

**L** LABORATOIRE

**O** ORLEANAIS

**G** DE GESTION

LABORATOIRE ORLEANAIS DE GESTION

I.A.E

Faculté de Droit d'Economie et de Gestion  
rue de Blois - B.P. 6739  
45067 Orléans Cedex 2

Tél. : 02 38 41 70 28

Fax : 02 38 41 73 60

E.Mail : [iae@univ-orleans.fr](mailto:iae@univ-orleans.fr)



INSTITUT  
D'ADMINISTRATION  
DES ENTREPRISES

**Document de  
recherche**

**N° 1997 - 4**

*Etude comparative  
de la classification  
ascendante hié-  
rarchique et de la  
classification floue  
pour identifier cinq  
familles de voitures.*

**Geneviève PAVIOT**

Etude comparative de la classification  
ascendante hiérarchique et de la classification  
floue pour identifier cinq familles de voitures

Geneviève PAVIOT  
Maître de Conférences en Gestion

## RESUME

Le marché automobile français est constitué de cinq catégories: les citadines, les compactes, les grandes routières et les monospaces, breaks. Notre but est de retrouver ces cinq classes à partir des caractéristiques techniques de plusieurs véhicules essence et diesel. Nous décrivons d'abord les données par des méthodes classiques (analyse en composantes principales, classification ascendante hiérarchique) puis à l'aide d'une typologie floue. Nous montrons que celle-ci permet d'identifier les voitures qui appartiennent à plusieurs classes et de mieux comprendre les raisons de ces appartenances multiples.

*Mots clés: typologie floue, voitures, analyse en composantes principales, classification ascendante hiérarchique, segmentation.*

## Comparative study of hierarchic classification and fuzzy clustering for identifying five types of cars

### SUMMARY

French car market include five types. The purpose of this paper is to find these five classes from technicals characteristics. Firstly, data are examined with classics analysis (principal components analysis, hierarchic classification) and with a fuzzy clustering. A fuzzy partition provides more insights on segments than its hard counterpart.

*Keywords: fuzzy clustering, cars, principal components analysis, hierarchical clustering methods, segmentation.*

## Introduction

Les **méthodes de classification « dure »** sont des méthodes de classification où chaque objet ne correspond qu'à un sous-ensemble. Ainsi chaque valeur d'appartenance ne peut prendre que la valeur 1 ou 0. Ces méthodes prévalent dans la détermination des segments de marché. Si pendant longtemps les méthodes de classification dure ont été principalement étudiées en marketing, les méthodes de **classification à recouvrement** (overlapping clustering ou recouvrement entre les groupes), afin de tenir compte de l'utilisation par le consommateur de plusieurs marques, ont suscité l'intérêt de certains chercheurs.

Dans la classification à recouvrement, des classes peuvent contenir les mêmes objets. Dans les études de segmentation par bénéfices, une personne peut désirer plusieurs bénéfices différents pour un produit particulier. Pour un dentifrice, un individu peut vouloir une haleine fraîche et la prévention contre les caries (Arabie, Caroll, DeSarbo et Wind, 1981). Une marque peut aussi être en concurrence avec plus d'un groupe de produits. Le chewing-gum peut être un substitut au dentifrice ou au bonbon. Le modèle ADCLUS (ADditive CLUStering) est le plus connu des modèles fournissant une partition à recouvrement (Shepard et Arabie, 1979). Une classification à recouvrement indique seulement si des objets appartiennent à plusieurs groupes sans mesurer leur degré d'appartenance aux classes. En revanche, la typologie floue estime, pour chaque objet, les valeurs d'appartenance dans un intervalle de 0 à 1. Le tableau 1 présente un exemple de ces trois méthodes de classification (dure, à recouvrement et floue).

**Tableau 1 Un exemple de trois méthodes de classification**

Objets	Classification dure			Classification à recouvrement			Classification floue		
	A	B	C	A	B	C	A	B	C
<b>1</b>	0	1	0	0	1	0	0,02	0,92	0,06
<b>2</b>	1	0	0	<b>1</b>	0	<b>1</b>	<b>0,54</b>	0,01	<b>0,45</b>
<b>3</b>	0	1	0	<b>1</b>	<b>1</b>	0	<b>0,35</b>	<b>0,62</b>	0,03
<b>4</b>	1	0	0	1	0	0	0,92	0,03	0,05
<b>5</b>	0	1	0	0	1	0	0,08	0,88	0,04
<b>6</b>	0	0	1	0	<b>1</b>	<b>1</b>	0,02	<b>0,42</b>	<b>0,,56</b>

Chaque classification fournit les valeurs d'appartenance des objets aux classes. Pour une classification dure les objets prennent seulement les valeurs 0 ou 1. Pour une classification à recouvrement plusieurs objets (n° 2, 3 et 6) sont en relation avec deux classes. Pour une typologie floue la force de l'appartenance de ces objets est mesurée (le n° 2 appartient plus à la classe A, le n° 3 à la classe B et le 6 à la classe C).

Nous présenterons tout d'abord la méthodologie suivie. Nous examinerons ensuite les résultats obtenus à l'aide d'une typologie classique, puis nous les comparerons à ceux d'une classification floue.

## **I METHODOLOGIE**

Nous étudierons des modèles de voiture de l'année 1997 selon leurs caractéristiques techniques: puissance, vitesse, consommation, cylindrée, longueur, largeur, hauteur, coffre. Le marché automobile est composé de cinq familles de voitures: les citadines, les compactes, les routières, les grandes routières et les breaks monospaces. Nous avons retenu 8 modèles par catégorie soit 41 voitures essence (une voiture de plus dans la dernière catégorie) et 40 voitures diesel<sup>1</sup>.

La problématique est la suivante: retrouve-t-on ces cinq familles de voiture à partir de leurs caractéristiques techniques?

Pour répondre à cette question nous procéderons à une analyse en composantes principales puis à une classification ascendante hiérarchique (critère de Ward) à partir des données centrées et réduites. Nous verrons si cette méthode de typologie classique répartit bien les voitures en 5 groupes. Si certaines voitures ne sont pas dans leur catégorie d'origine, comment ne pas retrouver cette ambiguïté dans l'esprit du consommateur? Or l'appartenance à une classe semble conditionner un prix particulier. Si des voitures ne sont pas dans leur groupe respectif, cette méthode de classification n'apportera pas plus d'informations. D'où l'intérêt d'avoir recours à la classification floue pour savoir quelles sont les voitures qui peuvent appartenir à plusieurs classes, et quel est leur degré d'appartenance aux classes.

---

<sup>1</sup> « L'auto-journal guide de l'acheteur 1997 ».

La typologie floue est issue de la notion de sous-ensemble flou introduite par Zadeh en 1965. L'idée retenue ici est l'appartenance partielle à une classe en remplaçant l'appartenance et la non-appartenance d'un objet à un sous-ensemble par une appartenance graduelle indiquant la proximité d'un objet à un groupe. L'introduction du flou en mathématique a permis de sortir de la logique formelle, la logique booléenne grâce à la notion d'appartenance pondérée (Kaufmann, 1977). Des applications de la logique floue dans de nombreux domaines se sont multipliées: économie, systèmes experts, reconnaissance de formes, classification.

Ruspini (1969, 1970) est le premier à avoir introduit le concept de sous-ensemble flou en classification. Dunn (1974) a étendu **l'algorithme c-moyens** (ou k-moyens) à la méthode de typologie floue, on parle d'algorithme c-moyens flous (**FCMA**).

Les principales étapes de l'algorithme c-moyens flous sont: la fixation arbitraire d'une matrice d'appartenance, le calcul des centroïdes des classes, le réajustement de la matrice d'appartenance suivant la position des centroïdes, l'algorithme doit converger vers un minimum local (Grabisch, 1994).

Une partition est représentée par une matrice  $U = | \mathbf{u}_{i,k} |$ .

$\mathbf{u}_{i,k}$  est la valeur d'appartenance de l'objet  $k$  dans la classe  $i$  (dans le cas d'une partition floue,  $\mathbf{u}_{i,k}$  appartient à l'intervalle  $[0, 1]$ )

$1 \leq i \leq c, 1 \leq k \leq n$        $c =$  nombres de classes       $n =$  nombre d'objets.

La technique de base est la minimisation de la variance intragroupes (Dunn, 1974):

$$\min \sum_{k=1}^n \sum_{i=1}^c (\mathbf{u}_{i,k})^m \|x_k - v_i\|^2.$$

$x_k =$  vecteur représentant l'objet  $k$

$v_i =$  centre de classe autour duquel les objets sont concentrés

$\| \cdot \| =$  norme d'un vecteur.

Le poids de l'exposant (ou exposant flou)  $m$  s'interprète comme un paramètre de distorsion. Il est supérieur à 1 et sa valeur maximale est de 2 (Dunn, 1974). Le flou de la partition augmente avec  $m$ . La valeur la plus utilisée de  $m$  dans les algorithmes semble être 2. De nombreuses mesures

ont été proposées pour valider les groupes obtenus. La première est celle du coefficient de partition F de Bezdek (1974) qui a été normalisé (NFI normalized partition coefficient). Le coefficient F a tendance à indiquer que la partition correcte est une partition en deux classes (Zimmermann, 1991). Pal et Bezdek (1995) ont testé la validité de plusieurs mesures. Selon eux, l'index de Xie-Beni est le plus sûr pour choisir le nombre de groupes.

$$v_{XB}(U, V; X) = \frac{\sum_{i=1}^c \sum_{k=1}^n u_{ik}^2 \|x_k - v_i\|^2}{n \left( \min_{i \neq j} \left\{ \|v_i - v_j\|^2 \right\} \right)} \quad (\text{Xie et Beni, 1991})$$

L'algorithme c-moyens flous de Dunn a été généralisé (Bezdek, 1981; Bezdek et Cie, 1981; Bezdek et Pal, 1992; Al-Sultan et Selim, 1993; Kamel et Selim, 1994; Hathaway et Bezdek, 1994). De nouveaux algorithmes ont été développés pour optimiser les c-moyens flous. Il semble que ceux-ci obtiennent de moins bons résultats que l'algorithme c-moyens flous de Bezdek (Pal, Bezdek et Hathaway, 1996).

Nous utiliserons un algorithme de typologie floue de type c-moyens après la standardisation des données puisque les unités de mesure sont différentes (CV= cheval fiscal, ch= cheval vapeur, km/h, l/100km, cm<sup>3</sup>, m, dm<sup>3</sup>).

## II LE CLASSEMENT DES VOITURES SELON LA METHODE CLASSIQUE

### 1 La dispersion des données

Nous avons tout d'abord cherché à savoir si le prix des voitures dépendait du type de famille auquel elles appartiennent. Cette relation est non seulement vérifiée pour les voitures essence mais aussi pour les voitures diesel (corrélations de rang de Spearman de 0,8572 et 0,8684 significatifs à 1%). Ceci est intéressant pour la suite car si une voiture est mal classée comment justifier son prix?

Nous avons établi les boxplots (boîtes à moustaches) de chaque variable pour chacune des 5 familles de voitures essence et diesel afin de visualiser la dispersion des données. Les moustaches vont du premier quartile au minimum et du troisième quartile au maximum. Elles ont une longueur qui

ne doit pas dépasser une fois et demie la distance interquartiles. Les points extrêmes inférieurs ou supérieurs sont isolés sur le graphique par un cercle si le dépassement est d'une fois et demie et par un astérisque si le dépassement est de trois fois. Le tableau 2 reprend les résultats des boxplots pour les voitures essence et diesel.

**Tableau 2 Résultats des boxplots par groupes d'observations**

	Prix	Pfiscale	Puissan.	Vitesse	Conso	Cyl.
<b>1</b> Citadines			<i>*Seat ibiza &gt;</i>	oNissan micra <  <i>*Seat ibiza &gt;</i>	oFiat cinquec <	
<b>2</b> Compactes	<i>oVW Golf &gt;</i>	*Honda Civic <	*Rover 214 > *Fiat brava <			<i>*Rover 220 &gt;</i>
<b>3</b> Routières						
<b>4</b> Grandes Routières		<i>*Renault safrane &lt;</i> <i>*Audi 6 &gt;</i>			<i>oRenault safrane &gt;</i>	<i>oRenault safrane &lt;</i> <i>*Mercedes E290 &gt;</i>
<b>5</b> Breaks et Mspaces	oLancia kappa > oMercedes E230Break >		*Lancia kappa >  <i>oFiat marea &gt;</i>	oLancia kappa >	<i>oAudi A4 &lt;</i>	

Légende: les voitures diesel sont en italique souligné;

Exemples: « \*Rover 214 > » pour la puissance veut dire que cette voiture essence a une puissance qui est trois fois supérieure (>) à la distance interquartile; « *oRenault safrane <* » pour la cylindrée indique que cette voiture diesel a une cylindrée qui est 1,5 fois inférieure (<) à la distance interquartile.

## Suite du tableau 2

	Long	Larg	Haut	Coffre
<b>1</b> Citadines	oFiat cinquec <	oFiat cinquec <	oFord fiesta<	
<b>2</b> Compactes		*Fiat brava >		oRenault mégane >
<b>3</b> Routières	oVW passat >			
<b>4</b> Grandes Routières				
<b>5</b> Breaks et Mspaces		*Chrysler voyager > *Hyundai lantra < *Vw golf Break <  <u>oChrysler voyager TD &gt;</u>		<u>oRenault espace &gt;</u>

Nous séparerons dans la suite des développements les résultats des voitures essence et diesel car des différences peuvent apparaître. Nous n'entrerons pas dans le détail de tous les résultats des analyses en composantes principales car le but est surtout de présenter les typologies qui en découlent.

## 2 Les résultats pour les voitures essence

### 2.1 L'analyse en composantes principales

Une analyse en composantes principales sur les 9 variables techniques a été réalisée en excluant le prix qui n'est pas une caractéristique technique. Seul le coffre n'a pas une communalité supérieure à 0,5. Une autre analyse, sans cette variable, donne un KMO de 0,80597 (test d'adéquation de l'échantillonnage de Kaiser, Meyer et Olkin) et des communalités supérieures à 0,5. Nous avons retenu deux axes (critère de Kaiser) qui

rendent compte de 88,6% de l'information. La projection simultanée des variables et des voitures sur les deux axes principaux est illustrée par la figure 1 (nous avons multiplié par 3 les variables dans un souci de lisibilité).

Il apparaît quatre groupes: à gauche les citadines, au milieu principalement les compactes et les routières, à droite les grandes routières et en haut les monospaces. Plus les voitures se situent du côté positif de l'axe 1 plus il s'agit de voitures puissantes, larges, longues, qui vont vite et qui consomment. L'axe 2 quant à lui exprime la hauteur des véhicules.

## **2.2 La typologie**

Une classification ascendante hiérarchique des données issues de l'analyse en composantes principales a été effectuée. Nous avons opté pour une typologie en quatre classes (figure 2). D'autres méthodes de classification (distance moyennes entre classes, nuées dynamiques) donnent les mêmes résultats.

Le premier groupe comprend les compactes et les routières, le deuxième les citadines, le troisième les grandes routières et le dernier les monospaces. Il s'avère que cette classification ne sépare pas deux catégories distinctes dans la réalité: les compactes et les routières. Nous tenterons avec la typologie floue d'apporter des précisions à ce sujet. Mais avant, nous examinerons les voitures diesel pour voir si nous pouvons faire le même constat.

## **3 Les résultats pour les voitures diesel**

### **3.1 L'analyse en composantes principales**

Le KMO est de 0,78238 et toutes les communalités sont supérieures à 0,5. Deux axes rendent compte de 75,7 % de l'information. La figure 3 illustre la projection simultanée des variables et des voitures diesel sur les deux axes. Il semble que, comme pour les voitures essence, quatre groupes apparaissent, des citadines aux monospaces.



Figure 2 Le dendrogramme des voitures essence

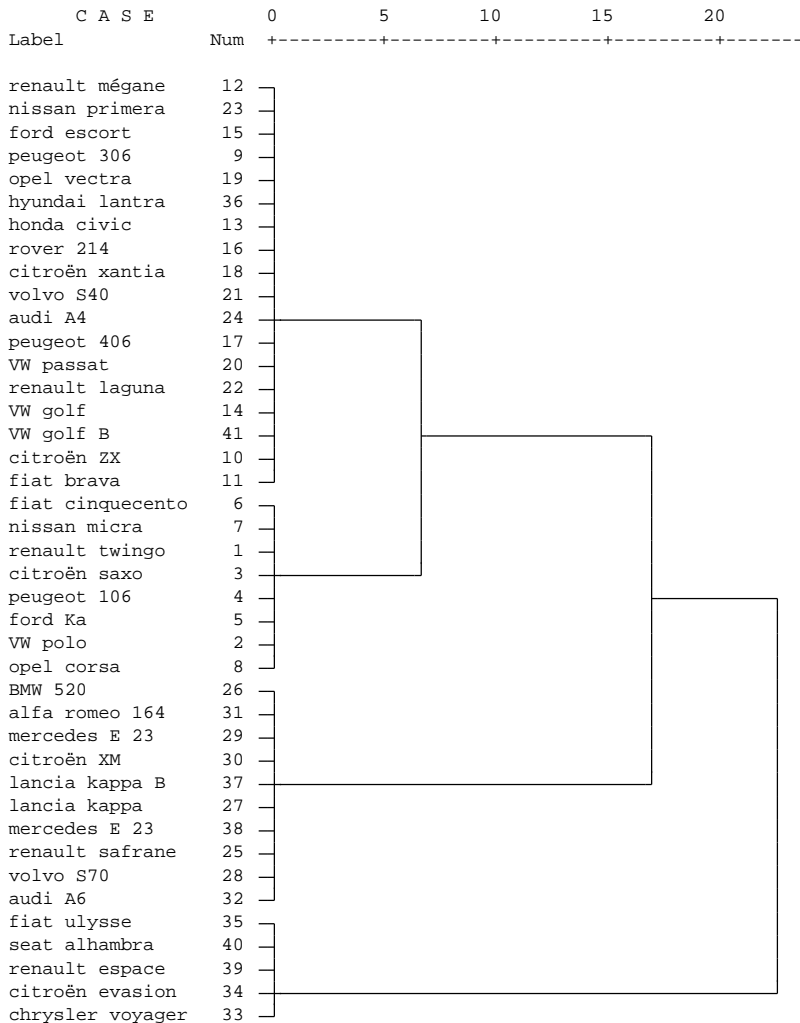
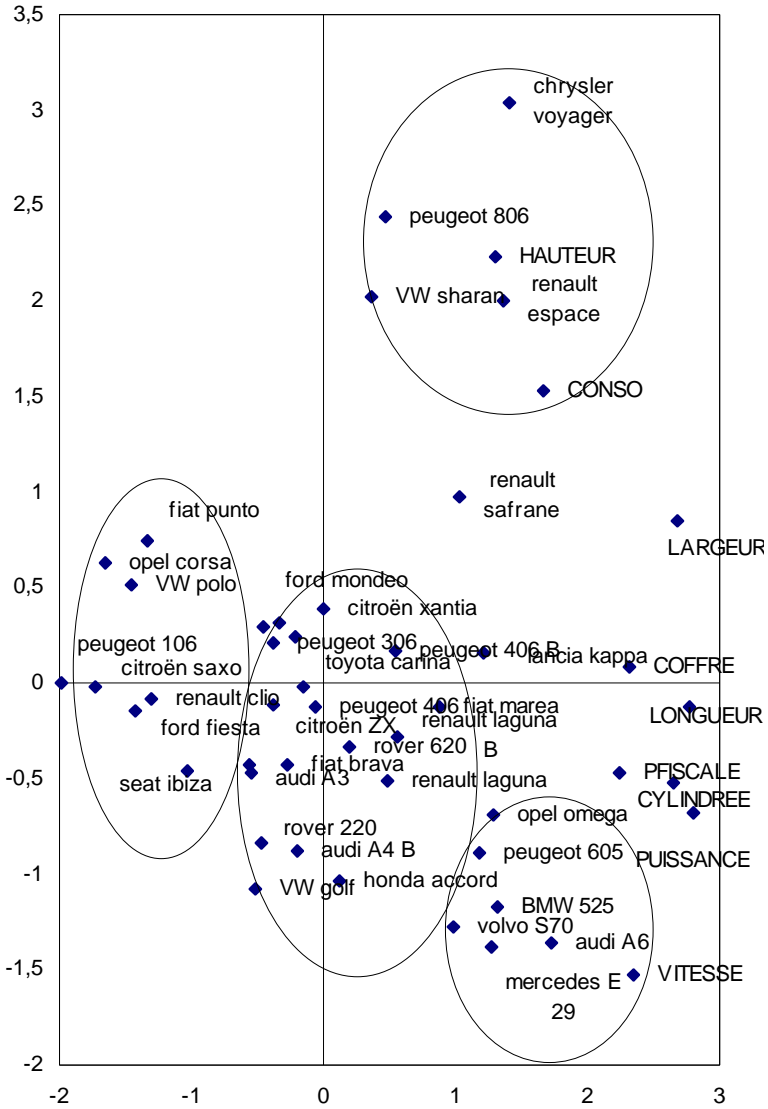


Figure 3 La projection des variables et des voitures diesel



### 3.2 La typologie

La figure 4 reprend l'arbre hiérarchique obtenu avec la méthode de Ward. Même avec d'autres méthodes de classification, nous pouvons retenir quatre classes: les compactes et routières, les grandes routières, les citadines et les monospaces. Comme pour les voitures essence, les compactes et les routières sont mélangées; ce qui semble vouloir dire qu'elles ont des caractéristiques proches. De plus les breaks se retrouvent parmi les compactes et routières.

## III LE CLASSEMENT DES VOITURES SELON LA TYPOLOGIE FLOUE

A partir du logiciel Matlab, nous avons réalisé une typologie floue de type c-moyens sur des données centrées et réduites en choisissant pour le paramètre flou  $m$  la valeur de 2. Nous avons effectué plusieurs typologies floues en faisant varier le nombre de groupes de 2 à 10 (le nombre maximum d'itérations a été fixé à 100). A chaque itération les valeurs d'appartenance approchent un minimum local de la fonction objective. Nous avons dû choisir entre 2 et 10 groupes, pour ce faire nous avons retenu l'index de Xie-Beni.

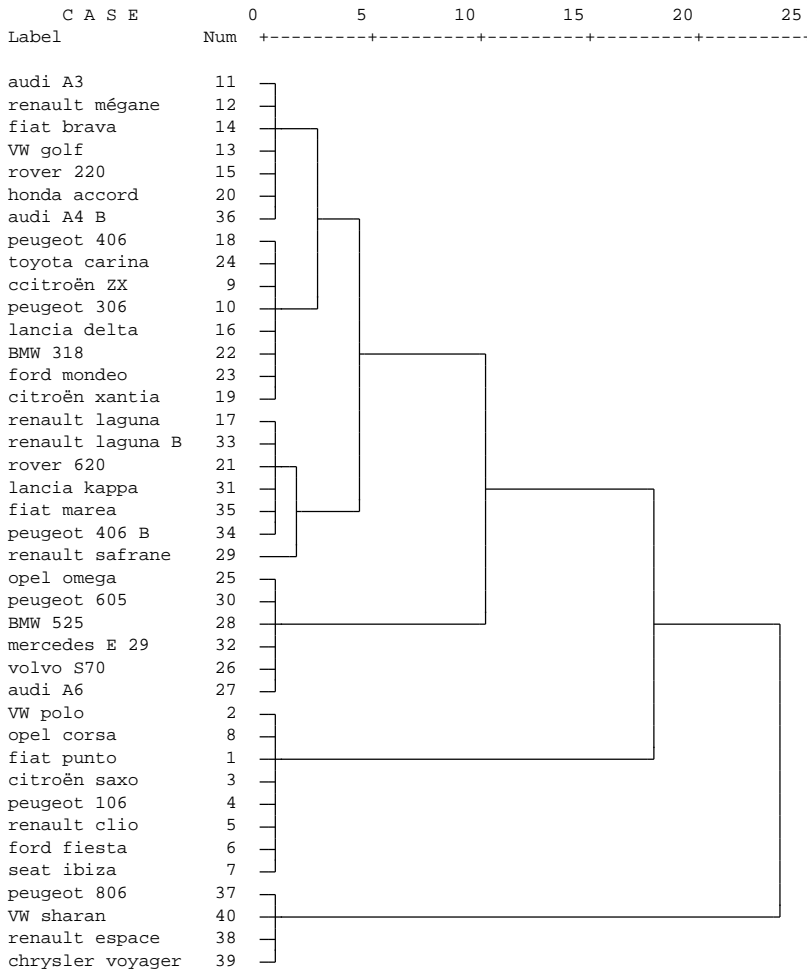
### 1 Les voitures essence

Les résultats concernant l'index de Xie-Beni, pour les voitures essence, sont regroupés dans le tableau 3. Nous optons pour cinq groupes.

**Tableau 3 L'index de Xie-Beni pour les voitures essence**

Nombre de groupes	Fonction objective	Index de Xie-Beni
2	142	3 463 415
3	79	285 034
4	51	1 943 598
<b>5</b>	<b>39,5</b>	<b>40 126</b>
6	33	78 601
7	29	133 708
8	23,7	200 017
9	19	321 816
10	17,4	424 390

**Figure 4 Le dendrogramme des voitures diesel**



Le tableau 4 reprend les groupes qui sont apparus avec la typologie floue par rapport aux cinq groupes du marché automobile. Certaines voitures ne sont que dans une seule classe (les 8 citadines et les 8 grandes routières), d'autres appartiennent à plus d'un groupe. Les monospaces sont ensemble mais les breaks sont dans d'autres catégories. Les compactes et les routières se retrouvent ensemble dans deux groupes.

**Tableau 4 Les résultats des cinq groupes flous de voitures essence**

<b>Groupes de voitures essence</b>	<b>Citadines (1 à 8)</b>	<b>Compactes (9 à 16)</b>	<b>Routières (17 à 24)</b>	<b>Grandes routières (25 à 32)</b>	<b>Breaks et Monospaces (33 à 41)</b>
<b>Groupe flou 1</b> Citadines	1, 2, 3, 4, 5, 6, 7, 8				
<b>Groupe flou 2</b> Compactes		9, 10, 11 <sup>*,+</sup> , 12 <sup>*</sup> , 13 <sup>*,+</sup> , 14, 15, 16 <sup>*,+</sup>	17 <sup>*</sup> , 22 <sup>*,+</sup>		36 <sup>*</sup> , 41 <sup>*,+</sup>
<b>Groupe flou 3</b> Routières		11 <sup>*</sup> , 12 <sup>*,+</sup> , 13 <sup>*</sup> , 16 <sup>*</sup>	17 <sup>*,+</sup> , 18, 19, 20, 21, 22 <sup>*</sup> , 23, 24		34 <sup>*</sup> , 36 <sup>*,+</sup> , 41 <sup>*</sup>
<b>Groupe flou 4</b> GRoutières				25, 26, 27, 28, 29, 30, 31, 32	37, 38
<b>Groupe flou 5</b> Monospaces					33, 34 <sup>*,+</sup> , 35, 39, 40

Légende: Les chiffres correspondent aux numéros de voitures, \* indique que la voiture appartient à plusieurs groupes, \*+ la voiture appartient plus à ce groupe.

Les routières qui se retrouvent avec les compactes ont des caractéristiques techniques qui se rapprochent plus ou moins d'elles avec une consommation autour de 10-11 l/100km. C'est le cas de la Renault laguna 1.8 RNE (n° 22) qui a une puissance et une vitesse inférieures à sa classe d'origine (les routières). Elle appartient plus aux compactes qu'aux routières, ce que ne laissent pas supposer les résultats des boxplots.

Les compactes qui sont avec les routières ont des caractéristiques qui rejoignent celles des routières (plus larges ou plus longues, plus de coffre, plus de vitesse, plus de puissance et une consommation autour de 8

l/100km). La Renault Mégane 1.6 e RTE (n° 12) est plus longue et a un coffre plus grand que les compactes.

Les voitures breaks qui ne sont pas dans le cinquième groupe sont souvent rattachées à la catégorie du modèle non break (n° 36, 37, 38 et 41). Un seul monospace se partage entre deux groupes (avec toutefois une valeur d'appartenance dominante dans sa catégorie) la Citroën évacion 1.8 (n° 34), car sa puissance et sa cylindrée sont inférieures aux autres monospaces.

## 2 Les voitures diesel

Comme pour les voitures essence, le choix de cinq groupes minimise de façon significative la fonction objective (tableau 5).

**Tableau 5 L'index de Xie-Beni pour les voitures diesel**

Nombre de groupes	Fonction objective	Index de Xie-Beni
2	145,7	1 124 228
3	87,7	13 703 125
4	62	351 474
<b>5</b>	<b>46,8</b>	<b>14 444</b>
6	37,5	81 099
7	32	5 000 000
8	27,5	305 556
9	24,5	362 426
10	22,3	55 750 000

Les voitures diesel compactes et routières ont des traits communs qui font que certaines d'entre elles se retrouvent ensemble (tableau 6). Le groupe des citadines est bien classé, à l'exception de la Seat Ibiza GT TDI (n° 7) qui, en raison d'une puissance, d'une vitesse et d'une cylindrée supérieures à sa catégorie, fait plus partie des compactes.

Du fait de certaines caractéristiques identiques (puissance, cylindrée, consommation), quatre routières se retrouvent plus ou moins avec les compactes et trois compactes avec les routières. La Peugeot 306 1.9 TD style (n° 10), du fait d'une consommation plus proche des routières, est plus dans cette catégorie. La Renault Safrane 2.2 dt RXT (n° 29), une grande routière, est plus classée avec les routières car son nombre de CV, sa

puissance, sa cylindrée, sa vitesse sont plus faibles que ceux des grandes routières. A noter que la Renault Laguna 2.2 dt RXT (n° 17) a quelques points communs avec les grandes routières: la vitesse et la cylindrée.

**Tableau 6 Les résultats des cinq groupes flous de voitures diesel**

<b>Groupes de voitures diesel</b>	<b>Citadines (1 à 8)</b>	<b>Compactes (9 à 16)</b>	<b>Routières (17 à 24)</b>	<b>Grandes routières (25 à 32)</b>	<b>Breaks et Monospaces (33 à 40)</b>
<b>Groupe flou 1</b> Citadines	1, 2, 3, 4, 5, 6, 7*, 8				
<b>Groupe flou 2</b> Compactes	7*+	9*, 10*, 11, 12, 13, 14, 15, 16*+	18*, 20*, 22*, 23*		36*
<b>Groupe flou 3</b> Routières		9, 10*+, 16*	17*+, 18*+, 19, 20*+, 21, 22*+, 23*+, 24	29*+	33*+, 34, 38*
<b>Groupe flou 4</b> 4GRoutières			17*	25, 26, 27, 28, 29*, 30, 31, 32	33*, 35, 38*
<b>Groupe flou 5</b> Monospaces				29*	37, 38*+, 39, 40

Légende: Les chiffres correspondent aux numéros de voitures, \* indique que la voiture appartient à plusieurs groupes, \*+ la voiture appartient plus à ce groupe.

Les voitures breaks sont rattachées à la même catégorie des voitures non breaks. La Renault Espace 2.2 dt RXT (n° 38) appartient à trois groupes: principalement aux monospaces puis aux grandes routières (coffre important) et aux routières (cylindrée, puissance).

### 3 Synthèse des résultats

La typologie floue est plus valide qu'une typologie classique car avec elle nous retrouvons les cinq catégories du marché de l'automobile tandis qu'avec la typologie ascendante hiérarchique seulement quatre groupes apparaissent. La typologie floue apporte des précisions sur le degré

d'appartenance des voitures à chaque classe tandis que la typologie classique donne un résultat global (compactes et routières mélangées). Les compactes et les routières sont très proches ce qui explique que la typologie classique ne les distingue pas. La typologie floue permet quant à elle d'affiner l'analyse. Certaines voitures appartiennent plus à une catégorie inférieure qu'à celle d'origine, d'autres à une catégorie supérieure. Or nous avons établi une relation entre le prix et le type de famille de voitures.

En n'affectant à chaque groupe que les voitures qui ont les plus fortes valeurs d'appartenance, il est possible de réaliser une analyse discriminante. 100% des voitures essence et diesel sont ainsi bien classées. L'analyse discriminante peut être utile pour l'affectation de nouveaux membres (Green et Tull, 1974). En entrant les caractéristiques techniques d'un nouveau modèle de voiture, l'analyse discriminante le classe dans l'un des cinq groupes déterminés à l'aide de la typologie floue. Si une voiture n'appartient pas à sa catégorie d'origine, l'analyse discriminante le montrera.

Les constructeurs automobiles doivent donc faire attention aux caractéristiques de leurs modèles car la frontière entre certaines catégories peut être si mince que cela peut provoquer une confusion dans l'esprit des consommateurs. Si des voitures ne sont pas clairement dans leur catégorie, ce flou peut se retrouver dans le choix du consommateur qui n'arrivera plus à reconnaître telle ou telle catégorie. C'est particulièrement le cas pour les routières et les compactes en raison de leurs nombreux points communs. Si un constructeur automobile ne propose pas une gamme cohérente il peut avoir des difficultés à vendre les modèles qui appartiennent à plusieurs catégories.

## **Conclusion**

Notre principal but a été de montrer la supériorité d'une technique de classification floue par rapport à une technique de classification classique pour identifier des groupes de voitures. Il s'avère que la typologie floue est plus riche en renseignements, fournit des groupes correspondant plus à la réalité d'un marché. Elle permet au manager d'identifier les modèles qui ont des caractéristiques techniques s'écartant de leur catégorie.

La typologie floue est un outil de décision car à partir des caractéristiques techniques d'un modèle de voiture, il est possible de déterminer si celui-ci

se rattache bien à sa catégorie ou si cette appartenance n'est pas aussi nette. Cela peut éviter aux constructeurs d'avoir des modèles se retrouvant dans une catégorie inférieure ou leur fournir des arguments de vente supplémentaires si le modèle se rapproche d'une catégorie supérieure. Bien sûr, il faudrait savoir dans quelle mesure le consommateur est sensible lors de son choix à ces critères. Il est clair que d'autres éléments sont pris en compte lors du choix parmi lesquels le confort, la tenue de route, la forme du modèle, la facilité de revente, le tarif d'assurance, les équipements et options proposés, le nom de la marque...

Ici l'objectif n'était pas d'étudier les critères de choix lors de l'achat d'une voiture mais de savoir si les cinq catégories de voitures existantes se retrouvaient à partir de leurs attributs techniques. Seule la typologie floue est en mesure de les faire ressortir. Elle révèle que les limites de certaines classes de voitures sont si floues qu'il est difficile de les différencier.

En résumé la typologie floue:

- fournit des groupes d'objets les plus proches possibles de la réalité en ne cantonnant pas un objet dans un seul groupe s'il appartient à plus d'un groupe;
- favorise la prise en compte du flou des limites d'un marché (multiplicité des produits, des marques, des usages...), ce qui facilite la compréhension de sa structure;
- détermine les membres types d'une classe, la force d'appartenance des objets aux classes, contrairement à une classification ascendante hiérarchique;
- donne des informations sur les produits en compétition;
- améliore les connaissances des managers.

La typologie floue ne doit pas être considérée comme une technique qui apporterait des solutions à toutes les interrogations. Il s'agit d'un outil parmi d'autres qui offre de nombreux avantages, qui améliore les connaissances mais qui comporte une part de subjectivité.

## Bibliographie

Al-Sultan K.S. et Selim S.Z. (1993), A global algorithm for the fuzzy clustering problem, Pattern Recognition, vol. 26, n° 9, 1357-1361.

Arabie P., Carroll J.D., deSarbo W. et Wind J. (1981), Overlapping clustering: a new method for product positioning, Journal of Marketing Research, vol. XVIII, august, 310-317.

Bezdek J.C. (1974), Cluster validity with fuzzy sets, Journal of Cybernetics, vol. 3, n° 3, 58-72.

Bezdek J.C. (1981), Pattern recognition with fuzzy objective function algorithms, New York: Plenum.

Bezdek J.C., Coray C., Gunderson R. et Watson J. (1981), Detection and characterization of cluster substructure, SIAM Journal of Applied Mathematics, vol. 40, n° 2, 339-372.

Bezdek J.C. et Pal S.K. (1992), Fuzzy models for pattern recognition, IEEE press, New York.

Dunn J.C. (1974), A fuzzy relative ISODATA process and its use in detecting compact well-separated clusters, Journal of Cybernetics, vol. 3, n° 3, 32-57.

Grabisch M. (1994), Classification et reconnaissance des formes, dans Logique floue, ARAGO 14, Observatoire Français des Techniques Avancées, Masson.

Green P.E. et Tull D.S. (1974), Recherche et décisions en marketing, Presses Universitaires de Grenoble.

Hathaway R.J. et Bezdek J.C. (1994), NERF c-means: non euclidean relational fuzzy clustering, Pattern Recognition, vol. 27, n° 3, 429-437.

Kamel M.S. et Selim S.Z. (1994), New algorithms for solving the fuzzy clustering problem, Pattern Recognition, vol. 27, n° 3, 421-428.

Kaufmann A. (1977), Introduction à la théorie des sous-ensembles flous 1. Éléments théoriques de base, Masson, 2e édition.

Pal N.R. et Bezdek J.C. (1995), On cluster validity for the fuzzy c-means model, IEEE transactions on fuzzy systems, vol. 3, n° 3, august, 370-379.

Pal N.R., Bezdek J.C. et Hathaway R.J. (1996), Sequential competitive learning and the fuzzy c-means clustering algorithms, Neural networks, vol. 9, n° 5, 787-796.

Ruspini E.H. (1969), A new approach to clustering, Information and control, 15, 22-32.

Ruspini E.H. (1970), Numerical methods for fuzzy clustering, Information Sciences, 2, 319-350.

Shepard R.N. et Arabie P. (1979), Additive clustering representation of similarities of discrete overlapping properties, Psychological Review, vol. 86, n° 2, march, 87-123.

Xie X.L. et Beni G.A. (1991), Validity measure for fuzzy clustering, IEEE Transactions on pattern analysis and machine intelligence, vol. 13, n° 8, august, 841-846.

Zadeh L.A. (1965), Fuzzy sets, Information and control, 8, 338-353.

Zimmermann H.J. (1991), Fuzzy set theory and its applications, Kluwer Academic Publishers, USA.