	4	6	3	3	4					
7	7	5	3	6	3	3	4					
8	8	7	3	7	4	1	4					
9	9	2	4	3	3	6	3					
10	10	3	5	3	6	4	6					
11	11	1	3	2	3	5	3					
12	12	5	4	5	4	2	4					
13	13	2	2	1	5	4	4					
14	14	4	6	4	6	4	7					
15	15	6	5	4	2	1	4					
16	16	3	5	4	6	4	7					
17	17	4	4	7	2	2	5					
18	18	3	7	2	6	4	3					
19	19	4	6	3	7	2	7					
20	20	2	3	2	4	7	2					
21												
22												
23												
24												
25												

Cliquez sur l’onglet « Affichage des variables » pour accéder à leurs propriétés.

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	Manquant	Colonnes	Align	Mesure	Rôle
1	caseno	Numérique	11	0		Aucun	Aucun	8	Droite	Echelle	Entré
2	v1	Numérique	11	0	Fun	Aucun	Aucun	8	Droite	Echelle	Entré
3	v2	Numérique	11	0	Bad for Budget	Aucun	Aucun	8	Droite	Echelle	Entré
4	v3	Numérique	11	0	Eating Out	Aucun	Aucun	8	Droite	Echelle	Entré
5	v4	Numérique	11	0	Best Buys	Aucun	Aucun	8	Droite	Echelle	Entré
6	v5	Numérique	11	0	Don't Care	Aucun	Aucun	8	Droite	Echelle	Entré
7	v6	Numérique	11	0	Compare Prices	Aucun	Aucun	8	Droite	Echelle	Entré
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

La colonne mesure permet de choisir le type de mesure utilisé pour chacune des variables :

- Echelle
- Ordonnées
- Nominales

La variable « caseno » représente le numéro de l'individu. C'est donc une variable Nominale. Cliquez sur la cellule correspondant à la mesure de « caseno » pour changer le type de mesure.

Denis Bories – Analyse Typologique - ACP

table_20.1_input[1].sav [Ensemble_de_données1] - PASW Statistics Editeur de données

FR Français (France)

Fichier Edition Affichage Données Transformer Analyse Graphes Utilitaires Fenêtre Aide

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	Manquant	Colonnes	Align	Mesure	Rôle
1	caseno	Numérique	11	0		Aucun	Aucun	8	Droite	Echelle	Entré
2	v1	Numérique	11	0	Fun	Aucun	Aucun	8	Droite	Echelle	Entré
3	v2	Numérique	11	0	Bad for Budget	Aucun	Aucun	8	Droite	Ordinales	Entré
4	v3	Numérique	11	0	Eating Out	Aucun	Aucun	8	Droite	Nominales	Entré
5	v4	Numérique	11	0	Best Buys	Aucun	Aucun	8	Droite	Echelle	Entré
6	v5	Numérique	11	0	Don't Care	Aucun	Aucun	8	Droite	Echelle	Entré
7	v6	Numérique	11	0	Compare Prices	Aucun	Aucun	8	Droite	Echelle	Entré
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

Affichage des données Affichage des variables

Document1 ... Sans titre - P... *Résultats1 [...] table_20.1_in... Downloads -... Solde pour B... 07:58

De même, vous pouvez modifier le nom des variables afin de rendre l'affichage de certains résultats fournis par SPSS plus compréhensible.

*table_20.1_input[1].sav [Ensemble_de_données1] - PASW Statistics Editeur de données

FR Français (France)

Fichier Edition Affichage Données Transformer Analyse Graphes Utilitaires Fenêtre Aide

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	Manquant	Colonnes	Align	Mesure	Rôle
1	caseno	Numérique	11	0		Aucun	Aucun	8	Droite	Nominales	Entré
2	Marrant	Numérique	11	0	Fun	Aucun	Aucun	8	Droite	Echelle	Entré
3	MauvaisBud...	Numérique	11	0	Bad for Budget	Aucun	Aucun	8	Droite	Echelle	Entré
4	MangerDeh...	Numérique	11	0	Eating Out	Aucun	Aucun	8	Droite	Echelle	Entré
5	Affaires	Numérique	11	0	Best Buys	Aucun	Aucun	8	Droite	Echelle	Entré
6	PasInteresse	Numérique	11	0	Don't Care	Aucun	Aucun	8	Droite	Echelle	Entré
7	ComparePrix	Numérique	11	0	Compare Prices	Aucun	Aucun	8	Droite	Echelle	Entré
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

Affichage des données Affichage des variables

Document1 ... Sans titre - P... *Résultats1 [...] *table_20.1_i... Downloads -... Solde pour B... 07:59

Cliquez sur l'onglet « Affichage des données » pour voir les changements effectués. Vous remarquerez que les en-têtes des colonnes ont été modifiés.

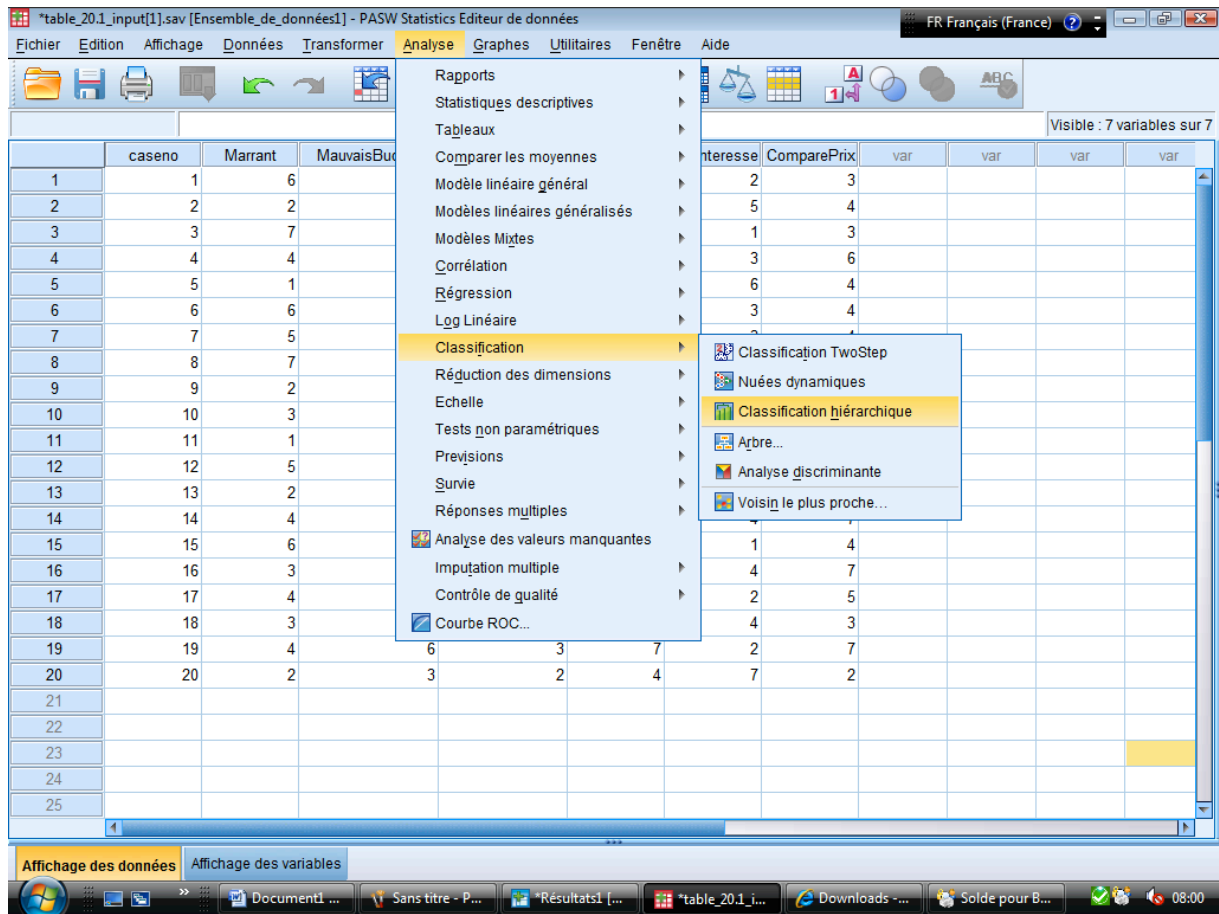
	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	var	var	var	var
1	1	6	4	7	3	2	3				
2	2	2	3	1	4	5	4				
3	3	7	2	6	4	1	3				
4	4	4	6	4	5	3	6				
5	5	1	3	2	2	6	4				
6	6	6	4	6	3	3	4				
7	7	5	3	6	3	3	4				
8	8	7	3	7	4	1	4				
9	9	2	4	3	3	6	3				
10	10	3	5	3	6	4	6				
11	11	1	3	2	3	5	3				
12	12	5	4	5	4	2	4				
13	13	2	2	1	5	4	4				
14	14	4	6	4	6	4	7				
15	15	6	5	4	2	1	4				
16	16	3	5	4	6	4	7				
17	17	4	4	7	2	2	5				
18	18	3	7	2	6	4	3				
19	19	4	6	3	7	2	7				
20	20	2	3	2	4	7	2				
21											
22											
23											
24											
25											

Analyse Typologique

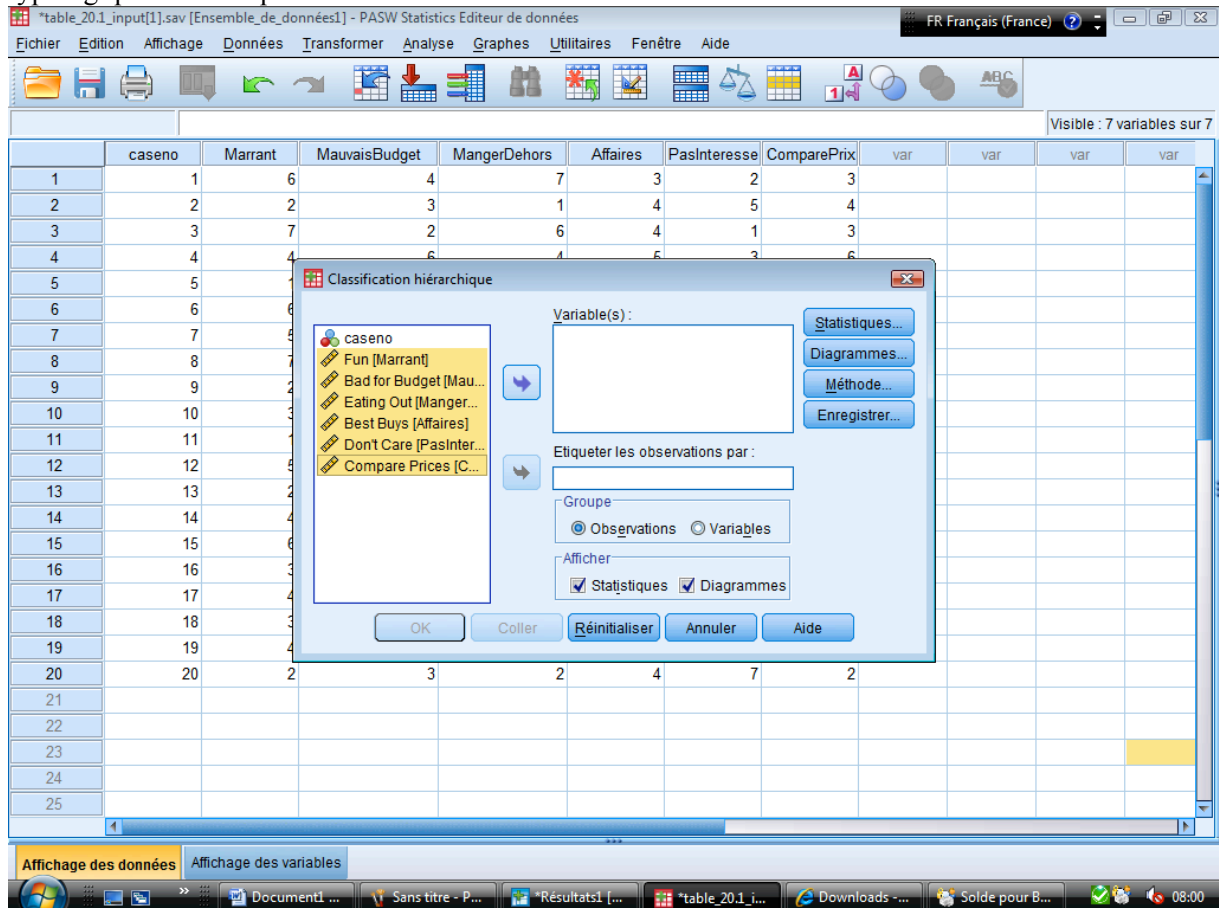
Classification Hiérarchique

Dans un premier temps, nous allons mettre en œuvre une classification hiérarchique. Pour sélectionner ce type d'analyse sous SPSS, suivez le chemin suivant :

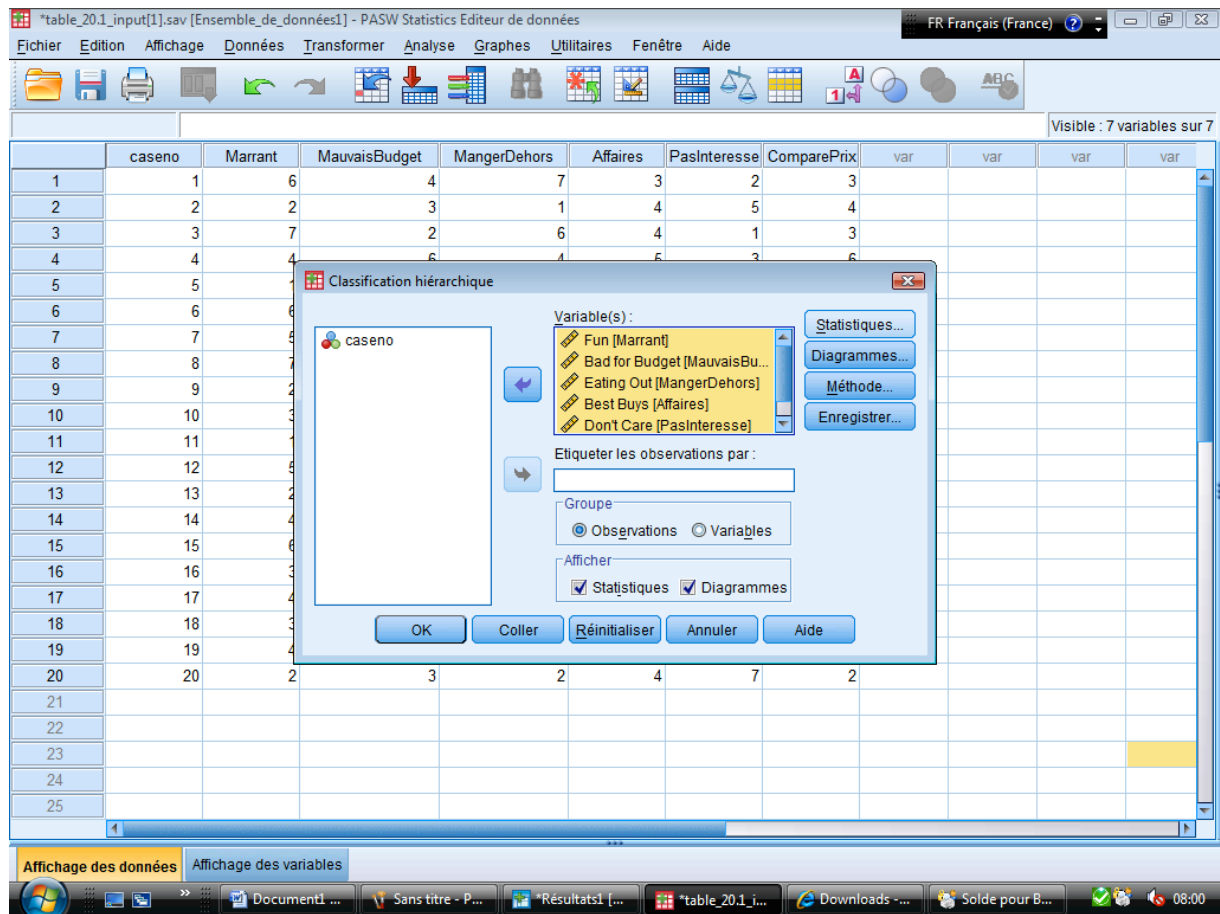
- Analyse > Classification > Classification hiérarchique



SPSS ouvre une nouvelle fenêtre présentant l'ensemble des options disponibles pour l'analyse typologique hiérarchique.



Dans un premier temps, vous devez sélectionner les variables que vous souhaitez analyser et les faire glisser sous la partie « Variables ».



Puis, cliquez sur le bouton « Statistiques » pour choisir les informations statistiques que vous souhaitez obtenir. Sélectionnez la chaîne des agrégations ainsi que la matrice des distances (Voir cours)

Sélectionnez ensuite le nombre de groupes que vous souhaitez obtenir (ici, de 2 à 6). SPSS vous proposera 5 solutions de classification :

- Une à 6 groupes
- Une à 5 groupes
- Une à 4 groupes
- Une à 3 groupes
- Une à 2 groupes

The screenshot shows the PASW Statistics interface with a data table and a dialog box for hierarchical classification. The data table has the following structure:

	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	var	var	var	var
1	1	6	4	7	3	2	3				
2	2	2	3	1	4	5	4				
3	3	7	2	6	4	1	3				
4	4	4	6	4	6	2	6				
5	5	6	4	6	4	6	6				
6	6	6	6	6	6	6	6				
7	7	5	6	6	6	6	6				
8	8	7	6	6	6	6	6				
9	9	7	6	6	6	6	6				
10	10	4	6	6	6	6	6				
11	11	3	6	6	6	6	6				
12	12	5	6	6	6	6	6				
13	13	2	6	6	6	6	6				
14	14	6	6	6	6	6	6				
15	15	4	6	6	6	6	6				
16	16	3	6	6	6	6	6				
17	17	4	6	6	6	6	6				
18	18	3	6	6	6	6	6				
19	19	4	6	6	6	6	6				
20	20	2	3	2	4	7	2				
21											
22											
23											
24											
25											

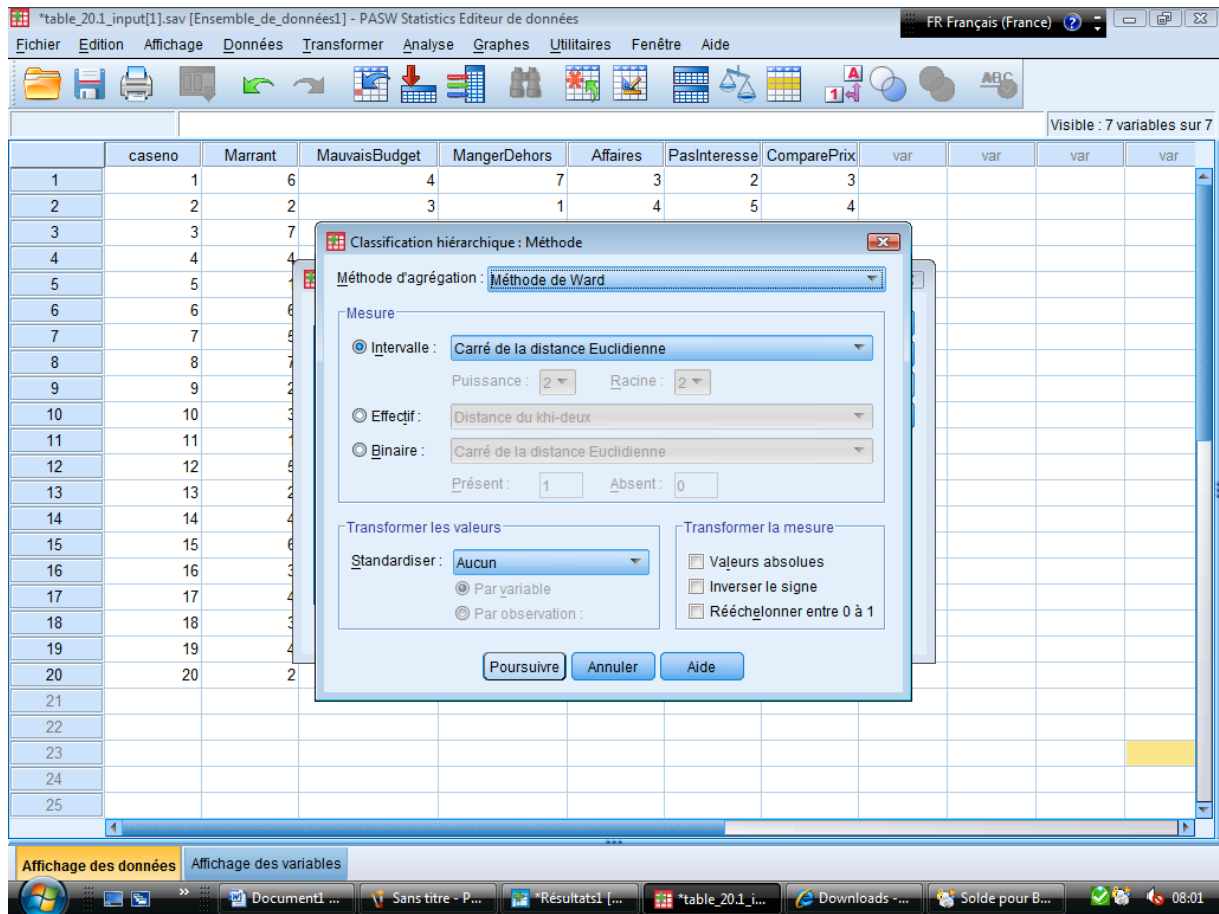
The dialog box 'Classification hiérarchique : Statistiques' is open, showing the following options:

- Chaîne des agrégations
- Matrice des distances
- Classe(s) d'affectation:
 - Aucun
 - Une seule solution
 - Plusieurs solutions
- Nombre de classes: []
- Nombre minimum de classes: [2]
- Nombre maximum de classes: [6]

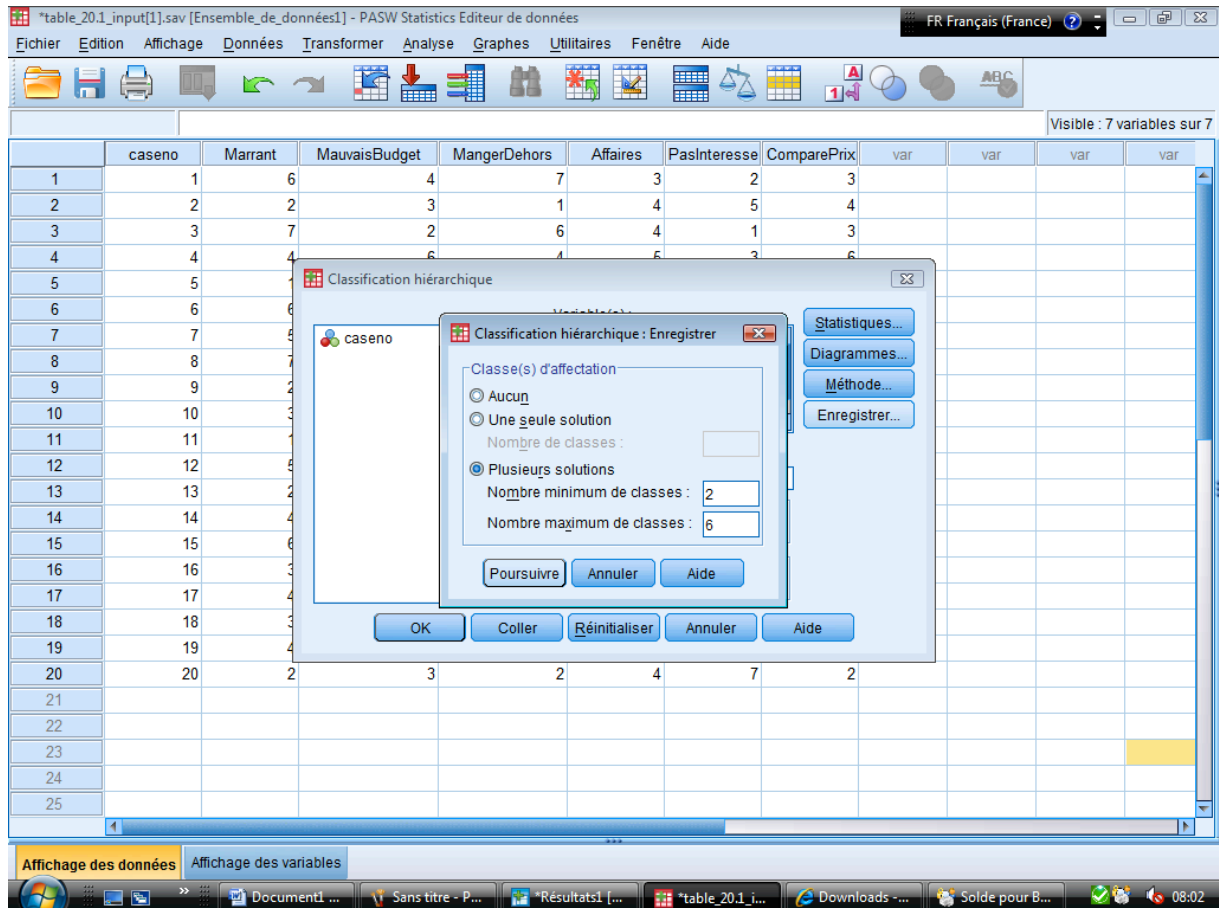
Buttons in the dialog: OK, Poursuivre, Annuler, Aide. Buttons on the right: Statistiques..., Diagrammes..., Méthode..., Enregistrer...

Cliquez ensuite sur le bouton « Diagrammes ». Choisissez l’affichage de l’arbre hiérarchique en cliquant sur « Arbre hiérarchique ».

Précisez le sens de l’affichage dans lequel vous souhaitez que l’arbre soit affiché, en vertical ou en horizontal. Ce choix n’a aucune incidence sur la classification réalisée, il s’agit juste d’une présentation graphique des associations entre individus.



Cliquez ensuite sur le bouton « Enregistrer ». Ici, vous pouvez sélectionner les résultats de la classification pour chacune des classifications que vous avez demandé, 2, 3, 4, 5, 6 groupes.



Cliquez sur le bouton « OK » pour lancer l'analyse.

Résultats de la classification hiérarchique

Ce tableau indique que le fichier contient 20 individus (N=20) et que le fichier ne contient pas de donnée manquante.

Récapitulatif de traitement des observations^a

Observations					
Valide		Manquante		Total	
N	Pourcentage	N	Pourcentage	N	Pourcentage
20	100,0	0	,0	20	100,0

a. Distance de Ward

La matrice de proximité présente les distances qui séparent les individus. Les lignes et les colonnes de cette matrice sont les individus. Le chiffre présenté est la distance calculée avec la méthode et la mesure sélectionnées. Les individus sont en lignes et en colonnes. Cette matrice est triangulaire, c'est à dire qu'elle est séparée par sa diagonale qui contient la valeur nulle (distance d'un individu par rapport à lui même égale à 0) et que la partie haute et la partie basse présentent les mêmes données.

Comment sont calculées ces distances : Les distances sont calculées par paire d'individus.

Ex : distance entre individu 1 et individu 2, mesure : distance euclidienne au carré. On utilise les réponses de ces deux individus pour chacune des questions posées pour calculer les distances.

caseno	v1	v2	v3	v4	v5	v6
1	6	4	7	3	2	3
2	2	3	1	4	5	4

$$\text{Distance euclidienne au carré} = (6-2)^2 + (4-3)^2 + (7-1)^2 + (3-4)^2 + (2-5)^2 + (3-4)^2 = 64$$

On fait ce même calcul pour l'ensemble des paires d'individus.

Matrice de proximité

Observation	Carré de la distance Euclidienne						
	1	2	3	4	5	6	7
1	,000	64,000	8,000	31,000	69,000	3,000	5,000
2	64,000	,000	68,000	31,000	7,000	47,000	39,000
3	8,000	68,000	,000	43,000	83,000	11,000	11,000
4	31,000	31,000	43,000	,000	44,000	20,000	22,000
5	69,000	7,000	83,000	44,000	,000	52,000	42,000
6	3,000	47,000	11,000	20,000	52,000	,000	2,000
7	5,000	39,000	11,000	22,000	42,000	2,000	,000
8	5,000	77,000	3,000	36,000	90,000	8,000	10,000
9	48,000	8,000	64,000	31,000	5,000	35,000	29,000
10	48,000	18,000	56,000	5,000	33,000	33,000	31,000
11	60,000	4,000	70,000	39,000	3,000	47,000	37,000
12	7,000	35,000	11,000	12,000	46,000	4,000	4,000
13	65,000	3,000	61,000	34,000	16,000	50,000	40,000
14	46,000	36,000	58,000	3,000	51,000	31,000	33,000
15	13,000	49,000	19,000	22,000	58,000	10,000	14,000
16	48,000	28,000	58,000	5,000	41,000	33,000	31,000
17	9,000	55,000	23,000	24,000	52,000	8,000	6,000
18	56,000	24,000	70,000	17,000	41,000	45,000	47,000
19	56,000	44,000	60,000	7,000	69,000	43,000	45,000
20	69,000	9,000	79,000	50,000	10,000	54,000	46,000

Matrice de proximité

Observation	Carré de la distance Euclidienne						
	1	2	3	4	5	6	7
1	,000	64,000	8,000	31,000	69,000	3,000	5,000
2	64,000	,000	68,000	31,000	7,000	47,000	39,000
3	8,000	68,000	,000	43,000	83,000	11,000	11,000
4	31,000	31,000	43,000	,000	44,000	20,000	22,000
5	69,000	7,000	83,000	44,000	,000	52,000	42,000
6	3,000	47,000	11,000	20,000	52,000	,000	2,000
7	5,000	39,000	11,000	22,000	42,000	2,000	,000
8	5,000	77,000	3,000	36,000	90,000	8,000	10,000
9	48,000	8,000	64,000	31,000	5,000	35,000	29,000
10	48,000	18,000	56,000	5,000	33,000	33,000	31,000
11	60,000	4,000	70,000	39,000	3,000	47,000	37,000
12	7,000	35,000	11,000	12,000	46,000	4,000	4,000
13	65,000	3,000	61,000	34,000	16,000	50,000	40,000
14	46,000	36,000	58,000	3,000	51,000	31,000	33,000
15	13,000	49,000	19,000	22,000	58,000	10,000	14,000
16	48,000	28,000	58,000	5,000	41,000	33,000	31,000
17	9,000	55,000	23,000	24,000	52,000	8,000	6,000
18	56,000	24,000	70,000	17,000	41,000	45,000	47,000
19	56,000	44,000	60,000	7,000	69,000	43,000	45,000
20	69,000	9,000	79,000	50,000	10,000	54,000	46,000

Ceci est une matrice de dissimilarité

Matrice de proximité

Observation	Carré de la distance Euclidienne						
	8	9	10	11	12	13	14
1	5,000	48,000	48,000	60,000	7,000	65,000	46,000
2	77,000	8,000	18,000	4,000	35,000	3,000	36,000
3	3,000	64,000	56,000	70,000	11,000	61,000	58,000
4	36,000	31,000	5,000	39,000	12,000	34,000	3,000
5	90,000	5,000	33,000	3,000	46,000	16,000	51,000
6	8,000	35,000	33,000	47,000	4,000	50,000	31,000
7	10,000	29,000	31,000	37,000	4,000	40,000	33,000
8	,000	69,000	53,000	79,000	10,000	72,000	49,000
9	69,000	,000	24,000	4,000	31,000	17,000	38,000
10	53,000	24,000	,000	28,000	21,000	19,000	4,000
11	79,000	4,000	28,000	,000	37,000	9,000	48,000
12	10,000	31,000	21,000	37,000	,000	34,000	23,000
13	72,000	17,000	19,000	9,000	34,000	,000	39,000
14	49,000	38,000	4,000	48,000	23,000	39,000	,000
15	18,000	45,000	39,000	51,000	8,000	52,000	39,000
16	51,000	32,000	2,000	38,000	23,000	29,000	2,000
17	16,000	41,000	39,000	49,000	10,000	58,000	37,000
18	71,000	24,000	14,000	30,000	31,000	29,000	22,000
19	53,000	56,000	8,000	60,000	27,000	41,000	6,000
20	90,000	5,000	35,000	7,000	48,000	16,000	55,000

Ceci est une matrice de dissimilarité

Matrice de proximité

Observation	Carré de la distance Euclidienne					
	15	16	17	18	19	20
1	13,000	48,000	9,000	56,000	56,000	69,000
2	49,000	28,000	55,000	24,000	44,000	9,000
3	19,000	58,000	23,000	70,000	60,000	79,000
4	22,000	5,000	24,000	17,000	7,000	50,000
5	58,000	41,000	52,000	41,000	69,000	10,000
6	10,000	33,000	8,000	45,000	43,000	54,000
7	14,000	31,000	6,000	47,000	45,000	46,000
8	18,000	51,000	16,000	71,000	53,000	90,000
9	45,000	32,000	41,000	24,000	56,000	5,000
10	39,000	2,000	39,000	14,000	8,000	35,000
11	51,000	38,000	49,000	30,000	60,000	7,000
12	8,000	23,000	10,000	31,000	27,000	48,000
13	52,000	29,000	58,000	29,000	41,000	16,000
14	39,000	2,000	37,000	22,000	6,000	55,000
15	,000	43,000	16,000	43,000	41,000	68,000
16	43,000	,000	35,000	24,000	8,000	47,000
17	16,000	35,000	,000	59,000	49,000	68,000
18	43,000	24,000	59,000	,000	24,000	31,000
19	41,000	8,000	49,000	24,000	,000	73,000
20	68,000	47,000	68,000	31,000	73,000	,000

Ceci est une matrice de dissimilarité

Distance de Ward

La chaîne des agrégations donne des informations sur les individus combinés à chaque étape du processus de classification hiérarchique

EX : Les individus 14 et 16 regroupés à l'étape 1. Un de ces deux individus a ensuite été classé à l'étape 6.

Chaîne des agrégations

Etape	Regroupement de classes		Coefficients	Etape d'apparition de la classe		Etape suivante
	Classe 1	Classe 2		Classe 1	Classe 2	
1	14	16	1,000	0	0	6
2	6	7	2,000	0	0	7
3	2	13	3,500	0	0	15
4	5	11	5,000	0	0	11
5	3	8	6,500	0	0	16
6	10	14	8,167	0	1	9
7	6	12	10,500	2	0	10
8	9	20	13,000	0	0	11
9	4	10	15,583	0	6	12
10	1	6	18,500	0	7	13
11	5	9	23,000	4	8	15
12	4	19	27,750	9	0	17
13	1	17	33,100	10	0	14
14	1	15	41,333	13	0	16
15	2	5	51,833	3	11	18
16	1	3	64,500	14	5	19
17	4	18	79,667	12	0	18
18	2	4	172,667	15	17	19
19	1	2	328,600	16	18	0

Le tableau d'appartenance à la classe présente le numéro de groupe auquel chaque individu a été affecté pour chacune des solutions de classification demandée (2, 3, 4, 5 et 6 groupes).

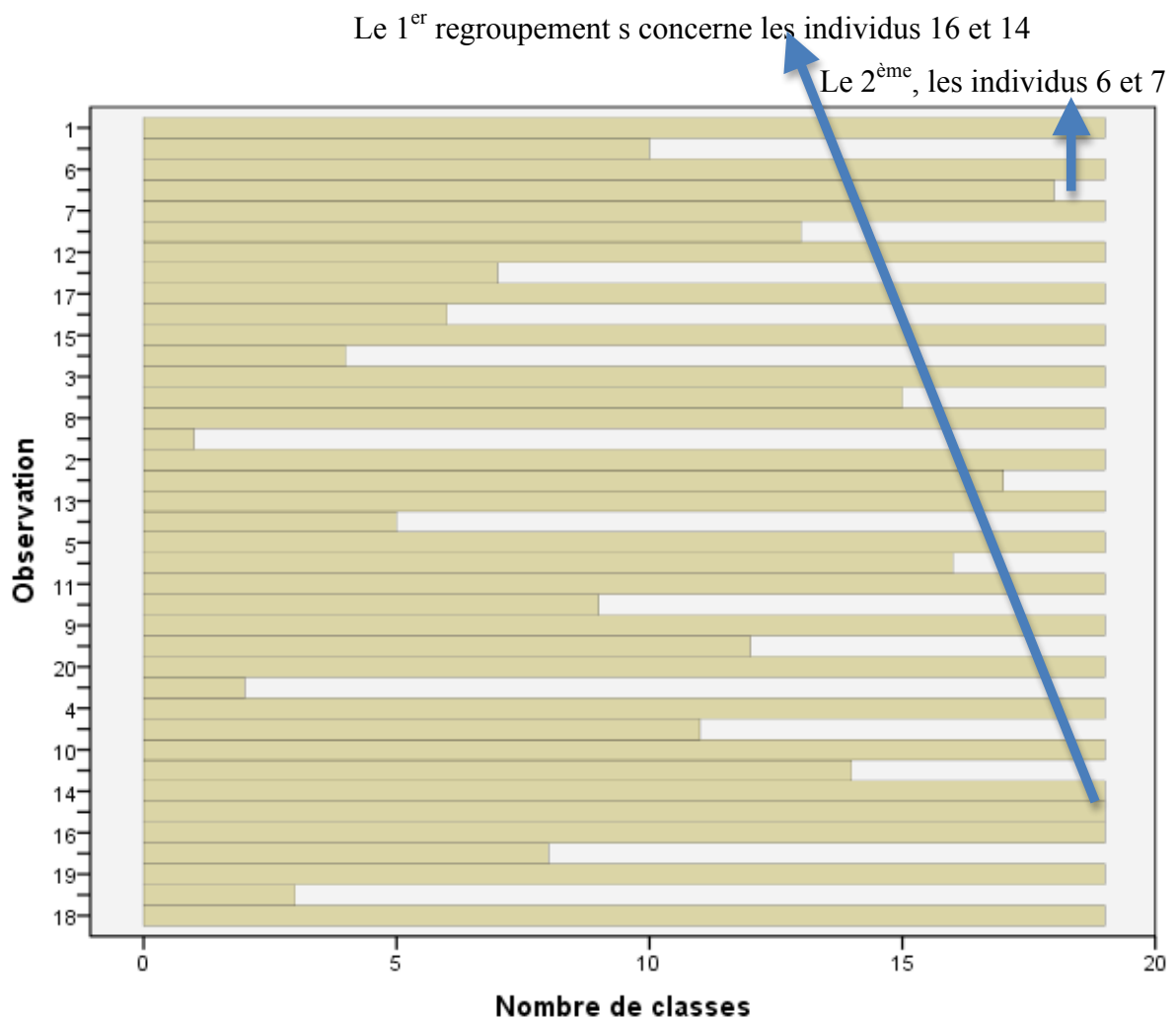
Appartenance à la classe

Observation	6 classes	5 classes	4 classes	3 classes	2 classes
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	1	1	1
4	4	4	3	3	2
5	5	2	2	2	2
6	1	1	1	1	1
7	1	1	1	1	1
8	3	3	1	1	1
9	5	2	2	2	2
10	4	4	3	3	2
11	5	2	2	2	2
12	1	1	1	1	1
13	2	2	2	2	2
14	4	4	3	3	2
15	1	1	1	1	1
16	4	4	3	3	2
17	1	1	1	1	1
18	6	5	4	3	2
19	4	4	3	3	2
20	5	2	2	2	2

Le diagramme en stalactite (icicle plot) est une représentation graphique des résultats de la classification. Il est appelé ainsi car, lorsqu'il est présenté verticalement, il ressemble à une rangée de stalactites. Lorsqu'il est présenté verticalement, il se lit de bas en haut. Les colonnes correspondent aux individus à classer et les rangs au nombre de groupes

Dans notre cas, il est présenté horizontalement et se lit donc de gauche à droite. Les individus sont en ligne (Observation). Le graphique présente l'ordre dans lequel les regroupements d'individus ont été réalisés.

Une des limites de ce graphique est constituée par la difficulté de lecture des classements réalisés.



Effectifs

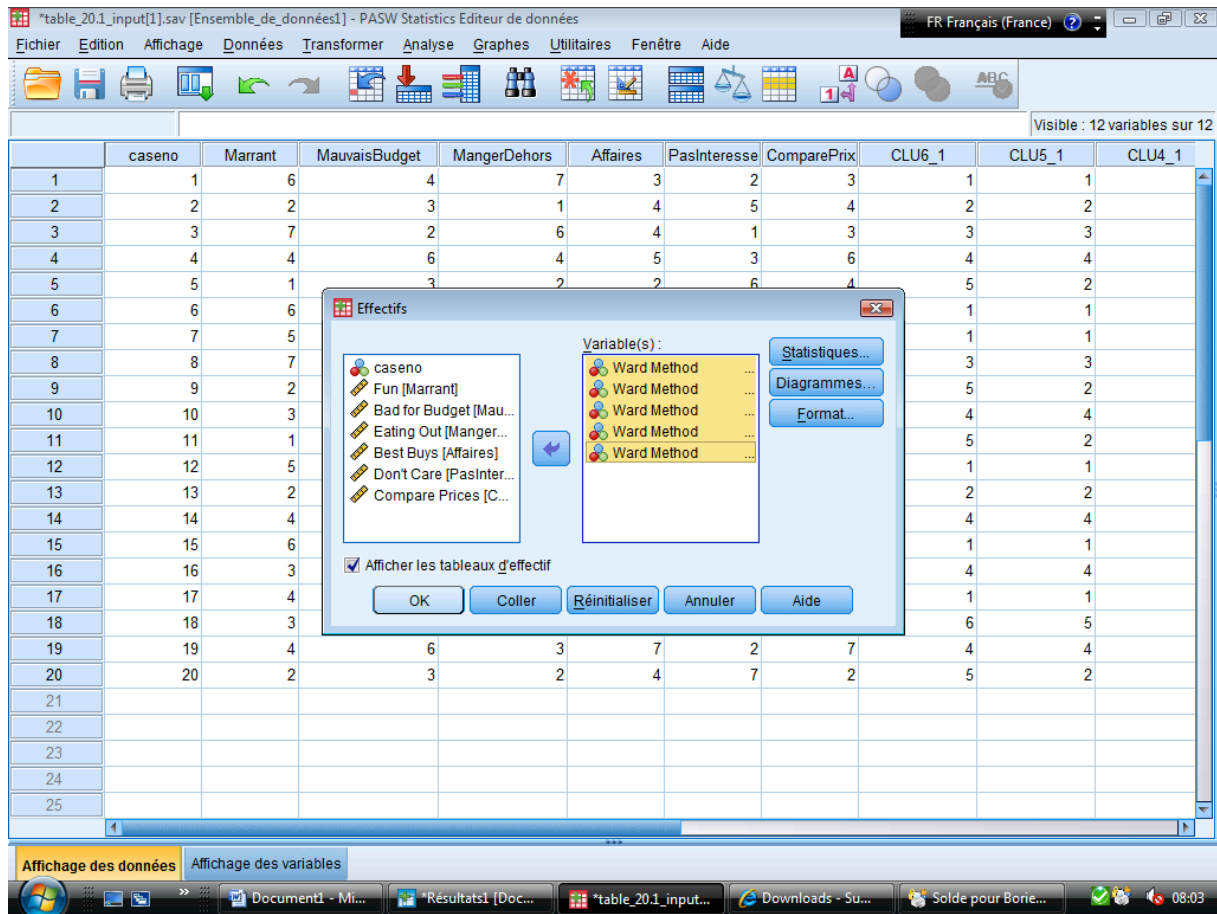
Les résultats de l'analyse typologique hiérarchique ne présentent pas les effectifs obtenus pour chaque classification.

Vous devez donc compter ces effectifs. Pour cela, utilisez le cheminement suivant : Analyse > Statistiques descriptives > Effectifs.

The screenshot shows the PASW Statistics interface. The 'Analyse' menu is open, and the path 'Statistiques descriptives' > 'Effectifs' is highlighted. The background shows a data table with the following visible columns: 'caseno', 'Marrant', 'MauvaisBuc', 'J6_1', 'CLU5_1', and 'CLU4_1'. The data rows are numbered 1 to 25.

	caseno	Marrant	MauvaisBuc	J6_1	CLU5_1	CLU4_1
1	1	6		1	1	
2	2	2		2	2	
3	3	7		3	3	
4	4	4		4	4	
5	5	1		5	2	
6	6	6		3	4	1
7	7	5		3	4	1
8	8	7		1	4	3
9	9	2		6	3	5
10	10	3		4	6	4
11	11	1		5	3	5
12	12	5		2	4	1
13	13	2		4	4	2
14	14	4		4	7	4
15	15	6		1	4	1
16	16	3		4	7	4
17	17	4		2	5	1
18	18	3		4	3	6
19	19	4	6	3	7	2
20	20	2	3	2	4	7
21				7	2	5
22						2
23						
24						
25						

Sélectionnez les variables contenant la classification des individus dans les groupes pour les solutions à 2, 3, 4, 5 et 6 groupes.



Les résultats obtenus sont les suivants.

Ces tableaux présentent les données valides et manquantes pour chacune des variables analysées

Statistiques

		Ward Method	Ward Method	Ward Method
N	Valide	20	20	20
	Manquante	0	0	0

Statistiques

		Ward Method	Ward Method
N	Valide	20	20
	Manquante	0	0

Tableau de fréquences

Les tableaux des effectifs indiquent les effectifs pour chaque solution de classification demandée, c'est à dire, 6, 5, 4, 3, 2 groupes.

Lecture du tableau :

Le groupe 1 a un effectif de 6 individus, ce qui représente 30% de l'échantillon total.

Le groupe 2 a un effectif de 2 individus, ce qui représente 10% de l'échantillon total.

Ward Method

		Effectifs	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	1	6	30,0	30,0	30,0
	2	2	10,0	10,0	40,0
	3	2	10,0	10,0	50,0
	4	5	25,0	25,0	75,0
	5	4	20,0	20,0	95,0
	6	1	5,0	5,0	100,0
	Total	20	100,0	100,0	

Ward Method

		Effectifs	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	1	6	30,0	30,0	30,0
	2	6	30,0	30,0	60,0
	3	2	10,0	10,0	70,0
	4	5	25,0	25,0	95,0
	5	1	5,0	5,0	100,0
	Total	20	100,0	100,0	

Ward Method

		Effectifs	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	1	8	40,0	40,0	40,0
	2	6	30,0	30,0	70,0
	3	5	25,0	25,0	95,0
	4	1	5,0	5,0	100,0
	Total	20	100,0	100,0	

Ward Method

		Effectifs	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	1	8	40,0	40,0	40,0
	2	6	30,0	30,0	70,0
	3	6	30,0	30,0	100,0
	Total	20	100,0	100,0	

Ward Method

		Effectifs	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	1	8	40,0	40,0	40,0
	2	12	60,0	60,0	100,0
	Total	20	100,0	100,0	

Combien de groupes retenir ?

Une classification des individus en 3 groupes semble être la solution la plus cohérente à retenir. En effet, d'une part, pour les classifications à 6, 5 et 4 groupes, 1 individu forme un groupe. Ce n'est qu'à partir d'une classification à 3 groupes que cet individu est affecté à un autre groupe. De plus, les groupes obtenus présentent des effectifs relativement équilibrés.

Pour conclure, l'observation de l'arbre hiérarchique confirme que cette classification a été obtenue relativement rapidement dans l'analyse, ce qui indique que les individus classés sont proches les uns des autres mais aussi suffisamment distincts pour ne pas être associés.

Arbre hiérarchique utilisant la Distance de Ward

Distance de combinaison des classes redimensionnée



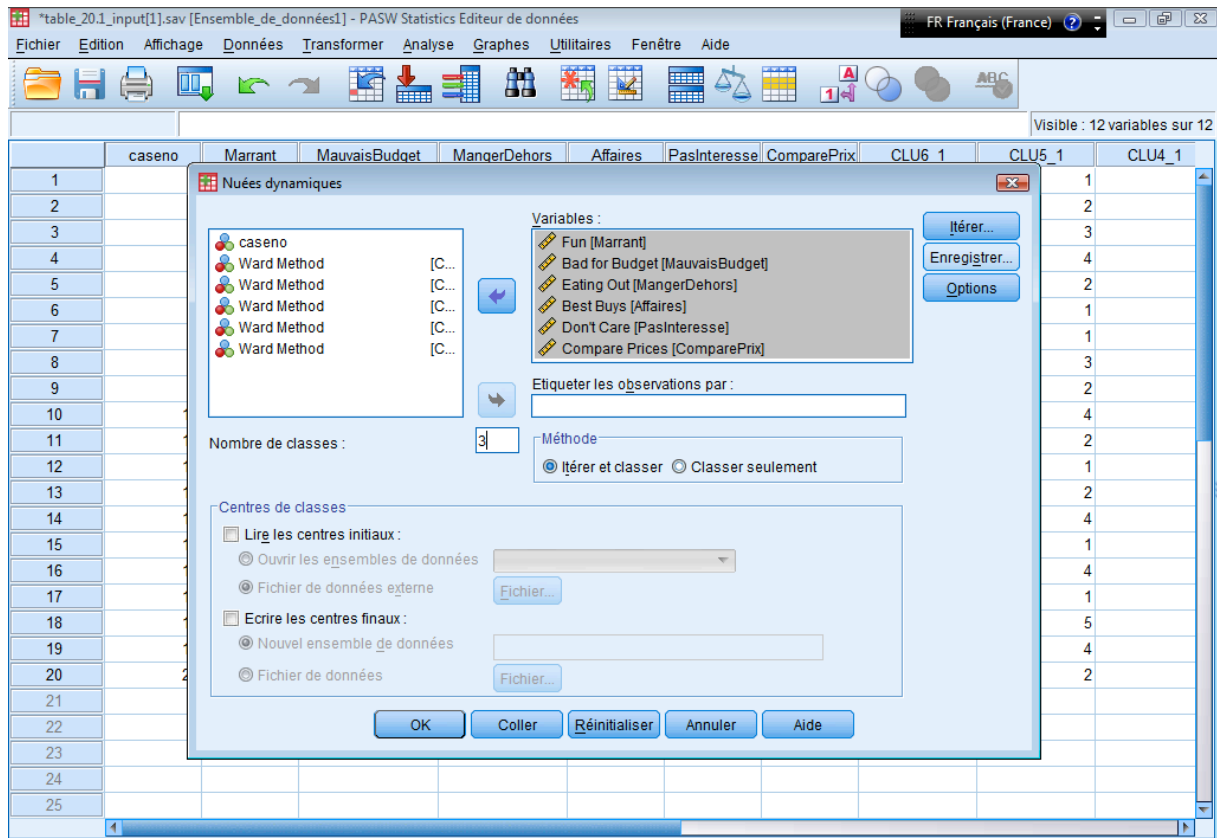
Classification Non Hiérarchique ou Nuées Dynamiques

Afin de vérifier les résultats obtenus grâce à la méthode de classification hiérarchique, nous allons maintenant analyser les données recueillies en utilisant une autre méthode de classification non hiérarchique, à savoir la méthode des nuées dynamiques en suivant le chemin Analyse > Classification > Nuées dynamiques

The screenshot shows the PASW Statistics interface. The 'Analyse' menu is open, and 'Nuées dynamiques' is highlighted. The data table in the background is as follows:

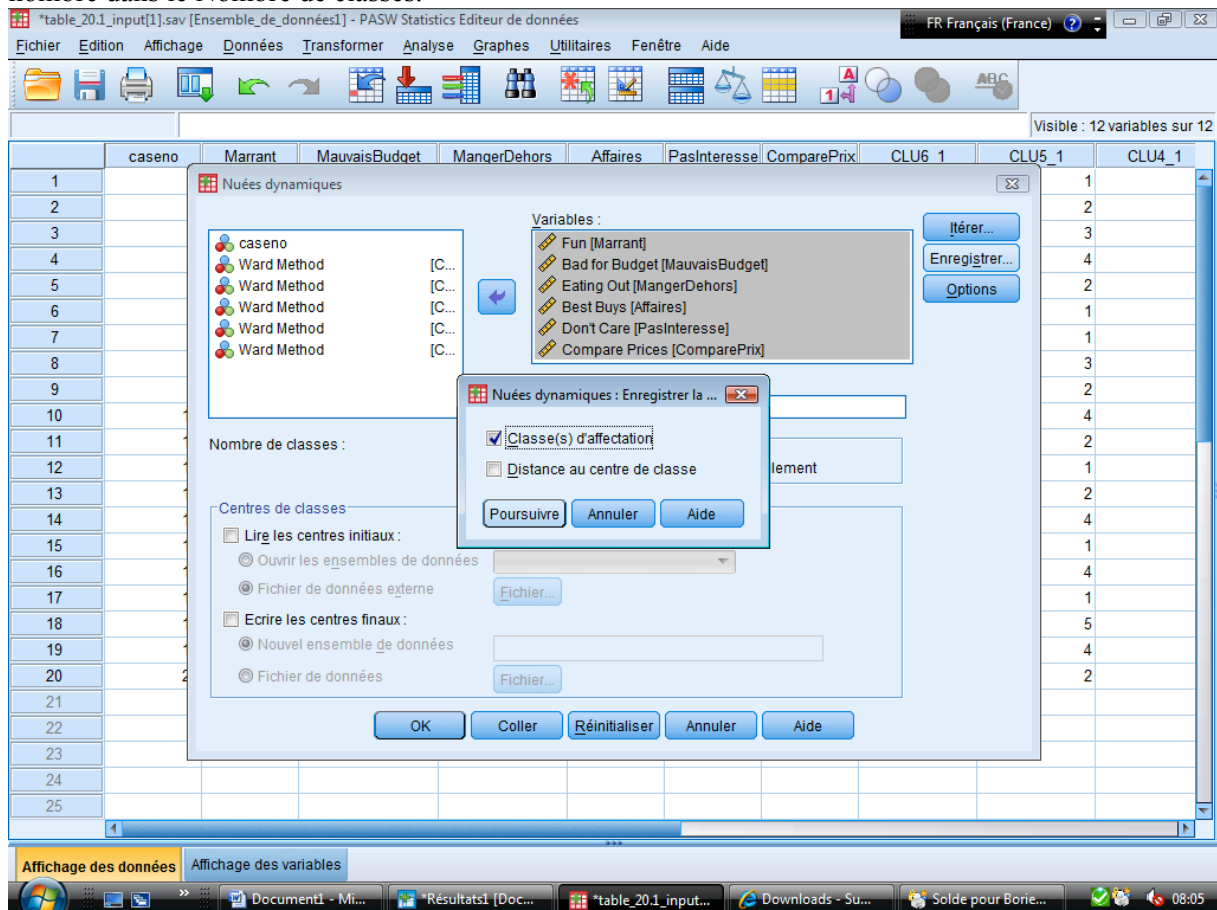
	caseno	Marrant	MauvaisBuc	Interesse	ComparePrix	CLU6_1	CLU5_1	CLU4_1
1	1	6		2	3	1	1	
2	2	2		5	4	2	2	
3	3	7		1	3	3	3	
4	4	4		3	6	4	4	
5	5	1		6	4	5	2	
6	6	6		3	4	1	1	
7	7	5				1	1	
8	8	7				3	3	
9	9	2				5	2	
10	10	3				4	4	
11	11	1				5	2	
12	12	5				1	1	
13	13	2				2	2	
14	14	4				4	4	
15	15	6		1	4	1	1	
16	16	3		4	7	4	4	
17	17	4		2	5	1	1	
18	18	3		4	3	6	5	
19	19	4		6	3	7	4	4
20	20	2		3	2	4	7	2
21								
22								
23								
24								
25								

Puis sélectionner les variables à partir desquelles la classification sera réalisée.



La méthode des nuées dynamiques nécessite de spécifier le nombre de groupes que nous souhaitons obtenir.

La classification hiérarchique nous ayant conduit à retenir une classification en 3 groupes, précisez ce nombre dans le Nombre de classes.



Demandez ensuite à SPSS d'enregistrer les affectations des individus à un groupe dans le fichier analysé en sélectionnant « Classes d'affectation ».

Cliquez sur « Poursuivre » puis sur « OK ».

Résultats de la classification nuées dynamiques

Les méthodes de classification non hiérarchique nécessite le choix de centres de classes à partir desquels la classification sera réalisée.

PASW précise dans le tableau ci-dessous les centres des classes initiaux retenus par le logiciel pour procéder à la classification en nuées dynamiques.

Ainsi, le centre de classe du groupe 1 est celui qui a répondu 7 à V1(Fun), 6 à V2(Bad for budget), 3 à V3(Eating out), 7 à V4 (Best Buys), 2à V5(Don't care) et 7 à V6(Compare Prices). En regardant le fichier initial disponible en cliquant sur l'onglet bas « affichage des données », vous remarquerez que cet individu est l'individu 19.

La même méthode nous indique que le centre de classe du groupe 2 est l'individu 20 et le centre de classe du groupe 3 est l'individu 3.

Centres de classes initiaux

	Classe		
	1	2	3
Fun	4	2	7
Bad for Budget	6	3	2
Eating Out	3	2	6
Best Buys	7	4	4
Don't Care	2	7	1
Compare Prices	7	2	3

caseno	v1	v2	v3	v4	v5	v6	
1	6	4	7	3	2	3	
2	2	3	1	4	5	4	
3	7	2	6	4	1	3	Centre Classe n°3
4	4	6	4	5	3	6	
5	1	3	2	2	6	4	
6	6	4	6	3	3	4	
7	5	3	6	3	3	4	
8	7	3	7	4	1	4	
9	2	4	3	3	6	3	
10	3	5	3	6	4	6	
11	1	3	2	3	5	3	
12	5	4	5	4	2	4	
13	2	2	1	5	4	4	
14	4	6	4	6	4	7	
15	6	5	4	2	1	4	
16	3	5	4	6	4	7	
17	4	4	7	2	2	5	
18	3	7	2	6	4	3	
19	4	6	3	7	2	7	Centre Classe n°1
20	2	3	2	4	7	2	Centre Classe n°2
	4	4	4	4	3	4	

Le tableau ci-après, l'historique des itérations, affiche pour chaque itération, la distance de déplacement de chaque centre. La dernière itération doit révéler une distance de déplacement minimale. Dans le cas contraire, on peut augmenter le nombre d'itérations ou modifier le nombre de groupes souhaités.

Historique des itérations^a

Itération	Changements dans les centres de classes		
	1	2	3
1	2,154	2,102	2,550
2	,000	,000	,000

a. La convergence obtenue est due à l'absence ou à la quasi-absence de modifications dans les centres de classes. La modification absolue maximale des coordonnées d'un centre est ,000. L'itération en cours est 2. La distance minimale entre les centres initiaux est 7,746.

Le tableau des centres de classes finaux affiche les coordonnées des centres de chaque classe finale. Ces centres (et donc les classes associées) sont interprétés par leur coordonnées sur les différentes variables. Une coordonnée importante (positive) sur une variable signifiera que le groupe correspondant est caractérisé par des valeurs fortes (au dessus de la moyenne) pour cette variable. On établit ainsi un profil de chaque classe.

Centres de classes finaux

	Classe		
	1	2	3
Fun	4	2	6
Bad for Budget	6	3	4
Eating Out	3	2	6
Best Buys	6	4	3
Don't Care	4	6	2
Compare Prices	6	3	4

SPSS présente ensuite le nombre d'individus que contient chaque groupe.

**Nombre d'observations dans chaque
classe**

Classe	1	6,000
	2	6,000
	3	8,000
Valides		20,000
Manquantes		,000

Valider la classification retenue

Les deux méthodes utilisées pour classer les groupes nous poussent à retenir une classification en 3 groupes.

Toutefois, avant de valider ces résultats, il nous faut vérifier que chaque méthode utilisée regroupe les mêmes individus au sein d'une même classe (groupe).

Pour se faire, il faut afficher un tableau croisé avec la classification obtenue par la méthode de classification hiérarchique et par la méthode des nuées dynamiques.

Utilisez le chemin suivant, Analyse > Statistiques descriptives > Tableaux croisés

The screenshot shows the SPSS interface with the 'Analyse' menu open, navigating to 'Statistiques descriptives' > 'Tableaux croisés'. The data table below shows the following structure:

	caseno	Marrant	MauvaisBuc	J6_1	CLU5_1	CLU4_1
1	1	6		1	1	
2	2	2		2	2	
3	3	7		3	3	
4	4	4		4	4	
5	5	1		5	2	
6	6	6		3	4	1
7	7	5		3	4	1
8	8	7		1	4	3
9	9	2		6	3	5
10	10	3		4	6	4
11	11	1		5	3	5
12	12	5		2	4	1
13	13	2		4	4	2
14	14	4		4	7	4
15	15	6		1	4	1
16	16	3		4	7	4
17	17	4		2	5	1
18	18	3		4	3	6
19	19	4	6	3	7	2
20	20	2	3	2	4	7
21						2
22						
23						
24						
25						

	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	CLU6_1	CLU5_1	CLU4_1
1	1	6	4	7	3	2	3	1	1	
2	2	2	3	1	4	5	4	2	2	
3	3									3
4	4									4
5	5									2
6	6									1
7	7									1
8	8									3
9	9									2
10	10									4
11	11									2
12	12									1
13	13									2
14	14									4
15	15									1
16	16									4
17	17									1
18	18									5
19	19									4
20	20									2

Affichage des données | Affichage des variables

Document1 - Mi... | *Résultats1 [Doc... | *table_20.1_input... | Downloads - Su... | Solde pour Borie... | 08:06

Sélectionnez, par exemple, pour un affichage en ligne, la variable regroupant les individus avec la méthode non hiérarchique en sélectionnant la variable « Ward Method – CLU_3_1 » et pour un affichage en colonne la variable « Classe d'affectation – QCL_1 » qui regroupe les affectations des individus aux groupes par la méthode des nuées dynamiques.

	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	CLU6_1	CLU5_1	CLU4_1
1	1	6	4	7	3	2	3	1	1	
2	2	2	3	1	4	5	4	2	2	
3	3									3
4	4									4
5	5									2
6	6									1
7	7									1
8	8									3
9	9									2
10	10									4
11	11									2
12	12									1
13	13									2
14	14									4
15	15									1
16	16									4
17	17									1
18	18									5
19	19									4
20	20									2

Cliquez sur OK.

Résultats

Récapitulatif du traitement des observations

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Ward Method * Classe d'affectation (nuées dynamiques)	20	100,0%	0	,0%	20	100,0%

Le tableau croisé ci-dessous confirme que les deux méthodes classent les individus dans les mêmes groupes.

En effet, tous les individus classés dans le groupe 1 par la méthode des nuées dynamiques sont classés dans le groupe 3 de la méthode hiérarchique (6 individus). De même, pour les groupes 2 et 3, tous les individus ont été classés dans les mêmes groupes indépendamment de la méthode utilisée.

Tableau croisé Ward Method * Classe d'affectation (nuées dynamiques)

Effectif

		Classe d'affectation (nuées dynamiques)			Total
		1	2	3	
Ward Method	1	0	0	8	8
	2	0	6	0	6
	3	6	0	0	6
Total		6	6	8	20

Nous pouvons conclure que les individus de la base de données étudiée peuvent être regroupés en 3 groupes car les deux méthodes utilisées fournissent les mêmes résultats.

Maintenant que le nombre de groupes est connu, il faut identifier le profil de ces groupes.

Ceci peut être fait en lisant le centre des groupes finaux obtenus par la méthode des nuées dynamiques (voir résultats de la classification en nuées dynamiques) ou en utilisant un tableau croisé présentant en colonne la classe d'affectation, c'est à dire, le numéro du groupe, et en ligne, les variables V1 à V6. La lecture de ces tableaux est un peu fastidieuse.

Sélectionnez les variables à inclure dans l'analyse, ici, toutes les variables à l'exception de la variable « caseno ».

The screenshot shows the 'Analyse factorielle' dialog box in PASW Statistics. The 'Variables' list contains the following items:

- Fun [Marrant]
- Bad for Budget [Mauv...
- Eating Out [Manger...
- Best Buys [Affaires]
- Don't Care [PasInter...
- Compare Prices [C...

The 'caseno' variable is listed in the left pane but is not selected. The 'Variable de filtrage' field is empty. Buttons for 'Descriptives', 'Extraction...', 'Rotation...', 'Facteurs...', and 'Options' are visible on the right.

Cliquez sur le bouton « Statistiques » et sélectionnez les statistiques descriptives à afficher.

The screenshot shows the 'Analyse factorielle : Caractéristiques' dialog box. The 'Statistiques' section has the following checked options:

- Caractéristiques univariées
- Structure initiale

The 'Matrice de corrélation' section has the following checked options:

- Coefficients
- Seuils de signification
- Indice KMO et test de sphéricité de Bartlett

Buttons for 'Poursuivre', 'Annuler', and 'Aide' are visible at the bottom.

Cliquez sur « Poursuivre » puis sur le bouton « Extraction ».

*table_20.1_input[1].sav [Ensemble_de_données] - PASW Statistics Editeur de données

FR Français (France)

Fichier Edition Affichage Données Transformer Analyse Graphes Utilitaires Fenêtre Aide

Visible : 13 variables sur 13

	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	CLU6_1	CLU5_1	CLU4_1
1	1	6	4	7	3	2	3	1	1	
2	2	2	3	1	4	5	4	2	2	
3	3	7	2	6	4	1	3	3	3	
4	4	4	6	4	5	3	6	4	4	
5	5	1	1	1	1	1	1	1	1	
6	6	6	1	1	1	1	1	1	1	
7	7	5	1	1	1	1	1	1	1	
8	8	7	3	3	3	3	3	3	3	
9	9	2	5	5	5	5	5	5	5	
10	10	3	4	4	4	4	4	4	4	
11	11	1	1	1	1	1	1	1	1	
12	12	5	1	1	1	1	1	1	1	
13	13	2	2	2	2	2	2	2	2	
14	14	4	4	4	4	4	4	4	4	
15	15	6	1	1	1	1	1	1	1	
16	16	3	4	4	4	4	4	4	4	
17	17	4	1	1	1	1	1	1	1	
18	18	3	6	6	6	6	6	6	6	
19	19	4	1	1	1	1	1	1	1	
20	20	2	3	3	3	3	3	3	3	
21										
22										
23										
24										
25										

Affichage des données Affichage des variables

Document1 - Mi... *Résultats1 [Doc... *table_20.1_input... Downloads - Su... Solde pour Borie... 08:07

Les axes seront extraits en utilisant le critère de Kaiser, c'est à dire, dès que la valeur propre de la composante sera supérieure à 1.

Cliquez sur « Poursuivre » puis sur le bouton « Rotation ».

Sélectionnez l'option « Carte factorielle » pour obtenir la représentation graphique de l'ACP.

*table_20.1_input[1].sav [Ensemble_de_données] - PASW Statistics Editeur de données

FR Français (France)

Fichier Edition Affichage Données Transformer Analyse Graphes Utilitaires Fenêtre Aide

Visible : 13 variables sur 13

	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	CLU6_1	CLU5_1	CLU4_1
1	1	6	4	7	3	2	3	1	1	
2	2	2	3	1	4	5	4	2	2	
3	3	7	2	6	4	1	3	3	3	
4	4	4	6	4	5	3	6	4	4	
5	5	1	1	1	1	1	1	1	1	
6	6	6	1	1	1	1	1	1	1	
7	7	5	1	1	1	1	1	1	1	
8	8	7	3	3	3	3	3	3	3	
9	9	2	5	5	5	5	5	5	5	
10	10	3	4	4	4	4	4	4	4	
11	11	1	1	1	1	1	1	1	1	
12	12	5	1	1	1	1	1	1	1	
13	13	2	2	2	2	2	2	2	2	
14	14	4	4	4	4	4	4	4	4	
15	15	6	1	1	1	1	1	1	1	
16	16	3	4	4	4	4	4	4	4	
17	17	4	1	1	1	1	1	1	1	
18	18	3	6	6	6	6	6	6	6	
19	19	4	1	1	1	1	1	1	1	
20	20	2	3	3	3	3	3	3	3	
21										
22										
23										
24										
25										

Affichage des données Affichage des variables

Document1 - Mi... *Résultats1 [Doc... *table_20.1_input... Downloads - Su... Solde pour Borie... 08:07

Cliquez sur « Poursuivre » puis sur le bouton « Facteurs ».

Sélectionnez l'enregistrement des coordonnées factorielles des individus sur les axes de l'ACP en utilisant la méthode de régression.

The screenshot shows the PASW Statistics interface with a data table and a dialog box for Factorial Analysis. The data table has 25 rows and 13 columns. The dialog box is titled 'Analyse factorielle' and has a sub-dialog 'Analyse factorielle : Facteurs'. In this sub-dialog, the 'Méthode' section has 'Régression' selected with a radio button. The 'Enregistrer dans des variables' checkbox is checked. The 'Afficher la matrice des coefficients factoriels' checkbox is also checked. Buttons for 'Poursuivre', 'Annuler', and 'Aide' are visible at the bottom of the sub-dialog. The main dialog also has 'OK', 'Coller', 'Réinitialiser', 'Annuler', and 'Aide' buttons.

	caseno	Marrant	MauvaisBudget	MangerDehors	Affaires	PasInteresse	ComparePrix	CLU6_1	CLU5_1	CLU4_1
1	1	6	4	7	3	2	3	1	1	
2	2	2	3	1	4	5	4	2	2	
3	3	7	2	6	4	1	3	3	3	
4	4	4	6	4	5	3	6	4	4	
5	5	1						5	2	
6	6	6						1	1	
7	7	5						1	1	
8	8	7						3	3	
9	9	2						5	2	
10	10	3						4	4	
11	11	1						5	2	
12	12	5						1	1	
13	13	2						2	2	
14	14	4						4	4	
15	15	6						1	1	
16	16	3						4	4	
17	17	4						1	1	
18	18	3						6	5	
19	19	4	6	3	7	2	7	4	4	
20	20	2	3	2	4	7	2	5	2	
21										
22										
23										
24										
25										

Cliquez sur « Poursuivre » puis sur OK pour lancer l'analyse.

Résultats

Le premier tableau présente une synthèse des statistiques descriptives, moyenne et écart-type pour chacune des variables observées.

Statistiques descriptives

	Moyenne	Ecart-type	n analyse
Fun	3,85	1,899	20
Bad for Budget	4,10	1,410	20
Eating Out	3,95	2,012	20
Best Buys	4,10	1,518	20
Don't Care	3,45	1,761	20
Compare Prices	4,35	1,496	20

SPSS affiche ensuite la matrice des corrélations ainsi que les coefficients de signification. Une corrélation mesure l'intensité d'une relation entre deux variables. Une corrélation peut prendre une valeur située entre -1 et 1. Une valeur proche de 1 indique que les variables sont fortement liées et évoluent dans le même sens. Une corrélation de 0 indique une absence de lien.

Si plusieurs variables sont corrélées (> 0.5), la factorisation est possible. Si non, la factorisation n'a pas de sens et n'est donc pas conseillée.

Matrice de corrélation

		Fun	Bad for Budget	Eating Out	Best Buys
Corrélation	Fun	1,000	,026	,838	-,104
	Bad for Budget	,026	1,000	-,054	,536
	Eating Out	,838	-,054	1,000	-,308
	Best Buys	-,104	,536	-,308	1,000
	Don't Care	-,875	-,104	-,721	,022
	Compare Prices	,019	,556	,041	,586
Signification (unilatérale)	Fun		,457	,000	,331
	Bad for Budget	,457		,411	,007
	Eating Out	,000	,411		,093
	Best Buys	,331	,007	,093	
	Don't Care	,000	,332	,000	,464
	Compare Prices	,468	,005	,432	,003

Matrice de corrélation

		Don't Care	Compare Prices
Corrélation	Fun	-,875	,019
	Bad for Budget	-,104	,556
	Eating Out	-,721	,041
	Best Buys	,022	,586
	Don't Care	1,000	-,203
	Compare Prices	-,203	1,000
Signification (unilatérale)	Fun	,000	,468
	Bad for Budget	,332	,005
	Eating Out	,000	,432
	Best Buys	,464	,003
	Don't Care		,196
	Compare Prices	,196	

Dans notre exemple, plusieurs variables sont corréllées entre elles :

Dans un deuxième temps, il faut observer l'indice de KMO (Kaiser-Meyer-Olkin) qui doit tendre vers 1. si ce n'est pas le cas, la factorisation n'est pas conseillée. Pour juger de l'indice de KMO, on peut utiliser l'échelle suivante :

- 0,50 et moins est misérable
- entre 0,60 et 0,70, c'est médiocre
- entre 0,70 et 0,80 c'est moyen
- entre 0,80 et 0,90 c'est méritoire
- et plus 0,9 c'est merveilleux.

Enfin, on utilise le test de sphéricité de Bartlett. : si la signification (Sig.) tend vers 0.000, c'est très significatif, inférieur à 0.05 significatif, entre 0.05 et 0.10 acceptable et au dessus de 0.10, on rejette.

Indice KMO et test de Bartlett

Mesure de précision de l'échantillonnage de Kaiser-Meyer-Olkin.		,554
Test de sphéricité de Bartlett	Khi-deux approximé	66,712
	ddl	15
	Signification de Bartlett	,000

Si l'ACP satisfait à au moins deux de ces trois conditions, on peut continuer. C'est le cas pour notre analyse. Pour choisir les variables à éliminer, on observe leur qualité de représentation : plus la valeur associée à la ligne « Extraction » est faible, moins la variable explique la variance.

Qualité de représentation

	Initial	Extraction
Fun	1,000	,924
Bad for Budget	1,000	,679
Eating Out	1,000	,858
Best Buys	1,000	,746
Don't Care	1,000	,873
Compare Prices	1,000	,735

Méthode d'extraction : Analyse en composantes principales.

Combien d'axes retenir

Trois règles sont applicables :

- 1ere règle : la règle de Kaiser qui veut qu'on ne retienne que les facteurs aux valeurs propres supérieures à 1.
- 2eme règle : on choisit le nombre d'axe en fonction de la restitution minimale d'information que l'on souhaite. Par exemple, on veut que le modèle restitue au moins 80% de l'information.

Pour ces deux premières règles, on examine le tableau « Variance Totale Expliquée ».

Dans notre cas, 2 composantes (axes) ont une valeur propre supérieure à 1 pour une variance totale expliquée de 80,238%.

Variance totale expliquée

Composante	Valeurs propres initiales			Extraction
	Total	% de la variance	% cumulés	Sommes des carrés des facteurs retenus
				Total
1	2,664	44,407	44,407	2,664
2	2,150	35,831	80,238	2,150
3	,473	7,878	88,116	
4	,433	7,216	95,332	
5	,217	3,608	98,940	
6	,064	1,060	100,000	

Méthode d'extraction : Analyse en composantes principales.

Variance totale expliquée

Composante	Extraction Sommes des carrés des facteurs retenus	
	% de la variance	% cumulés
1	44,407	44,407
2	35,831	80,238
3		
4		
5		
6		

Variance totale expliquée

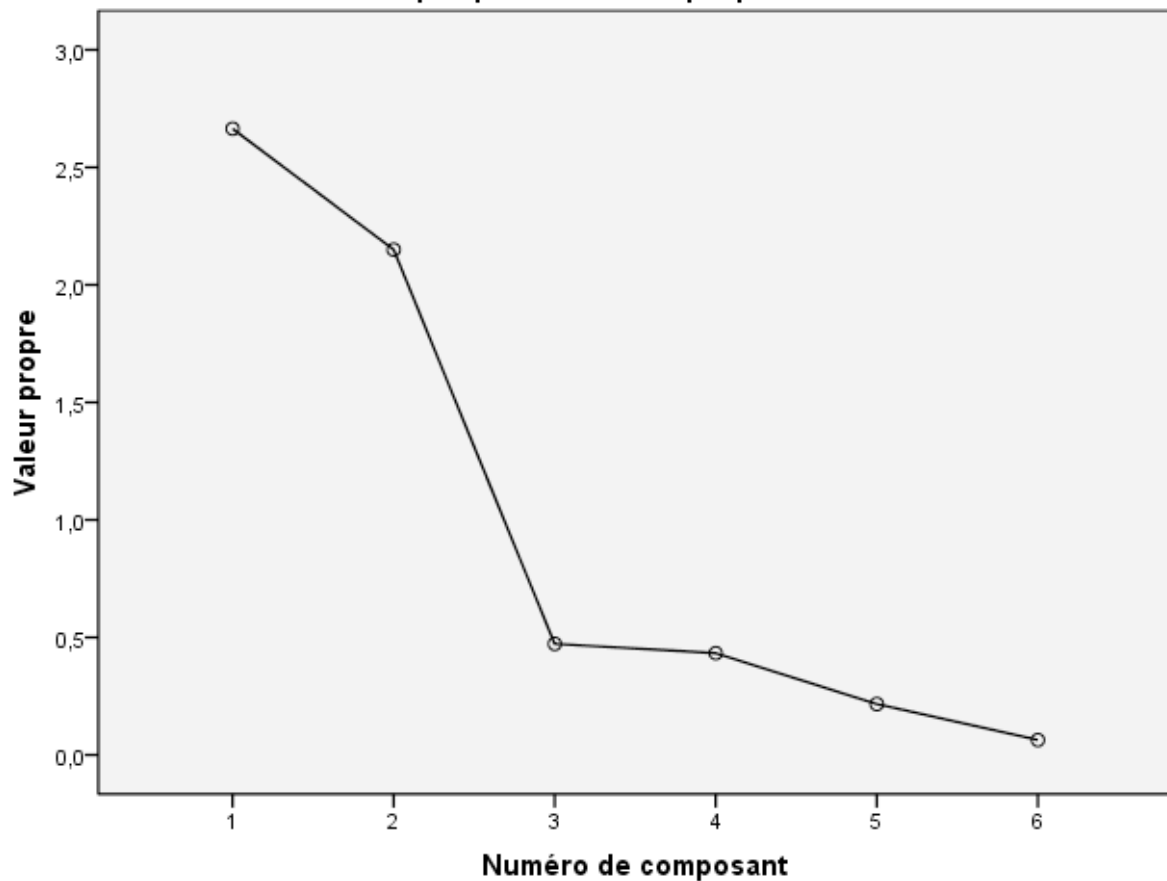
Composante	Extraction Sommes des carrés des facteurs retenus	
	% de la variance	% cumulés
1	44,407	44,407
2	35,831	80,238
3		
4		
5		
6		

Méthode d'extraction : Analyse en composantes principales.

3eme méthode : le « Scree-test » ou test du coude. On observe le graphique des valeurs propres et on ne retient que les valeurs qui se trouvent à gauche du point d'inflexion. Graphiquement, on part des composants qui apportent le moins d'information (qui se trouvent à droite), on relie par une droite les points presque alignés et on ne retient que les axes qui sont au dessus de cette ligne.

Ici, le point d'inflexion est entre les composantes 2 et 3. On retient donc 2 axes.

Graphique de valeurs propres



Interprétation des résultats

Ici, l'objectif consiste à donner un sens à un axe grâce à une recherche lexicale (ou recherche de mots) à partir des coordonnées des variables et des individus. Ce sont les éléments extrêmes qui concourent à l'élaboration des axes.

Matrice des composantes^a

	Composante	
	1	2
Fun	,960	,042
Bad for Budget	-,008	,824
Eating Out	,923	-,082
Best Buys	-,223	,834
Don't Care	-,915	-,190
Compare Prices	,064	,855

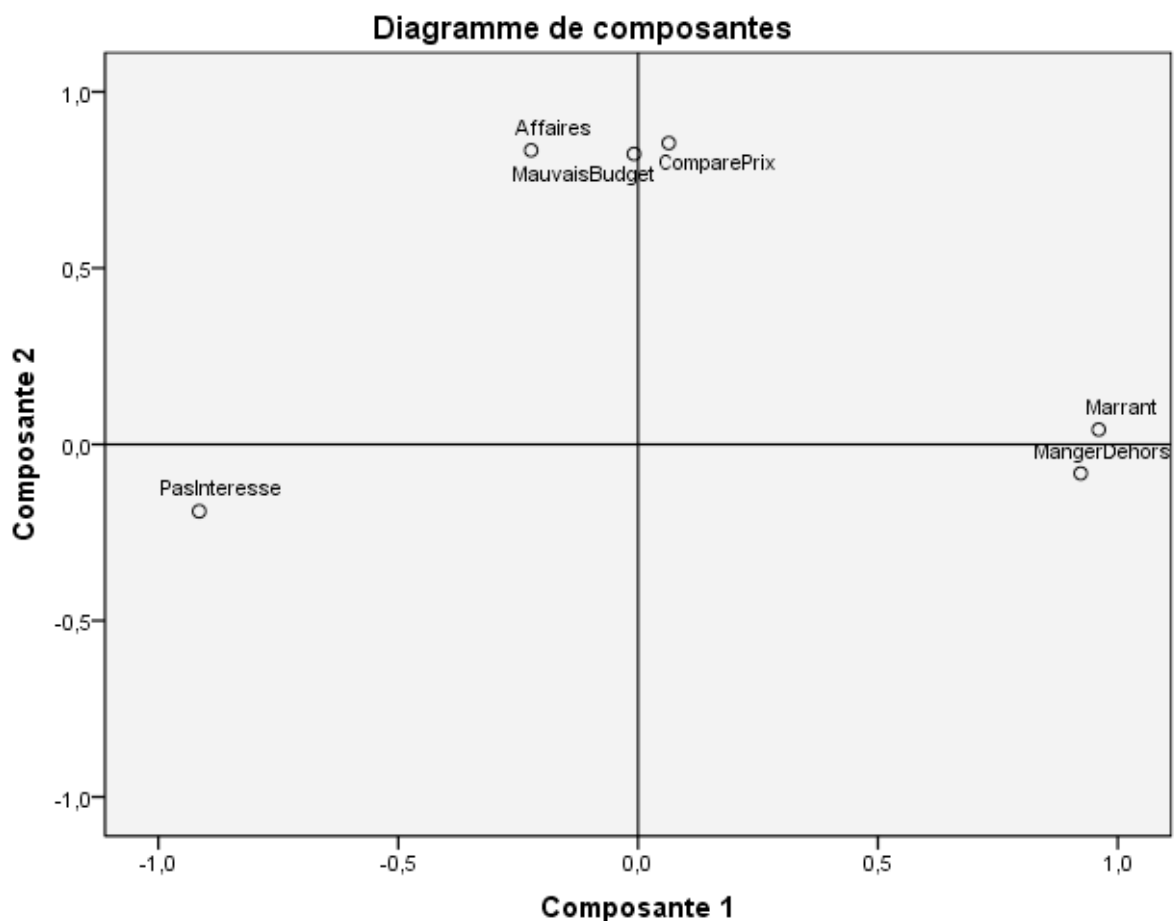
Méthode d'extraction : Analyse en composantes principales.

a. 2 composantes extraites.

Dans notre exemple :

- Les variables « Fun » et « Eating Out » sont celles qui concourent le plus à la formation de l'axe 1 pour sa portion positive et la variable « Don't care » pour sa partie négative.
- Les variables « Best Buys », « Bad for Budget » et « Compare Prices » contribuent à la formation de l'axe 2.

Le diagramme des composantes donne une représentation graphique des axes et des variables qui contribuent à leurs formations.



La partie positive de l'axe 1 représente l'intérêt pour le shopping.

La partie négative de l'axe 1 représente le désintérêt à l'égard du shopping.

La partie positive de l'axe 2 représente l'intérêt pour les affaires dans le cadre du shopping.

Matrice des coefficients des coordonnées des composantes

	Composante	
	1	2
Fun	,360	,020
Bad for Budget	-,003	,383
Eating Out	,346	-,038
Best Buys	-,084	,388
Don't Care	-,343	-,088
Compare Prices	,024	,398

Méthode d'extraction : Analyse en composantes principales.
Scores composante.

La matrice des co-variances indique que les deux composantes sont orthogonales.

Matrice de covariance des composantes

Composante	1	2
1	1,000	,000
2	,000	1,000

Méthode d'extraction : Analyse en composantes principales.
Scores composante.

L'ACP fournit aussi les coordonnées de chaque individu sur les deux axes retenus, FAC1_1 pour l'axe X et FAC1_2 pour l'axe Y (Voir Affichage des données).

Puis sélectionnez les coordonnées des individus, REGR FACTOR SCORE FAC1_1 pour l'axe X et REGR FACTOR SCORE FAC1_2 pour l'axe Y. Choisissez d'afficher les observations en utilisant le groupe auquel l'individu a été affecté en utilisant, par exemple, la méthode hiérarchique, en sélectionnant la variable « Ward Method CLU3_1 ».

The screenshot shows the SPSS 'Diagramme de dispersion simple' (Simple Scatter Plot) dialog box. The configuration is as follows:

- Axe des Y :** REGR factor score 2 for analysis 1 [FAC...]
- Axe des X :** REGR factor score 1 for analysis 1 [FAC...]
- Définir les marques par :** (Empty field)
- Etiqueter les observations par :** Ward Method [CLU3_1]
- Panel par :**
 - Lignes :** (Empty field)
 - Colonnes :** (Empty field)
- Modèle :** Utiliser les spécifications du diagramme de :

The background data table shows the following columns and rows:

	caseno	Marra
1	1	
2	2	
3	3	
4	4	
5	5	
6	6	
7	7	
8	8	
9	9	
10	10	
11	11	
12	12	
13	13	
14	14	
15	15	
16	16	
17	17	
18	18	
19	19	
20	20	
21		
22		
23		
24		
25		

Cliquez sur le bouton « Options » puis cochez la case « Afficher le diagramme avec les étiquettes d'observations ».

Diagramme de dispersion simple

Axe des Y : REGR factor score_2 for analysis_1/FAC2

Options

Valeurs manquantes

- Exclure toute observation incomplète
- Exclure les observations variable par variable
- Afficher les groupes définis par des valeurs manquantes
- Afficher le diagramme avec les étiquettes d'observations
- Afficher les barres d'erreur

Les barres d'erreur représentent

- Intervalles de confiance
- Niveau (%):
- Erreur standard
- Multiplicateur:
- Ecart type
- Multiplicateur:

Poursuivre Annuler Aide

OK Coller Réinitialiser Annuler Aide

Cliquez sur « Poursuivre » puis « OK ».

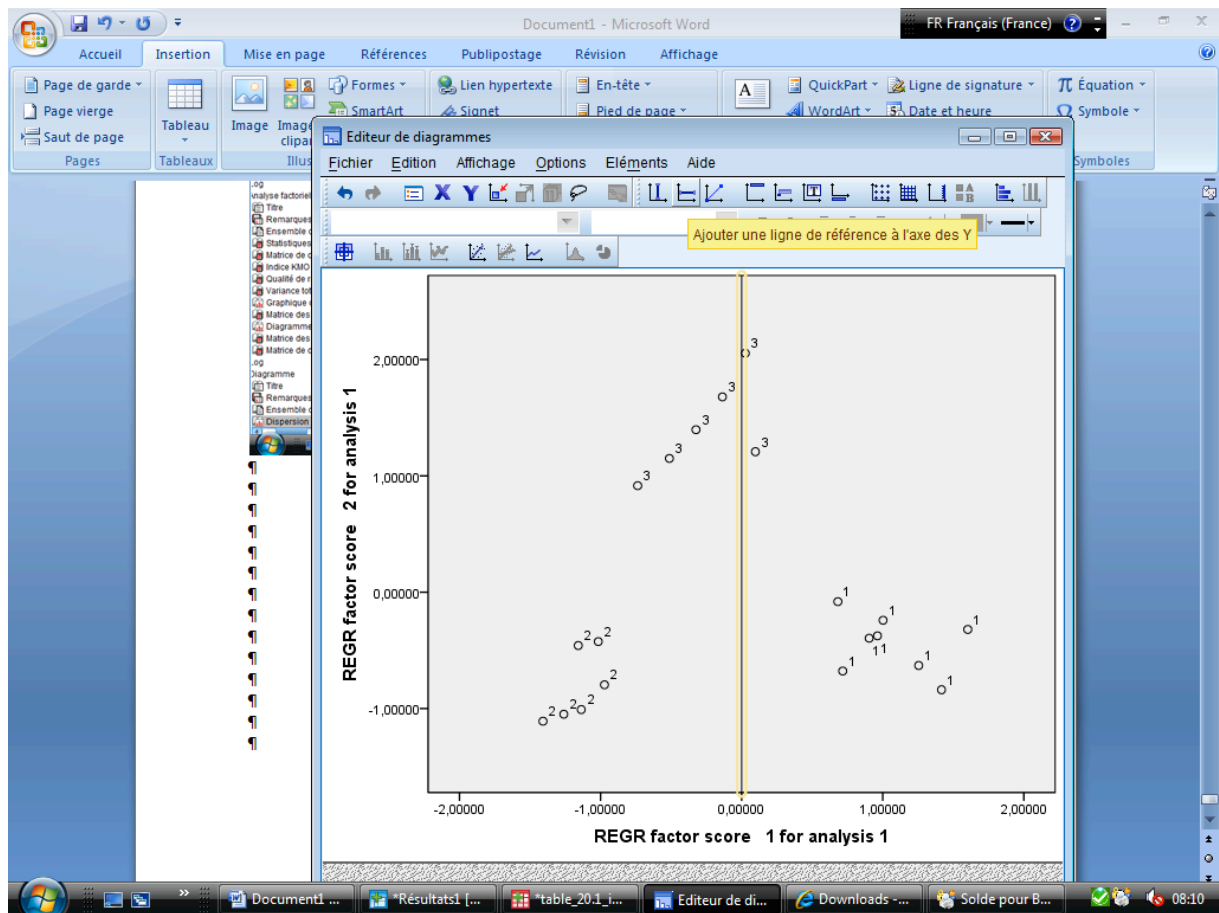
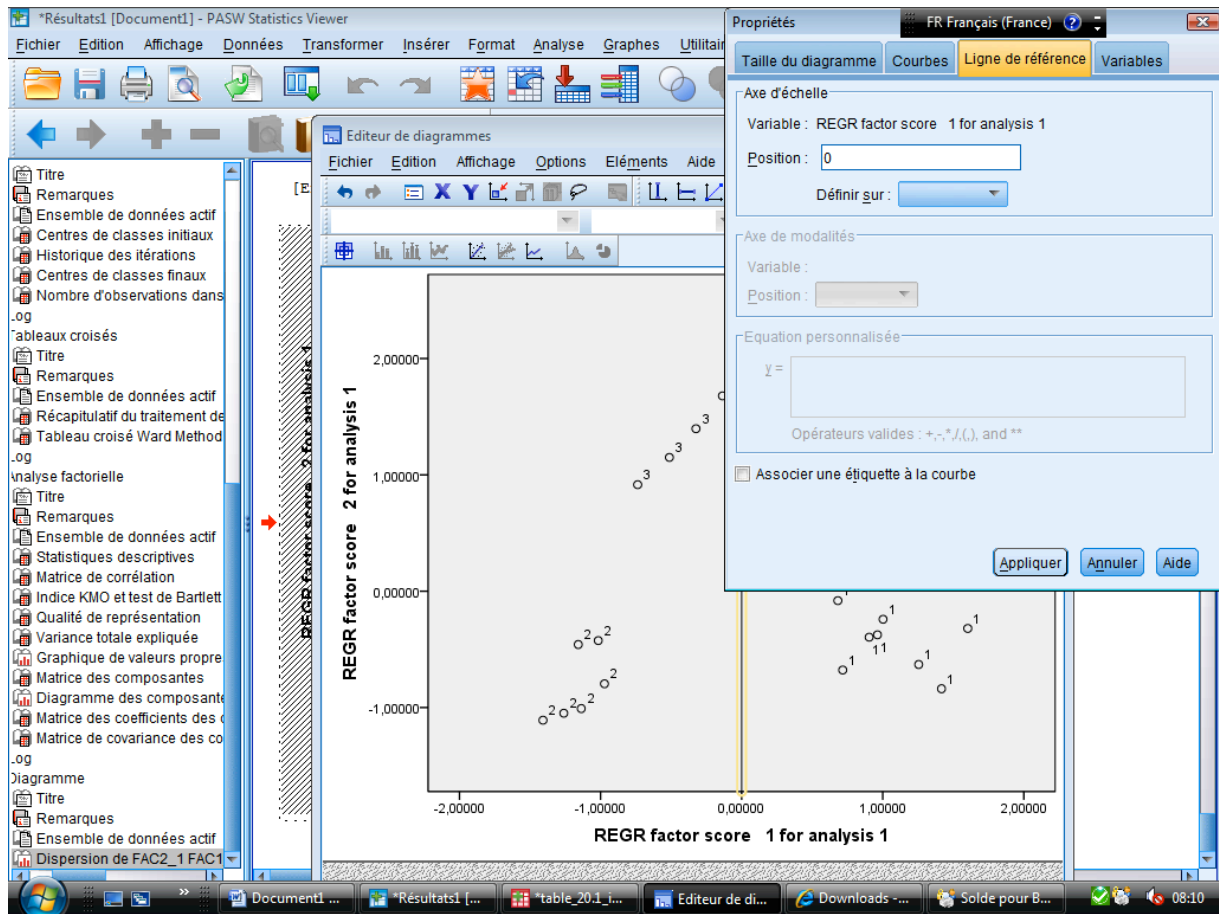
Lorsque le graphique s'affiche, double-cliquez sur celui-ci afin de faire apparaître des droites coupant l'axe des Y en 0 et l'axe des X en 0.

Editeur de diagramme

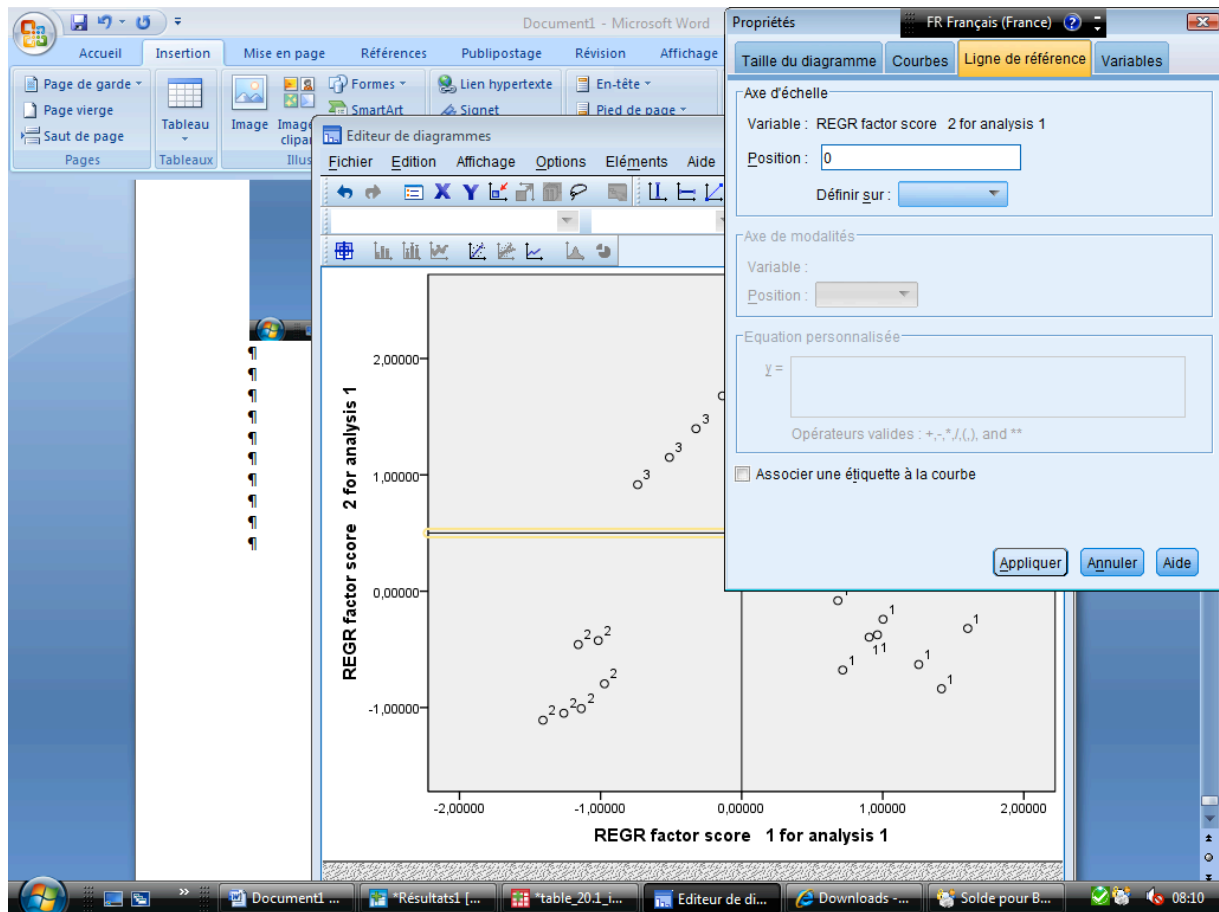
REGR factor score 2 for analysis 1

REGR factor score 1 for analysis 1

Précisez la position de la ligne verticale, à savoir 0.



De même, précisez la position de la ligne horizontale, à savoir 0.



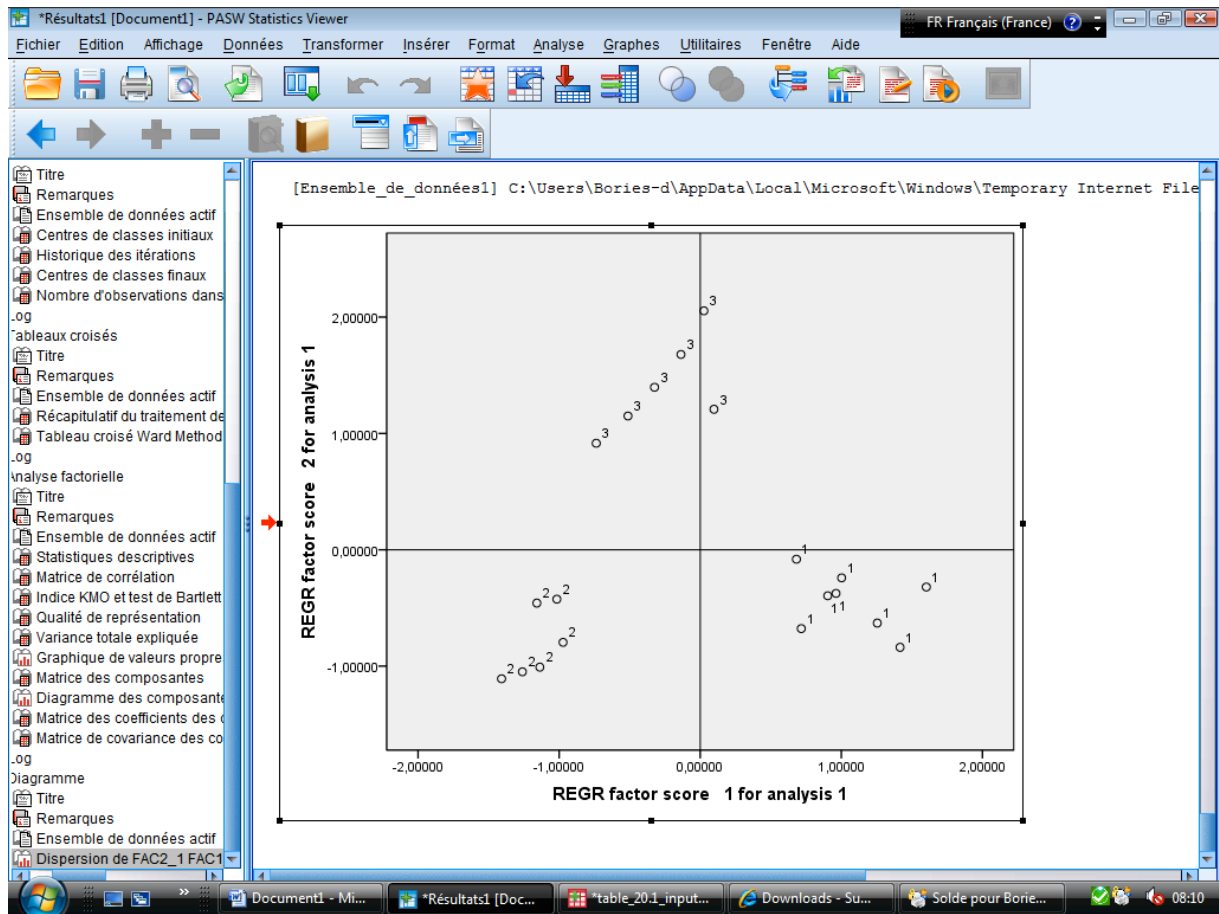
Les 3 groupes identifiés en utilisant la méthode de classification hiérarchique sont ainsi représentés sur le graphique ci-dessous.

Afin d'identifier le profil de chaque groupe, il suffit d'utiliser les noms retenus pour chaque axe.

Ainsi, les individus du groupe 1 sont placés sur la portion positive de l'axe X qui représente l'intérêt pour le shopping.

Les individus du groupe 2 sont proches de la portion positive de l'axe X qui représente l'intérêt pour les affaires dans le cadre du shopping.

Les individus du groupe 3 sont proche de la portion négative de l'axe X qui représente le désintéressement à l'égard du shopping.



L'Analyse en Composantes Principales a ainsi permis de présenter de manière graphique les regroupements des individus, facilitant ainsi la détermination du profil des groupes.