

J.-P. BENZÉCRI

Où vit-on le mieux en France ? Représentation cartographique de l'analyse de 33 variables

Les cahiers de l'analyse des données, tome 13, n° 3 (1988), p. 349-366

http://www.numdam.org/item?id=CAD_1988__13_3_349_0

© Les cahiers de l'analyse des données, Dunod, 1988, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

OU VIT-ON LE MIEUX EN FRANCE ? REPRÉSENTATION CARTOGRAPHIQUE DE L'ANALYSE DE 33 VARIABLES

[CARTE VIE FRANCE]

J.-P. BENZÉCRI

1 Des données aux analyses

1.1 Origine des données

Les amis des lettres russes connaissent le poème de H. A. Некрасов :

Кому на руси жить хорошо,

Qui vit bien en Russie, où l'on entend sept moujiks prôner successivement la félicité du propriétaire terrien, du fonctionnaire, du pope, du marchand, du boyar, du ministre et du Tsar...

Afin de déterminer où l'on vit le mieux en France, l'hebdomadaire *Le POINT* n'a pas sollicité la contribution des Muses, mais il a collationné, pour chacun des départements métropolitains, un ensemble de quelque cinquante variables, distribuées en huit chapitres. Fort opportunément, l'essentiel du tableau des données a été publié dans la revue.

Nous ne reproduirons pas ce tableau, ni la liste de références précises aux sources, dont il est accompagné. Sur la seule liste abrégée des variables, le lecteur pourra apprécier la diversité des aspects abordés; tout en s'interrogeant avec nous sur la valeur de certains nombres. Les *Cambriolages de lieux d'habitation*, par exemple, devraient être rapportés à un dénominateur; mais on hésite entre la population totale, le nombre des foyers et le nombre des résidences, qui comprend des résidences résidences secondaires, en proportion très variable suivant les départements. Les *Bouchons routiers* et les *Séjours de vacances* concernent directement les passagers plutôt que les résidents; dont ils conditionnent toutefois l'existence. Les *Dépenses en pharmacie par personne protégée* peuvent être regardées comme un indice de soins vigilants, ou, au contraire, de morbidité; ou, plus vraisemblablement, du taux de personnes âgées.

Ici, nous abordons le chapitre des jugements de valeur: si certaines variables, telles le *Taux de délinquance juvénile*, ont une polarité, du bien au mal, que chacun appréciera de la même façon, on peut s'interroger sur celle du *Nombre de radios locales*, ou de l'*Indice de fréquentation des cinémas*. L'influence de tels jugements dépend grandement du codage et de la méthode

Où vit-on le mieux en France: Le Point 29-2-88

Les départements Français caractérisés par des variables réparties en huit chapitres

Dynamisme	DJN pourcentage des < vingt-cinq ans
	DFC taux de FéCondité
	DES nombre de filières de l'Enseignement Sup
	DET pourcentage d'ETudiants parmi les dixhuit vingtcinq ans
	DRD nombre de RaDios locales
Hygiène	DSP nombre de licenciés en SPort pour cent mille habitants
	HMH Mortalité tte cause pour cent mille Hommes
	HMF Mortalité tte cause pour cent mille Femmes
	HCH mortalité Cancer pour cent mille Hommes
	HCF mortalité Cancer pour cent mille Femmes
Crise	HPH dépense PHarmacie/pers protégée
	CHT Taux de CHômage moyen
	CHD Durée moy du CHômage en j
	CHJ % Chômeurs < vigtcinq ans (Jeunes)
	CLG LoGements construits quatvingtsix/quatvingtdeux
Agrément	AST nombre de SiTes naturels protégés
	AVC séjours de VaCance dans le département en milliers
	AFR taux de boisement (FoRêt)
	ART taux de bouchons RouTiers (en heures.km)
	ASO pollution en tonnes de dioxyde de SOufre
Richesse	AAC nombre d'ACcidents de la route pour mille hab
	RIM montant moy de l'Impôt sur le revenu
	RRV indice de Richesse Vive
	RND investissement iNDustriel par salarié en mil franc
	REP EPargne par hab en francs
Criminalité	RPV % plus de soixante ans au min Vieillesse (Pauvre)
	RPJ % foyers au revenu < quatmilcinq (Pauvre Jeune)
	KDJ taux de Délinquance Juvéni'e/miljeunes
	KCR CRimes contre les personnes /mil habitants
	KVL crimes contre les biens /mil habitants (VoL)
Société	KBL criminalité en col BLanc/mil habit
	KHB cambriolages de lieux d'HaBitation
	KVT vols de VoiTure
	SCR places de CRêche/cent enfants de zéro à trois ans
	SEC nbre de trois à six ans/instituteur
Culture	SHF écart salaire Homme/Femme pour cadre moyen
	SRT places Maison ReTraite/cent > septante ans
	SSP nbre complexes SPortifs/centmil hab
	SAC nbre ACcidents travail/cent salariés
	QCI indice de fréquentation des CInémas
	QMU nombre de MUées
	QFT nombre de FesTivals
	QLY pourcent dixhuit ans en terminale (LYcée)
	QDP DéPense culturelle du DéPartement francs/hab

Tableau abrégé des variables publiées par *Le POINT*
 les sigles des variables non retenues par nous sont en italiques.

d'analyse adoptés pour donner des conditions de vie dans chaque département une évaluation globale.

1.2 Codage des données

Selon *Le POINT*, les valeurs des variables sont regardées comme des notes que l'on combine en les affectant de coefficients (éventuellement négatifs) afin d'obtenir, pour chaque département, une note dans chaque chapitre, puis une note globale. Si l'on procède ainsi, la polarité de chaque variable intervient de façon essentielle comme le signe de son coefficient. De plus, le codage étant linéaire, les 108 musées parisiens et les 64 radios locales recensées par *Le POINT* auront à charge de compenser 213065 heures.km de *Bouchons routiers* mises au débit de la capitale.

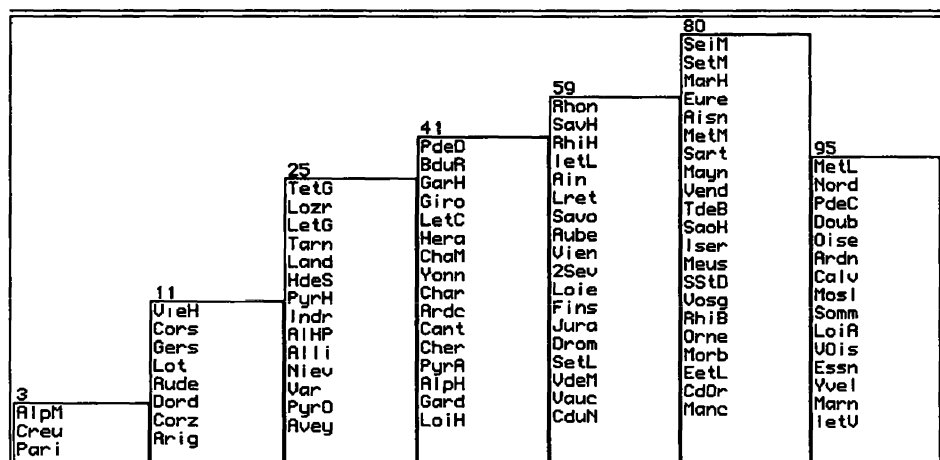
La critique est aisée: l'art est plus difficile! Nous sommes convaincu qu'il s'impose de rechercher dans ces données, de préférence à une ou plusieurs échelles de valeur, une image globale multidimensionnelle des conditions de vie selon les départements. Cette image elle-même dépendra nécessairement du choix des variables retenues, des poids inégaux que la redondance mal maîtrisée des données confère à chaque aspect de la vie. Et, selon le codage, les valeurs extrêmes, dont certains départements, Paris en tête, sont souvent crédités, auront un rôle majeur ou ne pèseront guère plus que les dix autres premières valeurs ou les dix dernières.

Nous avons choisi de découper chaque variable v en un ensemble J_v de modalités; et de coder les données sous forme binaire, en $(0,1)$: $k(d,j) = 1$ si le département d rentre dans la modalité j , et zéro sinon. Ce codage classique efface les effets massifs des valeurs extrêmes; ce qui nous paraît légitime si l'on veut obtenir une image globale où la diversité des départements moyens apparaisse nettement.

Reste que le découpage de chaque variable pourrait nous arrêter quelque temps; comme on en jugera d'après un exemple d'histogramme saisi en copie d'écran lors du déroulement du programme conversationnel *zrang* qui nous a servi à découper les variables. Une possibilité serait de découper chaque variable en un nombre assez élevé de modalités, par exemple cinq, et de recoder celles-ci suivant trois modalités, en perdant la forme simple en $(0,1)$ du tableau logique.

Nous avons été au plus simple, tant pour le codage que pour l'interprétation des résultats; et avons découpé chacune des variables en 3 modalités de même fréquence (à une unité près). La cohérence géographique des résultats nous semble justifier, *a posteriori*, ce procédé expéditif. On a obtenu, assurément, une image de la diversité des conditions de vie dans les départements; et, même si cette image n'est pas la seule qu'on puisse tirer des données, considérer sa représentation cartographique est riche de suggestions.

Lors de la saisie, moyennant un changement d'échelle éventuel, on a ramené toutes les données retenues à être des entiers positifs de moins de quatre chiffres; afin de faciliter la vérification, on a tapé par blocs les 95 valeurs de chaque variable; et le tableau a ensuite été transposé avant découpage.



faut il tracer un autre histogramme(H) pour DJN ou non(N)

Ci dessus, histogramme pour la variable DJN

Ci-dessous, bloc des valeurs de la variable DJN, tel qu'il a été saisi dans le tableau des données

DJN pour mil des < vingt-cinq ans

```

372 391 325 327 339 284 343 403 302 368 311 320 355 401 342 344 346 341 306
315 379 356 280 309 404 363 392 379 365 338 354 313 353 348 396 327 372 383
364 330 350 365 336 397 370 312 331 331 415 377 396 392 386 390 382 380 399
324 411 403 380 407 355 340 328 321 380 374 375 383 357 387 368 374 275 395
392 396 367 398 330 332 323 356 385 367 315 380 345 384 396 329 381 356 396

```

Quant au choix des sigles, nous pensons que les abréviations adoptées pour désigner les départements sont non ambiguës et assez évocatrices pour être aisément déchiffrées: nous signalons seulement que *Loie* est la Loire, et *Lret*, le Loiret. Les variables nous ont embarrassé davantage. Comme on le voit sur le tableau, on a d'abord associé à chacun des chapitres une lettre: C pour Crise, K pour Criminalité, Q pour Culture... On a ensuite choisi deux lettres pour chaque question: ainsi le taux de FéCondité, dans le chapitre du Dynamisme, est désigné par DFC. Enfin, les trois modalités ont été distinguées par les signes +, = et - : le + étant attribué à celle des deux extrémités de l'échelle qui est jugée la plus favorable, dans l'optique qui semble être celle du *POINT*. Ainsi DFC+ désigne une FéCondité élevée; et KDJ+, un taux de Délinquance Juvenile faible. Il importe de souligner qu'un tel choix des sigles ne présume en rien de

l'interprétation de l'analyse; (laquelle procède, en bref, en mettant en évidence des agrégats de départements associés à des agrégats de modalités); mais, sans imposer de jugement de valeur, il nous paraît aider au dépouillement des résultats.

1.3 Enchaînement des analyses et interprétation

Il y a 95 départements (la Corse du Nord n'étant pas distinguée de celle du Sud dans les données statistiques); et le découpage de 33 variables a produit 99 modalités. Pour mieux qu'aient été choisis les sigles de celles-ci, il n'est pas facile de les avoir ensemble présents à l'esprit. Quant aux départements, si les noms en sont bien connus de la plupart des lecteurs de langue française, l'entité géographique que le nom désigne est parfois seulement entrevue.

Dans ces conditions, l'étude directe des graphiques plans est fort laborieuse: et l'on s'aide opportunément des résultats de classifications automatiques. Ainsi, au lieu de deux centaines de points individuels, on a deux dizaines de classes.

La lecture rapide du contenu des classes de modalités met d'abord en évidence certaines concordances dans les signes {+, =, -} dont sont affectés les sigles; ainsi que dans les lettres initiales caractérisant les chapitres. En voyant ensemble {KDJ-, KCR-, KVL-} on songe immédiatement à un taux de délinquance élevé, même si l'on a oublié que DJ désigne la Délinquance Juvenile.

On peut ensuite substituer aux numéros des classes des sigles plus évocateurs; même si ceux-ci ne doivent leur valeur mnémotechnique qu'à une interprétation sommaire! Par exemple, la classe j185 qui, à côté de nombreux avantages signalés par des signes +, renferme les 3 modalités négatives du chapitre K de la criminalité est notée bbKm; où *b* signifie bien, et *m* mal.

Pour les départements, le médiocre géographe a besoin du secours d'une carte: en peignant les départements de chacune des classes par une trame propre à celle-ci, on peut avoir la satisfaction de voir apparaître des zones territoriales connexes.

Au delà de l'interprétation directe que suggère leur contenu, les classes doivent être considérées du point de vue de leurs proximités mutuelles. On voit clairement sur les plans (1,2) et (1,3) que la classe de modalités j185, qu'on a notée bbKm, est proches des classes de départements i140 (AlpM, ie. Nice,...) et i156 (Paris, Lyon, Marseille, Toulouse, etc...). Mais un système de relations complexe ne peut être saisi avec certitude sur deux ou trois plans: en effet, avec 99 modalités de 33 variables, le nuage est situé dans un espace de dimension 66 (99-33). La classification, effectuée dans l'espace engendré par les 20 premiers axes, prend en compte cette complexité; et les associations entre classes des deux

ensembles sont évaluées avec précision sur les listages d'aides à l'interprétation, utilisés pour l'étiquetage des arbres.

Nous avons choisi de commencer par exposer la classification des modalités: mais, quel que soit notre désir d'ordre, il faut accepter que, dans l'interprétation de chaque analyse, on renvoie sans cesse aux autres.

2 L'ensemble J des modalités

c Partition en 9 classes : Sigles des individus de la classe numéro c	
175	DJN- RIM= DFC- REP+ SHF+ HMH+ HMF+ {JmHb}
180	DSP+ QMU= SHF= AFR+ QDP+ CHD+ CHT+ {=CHb}
183	CLG- SEC= SAC= HMF= HMH= CHD= QDP= SHF- QLY- {===m}
189	DET= AAC= QCI= DSP= QLY= SCR= CHJ= RPJ= RRV= SSP= QFT= {====}
	CLG= AFR= DRD= DES= KVL= RND= KDJ= KCR=
188	SAC+ RPV- QMU- SSP+ REP= RND- CHT= KDJ+ KVL+ KCR+ DES- {KKKb}
	QFT- DRD- QCI-
182	CHJ- RIM- RPJ- RRV- SCR- DFC+ DET- CHD- AFR- SEC- AAC- {mFCb}
186	HMH- HMF- QDP- SSP- CHT- SAC- REP- DJN+ DSP- RPV+ {mJNb}
181	DFC= DJN= RND+ RPV= DES+ DRD+ QFT+ QMU+ {=Qbb}
185	RPJ+ RRV+ QLY+ DET+ CHJ+ QCI+ CLG+ RIM+ SCR+ AAC+ KDJ- {bbKm}
	KVL- KCR- SEC+

2.1 Classification de l'ensemble des modalités

Exprimée en dix millièmes de l'inertie, la suite des niveaux de dix nœuds les plus hauts est

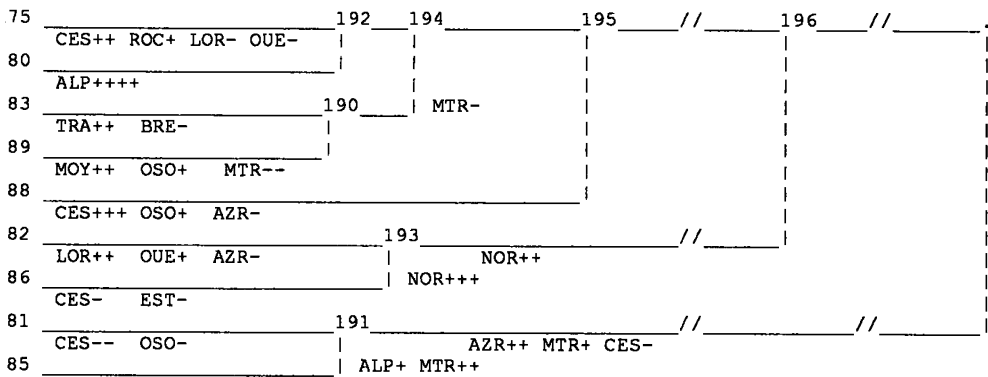
$$T = 1031 \ 774 \ 482 \ 330 \ 312 \ 259 \ 258 \ 255 \ 217 \ 190 \ e-4.$$

La dénivellation entre 255 et 217 nous a incité à demander d'abord une partition en 9 classes, définie par les 8 nœuds les plus hauts; même si certaines subdivisions de ces 9 classes sont interprétables, c'est cette partition que nous avons définitivement retenue.

Au §1.3, on a expliqué le choix du sigle bbKm pour la classe j185. À l'opposé, on a noté KKKb la classe j188 qui renferme les trois modalités positives du chapitre K de la criminalité; avec, comme autres notes +, l'équipement sportif, SSP+, et un faible taux d'accidents du travail, SAC+; mais aussi de nombreuses modalités négatives, ou jugées telles..., notamment en matière culturelle ou connexe: QCI-, faible fréquentation des cinémas, DRD-,

peu de radios libres; et aussi RND-, faibles investissements iNDustriels, DES-, peu de filières de l'enseignement supérieur.

Dans l'analyse factorielle, l'opposition entre j185 et j188 s'inscrit principalement sur l'axe 1 dont elle peut donner une interprétation.



Nombre de la partition des modalités en 9 classes

Non loin de la classe j185, et associée à celle-ci au sein de la classe j191, on trouve j181 notée =Qbb parce que, à côté de trois modalités moyennes, elle comprend quatre modalités + du domaine culturel, QMU+, musées, QFT+, festivals, DRD+, radios libres, et DES+, filières de l'enseignement supérieur; avec RND+, investissements industriels.

La classe j175 est notée JmHb, parce qu'elle comporte à la fois les deux modalités démographiques négatives DJN- et DFC- et les deux modalités positives de l'hygiène: HMF+ et HMH+, faible Mortalité des Femmes et des Hommes.

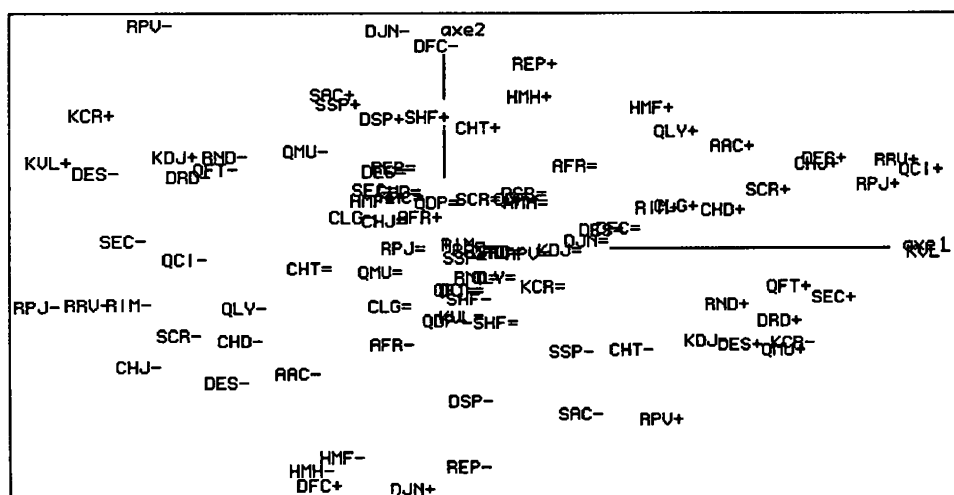
Les modalités positives de la démographie sont dans les deux classes jumelles j182 et j186, notées respectivement mFCb et mJNb. On se gardera d'assimiler DJN+, fort pourcentage de jeunes de 18 à 25 ans, et DFC+, FéCondité élevée: en effet, la population adolescente fuit certains lieux et en recherche d'autres; et une forte natalité ne prémunit pas contre l'exode. Sur l'axe 2, j186 s'oppose à j175.

La classe j180 a été notée =CHb parce que, à côté de quelques modalités moyennes elle contient les deux modalités CHT+ et CHD+, faible Taux et faible Durée du CHômage. La troisième modalité CHJ+, faible chômage des jeunes, a déjà été vue dans j185, bbKm, associée à la délinquance et à des avantages culturels; les modalités opposées, CHJ-, CHD- et CHT-, sont dans j182 et j186.

Enfin on ne s'étonnera pas de voir notées ===== et =====m les deux classes jumelles j189 et j183 où s'accumulent les modalités moyennes; avec, dans j183 deux modalités négatives signalées par la lettre m (mauvais).

L'arbre de la CAH des modalités a été étiqueté d'après la partition retenue pour l'ensemble des départements: pour apprécier cet étiquetage, il faut avoir lu le §3.

2.2 Représentation de l'ensemble des modalités d'après l'analyse factorielle



Ensemble des modalités : plan (1,2)

Au §2.1, on a considéré les modalités réparties en 9 classes: chacune de ces classes a une certaine dispersion, dont on doit tenir compte pour nuancer l'interprétation. Sur les graphiques obtenus par copie d'écran, certains sigles se superposent; mais on distingue bien la plupart des modalités écartées de l'origine, qui contribuent le plus à créer les axes.

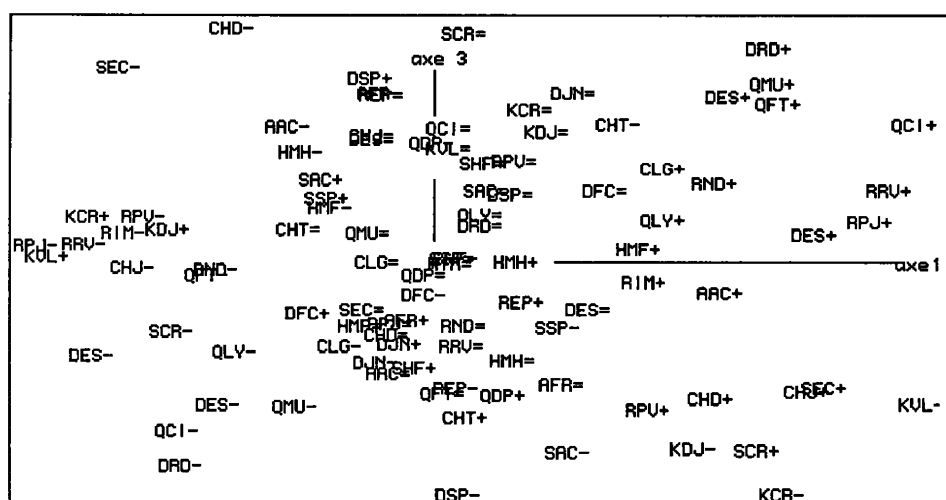
Commençons par ne considérer, sur le plan (1,2), que la distribution des signes {+, =, -}. Les signes = se concentrent autour de l'origine. Les + prédominent dans le quadrant ($0 < F_1$; $0 < F_2$); et les - dans le quadrant opposé ($F_1 < 0$; $F_2 < 0$). La deuxième bissectrice permettrait à peu près de séparer les + des -, n'étaient quelques exceptions majeures: celles des modalités de la délinquance et de la démographie.

Somme toute, les avantages recensés dans les données du *POINT* offrent un panorama des valeurs le plus souvent prônées aujourd'hui. À l'analyse, ce

panorama apparaît cohérent; et on ne s'étonnera pas de voir que dénatalité et délinquance, qui, selon des modes distincts, sont affaire de morale, portent la contradiction; respectivement sur l'axe 2 et sur l'axe 1.

rang :	1	2	3	4	5	6	7	8	9	10	
lambda :	2602	1860	967	922	848	797	701	654	607	560	e-4
taux :	1321	944	491	468	431	404	356	332	308	284	e-4
cumul :	1321	2265	2756	3224	3655	4059	4415	4747	5055	5339	e-4

Ci-dessus: Tableau des valeurs propres
Ci-dessous : Ensemble J dans le plan (1,3)



La suite des valeurs propres montre que les facteurs 1 et 2 sont bien séparés l'un de l'autre ainsi que de ceux qui les suivent. À partir du rang 3, la décroissance est lente et régulière: pour avoir un cumul de 75%, il faut aller jusqu'au 20-ème axe, ainsi que nous l'avons fait pour la classification automatique. On ne s'étonnera donc pas de trouver une interprétation des axes 1 et 2 ; tandis que la distribution des modalités sur l'axe 3 est difficile à résumer. On remarquera seulement la dispersion des modalités moyennes, non moins grande que celle des modalités extrêmes, + et -.

3 L'ensemble des départements

3.1 Partition des départements en 14 classes

Tandis que pour l'ensemble des modalités nous nous sommes arrêté à la partition en 9 classes considérée d'emblée d'après la décroissance des niveaux des nœuds, la répartition géographique des départements nous a incité à retenir

certaines subdivisions. De ce point de vue, seule la carte expliquée au §4 justifie le choix de la partition présentée ici. Ici, nous énumérerons les classes en citant leur contenu et quelques unes des caractéristiques que leur attribue l'étiquetage de l'arbre.

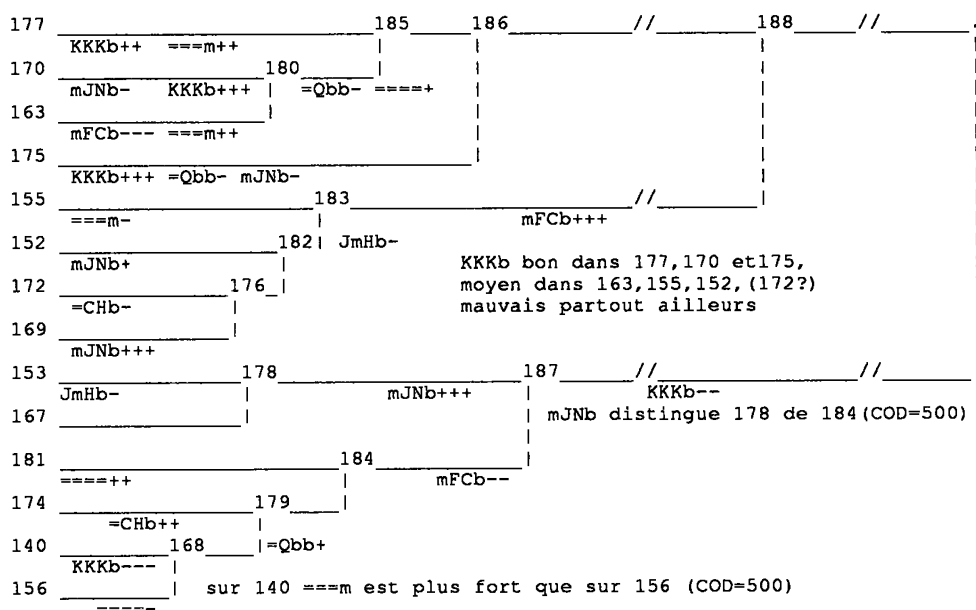
c Partition en 14 classes : Sigles des départements de la classe c											
177	Ain	LetC	Jura	Yonn	Aube	SetL	Niev	Cher	EetL	Eure	TRAnsversale
170	Indr	Vend	Char	LetG	Land	TetG	IetV				Ouest & Sud-Ouest
163	PyrA	PyrH	Corz	VieH	PdeD	Cors					ROC
175	Arig	Lot	Dord	Vien	Cant	2Sev	Gers	Avey			CENtre Sud
		Tarn	Mayn	Creu	Ardc	LoiH	Lozr				
155	Loie	CduN	Fins	Morb							BREtagne & Loire
152	Vosg	Orne	SaoH	Meus	MarH						LORraine & Orne
172	LoiA	Calv	Sart	Manc	MetL						OUEst
169	Nord	SeiM	Somm	Ardn	Aisn	PdeC					NORD
153	Mosl	TdeB	RhiH	MetM							EST
167	SStD	VdeM	VOis	SetM	Oise	Essn	Yvel				COUronne
181	Marn	Doub	CdOr	Drom	IetL	Lret	ChaM	Alli	Aude	Gard	MOYenne
174	Savo	AlpH	PyrO	AlHP	RhiB	SavH	Iser				ALPes... (Emploi)
140	Vauc	Giro	Hera	Var	AlpM						AZur (Délits...)
156	Rhon	GarH	BduR	HdeS	Pari						MÉTropoles

La classe i156 des MéTRopoles comprend les départements dont les chefs-lieux sont Paris, Lyon Marseille et Toulouse; avec Paris va le seul département des Hauts-de-Seine où est située La-Défense. Tous les autres départements de la petite et de la grande couronne, avec l'Oise et la Seine-et-Marne, composent la classe i167. Les Métropoles regroupent à tout ce qui est moyen (====-)!

Proche des Métropoles est i140, classe qui, la Gironde mise à part, suit à peu près le rivage Méditerranéen en contournant Marseille; i140 a le regrettable privilège d'un taux de délinquance maximum (KKKb---).

Dans la classe i174, les ALPes constituent un bloc de 5 départements, auxquels s'adjoignent les Pyrénées-Orientales et le Bas-Rhin; les caractères de l'emploi y sont favorables (=CHb++).

La classe i181, très dispersée sur la carte, du Loiret au Gard et de la Charente-Maritime au Doubs, est associée à l'ensemble ==== des modalités moyennes.



Partition des départements en 14 classes

Dans l'EST, 4 départements constituent une classe i153, jumelle de la COUronne de Paris, i167. Ces deux classes ont un pourcentage élevé d'adolescents (mJNb+++).

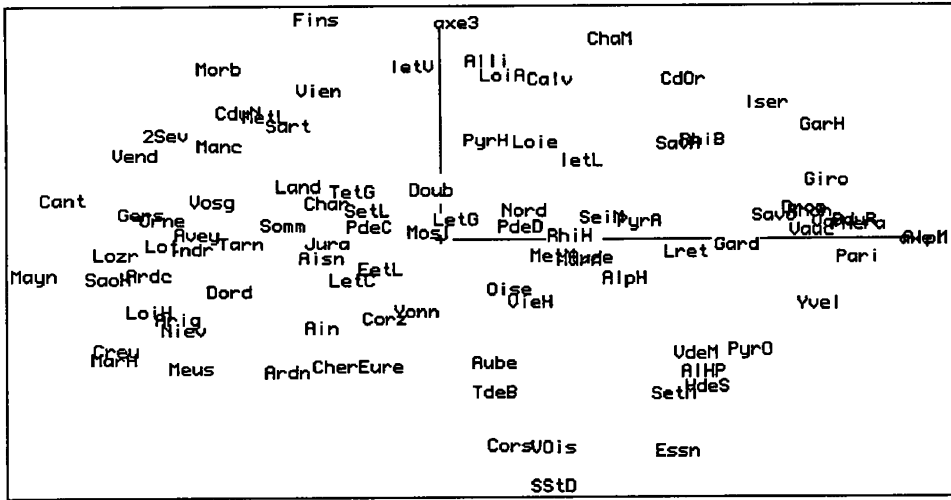
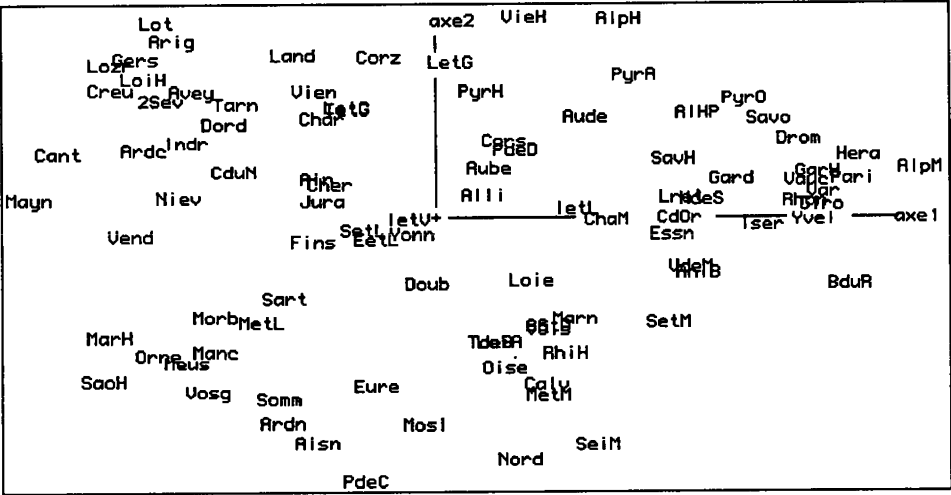
Dans la classe i183, caractérisée par une fécondité élevée qu'on ne trouve pas ailleurs, mFCb+++, nous avons distingué 4 subdivisions, en nous fondant sur la distribution géographique et les niveaux de délinquance, qu'on lira sur l'arbre.

Les 6 départements de la Pointe NORD de la France font la classe i169. Comme i153 et i167, i169 est étiquetée mJNb+++. Les classes i172, i152 et i155 ont pour dominantes géographiques respectives l'OUEst, la LORraine et la Bretagne.

Restent les 4 subdivisions de la classe i186. On remarquera que i177 dessine une bande TRANSversale continue de l'Eure au Jura; tandis que i175 occupe principalement le Sud du Massif Central; et i163 comprend des montagnes: Pyrénées, Corrèze ou Corse.

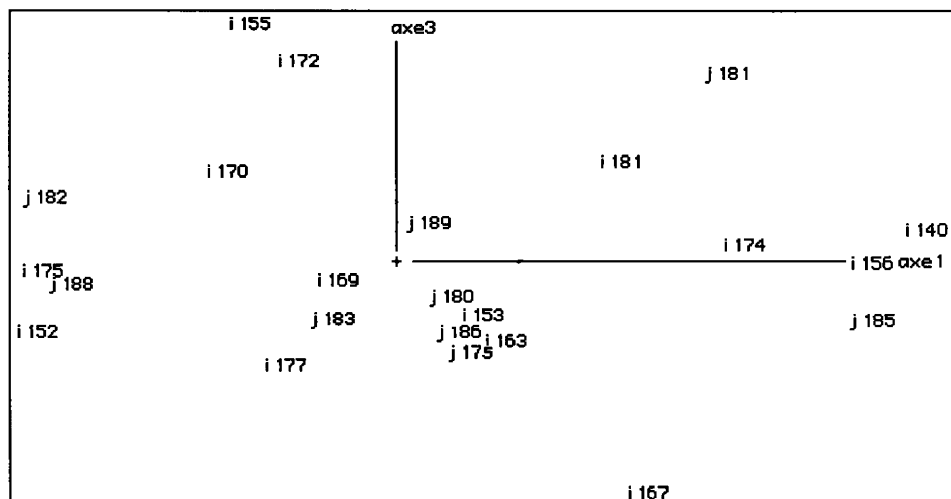
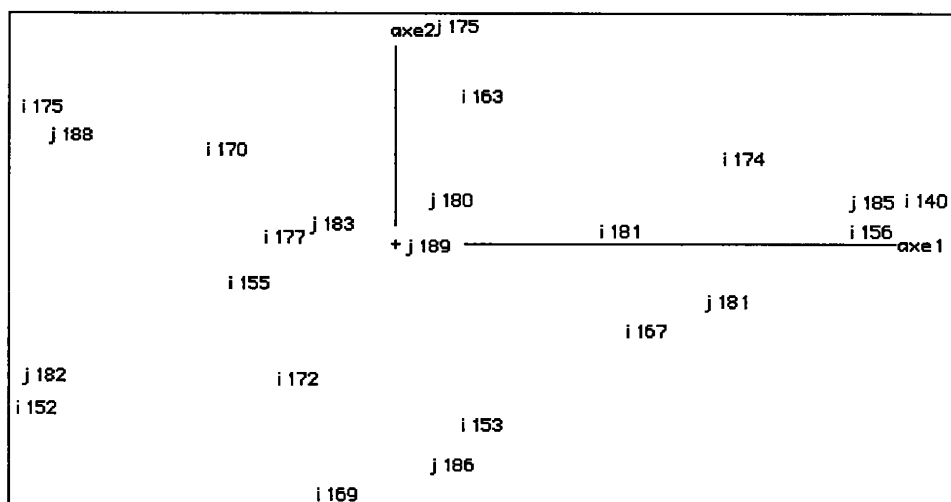
3.2 Représentation de l'ensemble des départements d'après l'analyse factorielle

Nous présentons les plans (1,2) et (1,3) d'une part pour le nuage des départements, et d'autre part pour les centres des classes des deux ensembles en correspondance: départements et modalités. On vérifiera la cohésion des classes



de départements, ainsi que la cohérence de l'étiquetage avec les proximités entre classes des deux ensembles.

Il vaut la peine de noter que l'axe 3, dont nous n'avons pas su, au §2.2, donner une interprétation concise en termes de modalités des variables apparaît ici comme caractérisé nettement, dans sa partie négative, par la classe i167 de la COUronne étendue de Paris.



4 Cartographie

Afin de cartographier les résultats de la présente étude, nous avons utilisé le programme Carthag, conçu par nous pour l'ordinateur Macintosh+, et déjà présenté par M.-M. Thomassin au congrès international de Morelia (Mexique) en 1987. Sans refaire une notice d'utilisation de ce programme et de ses annexes, nous énumérerons les étapes de notre travail.

Le fond de carte numérique de la France découpée en départements, a déjà servi, notamment à A. Moussaoui pour représenter les disparités de la consommation mensuelle des produits pétroliers en France. Outre ce fichier le programme Carthag a eu pour données les résultats de l'analyse factorielle et de la CAH; reste à décrire la méthode conversationnelle pour choisir les trames et à commenter la carte obtenue.

4.1 Le choix des trames

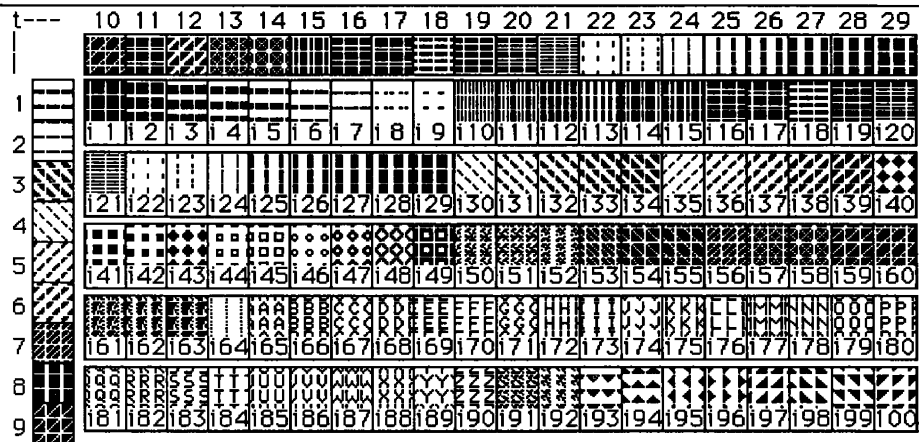
Dans sa version actuelle, Carthag offre à l'utilisateur un tableau de 100 trames, dont le dessin n'est pas définitivement arrêté mais où l'on trouve diverses hachures, quadrillages, ainsi qu'un alphabet de capitales... En entrant dans Carthag, l'utilisateur se voit proposer une palette de 29 trames de hachures horizontales et verticales: il peut modifier cette palette en puisant dans le tableau des trames, selon ce qu'il envisage d'utiliser pour sa carte.

Initialement, des trames sont attribuées aux classes d'unités territoriales dans l'ordre du numérotage de la palette, les classes étant prises de haut en bas sur le dessin de l'arbre: ainsi, la trame 1 est pour i177, la trame 2 pour i170,... la trame 14 pour i156. L'utilisateur peut composer sa palette en prévision de ces attributions initiales; mais le cours du programme lui offrira plus d'une occasion de revenir sur son choix.

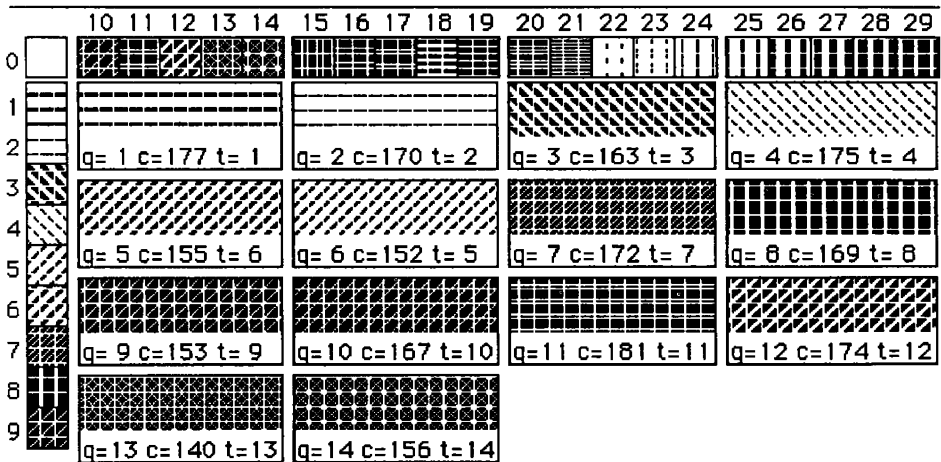
Après la palette s'affichent les résultats de l'analyse factorielle: on peut demander les plans (1,2) ou (1,3), ou tout autre plan... avec les ensembles I et J, ou les centres des classes, et, avec les classes i_q de départements, les trames qui leur sont présentement attribuées. Il est clair qu'on doit chercher à mettre en accord proximités entre classes et similitudes entre trames; l'idéal étant que les trames s'associent avec l'interprétation même des classes. Au vu des plans, on peut décider de modifier les attributions.

Et, en effet, en quittant l'affichage des plans, on entre dans un dialogue offrant de modifier l'attribution faite aux classes de trames de la palette. S'affiche alors l'arbre des classes avec leurs trames et enfin la carte. On peut ainsi apprécier l'harmonie du résultat; et, si l'on y trouve à redire, recommencer à partir du début pour perfectionner la carte avant de la saisir.

Dans le cas présent, le principe analogique qui nous a guidé a été de jouer sur l'intensité du noir (les cartographes parlent de *pouvoir couvrant*) et l'inclinaison des hachures. L'intensité du noir est, en général, d'autant plus forte

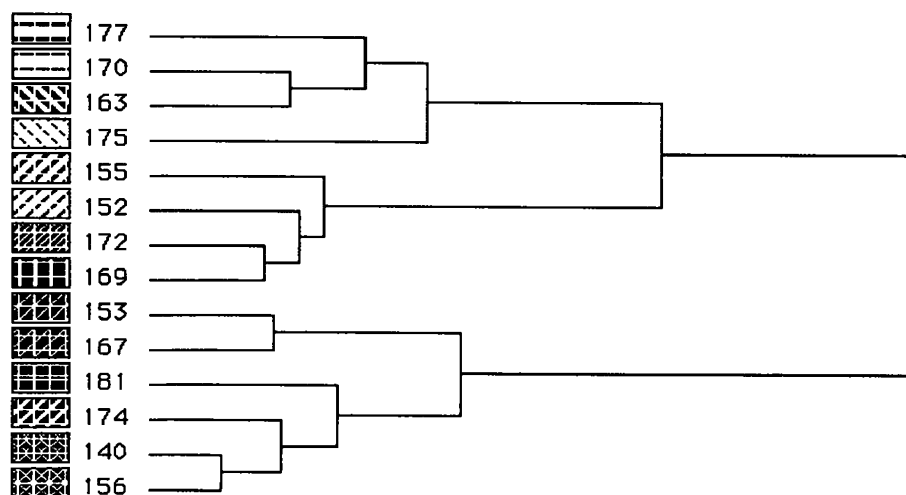
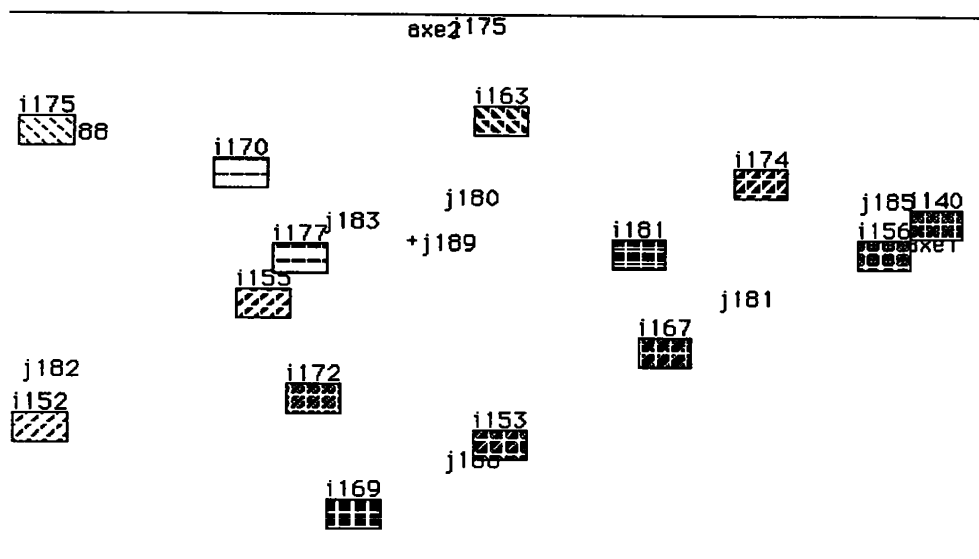


faut-il encore modifier la palette O ou N



on modifiera la trame de la classe q= 5

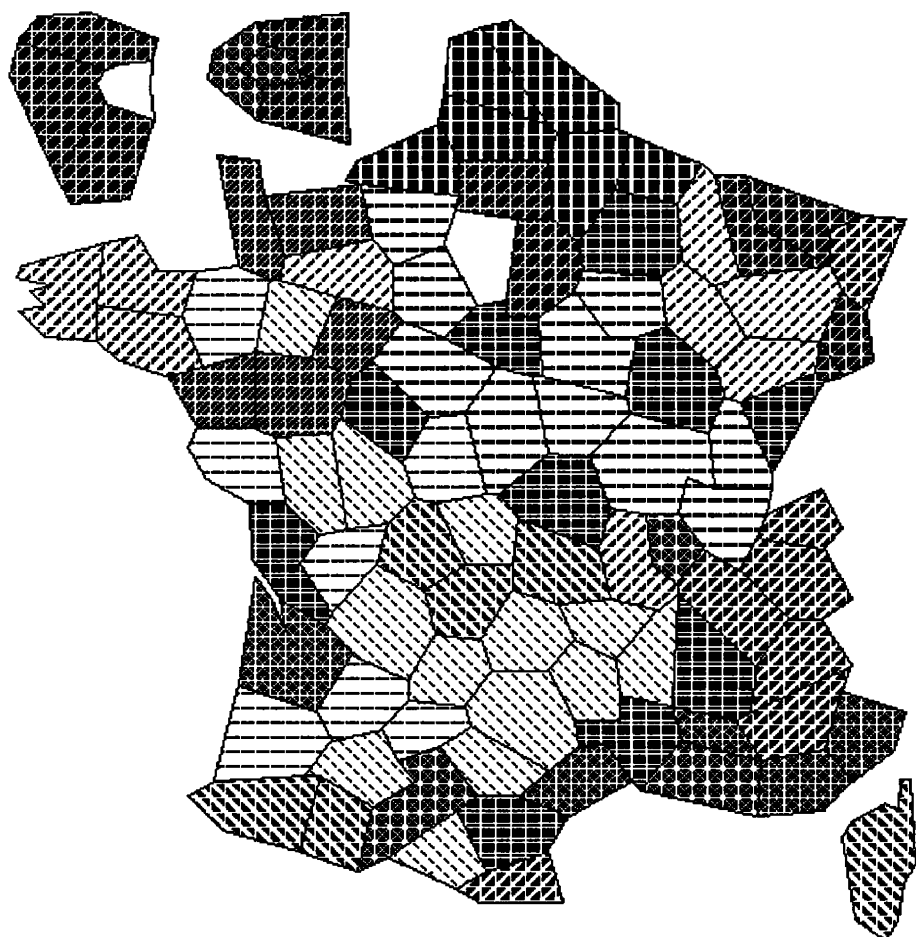
que le niveau de délinquance est plus élevé, information qui est reportée sur l'arbre étiqueté des classes. L'inclinaison des hachures est, de façon générale, liée d'abord à la démographie, une orientation ascendante de la gauche vers la droite symbolisant la fécondité et la présence de nombreux adolescents. D'autre part, dans la trame de la classe i181, fortement associées à des modalités moyennes, un dessin horizontal suggère le signe =; tandis que i140 et i156 (Nice



et Paris) reçoivent des hachures de toute direction, symbole d'aspects positifs et négatifs associés, avec, dans la trame de i156, un semis d'étoiles scintillantes.

4.1 Commentaire de la carte

Notre recherche d'un symbolisme ne dispense pas de recourir à l'étiquetage de l'arbre des classes de départements: il ne peut notamment rendre compte de la distinction entre fécondité et taux d'adolescents. En toute rigueur il conviendrait



même de ne se fier qu'au tableau des données pour affirmer des conclusions. (C'est ainsi que nous avons vérifié que, dans la quasi-totalité des départements, le niveau moyen de KKKb dans sa classe est un indice acceptable de la délinquance, basse si KKKb est élevé). Mais la cohérence du tableau obtenu ne peut être mise en doute. Sans répéter ce qui a été dit au §3, nous invitons le lecteur à considérer sur l'espace français, bientôt ouvert à l'Europe, ces grandes plages dont les plus claires sont marquées de hachures horizontales ou descendantes...

5 Conclusions

5.1 Cartographie et analyse des données multidimensionnelles

Dès 1970, on a soumis à l'analyse des correspondances des statistiques ventilées par unités territoriales, qu'il s'agisse des 95 départements français, des 20 arrondissements de Paris ou de leurs 80 quartiers, ou encore d'un ensemble de communes; les variables pouvant être les catégories socio-professionnelles, les classes d'âge-sexe, les causes de mortalité ou... les attitudes de votes aux grands scrutins nationaux. L. Lebart, au CREDOC, a fait dans ce domaine œuvre de pionnier; et M. Volle, à l'INSEE, a étudié avec finesse la démographie par cantons de la région Languedoc-Rousillon.

Il s'imposait de reporter sur une carte les résultats de telles analyses; la classification automatique permettant de constituer des groupes d'unités territoriales affectées d'une même couleur, ou d'une même trame en noir et blanc. Ainsi, on trouve de nombreuses cartes dans une étude, due à J. Tibeiro, des Disparités régionales entre les profils des prestations servies par les caisses d'assurance maladie de la sécurité sociale en France (in *CAD*, Vol IX, 1984, n°4, pp. 395-422). Mais le premier chef-d'œuvre de cartographie fondé sur l'analyse multidimensionnelle est la carte, imprimée pour l'ORSTOM à l'Institut Géographique Nationale, qui accompagne l'article de M.-M. Thomassin: Analyse des données et représentation cartographique; applications aux conditions physique et potentialités de la Sierra équatorienne (in *CAD*, Vol X, 1985, n°1, pp. 25-51). Cartographe de profession, Madame Thomassin a apporté à l'analyse des données les justes et strictes principes d'une méticuleuse tradition.

Quant à la réalisation technique, le chercheur isolé a d'abord utilisé la plume et l'encre pour couvrir ses cartes des trames choisies; depuis plus de 20 ans, on a, d'autre part, la ressource de découper et coller des feuilles transparentes, imprimées suivant une grande variété de dessins; gravure et impression n'étant pas ordinairement à la disposition du statisticien. Cependant, progressivement, l'informatique vulgarise les moyens typographiques les plus complexes: l'analyse multidimensionnelle bénéficie de ce progrès.

5.2 Géographie politique

Corrélation n'implique point causalité. On se gardera de rendre ce qu'on appelle les *avantages culturels* responsable de la délinquance concomitante; ni de déplorer que la probité ait pour dernier asile des lieux où coexistent toutes les récessions. Mais on ne peut refuser les suggestions d'une carte: autant que la volonté des hommes, l'avenir que la Providence réserve à la France suivra les pentes du terrain.