

LE "THEMASCOPE" OU ANALYSE STRUCTURELLE DES DONNEES D'ENQUETE

Alain Morineau

*Centre International de Statistique et d'Informatique Appliquées (C.I.S.I.A)
Saint-Mandé*

7.1 Introduction

Plus de vingt années de pratique ont démontré la puissance des méthodes d'analyses de données pour acquérir de l'information synthétique sur les gros ensembles de données complexes comme le sont les fichiers d'enquêtes. La souplesse des programmes de calcul et leur efficacité permettent d'y faire appel de façon répétée sur le même fichier. Une nouvelle pratique et, en quelque sorte, un nouvel outil se dessine dans le fait même de coordonner la répétition des analyses sur un fichier.

Comme il ne s'agit ni de techniques ni de méthodes nouvelles mais d'une certaine mise en pratique de méthodes connues, la présentation faite ici reposera essentiellement sur un exemple après le rappel succinct du contexte méthodologique.

7.2 Rappel du contexte méthodologique

Les questions apparaissant dans une enquête peuvent en général être regroupées selon plusieurs thèmes. Dans une enquête de type socio-économique par exemple, on classera dans le thème "signalétique" les variables à la base du plan de sondage (sexe, âge, nombre d'enfants, etc.) et les variables de structure socio-administrative (catégorie de revenu, profession du chef de famille, etc.). Les autres questions seront groupées dans un ou plusieurs thèmes selon le contenu même de l'enquête : opinions, attitudes, comportements.

Le traitement traditionnel de l'enquête consiste à utiliser les variables du signalétique pour définir des grilles de tabulations avec le reste des

variables. L'examen minutieux de ces tableaux doit fournir des présomptions d'explication où les variables du signalétique jouent le rôle de prédicteurs.

Élémentaire dans son application et dans son interprétation, ce traitement présente cependant des inconvénients pratiques. Il peut y avoir plusieurs milliers de tableaux à scruter et de plus leur examen séquentiel ne tient pas compte des dépendances entre les éléments du signalétique. Les problèmes sont présentés en détail dans Lebart et al. (1977, chap. IV). On y montre également comment l'analyse factorielle apporte une solution à la fois pour prendre en compte les interrelations entre variables du signalétique et pour sélectionner les tris croisés intéressants.

La notion de "thémascope" a été introduite à propos des stratégies de traitement statistique des données d'enquête (Lebart, 1989). Le thémascope y est présenté comme un enchaînement canonique de méthodes comprenant les étapes suivantes :

- Choix d'un thème "actif" c'est-à-dire d'une batterie de questions permettant de décrire les individus selon un certain point de vue. Un premier thème peut être le "signalétique" de l'enquête : âge, sexe, profession, etc. Les autres thèmes seront définis selon l'enquête, par exemple par un groupe de questions relatives aux habitudes alimentaires, ou par un groupe de questions relatives aux opinions politiques.
- Description "graphique" de l'échantillon des individus interrogés, la proximité entre individus étant fonction de la similitude des réponses aux questions du thème "actif". Les graphiques résultant des analyses factorielles (analyse en composantes principales ou correspondances multiples) peuvent fournir ces descriptions.
- Utilisation des questions ne faisant pas partie du thème actif pour interpréter les proximités entre individus. On parle ici du positionnement des éléments illustratifs sur les plans factoriels. Lorsque la lecture des résultats est gênée par l'abondance des éléments, des procédures de nature statistique seront utilisées pour sélectionner automatiquement les seuls éléments pertinents pour l'interprétation.
- Partitionnement de l'échantillon en groupes homogènes de réponses aux questions du thème actif. La technique utilisée pour réaliser ces classes sera en général une combinaison convenable de procédures de classification automatique permettant de surmonter les difficultés dues au volume des

calculs, à la non-optimalité des algorithmes et à l'ignorance du nombre de classes réelles.

- Description statistique du contenu de chaque classe homogène en utilisant tous les éléments disponibles dans le fichier de l'enquête. On sépare la description par les éléments actifs (utilisés pour la construction des classes) et la caractérisation par les éléments illustratifs. On met en œuvre ici aussi des procédures de sélection automatique des seuls éléments intéressants. La description obtenue est en général simple et facile à lire.
- Positionnement des centres des classes en éléments supplémentaires dans les plans factoriels. Les positions relatives des modalités de réponses et des classes de la typologie permettent d'enrichir les interprétations. La description des classes est en effet un exercice plus facile que la description d'un espace continu, mais la position des classes dans un espace factoriel pourra révéler les continuités de distribution cachées par l'arbitraire des frontières des groupes.

Le thémascope apparaît donc comme un outil permettant de décrire un thème, multidimensionnel par nature, en le situant dans le contexte global de l'enquête. Dans le domaine des enquêtes de type socio-économique, le premier thème soumis à analyse sera en général le thème "signalétique" : l'échantillon est décrit sous l'angle de ses caractéristiques socio-administratives et la grille obtenue sert à positionner les informations relatives aux thèmes concernant le contenu de l'enquête.

7.3 Variables actives et illustratives

Il existe une dichotomie fondamentale entre les éléments qui participent à la construction des axes ou des classes, appelés éléments actifs, et les éléments qui interviennent *a posteriori* pour caractériser ces axes ou ces classes et que l'on appelle illustratifs ou supplémentaires.

Les analyses factorielles permettent de trouver des sous-espaces pour représenter les proximités entre individus et entre variables et modalités. Elles permettent aussi de positionner dans ces sous-espaces des individus ou des questions (ou variables) supplémentaires.

On peut ainsi illustrer les plans factoriels par des informations n'ayant pas participé à la construction de ces plans, opération qui a des conséquences très importantes pour l'interprétation des résultats. Les éléments servant à

calculer les plans factoriels sont appelés éléments actifs : ils doivent former un ensemble homogène pour que les distances entre individus aient un sens. Ils sont relatifs à un même thème.

Il faut considérer qu'une analyse exploratoire ne fournit qu'un point de vue sur les données, point de vue déterminé par le choix des éléments actifs.

Il serait maladroit de mélanger, par exemple, des informations de type socio-économique (âge, sexe, profession, etc.) et des opinions ou des comportements. Comment interpréterait-on alors la proximité entre deux individus ? Si deux individus sont très éloignés, est-ce à cause de leurs opinions différentes ou parce qu'ils ont des profils socio-économiques opposés ?

Il est important d'obtenir une description des individus d'un seul point de vue, quitte ensuite à utiliser toute l'information disponible sur ces mêmes individus pour identifier et interpréter les proximités observées. On se restreint donc à prendre en compte un seul thème actif par analyse.

La figure 1 (cf. page suivante) illustre ce propos. La partie gauche schématise la description de la population enquêtée à partir des éléments du signalétique. On y voit par exemple une certaine répartition des âges le long d'une trajectoire remplissant le premier plan factoriel alors que les modalités *homme* et *femme* sont centrales.

Dans la description socio-administrative de l'échantillon, certaines modalités ou associations de modalités sont caractéristiques des jeunes, d'autres des vieilles personnes. Par contre il semble que l'ensemble des attributs socio-administratifs sont ici répartis indifféremment dans les deux sexes. Les variables illustratives, cachées ici en arrière plan, chercheront à associer à des zones socio-administratives bien définies certaines opinions ou certains comportements particuliers.

On pourrait y voir par exemple que certaines opinions (éléments illustratifs dans la grille socio-administrative) sont très souvent le fait de personnes âgées, d'autres sont caractéristiques de personnes plus jeunes.

Dans la partie droite du graphique se trouve représentée une autre analyse. Imaginons que les variables actives sont constituées ici d'une batterie homogène de questions d'opinion sur la famille et le mariage.

L'échantillon est cette fois décrit en fonction des associations structurelles (non aléatoires) entre les différentes opinions émises sur la famille et le mariage. Le graphique décrivant les proximités entre opinions se trouve ici

en arrière plan. En premier plan on a positionné en éléments illustratifs les attributs socio-administratifs des mêmes individus.

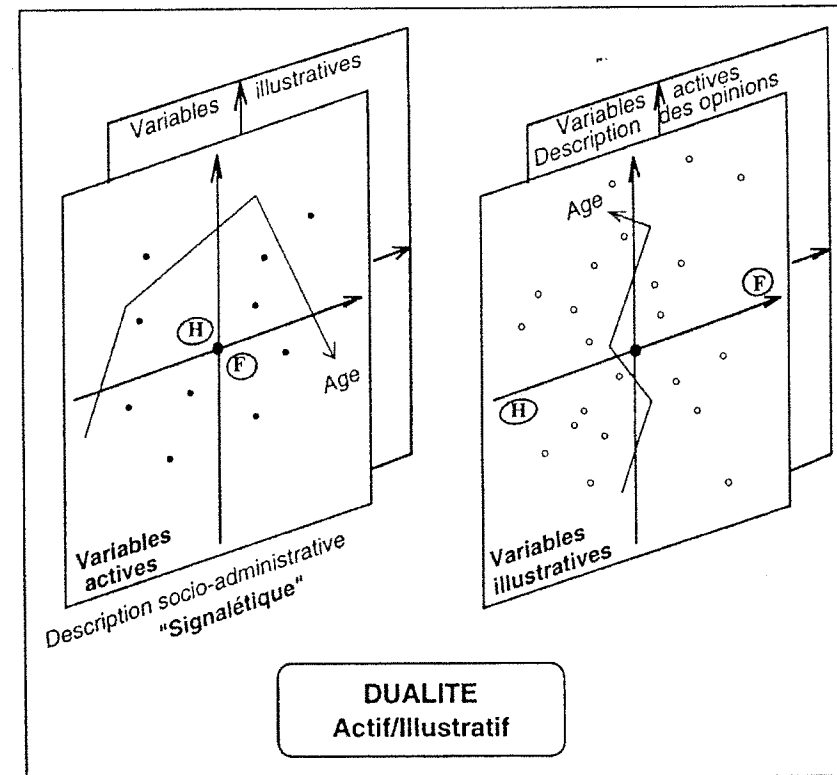


Figure 1

On voit ici que l'axe horizontal discriminant les opinions (thème actif) est aussi un axe qui sépare les sexes (variable illustrative). En "lisant" le graphique, on décrira donc comment les opinions sont liées au sexe. De même, la trajectoire des âges, apparaissant en élément illustratif le long de l'axe vertical, suggère que ces opinions (thème actif) sont liées aussi à l'âge. Plus précisément on *lira* comment les zones d'opinions visibles sur le plan des variables actives peuvent être reconstituées (ou prévues) en croisant le sexe et l'âge des enquêtés (dans cet exemple simpliste).

L'ensemble des éléments illustratifs n'a nul besoin d'être homogène. On a même intérêt à ce qu'il soit le plus large et hétéroclite possible. On

augmente ainsi les chances de découvrir les variables susceptibles d'*expliquer* les axes ou les classes.

La distinction entre variables actives et variables illustratives est fondamentale. Elle est analogue à la distinction, existant en régression multiple, entre les variables dites explicatives (x_1, x_2, x_3, \dots) et la variable à expliquer (y). D'un point de vue géométrique, les deux situations sont d'ailleurs très similaires. Dans une régression les variables (x_1, x_2, x_3, \dots) engendrent un sous-espace sur lequel on projette la variable y . Dans une analyse exploratoire, les variables actives engendrent aussi un sous-espace dont on réduit la dimensionnalité pour le visualiser et l'on y projette des variables illustratives (y_1, y_2, y_3, \dots) pour les situer par rapport aux variables actives.

On notera que l'analyse exploratoire met en jeu de façon essentielle les dépendances entre les variables actives. Par contre les variables illustratives sont introduites l'une après l'autre, sans tenir compte de leur éventuelle dépendance ou redondance. Toutes les variables illustratives seront essayées. Toutes les combinaisons, croisements ou recodages de variables sont des "explications" potentielles.

7.4 Les techniques factorielles

Les techniques mises en œuvre dans l'enchaînement canonique des opérations du thémascope appartiennent à deux familles de méthodes : les méthodes faisant appel à l'algèbre linéaire, appelées méthodes factorielles, et les techniques de classification automatique. Nous allons en rappeler brièvement les principes. Des exposés plus complets ont été présentés aux chapitres 4 "*Méthodes factorielles*" et 5 "*Classification des données d'enquêtes*" du présent ouvrage.

Il est toujours possible de calculer des distances entre les lignes et les colonnes d'un tableau rectangulaire de valeurs numériques. Ces distances peuvent s'interpréter en termes de corrélations ou de similarités. Mais il n'est pas possible de visualiser ces distances de façon immédiate : il est nécessaire de procéder à des transformations ou à des approximations pour en obtenir une représentation plane. L'opération de réduction de dimension de certaines représentations multidimensionnelles est l'une des tâches dévolues à l'analyse factorielle.

Les techniques factorielles usuelles reposent sur la même propriété mathématique des tableaux rectangulaires : la décomposition aux valeurs singulières (décomposition d'Eckart et Young). Cette propriété, qui concerne le tableau de données lui-même et non pas seulement la matrice

des corrélations ou une matrice de distances construite à partir des données, implique de façon analogue les lignes et les colonnes du tableau. Les analyses mettent donc en jeu simultanément dans des rôles semblables les individus (lignes) et les variables (colonnes).

Quelques propriétés simples vont gouverner les interprétations des proximités graphiques. Dans le sous-espace factoriel où l'on regarde la configuration du nuage des individus, tout groupe d'individus sera représenté par son point moyen. La proximité entre deux groupes doit donc être appréciée comme une proximité entre deux points moyens (moyenne sur les mêmes individus ou sinon sur des individus semblables puisqu'ils sont proches). Noter que toute modalité de réponse à une question définit un groupe d'individus : ceux qui ont choisi cette réponse. Naturellement une classe définit aussi un groupe d'individus. Quelle que soit l'analyse, les points-modalités (actifs et illustratifs) et les classes sont positionnés de cette façon.

Dans ce même espace contenant les individus, les variables continues (actives ou illustratives) peuvent être positionnées en prenant pour abscisse sur un axe la corrélation entre les valeurs individuelles et les coordonnées des individus sur cet axe. Alors un point-variable joint à l'origine dans l'espace des individus donne la direction des valeurs croissantes de cette variable et permet donc de caractériser les zones périphériques de la représentation. De plus la proximité avec le vecteur unitaire d'un axe, interprétée en terme de corrélation linéaire, peut conduire à *identifier* cet axe.

7.4.1 L'analyse en composantes principales

Elle s'applique aux tableaux de type "variables-individus", dont les colonnes représentent des variables à valeurs numériques, et les lignes des individus. Les proximités entre variables s'interprètent en termes de corrélation entre variables, les proximités entre individus s'interprétant en termes de similitudes globales des valeurs centrées réduites correspondantes.

L'analyse en composantes principales peut donner lieu à la projection de variables supplémentaires (numériques ou nominales) sur les plans factoriels selon les indications données ci-dessus. Ainsi dans l'analyse d'une batterie de notes de préférence attribuées à des produits de consommation, on pourra positionner en éléments supplémentaires le revenu des personnes interrogées (variable continue) et leur profession (variable nominale).

7.4.2 L'analyse des correspondances multiples.

L'analyse des correspondances multiples permet de décrire de grands tableaux de variables nominales, dont les fichiers d'enquêtes socio-

économiques constituent un exemple privilégié. Les lignes de ces tableaux sont en général les individus interrogés (il peut en exister plusieurs milliers) ; les colonnes sont les modalités de réponses aux questions de l'enquête.

Il s'agit en fait d'une extension du domaine d'application de l'analyse des correspondances, avec cependant des procédures de calcul et des règles d'interprétation spécifiques. Cette extension se fonde sur l'équivalence suivante. On dispose pour k individus des valeurs (réponses) prises par deux variables nominales ayant respectivement n et p modalités. Il est alors équivalent de soumettre à l'analyse des correspondances la table de contingence (n,p) croisant les deux variables, ou d'analyser le tableau binaire à k lignes et $(n + p)$ colonnes qui décrit les réponses. L'analyse de ce dernier tableau est plus coûteuse, mais plus intéressante, car elle se généralise immédiatement au cas de plus de deux variables.

7.4.3 Les techniques de classification

La classification consiste à faire des groupements, les moins arbitraires possibles, d'éléments qui sont en général les lignes d'un tableau (les personnes interrogées) à partir de leurs réponses au questionnaire. Ces groupements peuvent se faire par agglomération progressive des éléments deux à deux (classification hiérarchique) ou par recherche directe d'une partition : on affecte les individus à des centres provisoires de classes, puis on recentre ces classes et on ré-affecte de façon itérative les individus (méthode des nuées dynamiques, ou méthode "k-means").

L'algorithme de classification qui semble actuellement le mieux adapté au partitionnement d'un ensemble comprenant des milliers d'individus est un algorithme *mixte* procédant en quatre phases :

- 1) Partitionnement initial en plusieurs dizaines de classes (par une technique du type "nuées dynamiques" ou "k-means"). Cette étape fabrique de nombreux noyaux, plus ou moins artificiels, dans les zones denses de la configuration des individus.
- 2) Agrégation des noyaux obtenus. L'agrégation hiérarchique est assez coûteuse si elle s'applique à des milliers d'individus, c'est pourquoi il est commode de réduire la dimension du problème en opérant un regroupement préalable en quelques dizaines de noyaux.
- 3) Choix du nombre de classes par coupure de l'arbre (en général après une inspection visuelle) et optimisation de l'homogénéité des classes par réaffectations.

- 4) Caractérisation statistique des classes. Une fois les individus regroupés en classes, il est facile d'obtenir une description automatique de ces classes. On peut en effet, pour les variables numériques comme pour les variables nominales, calculer des statistiques d'écart entre les valeurs internes à la classe et les valeurs globales. On convertira ces statistiques en un critère appelé *valeurs-tests* pour opérer un tri sur ces éléments. On obtient finalement, pour chaque classe, les modalités et les variables les plus caractéristiques.

7.4.4 Complémentarité de la classification et des méthodes factorielles

Cette complémentarité a déjà été évoquée aux chapitres 1 et 5.

Les algorithmes de classification appliqués aux points-individus permettent d'alléger les sorties graphiques (en remplaçant ces points par des groupes de points, beaucoup moins nombreux) et de corriger certaines déformations dues à l'opération de projection dans un espace à deux dimensions.

De plus, il est plus facile de décrire automatiquement des groupes d'individus (les classes) qu'un continuum spatial (les plans factoriels).

Chaque analyse factorielle est donc avantageusement complétée par de telles classifications, avec identification automatique des classes par des caractérisations statistiques.

La notion de classe est intuitive (groupes d'individus les plus semblables possible). La description des classes est facile à lire car elle fait apparaître les classements des libellés en clair et ces classements sont fondés sur des comparaisons de pourcentages.

7.4.5 Tri systématique des caractéristiques.

La sélection des éléments les plus significatifs sur les plans factoriels et lors de la description des classes est faite automatiquement (elle est pilotée par des seuils).

Les éléments sélectionnés peuvent être décrits par des libellés longs explicites. Le lecteur dispose donc d'une information filtrée et parfaitement lisible.

Positionnement systématique des caractéristiques des individus

Les individus sont en général "anonymes" et ne nous intéressent que par leurs caractéristiques (âge, sexe, ou tout autre attribut). Au lieu de

représenter les positions des individus dans l'espace factoriel, il est plus utile de représenter les centres de gravité des groupes d'individus correspondant à leurs divers attributs.

Cependant il y a parfois plusieurs centaines (ou milliers) de modalités et il est exclu de faire une représentation graphique de l'ensemble des points correspondants. Le critère de sélection statistique des réponses significatives est la *valeur-test*.

Les valeurs-tests pour un axe factoriel ou une variable

C'est une quantité attachée (sur chaque axe factoriel) à chaque modalité de réponse. Dans le cadre de la théorie des tests ce critère devrait être supérieur à 2 ou inférieur à -2 pour que la modalité soit significativement liée à l'axe. Avec l'optique exploratoire des analyses, le seuil est indicatif : seules sont intéressantes les valeurs relatives du critère.

Précisons le principe de son calcul : si les individus d'une modalité (ou d'une classe) sont tirés au hasard dans la population, leur centre de gravité sera proche de celui de l'ensemble des individus (on tient évidemment compte de l'effectif de la modalité dans les calculs).

Ce que mesure la *valeur-test* est précisément la distance entre les centres de gravité dans le cas d'un tirage au hasard, distance qui est convertie en nombre d'écart-types d'une loi normale centrée réduite. La formule de calcul a déjà été présentée au chapitre 4.

Les valeurs-tests pour une classe ou un groupe

Un attribut est considéré comme caractéristique de la classe si son abondance dans la classe est jugée significativement supérieure à son abondance dans la population.

L'abondance de la modalité s'apprécie en premier lieu en comparant son pourcentage dans la classe ("MOD/CLA") à son pourcentage dans la population ("Global"). Mais cette comparaison est insuffisante : il faut aussi tenir compte du degré d'inclusion de la classe dans la modalité ("CLA/MOD").

Le critère appelé encore "valeur-test" tient compte de tous ces éléments. Il s'agit d'un critère statistique associé à la comparaison des effectifs dans le cadre d'une loi hypergéométrique et transformé en nombre d'écart-types d'une loi normale centrée réduite.

Les modalités sont d'autant plus intéressantes que les valeurs-tests sont fortes en valeur absolue. On peut alors trier ces valeurs-tests et n'imprimer que les éléments les plus intéressants, ce qui permet de caractériser très rapidement les classes.

7.5 Un exemple de découpage en thèmes

L'exemple est tiré de l'analyse d'un panel d'opinion sur la population Est-Allemande (Samson, 1992).

Nous n'aborderons pas le détail des résultats et des interprétations. Notre objectif est de donner une idée de la démarche associée à l'utilisation du thémascope comme instrument d'exploration des données d'une enquête réelle.

Le tableau 1 présente un extrait des quelques 250 questions soumises à chaque personne, certaines étant des questions posées lors de la première vague de l'enquête (notée v1), d'autres lors de la seconde vague (v2). Sur chacune des deux vagues nous comptons les mêmes 1 315 personnes enquêtées.

Dans le questionnaire, outre les variables du "signalétique" des enquêtés, on dégage plusieurs thèmes relatifs au contenu :

- la représentation de l'avenir,
- les attentes en matière socio-économique,
- les attitudes d'ouverture aux Allemands de l'Ouest, etc.

Nos commentaires sur les méthodes vont concerner deux thèmes seulement : le signalétique d'une part et un thème assez général appelé "politique" d'autre part.

7.5.1 Thème "Signalétique": Composition du Panel

Les 1 315 individus du panel sont comparés du point de vue de 12 paramètres les caractérisant sous l'angle socio-économique.

L'analyse des correspondances multiples conduit, dans le plan des deux premiers facteurs, à la description fournie par la figure 2.

Une procédure de classification mixte conduit à une partition en 10 classes homogènes vis à vis de ces paramètres.

Les 12 paramètres du signalétique sont :

- | | |
|---------------------|---------------|
| 1) sexe | (2 modalités) |
| 2) âge | (6 modalités) |
| 3) confession | (3 modalités) |
| 4) état matrimonial | (5 modalités) |
| 5) vit seul ou non | (3 modalités) |
| 6) niveau scolaire | (3 modalités) |
| 7) qualification | (5 modalités) |

- | | | |
|-----|------------------------|---------------|
| 8) | situation d'emploi | (6 modalités) |
| 9) | taille d'agglomération | (5 modalités) |
| 10) | land habité | (6 modalités) |
| 11) | revenu déclaré | (6 modalités) |
| 12) | syndiqué ou non | (3 modalités) |

Les items du signalétique apparaissent sur la figure 2 (à l'exception de la région d'habitation). Le centre du graphique représente le point moyen des 1 315 individus.

Chaque item est le point moyen des individus correspondant à cet attribut. La proximité des items est donc interprétée en terme de similitude plus ou moins grande des individus concernés.

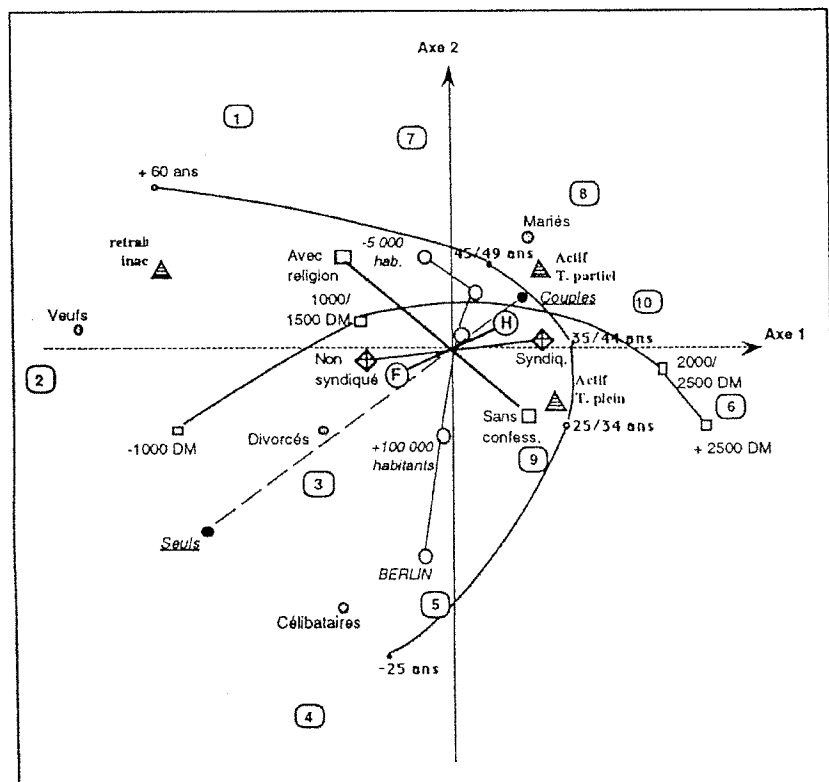


Figure 2

L'essentiel des liaisons entre les items du signalétique est représenté sur le graphique. Les numéros encadrés correspondent aux centres des classes d'une typologie des 1 315 individus, réalisée dans l'espace complet des questions actives du signalétique et positionnés ici "après coup" en éléments illustratifs.

La caractérisation statistique du contenu de certaines classes évoquée ci-dessous précise ce que le graphique suggère qualitativement.

Classe 2 97 individus 7 %

On ne trouve que des personnes âgées, retraitées et vivant seules ; tous les veufs et veuves. Que des bas revenus (vt15, ie : valeur-test égale à 15) et bas niveau scolaire (vt8). Non syndiqués (vt6). Légère sur-représentation des femmes (vt5) et des pratiquants (vt5).

Classe 5 122 individus 9 %

Sont célibataires (la classe contient la moitié des célibataires). Plutôt célibataires cohabitant (vt8). Très jeunes 18/24 ans (vt15) et vivant seuls (vt12). Niveau scolaire moyen (vt7). Légère sur-représentation des ouvriers qualifiés (vt4) et des bas salaires et chômeurs (vt3).

Classe 6 164 individus 13 %

Niveau scolaire élevé : 80 % ont suivi un enseignement supérieur. Revenu supérieur à 2 500 DM (vt8). Actifs plein temps (vt8). Mariés (vt6). Légère sur-représentation des hommes (vt4) et des 35/44 ans (vt4).

Classe 7 167 individus 13 %

Des 45/59 ans pour 85 % (vt16). Niveau scolaire bas (vt16). Mariés, pratiquants (vt8). Plutôt ouvriers non qualifiés ou semi qualifiés (vt7). Petite agglomération (vt6) et agglomération 5 000/20 000 (vt4).

Classe 8 193 individus 15 %

60 % chômeurs pour 22 % en moyenne (vt12). Mariés (vt9). Ouvriers qualifiés ou contremaîtres (vt10). Plutôt hommes : 80 % pour 50 % moyenne (vt9) et syndiqués : 70 % contre 45 % moyenne (vt8). Revenus plutôt moyens 1 500/2 000 DM (vt7). Plutôt en agglomération 20 000/100 000 (vt7). Légère sur-représentation des 35/44 ans (vt5).

Tableau 1

**Enquête PANEL RDA
Dictionnaire des variables**

3.	F1-v1.	Auquel de ces lieux avez-vs le sentiment d'appartenir	(8 Mod.)
4.	F2-v1.	De qui vous sentez-vous personnellement plus proche	(2 Mod.)
5.	F3-v1.	De qui vous sentez-vous plus proche	(2 Mod.)
6.	F4-v1.	Depuis Nov. 1989, avez-vous déjà dû changer de métier	(2 Mod.)
7.	F4b-v1.	Depuis Nov. 89, avez-vs déjà dû quitter votre région	(2 Mod.)
8.	F5-v1.	Pensez-vous personnellement avoir à changer de métier	(2 Mod.)
9.	F6-v1.	Pensez-vs personnellement avoir à quitter votre région	(2 Mod.)
10.	F7-v1.	Postes économiques en RDA occupés par RFA	(5 Mod.)
11.	F8-v1.	Postes politiques en RDA occupés par RFA	(5 Mod.)
12.	F11-v1.	Elément devant le PLUS intervenir dans prog. salaire	(7 Mod.)
13.	F12-v1.	Elément devant le MOINS intervenir dans prog. salaire	(7 Mod.)
14.	F13-v1.	Intention de vous installer en Allemagne de l'ouest	(4 Mod.)
15.	F14-v1.	Niveau d'intervention de l'état dans l'économie	(10 Mod.)
16.	F16-v1.	Rattraper compétitivité même si conséquences sociales	(4 Mod.)
17.	F17-v1.	Cette difficile période d'adaptation va durer pour vs	(9 Mod.)
18.	F18-v1.	A trav. égal, Salaire en RDA=quel % du salaire RFA	(6 Mod.)
19.	F19-v1.	Délai maximum acceptable pour rattrapage salaires	(7 Mod.)
20.	F26-v1.	Emploi : passé polit. plutôt que compétences	(4 Mod.)
21.	F27-v1.	Quelle est votre opinion sur le Japon	(5 Mod.)
22.	F28-v1.	Quelle est votre opinion sur la Pologne	(5 Mod.)
23.	F29-v1.	Quelle est votre opinion sur la France	(5 Mod.)
24.	F30-v1.	Quelle est votre opinion sur l'URSS	(5 Mod.)
25.	F31-v1.	Quelle est votre opinion sur la Suède	(5 Mod.)
... etc ...			
224.	F132-v2.	Intention de trav. en RFA comme "frontalier"	(2 Mod.)
225.	F151-v2.	Rôle de l'Etat: aider entrep. à investir en RDA	(4 Mod.)
226.	F152-v2.	Rôle de l'Etat: Rénov. des infrastructures	(4 Mod.)
227.	F153-v2.	Rôle de l'Etat: Assurer couv. sociale (5 lands)	(4 Mod.)
228.	F154-v2.	Rôle de l'Etat: Pr. en charge form. accélérée	(4 Mod.)
232.	F203-v2.	Montant global du loyer	(11 Mod.)
235.	F206-v2.	Situation d'emploi de l'enquêté(e)	(12 Mod.)
236.	F207-v2.	Avez-vous cette année pris une nouv. activ. prof	(2 Mod.)
237.	F208-v2.	Avez-vous peur de l'avenir	(2 Mod.)
238.	F209-v2.	Profession exercée	(8 Mod.)
239.	F210-v2.	Etes-vous adhérent d'un syndicat ?	(2 Mod.)
240.	F211-v2.	Revenu global	(8 Mod.)
244.		Vague d'enquête	(4 Mod.)
245.		Age de l'enquêté(e)	(6 Mod.)
246.		L'enquêté(e) a t-il une confession	(3 Mod.)
247.		Statut matrimonial	(5 Mod.)
248.		L'enquêté(e) vit-il(elle) seul(e)	(3 Mod.)
249.		Qualification de l'enquêté(e)	(5 Mod.)
250.		Situation d'emploi de l'enquêté(e)	(6 Mod.)
251.		Taille d'agglomération	(5 Mod.)
252.		Revenu Global	(6 Mod.)

A titre d'exemple, on va regarder la composition de la classe 5.

Le listage du tableau 2 fournit :

- 1) Sa caractérisation socio-économique (variables actives) ;
- 2) Les opinions caractéristiques pour la vague 1 ;
- 3) Celles devenues caractéristiques pour la vague 2.

Tableau 2

Typologie du "signalétique" : Caractérisation de la classe 5

MODALITES CARACTERISTIQUES		FOURCENTAGES		POIDS	V.TEST	
Variables actives du signalétique		CLASSE 5 / 10	GLOBAL			
CLASSE 5 / 10				9.28	122	
Statut matrimonial	Célibataire	59.79	92.62	14.37	189	21.15
Age de l'enquêté(e)	18--24 ans	55.56	61.48	10.27	135	15.12
L'enquêté(e) vit-il(elle) seul(e)	Oui	29.62	69.67	21.83	287	11.95
Niveau scolaire de l'enquêté(e)	10ème Classe	15.53	78.69	47.00	618	7.42
Revenu Global	Cuv. Qual. Cont.	12.50	74.59	55.36	728	4.50
Situation d'emploi de l'enquêté(e)	Moins 1000 DM	16.07	29.51	17.03	224	3.50
L'enquêté(e) a t-il une confession	Chômeur	14.43	35.25	22.66	298	3.24
	Non	11.19	77.05	63.88	840	3.17

Opinions à la Vague 1		FOURCENTAGES		POIDS	V.TEST	
		CLASSE 5 / 10	GLOBAL			
classe 5/10				9.28	122	
F4b-v1. Depuis Nov. 89, avez-vous déjà dû quitter votre région	Non	10.93	96.72	82.13	1060	4.96
	Oui	18.81	31.15	15.36	202	4.55
F4-v1. Depuis Nov. 1989, avez-vous déjà dû changer de métier	Non	11.30	86.89	71.33	938	4.13
	Oui	15.00	41.80	25.86	340	3.95
F13-v1. Intention de vous installer en Allemagne de l'ouest	Pas pour le moment	11.61	72.13	56.65	745	3.59
	Très grave	13.39	49.18	34.07	448	3.52
F84-v1. Les Roumains sont-ils trop nombreux en Allemagne	Oui	15.81	30.33	17.79	234	3.47
	Sans réponse	21.35	15.57	6.77	89	3.47
Q6-v1. Pour quel parti avez-vous voté le 14 Octobre	Sans réponse	18.60	19.67	9.81	129	3.38
F17-v1. Cette difficile période d'adaptation va durer pour vs	9 ans et plus	15.49	28.69	17.19	226	3.23
	Pas fait de choix	16.56	22.13	12.40	163	3.07
Q37-v1. Pour quel parti pensez vous voter le 2 décembre	Niveau 1 (bas)	18.69	16.39	8.14	107	3.05
	Sans réponse	13.23	42.62	29.89	393	3.05
Q35-v1. Niv. Soc. estimé dans société de l'Allemagne unifiée	Plus de 1000 DM	13.96	35.25	23.42	308	3.01
F29b-v1. Aux élections, la force de l'Allemagne unie comptera peu	Plus de 1000 DM	24.00	9.84	1.80	50	3.00
F16-v1. Rattraper compétitivité même si conséquences sociales	Plutôt pas d'accord					
	Plutôt pas					
Q37-v1. Pour quel parti pensez vous voter le 2 décembre	Sans réponse					

Opinions à la Vague 2		FOURCENTAGES		POIDS	V.TEST	
		CLASSE 5 / 10	GLOBAL			
classe 5/10				9.28	122	
F6-v2. Pensez-vs personnellement avoir à quitter votre région	Oui	22.86	25.41	10.27	135	4.98
	Non	20.24	27.87	12.78	168	4.62
F18-v2. A travail égal, Salaire en RDA quel % du salaire RFA	90 %	11.72	84.43	66.84	879	4.47
F207-v2. Avez vs cette année pris une nouvelle activité prof	Non	14.26	54.92	35.74	470	4.45
F84-v2. Les Roumains sont-ils trop nombreux en Allemagne	Oui	14.88	44.26	27.60	363	4.06
	Non	14.22	49.18	32.09	422	4.03
F41-v2. Avez vous confiance dans le Bundestag	Plutôt pas	21.50	18.85	8.14	107	3.90
	Si aucun	17.75	24.59	12.85	169	3.63
F87-v2. Pour quel parti avez vs voté aux élections du 2 déc.	N'a pas voté	11.34	80.33	65.70	864	3.61
F25-v2. Depuis Nov. 1989, avez-vous déjà dû changer de métier	Non	14.06	43.44	28.67	377	3.57
	normale	13.83	39.34	26.39	347	3.19
F13-v2. Intention de vous installer en Allemagne de l'ouest	Pas pour le moment	15.09	28.69	17.64	232	3.08
	Niveau 1	12.71	49.18	35.89	472	3.06
F39-v2. Niv. Soc. estimé dans société de l'Allemagne unifiée	Niveau 1					
	Plutôt pas					
F42-v2. Avez vous confiance dans le gouvernement fédéral	Plutôt pas					

En étudiant la partition complète, on constate que les groupes structurels de l'échantillon ont des opinions assez typées dans certains domaines. Cependant la segmentation de la population sur ces descripteurs socio-économiques ne permet pas de définir des profils d'opinions complets sur toute la gamme des questions de l'enquête.

On note aussi qu'il y a assez peu de modifications très typiques des opinions d'une vague à l'autre en se référant à cette segmentation de la population. Quelques faits sont notables cependant pour la classe 5 par exemple : alors qu'ils n'envisageaient pas de s'installer à l'Ouest en vague 1 ("vt4", ie : valeur-test égale à 4), ils envisagent de le faire s'il n'y a aucun mieux en vague 2 (vt4). En vague 2 apparaît l'item "plutôt pas confiance Bundestag" (vt4) ni "pour le gouvernement fédéral" (vt3).

En conclusion la typologie socio-administrative en 10 classes basée sur le signalétique fournit une bonne description de la structure du panel. Mais les changements d'opinions entre les deux vagues ne semblent pas pouvoir s'analyser en mouvements cohérents internes à ces segments de population.

On en conclut en particulier qu'il y aura certainement avantage à étudier les mouvements d'opinion de façon intrinsèque, sans s'appuyer sur la typologie exhibée. Nous allons prendre pour exemple le thème "Politique".

7.5.2 Analyse du thème "Politique" et de son évolution

Le tableau 3 donne la liste des variables choisies pour définir le thème "POLITIQUE". Pour réaliser les opérations du thémascope (analyse des correspondances multiples et typologie), on constitue provisoirement un échantillon fictif qui rassemble les 1315 enquêtés de la 1ère vague et les mêmes 1315 individus de la 2ème vague, comme s'il s'agissait d'échantillons indépendants. On dispose donc de 2630 individus dont on va comparer l'ensemble des réponses sur les variables du thème "politique".

Concrètement il s'agit d'un tableau numérique à 2630 lignes et 11 colonnes. Géométriquement il s'agit d'un nuage de 2630 points dans l'espace de toutes les réponses possibles aux questions "politiques".

Mr "Müller-vague 2" se trouve près du même Mr "Müller-vague 1" s'il a donné les mêmes réponses sur le thème politique dans les deux vagues. Sinon il s'en éloigne en vague 2 et se trouvera plus proche des individus qui ont répondu comme lui en vague 2.

L'analyse des Correspondances Multiples du tableau à 2630 lignes élabore la configuration du nuage des 2630 points.

On synthétise cette configuration en constituant des classes homogènes d'individus par classification automatique. La typologie fera apparaître 6 classes de réponses homogènes parmi les 2630 réponses.

Tableau 3

Variables actives du thème "Politique"

- . Postes politiques en RDA occupés par All. de l'ouest
- . Vous intéressez-vous à la politique
- . Etes-vous habituellement proche d'un parti
- . Niveau intervention Etat dans économie
- . Opinion sur l'URSS
- . Opinion sur les Etats-Unis
- . Confiance dans le Bundestag
- . Confiance dans le gouvernement du Land
- . Confiance dans la cour constitutionnelle
- . Confiance dans le syndicat
- . Opinion sur les 40 ans dévelop. en RDA

La figure 3 montre comment est réalisée la première étape du calcul qui consiste à faire l'analyse et la typologie sur l'empilement des deux vagues du panel.

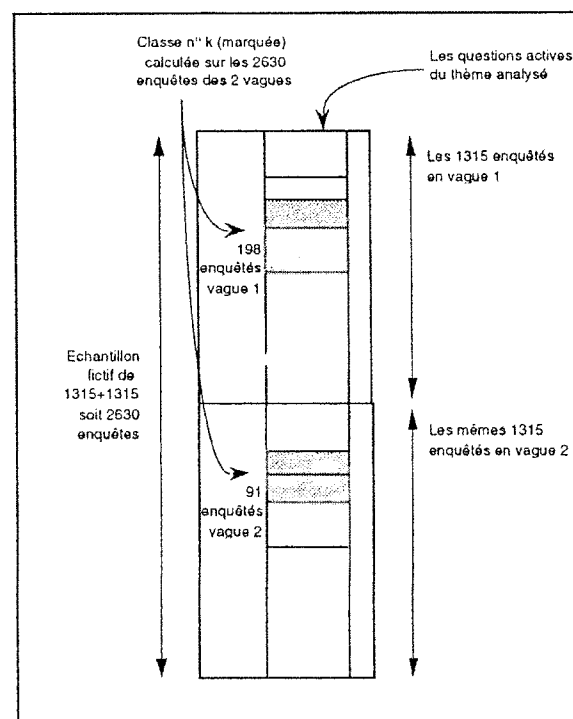


Figure 3

Analyse et typologie

La figure 4 illustre la position de la classe n° k dans cet échantillon fictif. Il y a 198 lignes tirées du tableau de la vague 1 auxquels s'ajoutent 91 lignes tirées du tableau de la vague 2. Mais en réalité les individus correspondant à ces lignes, en tant que personnes, ne sont pas tous distincts.

Lorsque les classes sont constituées, les réponses peuvent en effet être distinguées selon qu'elles ont été données en vague 1 ou en vague 2. La figure 5 en montre le mécanisme. Parmi les 289 individus de la classe, il y a 198 réponses émises à la vague 1. Certains de ces 198 individus se retrouvent dans la même classe en vague 2 et ils sont donc comptés parmi les 91 de la vague 2 (les autres se sont répartis dans les classes 2 à 6 pour la vague 2). Et parmi les 91 de la vague 2 se trouvent aussi des individus qui étaient dans une des classes 2 à 6 lors de la vague 1.

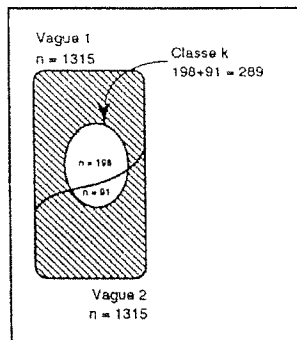


Figure 4
Fabrication d'une classe

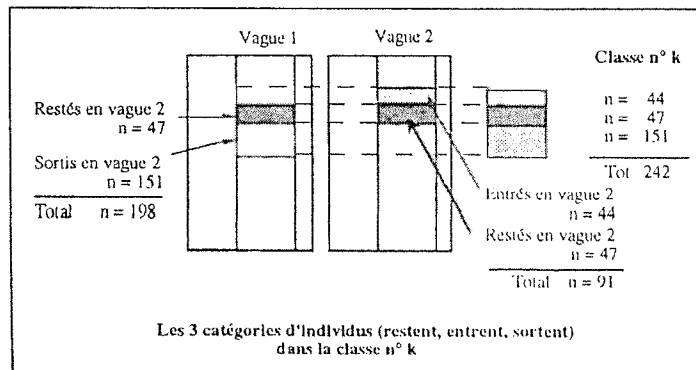


Figure 5
Composition d'une classe

7.5.3 Caractérisation des classes du thème "politique"

Le tableau 4 donne les éléments de fabrication des 6 classes de la typologie "politique".

Tableau 4

Les 6 classes de la typologie "politique"
établies sur l'échantillon empilé des 2 vagues

	Effectifs des classes		Répartition Vague 1		Répartition Vague 2	
classe 1	289	11%	198	15%	91	7%
classe 2	918	35%	492	37%	426	32%
classe 3	596	23%	222	17%	374	28%
classe 4	216	8%	73	6%	143	11%
classe 5	327	12%	124	9%	203	15%
classe 6	284	11%	206	16%	78	6%
TOTAL	2630	100%	1315	100%	1315	100%

En réalité, les classes qui nous intéressent sont des classes définies par des individus *physiques* et non par les *lignes* des tableaux de données. Chaque classe d'individus sera définie par les individus qui ont appartenu à la classe de lignes concernées au moins une fois : soit l'individu y était en vague 1, soit il y est entré en vague 2, soit il y était aux deux vagues.

La définition de ces classes résulte facilement du tableau 5 qui détaille les échanges entre les deux vagues.

Tableau 5

Les échanges entre vagues

EN LIGNE : Partition selon comportement politique en vague 1
EN COLONNE : Partition selon comportement politique en vague 2

	Classe 1/6 (vague 1)	Classe 2/6 (vague 1)	Classe 3/6 (vague 1)	Classe 4/6 (vague 1)	Classe 5/6 (vague 1)	Classe 6/6 (vague 1)	ENSEMBLE (vague 1)
Classe 1/6 (vague 2)	47	78	26	14	25	8	198
Classe 2/6 (vague 2)	25	240	120	32	57	18	492
Classe 3/6 (vague 2)	2	35	121	45	17	2	222
Classe 4/6 (vague 2)	5	8	27	25	8	2	73
Classe 5/6 (vague 2)	6	27	23	12	44	12	124
Classe 6/6 (vague 2)	8	38	57	15	52	36	206
TOTAL (vague 2)	91	426	374	143	203	78	1315

On résume dans le tableau 6 le volume des classes obtenues.

Tableau 6
Définition des classes à caractériser

	Effectif	% de 1 315
classe 1	242	18.4%
classe 2	678	51.6%
classe 3	475	36.1%
classe 4	191	14.5%
classe 5	283	21.5%
classe 6	248	18.9%

Ainsi définies, on remarquera que les 6 classes ne constituent plus une "partition" des 1 315 individus. En effet les individus qui changent de classe sont comptés deux fois.

Dans la dernière phase du thémascope, on se propose d'étudier les caractéristiques statistiques des classes homogènes de réponses obtenues dans cette typologie.

Pour caractériser chaque classe, on définit la variable d'appartenance à la classe (oui/non) et on compare les particularités de ceux qui sont dans la classe avec celles des autres individus (procédure DEMOD sur chaque variable d'appartenance, réalisée sur les 1 315 individus).

Pour la caractérisation de la classe 1 par exemple (242 individus soit 18.4 % du panel), on trouvera successivement :

- Caractérisation par les questions "actives" ;
- Caractérisation par le signalétique ;
- Caractérisation par les opinions vague 1 ;
- Caractérisation par les opinions vague 2.

Thème Politique (questions actives)

Ils sont "très favorables" à l'occupation de postes politiques par les Allemands de l'Ouest (cette question est très déterminante dans la typologie globale). Ils ont tout à fait confiance dans le Bundestag, la cour constitutionnelle, le gouvernement du Land. Ils considèrent que les 40 ans de RDA sont années perdues. Ils sont très proches d'un parti (le CDU). Ils ont bonne opinion des USA, et mauvaise de l'URSS. L'état doit intervenir "peu".

Thème Signalétique

Niveau scolaire et qualification basse. Sur-représentation des ouvriers, des hommes, et quelques retraités, des inactifs. Classe du signalétique : 1/10.

On voit ici comment les deux typologies s'enrichissent en montrant ici la conjonction de deux classes réalisées sur des thèmes différents.

Thème des opinions : Opinions vague 1 et vague 2

Proches du CDU et éloignés du PDS. Confiance très marquée dans le gouvernement fédéral. Confiance dans les valeurs traditionnelles et les valeurs "de l'ouest".

Tableau 7

CLASSE 1 : Caractérisation par le thème Actif

MODALITES CARACTERISTIQUES	POURCENTAGES			POINTS	V.TEST	
	CLA/MOD	MOD/CLA	GLOBAL			
F8-v1. Postes politiques en RDA occupés par All. de l'ouest	Favorable (très)	90.61	79.75	16.20	213	26.77
F41-v1. Avez vous confiance dans le Bundestag	Tout à fait	46.77	35.95	14.14	186	9.69
F41-v1. Avez vous confiance dans le Bundestag	Oui	24.94	89.26	65.86	866	9.08
F44-v1. Avez vous confiance dans la cour constitutionnelle	Tout à fait	34.24	36.36	19.54	257	6.81
F86-v1. Etes vous habituellement proche d'un parti	Très proche	34.96	32.64	17.19	226	6.54
F74-v1. Opinion à l'égard des 40 ans de dévelop. de la RDA	Années perdues	35.64	29.75	15.36	202	6.33
RF33-v1 Opinion sur les Etats-Unis	Bonne	23.48	75.21	58.94	775	5.76
F43-v1. Avez vous confiance dans le gouvernement du Land	Tout à fait	33.67	27.27	14.90	196	5.53
RF74-v1 Opinion sur les 40 ans dévelop. en RDA	Négatif	24.83	61.16	45.32	596	5.40
RF44-v1 Confiance dans la cour constitutionnelle	Oui	22.76	72.31	58.48	769	4.85
F33-v1. Quelle est votre opinion sur les Etats-Unis	Très bonne	33.58	19.01	10.42	137	4.43
RF30-v1 Opinion sur l'URSS	Mauvaise	25.77	38.02	27.15	357	4.03
F8-v2. Postes politiques en RDA occupés par All. de l'ouest	Favorable (très)	91.84	37.19	7.45	98	16.72
F41-v2. Avez vous confiance dans le Bundestag	Tout à fait	59.78	22.73	7.00	92	9.14
F41-v2. Avez vous confiance dans le Bundestag	Oui	27.15	76.86	52.09	685	8.67
F74-v2. Opinion à l'égard des 40 ans de dévelop. de la RDA	Années perdues	43.92	26.86	11.25	148	7.62
RF33-v2 Opinion sur les Etats-Unis	Bonne	24.08	78.10	59.70	785	6.59
F33-v2. Quelle est votre opinion sur les Etats-Unis	Très bonne	40.88	23.14	10.42	137	6.45
F44-v2. Avez vous confiance dans la cour constitutionnelle	Tout à fait	33.90	33.06	17.95	236	6.31
RF74-v2 Opinion sur les 40 ans dévelop. en RDA	Négatif	26.73	54.13	37.26	490	5.85
F86-v2. Etes vous habituellement proche d'un parti	Très proche	34.87	21.90	11.56	152	5.10
F43-v2. Avez vous confiance dans le gouvernement du Land	Tout à fait	24.19	21.90	11.79	155	4.95
RF43-v2 Confiance dans le gouvernement du Land	Oui	22.47	74.38	60.91	801	4.78

CLASSE 1 : Caractérisation par le SIGNALÉTIQUE

MODALITES CARACTERISTIQUES	POURCENTAGES			POINTS	V.TEST	
	CLA/MOD	MOD/CLA	GLOBAL			
Niveau scolaire de l'enquêté(e)	Inf. 10ème CL.	24.89	47.11	34.83	458	4.30
F209-v2. Profession exercée	Ouvrier	25.18	43.80	32.02	421	4.20
Qualification de l'enquêté(e)	Ouv. Qual. Cont.	21.98	66.12	55.36	728	1.69
F209-v1. Profession exercée	Ouvrier	23.92	41.32	31.79	418	3.40
Sexe de l'enquêté(e)	Homme	21.96	57.44	48.14	633	3.14
partition socio-dém en 10 classes	classe 1/10	29.91	14.46	8.90	117	3.08
Situation d'emploi de l'enquêté(e)	Retraité inactif	24.36	27.69	20.91	275	2.72

CLASSE 1 : Caractérisation par les opinions vague 1

MODALITES CARACTERISTIQUES	POURCENTAGES			POINTS	V.TEST	
	CLA/MOD	MOD/CLA	GLOBAL			
F7-v1. Postes économiques en RDA occupés par All. de l'ouest	Favorable (très)	55.64	61.16	20.23	266	15.97
F42-v1. Avez vous confiance dans le gouvernement fédéral	Tout à fait	47.92	38.02	14.60	192	10.26
F16-v1. Rattraper compétitivité même si conséquences sociales	Tout à fait d'accord	42.23	43.80	19.09	251	9.95
Q32-v1. Quel coalition gouvernementale préféreriez vous	CDU/CSU/FPD	30.59	55.37	33.31	438	7.79
F90-v1. De quel parti vous sentez vous le plus éloigné	PDS	29.57	56.83	31.63	416	6.84
F89-v1. De quel parti vous sentez vous le plus proche	CDU	29.43	48.76	30.49	401	6.56
Q37-v1. Pour quel parti pensez vous voter le 2 décembre	CDU	41.41	21.90	9.73	128	6.34
F47-v1. Avez vous confiance dans la télévision	Tout à fait	41.41	21.90	9.73	128	6.34
Q35-v1. Pour quel parti avez-vous voté le 18 Mars	CDU	28.05	51.24	33.61	442	6.21
F58-v1. Le respect des personnes âgées est une valeur qui ...	Comptera plus	25.75	60.33	43.12	567	5.89
F63-v1. La sécurité sociale des citoyens est une valeur qui ...	Comptera plus	25.39	61.16	44.33	583	5.75
Q36-v1. Pour quel parti avez-vous voté le 14 Octobre	CDU	28.95	40.91	26.01	342	5.58
F1-v1. Auquel de ces lieux avez-vous le sentiment d'appartenir	Europe	33.16	26.86	14.90	196	5.35
F27-v1. Quelle est votre opinion sur le Japon	Très bonne	29.62	35.12	21.80	287	5.24
F75-v1. Selon vous, Y-a-t-il trop d'immigrés en Allemagne	Oui	22.18	80.58	66.84	879	5.12
F40-v1. Conséquences changements dans le pays sur votre vie	Je serais mieux	23.70	69.42	54.83	721	5.04

CLASSE 1 : Caractérisation par les opinions vague 2

MODALITES CARACTERISTIQUES	----- POURCENTAGES -----			POIDS V. TEST		
	CLA/MCO	MCO/CLA	GLOBAL			
F7-v2. Postes économiques en RDA occupés par All. de l'Ouest	Favorable (très)	62.71	30.58	8.97	138	11.25
F12-v2. Avez-vous confiance dans le gouvernement (fédéral)	Tout à fait	60.20	24.38	7.45	98	9.57
F89-v2. En cas d'élections dimanche prochain, votez-vous pour	CDU	37.40	39.26	19.32	254	8.06
F89-v2. De quel parti vous sentez-vous le plus proche	CDU	35.83	35.54	18.25	240	7.15
F92-v2. De quel parti vous sentez-vous le plus éloigné	PDS	31.72	43.39	25.17	331	6.85
F92-v2. Financement de l'Allemagne à la guerre du Golfe	Tout à fait normal	47.31	18.18	7.07	93	6.56
F87-v2. Pour quel parti avez-vous voté aux élections du 2 oct.	CDU	28.95	50.00	31.79	418	6.49
F24BIS-v2. Pour les chômeurs, efficace de s'adresser au	CDU	32.62	37.60	21.22	279	5.48
F17-v2. Avez-vous confiance dans la télévision	Tout à fait	49.38	16.53	6.16	81	6.48
F1-v2. Auquel de ces lieux avez-vous le sentiment d'appartenir	Allemagne unifiée	37.34	24.38	12.02	158	5.97
F17-v2. Opinion sur réalisation du 1er marché européen en 93	Très bonne chose	29.67	41.32	25.63	337	5.89
F18-v2. Avez-vous confiance dans le patronat	Tout à fait	58.97	9.50	2.97	39	5.59
F58-v2. Le respect des personnes âgées est une valeur qui	Comptera plus	26.41	50.41	35.13	462	5.35
F15-v2. Avez-vous confiance dans la police	Tout à fait	40.00	17.36	7.98	105	5.33

7.5.4 Thémascope et panel

On inclut ici l'étude des mouvements des individus d'une classe dans une autre entre les deux vagues.

Le tableau 5 donne les flux d'entrée et de sortie de chaque classe. On y voit que les éléments diagonaux (individus restés dans la même classe) ne sont pas toujours importants vis-à-vis des marges (volume des classes dans chaque vague) : on atteint très rarement 50 %. Il y a donc des mouvements considérables d'entrée/sortie des classes d'opinion.

Tableau 8

Synthèse des entrées/sorties

	Classe 1/6	Classe 2/6	Classe 3/6	Classe 4/6	Classe 5/6	Classe 6/6
Restent (V1 et V2)	47	240	121	25	44	36
Sortent après V1	151	252	101	48	80	170
Entrent en V2	44	186	253	118	159	42
cumul "concernés"	242	678	475	191	283	248
"non concernés"	1073	637	840	1124	1032	1067
TOTAL	1315	1315	1315	1315	1315	1315

Le tableau 8 donne une idée plus précise du volume de ces mouvements. Pour la classe 1 par exemple, chacun des individus est affecté dans une des catégories suivantes :

- "Reste" En vague 2, il est resté dans cette classe ;
- "Sort" En vague 2, il est sorti de cette classe ;
- "Entre" En vague 2, il est entré dans cette classe alors qu'il n'y était pas en vague 1 ;
- "Non concerné" Il n'a jamais été dans cette classe.

L'ensemble des individus "concernés" par cette classe est cumulé pour définir la classe en question parmi les 1 315 individus du panel. Caractériser une classe, c'est donc trouver les particularités statistiques des individus "concernés" par la classe. On en trouve 242 dans la classe 1. On peut, de la même façon, caractériser les sous-groupes d'individus constitués par ceux qui "entrent", "restent" ou "sortent" d'une classe.

7.5.5 Caractérisation des mouvements dans la classe 1

Caractérisation globale de la classe "Politique" numéro 1

Rappelons qu'ils sont "très favorables" à l'occupation de postes politiques par les Allemands de l'Ouest (cette question est très déterminante dans la typologie globale). Ils ont tout à fait confiance dans le Bundestag, la cour constitutionnelle le gouvernement du Land. Ils considèrent que les 40 ans de RDA sont des années perdues. Ils sont très proches d'un parti (CDU). Ils ont bonne opinion des USA et mauvaise de l'URSS. L'état doit intervenir "peu".

Sont restés en classe 147 individus 19.4 %

Comparés aux autres individus de la classe 1, ils sont "très favorables" à l'occupation par l'Ouest de postes politiques en RDA. Ils ont "tout à fait" confiance dans le Bundestag et le gouvernement fédéral (surtout en vague 2). Ils se sentent éloignés du PDS (en vague 2).

Sont sortis de la classe 1151 individus 62.4 %

Ils vont essentiellement aller en classe 2, et certains en classe 3 et 5 (les classes des "sans réponse"). Parmi ceux de la classe 1, on trouve une légère sur-représentation des personnes âgées. Comparés aux autres de la classe 1, ils étaient "très favorables" en vague 1 à l'occupation par l'Ouest de postes politiques en RDA, et sont devenus "plutôt favorables" ou même "défavorables" en vague 2. Sont sortis ceux qui, bien que dans la classe 1, ont en vague 2 une opinion mauvaise des USA, ou ont peu confiance dans le gouvernement fédéral.

Sont entrés en classe 144 individus 18.2 %

Ils étaient essentiellement en classe 2 en vague 1, et un peu des "sans réponses" dans la vague 1. Comparés aux autres de la classe 1, ils étaient plus ou moins "favorables" en vague 1 à l'occupation de postes politiques par les Allemands de l'Ouest. Ils sont davantage favorables en vague 2.

Tableau 9

Caractérisation des changements dans la classe 1 de la typologie "Politique"

MODALITÉS CARACTÉRISTIQUES		----- POURCENTAGES -----		POIDS	V. TEST	
		CLA/MED	MED/CLA	GLOBAL		
Reste				19.42	47	
Mouvement de la classe 1 (typologie politique)	Reste	100.00	100.00	19.42	47 99.99	
F8-v1. Postes politiques en RDA occupés par All. de l'ouest	Favorable (très)	23.83	97.87	79.75	193 3.70	
F12-v1. Élément devant le HDIRS intervenir dans prog. salaire	Diplômes	56.25	19.15	6.61	16 3.17	
F8-v2. Postes politiques en RDA occupés par All. de l'ouest	Favorable (très)	48.89	93.62	37.19	90 8.95	
F7-v2. Postes économiques en RDA occupés par All. de l'ouest	Favorable (très)	40.54	63.83	30.58	74 5.15	
F11-v2. Avez-vous confiance dans le Bundestag	Tout à fait	41.82	48.94	22.73	55 4.33	
F90-v2. De quel parti vous sentez-vous le plus éloigné	FDS	32.38	72.34	43.39	105 4.31	
F42-v2. Avez-vous confiance dans le gouvernement (fédéral)	Tout à fait	35.59	44.58	24.38	59 3.28	
Sort				62.40	151	
Mouvement de la classe 1 (typologie politique)	Sort	100.00	100.00	62.40	151 17.52	
Age de l'enquêté(e)	60 ans et plus	76.79	28.48	23.14	56 2.43	
F8-v1. Postes politiques en RDA occupés par All. de l'ouest	Favorable (très)	76.17	97.25	79.75	193 8.78	
F206-v1. Situation d'emploi de l'enquêté(e)	Pétraité	77.78	27.81	22.31	54 2.55	
F46-v1. Avez-vous confiance dans les banques	Plutôt	70.49	56.95	50.41	122 2.49	
F8-v2. Postes politiques en RDA occupés par All. de l'ouest	Favorable (plutôt)	100.00	52.98	33.06	80 9.58	
F7-v2. Postes économiques en RDA occupés par All. de l'ouest	Favorable (plutôt)	81.98	60.26	45.87	111 5.78	
F8-v2. Postes politiques en RDA occupés par All. de l'ouest	Défavorable (plutôt)	95.00	25.17	16.53	40 4.96	
F33-v2. Quelle est votre opinion sur les États-Unis	Plutôt mauvaise	100.00	13.25	8.26	20 3.90	
F833-v2. Opinion sur les États-Unis	Mauvaise	92.11	15.88	10.74	26 3.38	
F23-v2. Grève efficace en matière d'emploi avec les salariés	Des cinq	Landers	74.07	52.98	44.63	108 3.26
F8-v2. Postes politiques en RDA occupés par All. de l'ouest	Vraiment contre	100.00	9.27	5.79	14 3.07	
Entre				18.18	44	
Mouvement de la classe 1 (typologie politique)	Entre	100.00	100.00	18.18	44 99.99	
F8-v1. Postes politiques en RDA occupés par All. de l'ouest	Favorable (plutôt)	94.12	72.73	14.05	34 10.81	
F7-v1. Postes économiques en RDA occupés par All. de l'ouest	Favorable (plutôt)	36.84	61.36	30.17	73 4.62	
F8-v1. Postes politiques en RDA occupés par All. de l'ouest	Défavorable (plutôt)	77.78	15.91	3.72	9 3.68	
F57-v1. La suppression des crèches d'entreprises serait ...	réponse manquante	100.00	11.36	2.07	5 3.59	
F8-v2. Postes politiques en RDA occupés par All. de l'ouest	Favorable (très)	47.78	97.73	37.19	90 9.38	
F7-v2. Postes économiques en RDA occupés par All. de l'ouest	Favorable (très)	37.84	63.64	30.58	74 4.90	

7.6 Conclusion : validation par redondance et accumulation

A l'issue des exploitations statistiques d'un fichier d'enquête, on se trouve devant des groupes d'individus dont on commence à comprendre le comportement, les opinions, les motivations ou les intentions.

Ces groupes se trouvent caractérisés par plusieurs batteries d'éléments :

- le rangement des éléments du signalétique qui les différencie statistiquement du reste de la population (différences significatives de proportions, écarts significatifs de moyennes),
- le tri complet des informations subjectives et réponses aux questions du questionnaire, sur tous les thèmes de l'enquête, en commençant par les plus caractéristiques du groupe.

Une source importante de richesse des données d'enquête est à attendre de leur volume. La cohérence interne émerge peu à peu de l'accumulation d'informations redondantes.

Le critère des *valeurs-tests* participe lui-même à l'interprétation des résultats des analyses exploratoires et simultanément à leur classement.

La conviction du statisticien devant ce matériau ne s'appuie plus sur un ou quelques tests statistiques isolés. On connaît d'ailleurs la fragilité des tests statistiques usuels sur un corpus amorphe de données d'enquête.

Une forme plus convaincante de la "vérité" de l'interprétation émerge de l'accumulation d'informations cohérentes, complémentaires, qui s'éclairent les unes les autres selon les différents angles-de-vue ou thèmes scrutés avec le thémascope.