

UNE UTILISATION DE LA METHODE DES NUEES
DYNAMIQUES LORS D'UNE ENQUETE SOCIO-DEMOGRAPHIQUE

Par Philippe ANTOINE
et Agnès GUILLAUME

La tentation est toujours grande, lors d'une enquête statistique, de multiplier les questions afin de saisir le mieux possible la réalité sociale. Or, il est souhaitable qu'une enquête portant sur un échantillon conséquent d'individus ne comporte qu'un nombre restreint de questions (contraintes de temps, de coût, de fiabilité des réponses, ...). Afin de mieux comprendre et expliquer les données recueillies, il semble préférable de se livrer à des investigations plus approfondies auprès de certains enquêtés. Divers critères peuvent être envisagés : tirer au hasard un individu sur 100 ou 1 000, retenir les personnes les plus communicatives, procéder à un choix raisonné, etc. Ces différents modes de tirage d'un sous-échantillon ne sont guère satisfaisants car aucun moyen ne permet de connaître le degré de représentativité des individus sélectionnés.

Certaines méthodes d'analyse de données peuvent aider dans ce choix et la méthode des nuées dynamiques utilisée dans cet exemple présente de nombreux avantages qui justifieraient son utilisation plus fréquente lors de l'exploitation et de l'analyse des enquêtes démographiques et statistiques portant sur des effectifs importants.

1 - OBJECTIF INITIAL : PASSER DU QUANTITATIF AU QUALITATIF

A partir des données recueillies à Abidjan lors de l'enquête ivoirienne à passages répétés (EPR) (1) nous comptons prolonger

(1) Cette enquête a été réalisée et financée par la Direction de la Statistique de Côte d'Ivoire (01 BP V 55 - ABIDJAN 01). Les premiers résultats sont disponibles auprès de ce service. Pour une présentation de l'enquête d'Abidjan, voir aussi ANTOINE **O.R.S.T.O.M. Fonds Documentaire** (1983).

B 15.972

Paris, ORSTOM, 1984

30 OCT. 1984 No : 15.972

Cote : B

Statéco Sept. 84 ?

l'analyse des relations familiales en essayant de mieux cerner l'espace de vie des chefs de ménage tant dans les liens entretenus avec le milieu d'origine que dans le champ des relations familiales en ville. Une démarche plus anthropologique s'imposait, il restait donc à choisir les individus susceptibles d'être enquêtés. L'hypothèse initiale émise est la suivante : il existe des comportements différenciés selon les types d'habitat et il convient donc de distinguer les chefs de ménage selon cette catégorie et également selon le type de famille auquel ils appartiennent. L'habitat nous semble être une variable assez synthétique traduisant bien la diversité du fait abidjanais. Quatre types peuvent être distingués :

- l'habitat de haut et moyen standing (ou résidentiel) qui regroupe les catégories sociales les plus aisées ;

- l'habitat économique moderne, où résident surtout des salariés qualifiés du secteur moderne ;

- l'habitat évolutif, ou cour, constitue l'habitat le plus répandu et le plus ancien de l'agglomération ; chaque cour comprend plusieurs logements ;

- l'habitat sommaire, comprend des constructions en bois de palme et surtout des baraques. Il est malgré sa précarité relativement bien structuré et doit être considéré comme un habitat de substitution à l'habitat évolutif.

Les individus retenus doivent donc être représentatifs des différentes modalités d'une typologie combinant catégories d'habitat, formes de ménage (famille nucléaire, famille élargie, groupement de célibataires et ménage d'une personne) et incluant d'autres variables pouvant influencer la typologie (CSP, âges, ...). Notre but est de sélectionner des individus, a priori représentatifs des différentes classes, que nous devons ensuite retrouver sur le terrain et interviewer pour mieux connaître à la fois l'évolution de la composition de leur famille, leur réseau familial en ville et les modalités de leur installation à Abidjan. Par ailleurs, cette nouvelle investigation amène à appréhender le devenir de la population cinq ans plus tard (1978 - 1983).

2 - LE CHOIX DE LA METHODE

Les méthodes de classification automatique effectuent une partition d'un ensemble de données mais peu d'entre elles permettent de dégager les individus représentatifs de cette typologie. Différentes approches sont possibles comme rechercher une partition qui minimise un certain critère, selon un certain seuil ; ou bien retenir des techniques basées sur l'agrégation de groupes par itérations successives.

Nous avons retenu les méthodes de partitionnement par agrégation à partir de noyaux (1). En effet, elles permettent de traiter de grands ensembles de données (plusieurs milliers d'individus).

Deux méthodes de "partitionnement" ont été prises en considération :

- la méthode des boules optimisées ;
- la méthode des nuées dynamiques.

Ces deux méthodes sont des algorithmes d'agrégation autour de noyaux variables. La différence essentielle entre elles étant le mode de choix des noyaux à partir duquel on agrège.

Dans le cas des boules optimisées, l'agrégation se fait en fonction d'une distance inférieure à un rayon fixé a priori (mais dont le choix peut être guidé). Si la méthode dégage facilement des typologies, elle ne permet pas d'extraire automatiquement les individus les plus représentatifs. Le nombre de classes de la typologie est très sensible à une variation même minime du rayon.

(1) Les méthodes de classification hiérarchique (méthodes ascendantes) procèdent par agrégations successives (selon un critère de distance) d'éléments isolés au départ, à la différence d'autres méthodes de classification qui agrègent selon un noyau central fixé a priori. Ces méthodes hiérarchiques n'ont pu être utilisées pour un fichier aussi important que le nôtre (7 800 individus x 60 modalités).

La méthode des nuées dynamiques mise au point par E. DIDAY (1971) est la mieux adaptée à nos objectifs. Elle permet de "générer" une partition à partir d'un corps de données sur lequel on ne demande pas de faire d'hypothèses a priori. Cette partition doit réaliser les deux propriétés suivantes :

- les individus de chaque partie se ressemblent le plus possible ;
- les individus de deux parties différentes se ressemblent le moins possible (DIDAY, p. 20).

Le processus est itératif et à chaque étape la qualité de la partition s'améliore. Le nombre de classes souhaité est déterminé a priori ainsi que le nombre d'éléments centraux désiré, c'est-à-dire le nombre d'éléments au centre du noyau qui seront énumérés. Au départ, un ensemble de points ou noyaux d'une classe peut être tiré au hasard. Autour de ces points se regroupent les éléments les plus proches pour former une partition. La distance calculée par rapport au centre de classe est la distance du Chi-Deux. A partir de cette partition créée, une autre famille de noyaux est définie, elle regroupe les points les plus proches formant une nouvelle classe et ainsi de suite jusqu'à obtention d'un nombre fini de classes. Si, après un certain nombre d'itérations, les classes formées sont stables, les données sont dites "classifiables" et constituent des "formes fortes". Les individus qui changent de classes selon les tirages sont les "individus charnières".

3 - L'APPLICATION

L'application présentée ici a été réalisée à l'aide des logiciels mis à la disposition des utilisateurs par l'ADDAD (1).

(1) ADDAD : Association pour le Développement et la Diffusion de l'Analyse des Données

Laboratoire du Professeur BENZECRI.

Siège social : Laboratoire de Statistique
4 rue André Chenier
92130 ISSY-les-MOULINEAUX

Un fichier "chef de ménage" a été établi à partir des données de l'EPR (strate Abidjan). Il comprend environ 7 800 chefs de ménage africains résidents au 1er passage de l'enquête et seulement une dizaine de variables démographiques et socio-économiques parmi les plus discriminantes (âge, situation matrimoniale, instruction, profession, habitat, type de famille, ...) sont retenues pour l'analyse typologique. Les modalités de ces variables sont tout d'abord regroupées et transformées en variables disjonctives (code 0 ou 1), c'est-à-dire qu'une variable ayant n modalités a donné naissance à n variables disjonctives. Cette transformation s'opère très facilement avec le programme RECODE du logiciel utilisé. A l'issue de cette opération le fichier comprenait 60 variables disjonctives.

Lors de l'utilisation du programme intitulé NUEDYN, l'utilisateur doit seulement opérer le choix de quelques paramètres comme le nombre de classes désiré, le mode de tirage des individus étalons (tirage au hasard ou numéros introduits par l'utilisateur).

Nous avons, dans notre cas, procédé en deux temps. Dans une première phase, très empirique, les éléments étalons (au nombre de 10 pour chacune des classes) sont tirés au hasard ; différentes partitions sont ainsi obtenues permettant de dégager diverses catégories de chefs de ménage. Les individus sélectionnés au cours de ces différents essais servent de base au choix des éléments étalons utilisés lors de la phase ultérieure. Nous avons, dans un second temps, constitué 9 classes, de chacune 10 éléments étalons. A priori les définitions de chacune de ces phases sont les suivantes :

Type 1 : Chef de ménage famille élargie de l'habitat résidentiel

Type 2 : Chef de ménage masculin de plus de 65 ans résidant dans les cours

Type 3 : Chef de ménage famille élargie résidant dans l'habitat moderne

Type 4 : Chef de ménage famille élargie résidant dans les cours

Type 5 : Chef de ménage féminin

Type 6 : Chef de ménage famille nucléaire résidant dans l'habitat de cour

Type 7 : Groupement de célibataire cour et sommaire

Type 8 : Chef de ménage famille nucléaire résidant dans l'habitat sommaire

Type 9 : Ménage de une personne résidant dans l'habitat de cour ou le sommaire.

Cette typologie est loin d'être fine mais elle couvre les principales catégories perceptibles à Abidjan. Les résultats obtenus se présentent sous forme de tableaux (Cf page suivante). Le premier donne la liste des individus étalons des noyaux. Le second tableau, le plus complet, donne pour chaque classe son effectif et la liste exhaustive des numéros des individus appartenant à chacune des classes. Le troisième tableau présente la liste des individus centraux de chacune des classes, correspondant à la sélection désirée (10 individus par classe dans notre cas). Deux tableaux d'aide à l'interprétation sont fournis également (non présentés ici) : une mesure d'homogénéité des classes (variance) et une mesure de la valeur du critère pour chaque partition obtenue. Enfin, quand la convergence est atteinte, un tableau des individus des formes fortes figure en fin de listing.

Il faut bien distinguer entre les éléments centraux obtenus en sortie et l'ensemble des individus composant la classe. Un simple retour au fichier permet de connaître les caractéristiques des individus tirés, qui ont pour une même classe des caractéristiques voisines. Par contre, l'ensemble des individus composant une même classe peut être plus hétérogène.

Le tableau 2 donne la répartition des 9 classes selon l'habitat et selon la catégorie socio-professionnelle. Ainsi le type 1 regroupe essentiellement des cadres, indépendamment de l'habitat où ils résident (une minorité logeant dans les cours plus traditionnelles). Dans ce cas, le statut professionnel prime sur le type de logement de résidence. De même, pour la classe 7 qui regroupe exclusivement des chefs de ménage apprentis ou aides familiaux. Il s'agit le plus souvent dans ce cas de jeunes qui logent dans la boutique de leur patron. A l'inverse, la classe 8 rassemble essentiellement des chefs de ménage résidant dans l'habitat sommaire quel que soit le type de famille auquel ils appartiennent. Après analyse des différents tableaux, il s'avère que certains types doivent comporter des intitulés différents. C'est le cas pour les types :

1.1.	10.1	5049	5965	5953	5755	5186	3409	5750	5188	6638	5212
2.1.	10.1	653	1566	7894	593	6079	1554	1671	9405	6662	2698
3.1.	10.1	3610	2142	3594	7221	3617	2361	7202	3613	2130	2464
4.1.	10.1	9282	8606	8260	7890	8012	9179	654	21	179	8567
5.1.	10.1	8872	8785	5499	605	8194	184	2917	1550	1488	8532
6.1.	10.1	2738	4456	3116	2991	2988	2795	2982	3098	3130	9321
7.1.	10.1	3464	2168	4470	3067	3065	1641	1416	1510	4456	4474
8.1.	10.1	4165	4030	4020	3899	3847	3831	4022	3809	4173	3966
9.1.	10.1	3481	3171	1542	3062	1511	2951	1676	2039	8951	2986

En entrée numéro des éléments étalons

* ITERATION NUMERO 1

INDIVIDUS DES CLASSES

1	266	72	91	99	167	221	237	243	254	256	257	259	260	264	274	279	285	287	293	295	298
		303	371	373	381	385	403	403	424	424	424	424	424	424	424	424	424	424	424	424	424
		1851	1929	1939	1945	1952	1962	2037	2139	2215	2338	2415	2466	2483	2486	2490	2510	2511	2513	2515	2519
		2520	2522	2565	2579	2595	2600	2603	2639	2662	2681	3009	3011	3013	3019	3051	3073	3316	3319	3320	3321
		3324	3337	3401	3403	3404	3408	3409	3414	3416	3603	3620	4275	4280	4292	4360	4376	4380	4494	4495	4504
		4506	4507	4508	4512	4519	4527	4539	4812	4815	4816	4817	4819	4820	4826	4827	4828	4831	4834	4836	4837
		4838	4844	4893	4896	4905	4907	4912	4917	4922	4923	4935	4943	5010	5020	5024	5025	5026	5030	5037	5039
		5047	5048	5049	5052	5053	5055	5059	5066	5073	5074	5076	5086	5094	5131	5133	5137	5138	5142	5144	5153
		5160	5164	5168	5170	5174	5176	5177	5183	5184	5186	5187	5188	5189	5197	5200	5202	5203	5205	5206	5207
		5210	5212	5213	5216	5217	5218	5219	5220	5222	5337	5688	5704	5705	5716	5721	5724	5727	5728	5738	5750
		5751	5753	5754	5757	5759	5775	5779	5780	5788	5790	5798	5919	5938	5948	5949	5953	5954	5956	5965	5969
		5970	5979	6494	6596	6598	6599	6605	6609	6610	6611	6618	6619	6622	6623	6624	6628	6633	6635	6636	6638
		6639	6806	7068	7220	7436	7471	7477	7501	7712	7746	7763	7765	7767	8016	8244	8296	8318	8319	8320	8321
		8322	8326	8327	8329	8727	8878														
2	124	126	171	196	308	355	426	452	593	623	643	647	653	667	684	692	712	719	765	903	1161
		1169	1175	1189	1193	1213	1220	1235	1247	1262	1270	1391	1396	1423	1476	1493	1496	1554	1566	1586	1632
		1639	1659	1671	1770	1782	1992	2021	2038	2060	2077	2232	2378	2684	2698	2763	2786	2851	2943	3426	3471
		3512	3518	3520	4127	4324	4387	4525	4559	4683	4974	5102	5132	5308	5309	5402	5607	5656	5714	5861	
		6079	6172	6189	6196	6198	6359	6418	6423	6428	6501	6515	6537	6538	6656	6662	6705	6711	6731	6737	
		6767	7294	7349	7421	7446	7857	7894	7934	7967	8098	8408	8419	8497	8738	8744	9013	9330	9378	9386	9396
		9397	9403	9405	9411																

En Sortie:
1) Individus par classe à chaque itération

INDIVIDUS DES NOYAUX

1.1.	10.1	5965	5953	4506	4504	5049	5753	5186	3409	8326	1952
2.1.	10.1	653	1566	7894	593	6079	1554	8419	9330	1671	1391
3.1.	10.1	2503	2501	2127	3613	2464	3580	2143	2130	7202	2330
4.1.	10.1	9282	8606	8260	7890	7800	654	179	21	8012	7983
5.1.	10.1	8872	8785	2917	5499	2291	8192	8194	8959	10	605
6.1.	10.1	1861	1469	1275	3116	2991	2988	2795	1881	9321	3130
7.1.	10.1	3464	2168	4470	3067	3075	8411	7840	1998	1846	1443
8.1.	10.1	4165	4030	4020	3899	3847	3831	3237	4173	4151	4103
9.1.	10.1	2018	2986	3481	1511	2039	1542	3062	8951	1676	3171

En sortie:
2) Numéro des individus sélectionnés dans chacune des classes, lors de la dernière itération.

TABLEAU 2

REPARTITION DE LA TYPOLOGIE SELON L'HABITAT ET LES CATEGORIES SOCIO-PROFESSIONNELLES

HABITAT \ TYPOLOGIE N. DYN.	1	2	3	4	5	6	7	8	9	TOTAL
Résidentiel	163	5	154	27	13	23	2	1	17	405
Economique Moderne	76	5	971	215	42	86	3	7	44	1 449
Cour	25	106	390	1 905	230	1 618	90	0	623	4 987
Sommaire	1	8	14	14	22	0	19	731	144	953

CATEGORIES
SOCIO-PROFESSIONNELLES

Sans emploi	4	61	9	4	280	2	0	6	27	393
Cadres - Employeur	261	2	14	41	0	24	0	3	7	352
Exploitant indépendant	0	24	68	579	27	563	0	206	296	1 763
Salarié qualifié commerce - service	0	3	1 157	161	0	105	0	35	93	1 554
Salarié non qualifié commerce - service	0	5	15	139	0	255	0	133	83	630
Salarié qualifié industrie	0	0	250	383	0	194	0	68	121	1 016
Salarié non qualifié industrie	0	6	5	802	0	537	0	240	186	1 776
Aide familial - Apprenti	0	0	0	0	0	0	114	0	0	114
Travailleur de l'agriculture	0	23	11	52	0	47	0	48	15	196
TOTAL	265	124	1 529	2 161	307	1 727	114	739	828	7 794

Type 1 : Chef de ménage cadres ou employeur

Type 3 : Chef de ménage, famille élargie, salarié des secteurs commerce et service moderne

Type 5 : Chef de ménage sans emploi (ce type comprend aussi des hommes, alors que les individus étalons et les individus représentatifs sont des femmes)

Type 6 : Chef de ménage, famille nucléaire ou groupement de célibataire résidant dans l'habitat de cour

Type 7 : Chef de ménage apprenti ou aide familial

Type 8 : Chef de ménage résidant dans l'habitat sommaire.

Ainsi pour certains types la définition doit être revue et l'hypothèse avancée doit donc être modifiée. Il n'en demeure pas moins, dans notre cas, que les types retenus se stabilisent rapidement, 91 % des individus se positionnent dès la première itération et dès la troisième itération chaque individu a rejoint sa partition. A titre d'exemple nous donnons les effectifs de chacune des combinaisons possibles (tableau 3). Le code correspond au type affecté à chacune des trois itérations : le code 463 par exemple correspond à un individu qui rejoint le type 4 à la première itération, le type 6 à la seconde et le type 3 à la dernière itération. Ce tableau permet d'évaluer le nombre des individus charnières, ainsi par exemple 31 individus passent de la classe 9 (ménage 1 personne) à la classe 7 (apprenti ou aide familial). Il n'y a aucune difficulté à lister ces individus afin de les identifier, et éventuellement les retrouver ensuite sur le terrain.

Notre objectif premier était d'obtenir une liste d'individus appartenant à chacun des types. La méthode de sélection a répondu à notre attente, toutefois dans notre cas le temps écoulé entre l'enquête initiale et l'investigation sociologique est trop long : cinq ans, mais malgré ce handicap, nous avons pu retrouver de nombreux chefs de ménage. Reste à s'interroger sur la notion d'individu central d'une classe, d'individu moyen. Effectivement, pour une classe donnée, les principales caractéristiques retenues sont semblables (par exemple pour les vieux : l'âge, l'habitat, la date d'installation à Abidjan, ...), mais derrière ces indications ponctuelles se cachent des vies bien singulières. Il est évident que la méthode ne permet de sélectionner les individus qu'en fonction des variables introduites et si les personnes sont interrogées sur un autre sujet (les conditions de leur insertion à Abidjan par

INDIVIDUS STABLES ET INDIVIDUS CHARNIERES

CODE	FREQ	EFF. DEF.	CODE	FREQ	EFF. DEF.	CODE	FREQ	EFF. DEF.
<u>111.</u>	<u>262</u>	265	564.	1		855.	10	
133.	1		566.	2		866.	10	
155.	3		599.	2		877.	1	
<u>222.</u>	<u>124</u>	124	611.	1		886.	1	
<u>333.</u>	<u>1 386</u>	1 529	633.	33		<u>888.</u>	<u>718</u>	739
336.	4		636.	2		911.	2	
344.	30		644.	43		933.	11	
355.	7		646.	2		939.	1	
363.	2		655.	16		944.	4	
366.	11		663.	2		949.	1	
399.	1		664.	74		955.	4	
433.	82		665.	2		959.	1	
<u>444.</u>	<u>1 979</u>	2 161	<u>666.</u>	<u>1 528</u>	1 727	966.	7	
446.	15		686.	1		969.	2	
455.	36		688.	3		977.	31	
463.	1		699.	1		988.	6	
464.	26		766.	1		993.	3	
466.	142		<u>777.</u>	<u>82</u>	114	995.	1	
488.	7		833.	8		998.	5	
544.	1		844.	3		<u>999.</u>	<u>819</u>	828
<u>555.</u>	<u>228</u>	307	846	1				

VALID CASES 7 794

MISSING CASES 0

exemple) les réponses divergent. Mais même si le passé diffère, des traits saillants communs apparaissent comme l'importance du tuteur au moment de l'adolescence, le rôle joué par l'épouse, etc.

4 - INTERET ET LIMITES DE LA METHODE

Cette méthode permet d'obtenir une représentation synthétique de grands ensembles de données en utilisant un temps de calcul relativement limité. Elle fournit de façon plus lisible le classement de chaque individu au sein de la typologie, ce qui constitue, pour cet aspect, un avantage par rapport aux méthodes d'analyse factorielle des correspondances. En effet, la projection des individus sur le plan factoriel est bien souvent d'une lecture difficile lorsque de grands tableaux de données sont pris en considération, la superposition des points constituant une perte d'information.

La partition peut être réalisée, on l'a vu, de deux façons :

- soit à partir d'éléments tirés au hasard, ce qui est utile lorsque aucune information n'est disponible a priori sur le corps des données ; cette procédure semblant un bon préalable à un dépouillement classique ;

- soit à partir d'éléments étalons : la typologie obtenue vérifie ou non les hypothèses préalablement établies. Cette deuxième solution est préférable car, en principe, tout utilisateur a un corps d'hypothèses qu'il désire tester. La typologie peut être réalisée en introduisant des individus fictifs possédant des caractéristiques bien déterminées de façon à obtenir ainsi les individus réels qui s'agrègent à chacun de ces modèles.

Un autre intérêt de la méthode réside dans la possibilité d'extraire la classe affectée à chacun des individus et l'adjoindre au fichier initial. Cette typologie devenant elle-même une nouvelle variable, plus synthétique, permettant d'autres analyses.

A notre sens, la possibilité d'extraire nominalement les individus centraux de chacun des types constitue l'avantage principal de la méthode. Ces individus centraux possèdent en principe les

caractéristiques moyennes de la classe. Le nombre de ces éléments est fixé a priori par l'utilisateur. Une identification rapide de ces individus à l'issue de l'enquête permet de réaliser conjointement une investigation plus approfondie sur un sous-échantillon représentatif de chacune des classes obtenues. A l'inverse, les cas marginaux, révélés par leur plus faible capacité à se stabiliser rapidement dans le modèle, méritent aussi d'être étudiés : leur marginalité statistique traduit peut-être une marginalité sociale.

Comme toute méthode d'analyse des données, celle des nuées dynamiques présente quelques limites.

Les modalités comprenant des effectifs peu nombreux tendent à constituer des classes parfois non pertinentes ; mais cette limite est aussi un avantage dans la mesure où certains cas aberrants peuvent être décelés de cette façon, surtout si un fichier de données brutes est utilisé.

Une modalité composant majoritairement une variable occulte les autres possibilités : ainsi nous avons dû constituer, lors du travail préparatoire, un second fichier ne comprenant pas l'habitat de cour, afin de mieux dégager les types intégrant les populations résidant dans les trois autres formes d'habitat .

Des variables trop redondantes risquent de masquer l'influence des autres caractéristiques .

Le choix des étalons laisse une grande part d'arbitraire dans la mesure où certaines variables peuvent être privilégiées à tort. Mais l'utilisation dans un premier temps d'un tirage au hasard des étalons, puis l'élaboration de la typologie à partir des premiers individus tirés limite les risques de manipulation inconsidérée des données.

Cette méthode comporte une part d'arbitraire mais il faut surtout concevoir son utilisation comme le moyen de tester un corps d'hypothèses. Si le choix des étalons est judicieux, ces derniers se retrouvent en grand nombre comme éléments centraux ou pour le moins dans la classe qu'ils ont aidé à déterminer. Si l'hypothèse est fautive, ces individus étalons se dispersent dans diverses classes à l'issue de la partition.

BIBLIOGRAPHIE

ANTOINE Ph., HERRY Cl. - 1983

Quelques éléments pour l'évaluation des résultats de l'enquête
à passages répétés menée à Abidjan

STATECO n° 34, Juin 1983, INSEE, pp. 15-38

DIDAY E. - 1971

Une nouvelle méthode en classification automatique et reconnaissance
des formes. La méthode des nuées dynamiques

Revue de Statistiques Appliquées, vol. XX, n° 2, pp. 19-33

FLAMENBAUM G., THIERY J., BENZECRI J.P. - 1979

Agrégation en boules de rayon fixe et centre optimisés
(boules optimisées)

Les cahiers de l'analyse des données, n° 3, pp. 357-364

LEBART L., MORINEAU A., TABARD N. - 1977

Techniques de la description statistique

Dunod, Paris, 351 pages

VOLLE M. - 1981

Analyse des données

Economica, Paris, 317 pages