

M. ZERROUG

## **Hommes espace et dépenses : le profil des premières usines de France**

*Les cahiers de l'analyse des données*, tome 14, n° 4 (1989),  
p. 425-438

[http://www.numdam.org/item?id=CAD\\_1989\\_\\_14\\_4\\_425\\_0](http://www.numdam.org/item?id=CAD_1989__14_4_425_0)

© Les cahiers de l'analyse des données, Dunod, 1989, 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/>

# HOMMES ESPACE ET DÉPENSES: LE PROFIL DES PREMIÈRES USINES DE FRANCE

[USINES FRANCE]

M. ZERROUG\*

## 1 Origine des données; leur élaboration

### 1.1 Le tableau de base

Au printemps de 1989, le périodique *L'USINE NOUVELLE* a publié un numéro spécial intitulé *L'atlas des usines 1989*, avec en sous-titre la mention: *1000 établissements industriels répertoriés par groupes par secteurs et par régions.*

L'idée vient naturellement à l'esprit de constituer à partir de cet *Atlas* un tableau de données propre à être soumis à l'analyse des correspondances et à la classification ascendante hiérarchique. Mais ici, de nombreuses difficultés se présentent. Chacune des plus grandes entreprises a des qualités qui lui sont propres et n'offrent donc pas matière à comparaison statistique. *A fortiori* trouve-t-on dans chaque secteur des caractères de production que les autres ne connaissent pas. Enfin, dans les puissantes entreprises du secteur de l'armement, beaucoup de données sont tenues secrètes.

Pourtant, pour quelque huit-cents usines, l'*Atlas* propose un tableau de *Chiffres clés*, avec trois colonnes, afférentes respectivement aux années 1987, 1988 et 1989. De ces *Chiffres*, il faut éliminer celui de la production exprimée tantôt en valeur, tantôt en nombre d'unités produites; de plus, il ne faut pas songer à constituer un seul individu statistique avec les informations relatives à trois années d'activité d'une même usine; car, dans trop de cas l'on n'a d'informations complètes que pour une ou deux années.

Finalement, on a pu retenir un tableau de base comportant 287 lignes, ou couples (année, usine) provenant de 118 usines différentes, avec 8 colonnes de chiffres, suivies de deux colonnes donnant respectivement l'année (0, 1, 2 pour 1987, 1988 et 1989) et un numéro de secteur. Toutes ces informations ont été

---

(\*) Étudiant en doctorat d'université.

## l'Usine Nouvelle: atlas des usines 1989

unités:	cent	mc	nbr	nbr	MF	MF	MF	%xcent	MF	Date	Prod
10	Eten	Ouvr	Cadr	Budg	Acha	Sala	form	Invs			
AVI7 latecoere	324	711	67	375	170	173	150	20	0	0	
ALa7 bocaouest	58	486	30	677	524	41	120	17	0	1	
ALa8 bocaouest	58	448	29	715	545	35	120	32	1	1	
ALb7 brassarto	1170	825	107	401	318	101	278	90	0	1	
ALc7 geo	200	240	5	204	174	30	189	26	0	1	
ALc8 geo	200	230	5	199	169	30	170	23	1	1	
ALc9 geo	200	240	6	212	177	35	200	20	2	1	
ALd7 lerial	190	633	19	398	300	55	120	11	0	1	
ALd8 lerial	190	620	17	399	305	55	120	6	1	1	
ALe7 vico	140	450	12	250	150	44	150	20	0	1	
ALe8 vico	165	480	12	270	165	46	150	21	1	1	
ALe9 vico	180	470	12	280	170	47	150	20	2	1	
TRa7 behrfrance	312	683	20	505	351	58	250	19	0	2	
TRa8 behrfrance	318	660	20	401	292	57	667	7	1	2	
TRa9 behrfrance	318	680	20	440	320	59	487	20	2	2	
TRb7 heuliezaut	900	1073	56	688	460	113	120	17	0	2	
TRb8 heuliezaut	900	1175	57	851	490	140	120	40	1	2	

*Le tableau de base, tel qu'il a été saisi*

ensuite converties en un tableau de rapports aptes à être découpés en classes puis analysés.

Avant d'expliquer les transformations effectuées, il vaut la peine de soumettre d'abord au lecteur les premières lignes du tableau primaire, ainsi qu'un bilan par secteur des données retenues. Après une ligne de titre (terminée par un retour de chariot) le tableau comporte, avant le nombre des colonnes (ici 10) quelques indications qui, ne comportant aucun chiffre, sont considérées par la procédure de lecture comme des commentaires; il s'agit ici des unités utilisées pour exprimer les variables. Suivent le nombre des colonnes et les sigles de celles-ci, dont nous expliquerons le contenu en lisant la première ligne (année, usine).

Nous lisons d'abord le sigle AVI7: il s'agit en effet des informations relatives à l'activité en 1987 d'une entreprise du secteur de l'AViation (la seule, en fait, de ce secteur, qui ait consenti à lever le secret d'une colonne complète de chiffres!); avant les chiffres, on lit le nom de l'entreprise "Latécoère" (mention qu'ici encore la procédure de lecture assimile à un commentaire).

Le premier chiffre, Eten = 324, est la superficie de l'usine exprimée en centaines de mètres carrés, soit 32400 m<sup>2</sup>; viennent ensuite le nombre total d'ouvriers et employés de toutes catégories, Ouvr = 711; et le sous-total des individus recensés dans les catégories "ingénieurs et cadres", Cadr = 67. Suivent trois colonnes afférentes à l'activité annuelle totale, et exprimées en MF (millions de Francs): Budget annuel (Budg = 375), volume d'achats (Acha = 170) et masse salariale (Sala = 173). Il faut noter, au passage, que les entreprises qui ont donné de tels chiffres pour 1989 n'ont pu fournir que des prévisions.

Sigl	Prod	Nom du secteur	nli	usines
AV	0	aéronautique, c. navale, armements (AVion)	1	1 ( 60)
AL	1	agroALimentaire	11	4 ( 52)
TR	2	automobile, matériel ferroviaire (TRansport)	10	5 ( 65)
BT	3	Bois - Textile	27	10 ( 44)
CH	4	CHimie, parachimie, matériaux	23	9 (147)
VR	5	consommation, loisirs, édition (VaRiété)	15	6 ( 32)
EL	6	ELEctronique	39	14 ( 85)
EN	0	ENergie	1	1 ( 27)
EQ	7	EQUIpements industriels	60	24 ( 59)
IN	8	INformatique - bureautique	8	3 ( 20)
TP	9	manutention, matériels TP et agricoles	20	9 ( 26)
MM	10	Matériel électrique et équipement Ménager	36	13 ( 72)
MC	11	MéCannique, transformation des métaux	21	11 ( 65)
SD	12	SidÉrurgie	15	8 ( 47)

### *Inventaire des données par secteurs*

Le chiffre suivant est le pourcentage de la masse salariale consacré à la formation; afin d'éviter de quitter le format entier (qui permet d'introduire au sein des lignes des commentaires littéraux tels que le nom de l'entreprise) nous avons multiplié le pourcentage par 100: form = 150, signifie donc 1,5%, ou encore 150/10000. Enfin le montant global des investissements est en MF: Invs = 20.

Nous avons déjà dit que la Date est donnée par un seul chiffre; nous en expliquerons l'usage pour le codage des données. Quant au numéro de secteur, il se lit dans la colonne "Prod" de l'inventaire des données; on notera que les secteurs Énergie et AViation, qui n'ont fourni chacun qu'une seule ligne, sont tous deux affectés du numéro 0.

Avec la 2-ème ligne, on passe dans le secteur de l'ALimentation: il s'agit de l'entreprise "BOCA OUEST"; laquelle, étant, dans l'ordre alphabétique, la première de son secteur à nous offrir une ligne de données complètes, reçoit dans son sigle la lettre 'a'; avec, en plus, pour la ligne, le chiffre '7' de l'année considérée. De façon précise, on lit, sur les dernières colonnes de l'*Inventaire par secteur*, que l'ALimentation a fourni 11 lignes complètes (usine, année) provenant de 4 usines différentes sur les 52 de ce secteur pour lesquelles l'*Atlas* donne des chiffres, au moins partiels. Etc...

## 1.2 Calcul et découpage des variables quotient

L'hétérogénéité des données impose d'effectuer un découpage des variables en classes. Mais il faut d'abord noter que les variables fournies par les huit premières colonnes ne sont elles-mêmes pas propres à être soumises telles quelles à un découpage. En effet, toutes, à l'exception de *form*, dépendent directement de la taille de l'entreprise: il faut donc calculer des variables quotient. (Nous nous bornerons à donner les formules, sans reproduire le programme écrit pour transformer le tableau de base, en utilisant les procédures du logiciel MacSAIF).

Ceci admis, le choix même du dénominateur ne s'impose pas. Comme l'indique le titre de l'article, on a considéré les hommes, l'espace et les dépenses. Du point de vue des hommes, on a trois quotients:

$$\begin{aligned} \text{REM} &= \text{Sala} / \text{Ouvr} ; && \text{“Rémunération”} \\ \text{CDR} &= \text{Cadr} / \text{Ouvr} ; && \text{“extension du statut de Cadre”} \\ \text{FRM} &= \text{form} ; && \text{“Formation”} \end{aligned}$$

(auxquels il faut adjoindre l'une des variables d'espace: SPC). On notera que, de divers points de vue, chacun de ces quotients est d'autant plus élevé que la situation des hommes est meilleure. Il importe de préciser ici que dans le quotient REM, la donnée brute Sala a été corrigée en fonction de l'année, en admettant des indices de valeur respectifs de 103, 100 et 97 pour les sommes en F afférentes à 1987, 1988 et 1989.

La variable primaire de l'étendue a été considérée de deux points de vue: selon qu'on la rapporte aux hommes ou à l'argent. On distinguera donc:

$$\begin{aligned} \text{SPC} &= \text{Eten} / \text{Ouvr} ; && \text{“Spaciosité”} \\ \text{SRF} &= \text{Eten} / \text{Budg} ; && \text{“Surface”} \end{aligned}$$

Quant aux dépenses, les valeurs en MF ont été rapportées au budget annuel, exprimé dans la même unité. On a donc:

$$\begin{aligned} \text{SLR} &= \text{Sala} / \text{Budg} ; && \text{“dépenses Salariales”} \\ \text{ACH} &= \text{Acha} / \text{Budg} ; && \text{“dépenses d'Achat”} \\ \text{INV} &= \text{Invs} / \text{Budg} ; && \text{“dépenses d'Investissement”} \end{aligned}$$

On peut voir dans SLR et INV des indices de valeur ajoutée; la variable quotient ACH jouant dans un sens opposé à celui des deux autres.

Les variables quotient ainsi définies ont été découpées par le programme 'zrang', sous l'option 'Q', *barycentrique*, qui effectue un codage linéaire par morceaux (cf. [CODAGE LIN.], in *CAD*, Vol XIV, n° 2, pp. 203-210). Il vaut la peine de publier le listage 'Dcodx' des valeurs pivots prises pour découper chaque quotient en 3 classes {-, =, +}; car ces valeurs ont été choisies d'après des histogrammes pour représenter respectivement une tendance faible, une moyenne et une tendance forte.

De façon précise, on sait que, dans le codage barycentrique sur 3 modalités d'une variable  $v$ , les valeurs inférieures ou égales à la valeur pivot  $v^-$  sont codées {1, 0, 0}; les valeurs comprises entre  $v^-$  et  $v^=$  sont codées sur les deux premières modalités { $x$ ,  $1-x$ , 0}, avec une valeur de  $x$  d'autant plus élevée qu'elles sont plus écartées de la valeur pivot  $v^-$ ; semblablement, les valeurs comprises entre  $v^=$  et  $v^+$  sont codées {0,  $1-y$ ,  $y$ }; et celles au-delà de  $v^+$  sont codées {0, 0, 1}.

1'Usine Nouvelle: atlas des usines 1989  
 9:USISDcodx: bornes pour le découpage des variables  
 le nombre des variables est 8

SPC a	3	modalités dont les sigles et valeurs pivot sont			
		SPC-	SPC=	SPC+	
		2.05e-1	6.82e-1	1.53e+0	
REM a	3	modalités dont les sigles et valeurs pivot sont			
		REM-	REM=	REM+	
		8.75e+0	1.34e+1	1.89e+1	
CDR a	3	modalités dont les sigles et valeurs pivot sont			
		CDR-	CDR=	CDR+	
		2.50e-2	5.88e-2	1.40e-1	
SRF a	3	modalités dont les sigles et valeurs pivot sont			
		SRF-	SRF=	SRF+	
		3.10e+1	1.06e+2	2.92e+2	
ACH a	3	modalités dont les sigles et valeurs pivot sont			
		ACH-	ACH=	ACH+	
		2.00e-1	4.81e-1	7.98e-1	
SLR a	3	modalités dont les sigles et valeurs pivot sont			
		SLR-	SLR=	SLR+	
		1.21e-1	2.50e-1	5.00e-1	
FRM a	3	modalités dont les sigles et valeurs pivot sont			
		FRM-	FRM=	FRM+	
		1.00e+2	2.70e+2	6.67e+2	
INV a	3	modalités dont les sigles et valeurs pivot sont			
		INV-	INV=	INV+	
		1.71e-2	5.91e-2	1.08e-1	

### *Bornes et valeurs typiques des quotients*

Il faut seulement prendre garde au fait que toutes les considérations numériques de la présente étude ne concernent que les entreprises de l'*Atlas* ayant fourni des données complètes, au moins pour une année; et que ces entreprises, toutes de grande taille, se distinguent par une propension à la communication qui interdit de les prendre pour représentants de l'ensemble des usines de France.

On a ainsi, par exemple pour tendance moyenne de la "Spaciosité", SPC=, la valeur 6.82e-1, soit 68,2 m. carrés par personne. Pour la "Rémunération", REM=, (compte tenu des coefficients {97, 100, 103} introduits pour ramener l'évaluation des dépenses à 1988), donne un salaire annuel de 134 kF par personne, en 1988; ceci nous confirme qu'il ne s'agit pas d'entreprises ordinaires, mais que qualification et salaire sont, en général, élevés. Le taux moyen de cadre approche de 6% (5,88 pour CDR=). SURF= correspond à 106 m carrés pour un budget de 1 MF. Les achats constituent moins de la moitié du budget (0,48 pour ACH=). Etc...

## **2 Analyse des données après codage**

On a soumis à l'analyse des correspondances le tableau (287 × 24) croisant l'ensemble des 287 couples (usine, année) pour lesquels on dispose d'informations complètes avec l'ensemble des 24 modalités {-,+} obtenues par codage barycentrique à partir des 8 variables quotient:

Qww:1'Usine Nouvelle: atlas des usines 1989

trace : 1.156e+0  
rang : 1 2 3 4 5 6 7 8 9 10  
lambda : 2004 1710 1284 947 901 820 670 596 540 507 e-4  
taux : 1734 1479 1111 819 779 709 580 515 467 439 e-4  
cumul : 1734 3212 4323 5142 5921 6630 7210 7725 8192 8631 e-4

	SIGJ	QLT	PDS	INR	F 1	CO2	CTR	F 2	CO2	CTR	F 3	CO2	CTR	F 4	CO2	CTR
SPC-	707	44	41	-568	300	71	646	388	108	-144	19	7	12	0	0	
SPC=	248	55	26	102	19	3	-74	10	2	-235	104	24	-248	115	36	
SPC+	860	25	61	769	214	75	-967	337	138	766	212	116	521	98	73	
REM-	645	32	51	-350	68	20	-711	279	96	-733	296	136	70	3	2	
REM=	260	57	25	181	64	9	-16	0	0	47	4	1	-313	191	59	
REM+	610	36	50	30	1	0	667	278	94	588	216	97	432	116	71	
CDR-	588	37	51	-252	39	12	-372	86	30	-850	451	207	137	12	7	
CDR=	186	55	27	130	30	5	-69	8	2	158	45	11	239	102	33	
CDR+	593	33	52	63	2	1	536	158	55	690	261	121	-559	171	108	
SRF-	777	35	51	-994	593	175	476	136	47	286	49	23	-4	0	0	
SRF=	260	57	26	74	10	2	46	4	1	-276	145	34	-230	100	31	
SRF+	805	33	50	941	503	146	-589	197	67	165	15	7	397	89	55	
ACH-	492	27	52	644	188	57	583	154	54	-549	137	64	167	13	8	
ACH=	88	62	21	17	1	0	68	12	2	53	7	1	165	69	18	
ACH+	610	36	44	-525	191	49	-567	223	67	330	76	30	-416	120	65	
SLR-	729	35	52	-813	381	114	-723	302	106	151	13	6	-238	33	21	
SLR=	127	59	24	176	67	9	33	2	0	160	55	12	36	3	1	
SLR+	638	31	47	563	182	50	733	309	99	-467	125	53	194	22	12	
FRM-	210	39	45	249	47	12	-261	51	16	98	7	3	-373	105	58	
FRM=	50	63	22	53	7	1	113	31	5	-52	7	1	-47	5	1	
FRM+	360	23	52	-579	125	38	138	7	3	-23	0	0	778	227	145	
INV-	370	43	42	-492	216	52	-152	21	6	277	68	26	270	65	33	
INV=	146	50	27	-77	9	1	31	2	0	-227	82	20	184	53	18	
INV+	493	32	61	797	285	100	159	11	5	-18	0	0	-662	196	146	

*Analyse de correspondance: ensemble des 24 modalités*

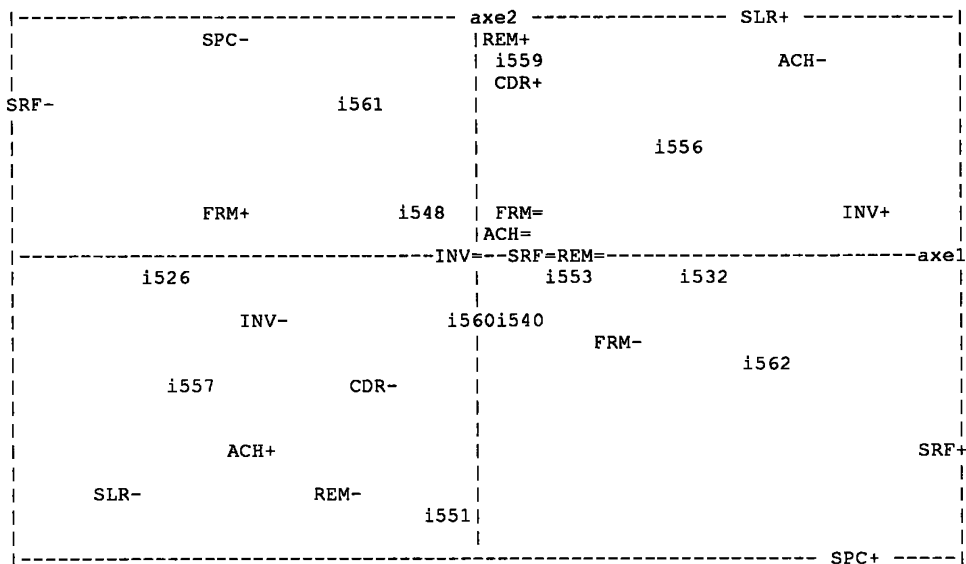
{SPC, REM, CDR, SRF, ACH, SLR, FRM, INV}.

On a ensuite effectué des classifications ascendantes hiérarchiques sur les deux ensembles en correspondance; en créant des listages d'aide à l'interprétation Facor et Vacor.

Enfin, à titre complémentaire, on a effectué des analyses et classifications en mettant en éléments supplémentaires les modalités des deux variables {SPC, SRF} afférentes à la superficie.

**2.1 Analyse de correspondance**

Avec 24 modalités qui, par groupes de 3, sont assujéties à la relation d'avoir pour somme 1 sur chaque ligne, on a 16 facteurs non triviaux. Le listage montre le rôle prépondérant des 3 premiers facteurs qui expliquent bien toutes les variables, à l'exception de la FoRMation . Comme il est impossible de figurer le détail des 287 individus (usine, année), on a projeté sur les graphiques des plans (1,2) et (1,3) les centres de gravité des classes obtenues ultérieurement (cf. §2.2).



On note d'abord que les modalités v= sont groupées au voisinage de l'origine: certaines d'entre elles sont même omises des graphiques du fait de la densité de l'accumulation des sigles. La classification automatique range dans une seule classe les huit modalités v=; et le centre de cette classe ne s'éloigne sensiblement de l'origine que sur le demi-axe ( $F_5 > 0$ ).

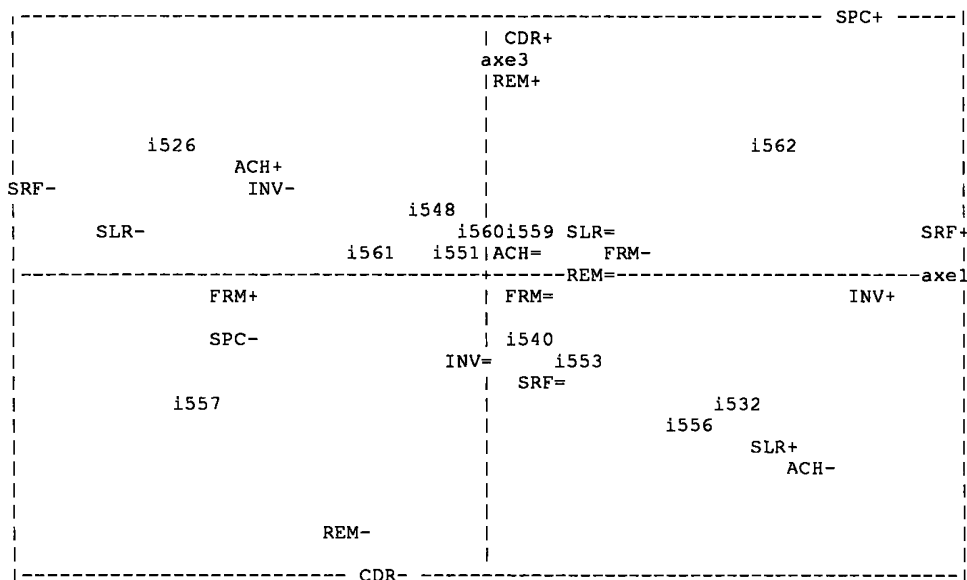
Les deux variables SPC, spaciosité, et SRF, calculées à partir de la superficie mais avec des dénominateurs différents (respectivement le nombre d'ouvriers, Ouvr, et le budget, Budg), ne se séparent guère l'une de l'autre dans le plan (1,2); mais, sur l'axe 3, SPC+ est associé à CDR+: ce qui suggère que les entreprises qui emploient le personnel le plus qualifié sont aussi celles qui offrent à chacun le plus d'espace.

SLR et REM sont également deux variables calculées avec un même numérateur (la masse salariale) mais des dénominateurs différents (Ouvr et Budg); proches dans le plan (1,2), elles se séparent sur l'axe 3, où on ne s'étonnera pas de voir REM+ associé à CDR+.

Dans le quadrant ( $F_1 < 0$ ;  $F_2 < 0$ ), on trouve ACH+ avec {SLR-, REM-, CDR-, INV-}: cette constellation caractérise des entreprises réalisant une faible valeur ajoutée: des ACHats, mais peu d'INvestissements; et un personnel peu qualifié. Dans le quadrant opposé, ( $F_1 > 0$ ;  $F_2 > 0$ ) on trouve les modalités de signe contraire, donc le tableau d'une forte valeur ajoutée.

Il est difficile de placer la FoRMation dans le tableau global sans recourir à la CAH; en saisissant les données, nous avons noté que la formation pouvait





aller avec un haut niveau de technique; mais qu'elle pouvait aussi être développée par une entreprise dont les effectifs décroissent, pour permettre à des ouvriers licenciés de se placer ailleurs.

L'examen du plan (1,3), complété par la lecture de la colonne CTR3 du listage, montre que que l'axe 3 est créé par l'opposition entre les deux groupes de modalités {CDR+, REM+, SPC+} ( $F3 > 0$ ) et {CDR-, REM-} ( $F3 < 0$ ).

Quant à l'ensemble des individus, la lecture attentive du listage (non publié) atteste que les usines pour lesquelles on dispose des chiffres de trois années successives ont généralement des profils stables; bien que certaines données primaires, comme les investissements, varient fortement avec les années; et que de nombreuses entreprises soient en pleine restructuration... On conjecturera donc que la typologie qui ressort de nos données représente des formes d'activité difficiles à modifier; même si, d'autre part, un bilan précis montre que les classes issues de la CAH ne correspondent guère à la répartition des usines par secteur.

## 2.2 Classification ascendante hiérarchique

### 2.2.1 L'ensemble des modalités des variables quotient

La classification des 24 modalités s'interprète parfaitement en des termes que nous a déjà suggérés l'interprétation de l'analyse factorielle: valeur ajoutée, potentiel humain.

SPC-	SRF-	39	42	46
FRM+	F4++			potentiel humain
REM+		38		F1--, F2+++
CDR+	F4--		F3+++	
INV-		37	41	43
FRM-				valeur ajoutée faible
SLR-	ACH+			F1--, F2---
REM-	CDR-			
{==}	F5+++	36		45
INV+	F4--	40	44	
SLR+	ACH-		v. ajoutée maxima	F1++++
SRF+	SPC+			

Sur l'arbre de la partition en douze classes, il a été facile de marquer la composition des classes dont chacune compte une ou deux modalités; à l'exception de la classe 36, notée {==}, qui comprend les modalités centrales des 8 variables quotient. De plus, en reportant quelques indications du listage Facor, on a pu situer les centres de gravité des classes sur le système des axes factoriels.

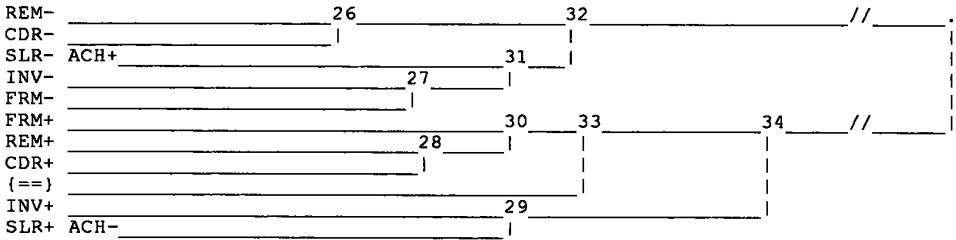
La classe {==} ne se sépare de l'origine que sur l'axe 5. Elle s'agrège à un haut niveau à la classe 44, étiquetée F1++++, et qui comprend en effet un ensemble de modalités se projetant sur le demi-axe (F1>0); avec l'interprétation claire (fondée sur ACH-): "valeur ajoutée maxima".

Il faut toutefois noter qu'avec des investissements et une masse salariale relativement élevée, {INV+, SLR+} on ne trouve pas dans cette classe les caractères d'un personnel hautement qualifié: REM-, CDR+: ces modalités, avec FRM+, constituent une autre classe qu'on a pu étiqueter: "potentiel humain". En effet, une valeur ajoutée élevée peut résulter de la mise en œuvre d'équipements perfectionnés plus encore que de la qualification du personnel.

Il vaut la peine de considérer comment la superficie s'agrège aux autres variables: {SRF+, SPC+} vont avec "valeur ajoutée maxima"; tandis que {SRF-, SPC-} s'agrègent avec "potentiel humain". En effet, un personnel hautement qualifié peut mettre en œuvre un matériel perfectionné mais léger dans des laboratoires qui ne s'étendent pas dans l'espace comme le fait une usine ordinaire.

Enfin, dans la classe 43, étiquetée "valeur ajoutée faible", on trouve, associées à ACH+, les modalités négatives de toutes les variables autres que celles de la superficie (et que ACH-, cela va sans dire!).

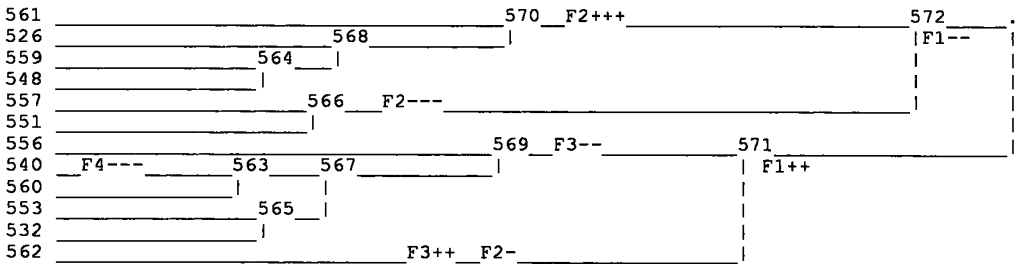
On remarquera que FRM+, (agrégée à {SRF-, SPC-}), est dans la classe 42, "potentiel humain"; tandis que FRM- s'agrège d'abord à INV-, au sein de 43, "valeur ajoutée faible". On notera encore la parfaite opposition entre ACH et SLR: leurs modalités de signes opposés s'agrègent d'abord entre elles avant de s'agréger à toute autre modalité: {ACH-, SLR+} et {ACH+, SLR-}. Tandis que REM et CDR vont ensemble {REM+, CDR+}, {REM-, CDR-}.



30=42 de la précédente; 29=44 ; 32=43

Si dans l'analyse factorielle on met en supplémentaire les modalités des variables de superficie, on aboutit à une classification des 18 modalités restantes qui s'accorde pleinement avec celle que nous venons de commenter; avec une classe {==}, et trois classes principales dont la composition est la même, comme on l'a noté sous le graphique de l'arbre.

### 2.2.2 Classification des entreprises



Effectuée sur un ensemble de 287 éléments (usine, année), cette classification n'est pas d'une interprétation directe. Nous avons d'abord étiqueté l'arbre de la CAH en terme de facteurs, grâce au listage Facor, ce qui permet de compléter, en tenant compte de l'ensemble des dimensions de l'espace, les représentations données au §2.1 dans les plans (1,2) et (1,3).

Puis nous avons eu recours à la procédure d'aide à l'interprétation Vacor. Plus précisément, nous avons demandé l'interprétation de la CAH sur I, non en terme des 24 modalités des variables quotient, mais (pour plus de clarté) relativement à la partition de J en 12 classes adoptée au §2.2.1.: c'est ce qu'on appelle: faire Vacor sur Jq pour I.

Dans sa version actuelle, le programme de CAH produit, en même temps que le listage Vacor, un tableau cumulé 'tiqjq' croisant les deux partitions (dans notre cas en 12 classes) adoptées pour les ensembles I et J. Le format de ce tableau est tel qu'il peut être soumis à l'analyse des correspondances.

c	Partition en 12 classes : Sigles des individus de la classe numéro c														
561	AVI7	VRa9	ELc7	ELn8	ELn7	VRa8	VRa7	CHa7	CHa9	CHa8	CHf7	CHf8	CHf9		
		MM17	ELn9	MM18	MMe7	MMe9	MMe8	MM19	MMg9	MMf7	MMf8	MMf9	MMg8	MMg7	
		ELj7	ELj8	TPe7	EQv7	EQv9	EQv8	MMh9	MMh8	MMh7	MMk7	MMk8	MMk9	EL18	
		EL17	MCA8	MCA9	VRb7	VRb8	VRb9								
526	INc8	INc9	CHI7	CHI9	CHI8	ELh9	INb9	ELh8	ELh7	EQq9	INb8	INb7			
559	ELf8	ELf9	INa8	ELf7	INa9	INa7	MMj9	MMj7	MMj8	ELk7	ELk9	ELk8			
		EQf7	EQf8	ELe9	ELe8	ELe7									
548	EQp7	EQp8	EQp9	EQb7	EQb8	TPc7	TPc9	TPc8	EQx9	EQx8	TPf8	TPf9			
		MMi7	MMi9	TPb7	TPb9	TPb8	MCb7	MCb8	MMd7	MMd8	TPf7				
557	EQh7	BTe7	EQh9	BTe8	BTe9	EQh8	ELd8	ALe7	ALe8	ALe9	EQk7	ELg7	MMb7	TRa7	
ALd7	ALd8	ALa8	ALA7	TRa8	MMc9	MMc8	TRd7	TRa9	TRd8	MMc7	MMb9	ELg9	MMb8	ELg8	
551	SDg8	SDg7	SDh7	SDh9	SDh8	SDF7	SDe8	SDe7	BTh8	BTh9	BTh7				
556	MCk7	MCk8	ELc9	ELc8	MCf7	MCf9	MCf8	VRe8	VRe9	EQo9	EQo8	ELb7	MMm9	MMm8	VRe7
MCd7	SDc7	ELa9	ELa8	ELA7	EQt8	EQt7	EQt9	ELi7	ELi8	ELi9	EQn8	EQn9	BTA9	BTA8	BTA7
540	TRc7	BTi8	MCg7	MCg8	EQa7	EQa8	EQq8	MMi8	EQq7	ALc9	ALc7	ALc8			
560	CHh9	CHh7	CHh8	SDa7	SDa9	SDa8	EQs8	EQs9	CHb7	CHb8	CHb9	TPh8	SDd7	SDd8	
TRe9	TRe7	TRb8	TRb7	BTj7	BTj8	BTj9	EQl8	TPg7	TPg9	TPg8	MCh8	VRf8	VRf7	BTd7	
553	CHd7	MMA8	ELd7	MMA7	EQw7	EQw9	EQw8	ELb8	ELb9	EQd7	EQd9	EQd8	VRd7	VRd9	
		VRd8	EQu8	MCC8	MCC7	EQu9	ELm7	ELm8	ELm9	TPd8	TPd9	TPd7	EQj7	EQj8	
532	MCE9	MCE7	EQu7	EQi7	MCj7	MCj8	EQi8	EQi9	MCE8	EQe8	EQe9	EQe7			
562	EQg7	EQg8	EQg9	EQr7	EQr9	EQr8	ENa8	TPi9	TPi8	SDb7	CHg7	CHg9	CHg8	MCi8	
		BTg9	BTg8	BTf7	EQc8	EQc9	TPa7	EQm8	EQm7	BTd9	BTd8	BTf9	BTf8		
		EQm9	EQc7	VRc7	BTc7	BTb7	BTb8	BTb9	CHc9	BTc8	BTg7	ALb7	CHe9	CHe7	CHe8

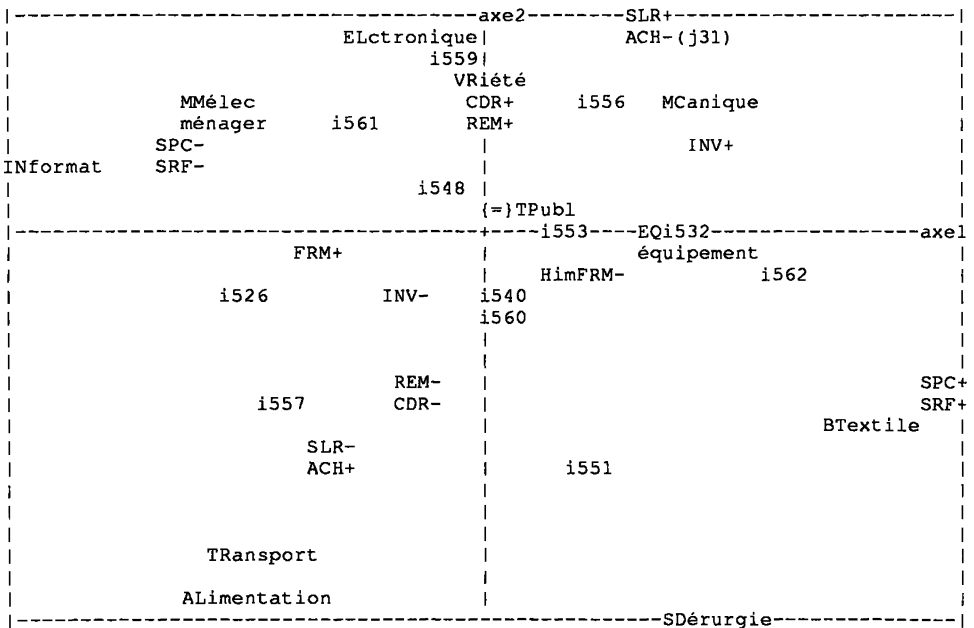
Par rapport à la lecture fastidieuse des tableaux du listage Vacor, cette analyse a l'avantage de construire des graphiques plans (directement interprétables, même s'il s'impose d'en vérifier les suggestions sur le listage Vacor) où figurent les deux ensembles Iq et Jq, caractérisés par la correspondance qu'ils ont entre eux d'après le tableau cumulé 'tiqq'; ce qui représente une simplification par rapport à l'analyse du tableau initial croisant I et J. Dans le cas présent, le plan (1,2) fournit une représentation satisfaisante.

Mais l'interprétation des classes en termes de variables (ou d'agrégats de modalités de variables) ne suffit pas: il faut tenter de retrouver les entreprises elles-mêmes, dans leur activité industrielle. Le listage du contenu des classes montre d'abord la stabilité des profils des entreprises au cours du temps: non seulement les profils successifs sont dans une même classe de la partition retenue, mais, au sein de cette classe, ils sont très proches; ce que l'on voit sur l'arbre (non publié) de la CAH générale; et également sur le tableau du contenu

Qww: l'Usine Nouvelle atlas des usines 1989  
 cariq = 12 ; carjq = 12

	SPC-	FRM+	REM+	CDR+	INU-	FRM-	SLR-	REM-	(==)	INU+	SLR+	SRF+
12	j26	j21	j6	j9	j22	j19	j29	j35	j36	j24	j31	j28
i561	66	18	30	8	19	5	9	14	165	5	21	1
i526	18	1	3	5	8	5	21	2	33	0	0	0
i559	19	5	7	17	4	1	6	1	47	10	18	1
i548	14	2	3	19	11	8	12	3	97	1	3	2
i557	31	10	0	0	12	9	40	49	75	3	0	3
i551	0	5	2	0	7	0	15	12	27	0	3	17
i556	18	1	7	5	4	15	0	18	111	16	46	8
i540	4	2	2	2	0	4	12	11	46	10	2	2
i560	6	1	13	4	15	15	24	9	129	2	3	11
i553	6	2	1	2	5	2	4	19	150	4	10	12
i532	0	2	2	0	2	6	0	10	48	3	12	10
i562	0	3	15	13	11	21	16	11	127	19	16	67

trace = 1.1562645e+0  
 fin d'écriture de 9:usi\$QjUacoriq



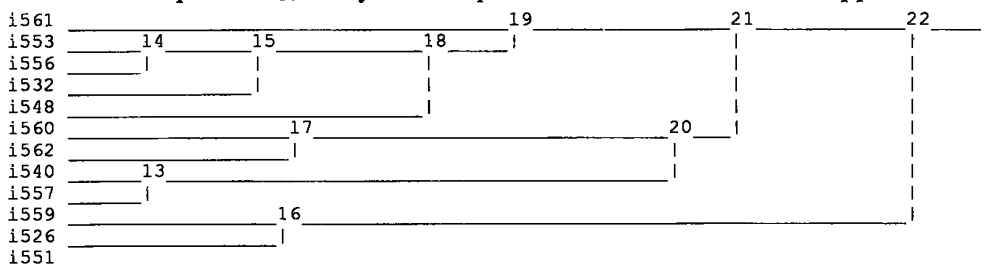
des classes où sont presque toujours groupés les sigles qui ne diffèrent que par le 4-ème caractère (chiffre de l'année après le sigle propre à l'usine; cf. §1.1: e.g., dans la classe 526, CHi7, CHi8, CHi9; les trois années de l'entreprise 'i' du secteur 'CH': "CHimie, parachimie et matériaux").

Reste à placer les secteurs dans notre typologie. À cette fin, on a construit un tableau de contingence croisant secteurs et classes de la partition. Ce tableau donne, e.g., à l'intersection de la ligne i561 et de la colonne EL, le nombre (8)

c	Partition en 12 classes : composition de la classe numéro c											
secteurs: 12	AL	TR	BT	CH	VR	EL	EQ	IN	TP	MM	MC	SD
i561   AV	0	0	0	6	6	8	3	0	1	18	2	0
i526	0	0	0	3	0	3	1	5	0	0	0	0
i559	0	0	0	0	0	9	2	3	0	3	0	0
i548	0	0	0	0	0	0	7	0	9	4	2	0
i557	7	5	3	0	0	4	4	0	0	6	0	0
i551	0	0	3	0	0	0	0	0	0	0	0	8
i556	0	0	3	0	3	9	7	0	0	2	6	1
i540	3	1	1	0	0	0	4	0	0	1	2	0
i560	0	4	4	6	2	0	3	0	4	0	1	5
i553	0	0	0	1	3	6	10	0	3	2	2	0
i532	0	0	0	0	0	0	7	0	0	0	5	0
i562   EN	1	0	13	7	1	0	12	0	3	0	1	1

des couples (usine, année) afférents au secteur "Electronique - bureautique" et rentrant dans la classe i561.

La lecture directe de ce tableau ne suggère pas de conclusion tranchée. C'est pourquoi on l'a d'abord soumis aux programmes d'analyse des correspondances et de CAH. On obtient ainsi une nouvelle structure arborescente sur l'ensemble des 12 classes iq (de i561 à i562): or celle-ci ne concorde pas avec celle publiée au début du présent §; il n'y a donc pas de lien étroit entre les rapports de



proximités entre classes iq du point de vue de leur description par les variables et du point de vue de leur composition sectorielle.

Pourtant, cette composition sectorielle n'est pas dépourvue de sens. On le voit en adjoignant les profils des secteurs sur l'ensemble Iq des 12 classes (colonnes du tableau ci-dessus) en colonnes supplémentaires à l'analyse du tableau tiqjq. Un tableau cohérent apparaît alors dans le plan (1,2). Dans le quadrant ( $F1 < 0$ ,  $F2 > 0$ ), on a IN, MM et EL associés à faible superficie {SPC-, SRF-}; plus précisément, sur le demi-axe ( $F2 > 0$ ) EElectronique est associé à {CDR+, REM+}: personnel, hautement qualifié. A l'opposé, on a Bois-Textile et SiDérurgie. Dans le quadrant ( $F1 < 0$ ,  $F2 < 0$ ), TTransport et ALimentaire sont associés à une faible valeur ajoutée. Il s'agit en fait de secteurs qui élaborent des produits de base de valeur élevée. Le cas de TR est particulier: du fait de la sous-traitance, une usine travaillant dans l'industrie du matériel de transport effectue des ACHats de valeur élevée.

**N.B.** Pour adjoindre en supplémentaire au tableau 'tiqq' le tableau de contingence croisant Iq avec l'ensemble des secteurs, deux voies s'offrent à l'utilisateur du logiciel MacSAIF. On peut, après avoir recopié ce dernier tableau en format 'ww' par l'option 'C' de 'zrang', juxtaposer les deux tableaux par 'juxtab' (cf. *CAD*, Vol XIV, n°3, pp. 347-354); et effectuer l'analyse du tableau ainsi construit en spécifiant que les profils des secteurs sont en supplémentaire. On peut encore transposer le tableau croisant Iq et les secteurs (par l'option 'T' de 'zrang'); et, après avoir donné au tableau obtenu le même nom qu'à 'tiqq' avec en plus le suffixe 'b', utiliser le programme 'qorlsup' (cf. *CAD*, Vol XIV, n°1, pp. 48-50) pour adjoindre à l'analyse de 'tiqq' un tableau de colonnes supplémentaires extérieures.

### 3 Conclusions

Nous pensons avoir montré que l'analyse des données chiffrées publiées dans l'*Atlas des Usines* révèle une structure cohérente et interprétable. Il faut seulement prendre garde à ce que les données sont trop lacunaires pour être représentatives, mêmes des grandes entreprises qui sont seules recensées par l'*Atlas*.

Quant à la méthode, on notera que les traitements ont été effectués avec le seul logiciel MacSAIF; à l'exception du calcul des variables quotient; pour lequel on a écrit un programme particulier, mais qui aurait pu être fait sur un tableur usuel.

### Référence Bibliographique

*L'atlas des Usines, 1000 établissements industriels répertoriés par groupes par secteurs et par régions*; Numéro Spécial de *L'USINE NOUVELLE*; 1989.