

A. KOUTSANDRÉAS

F. DOSSOU

Choix des individus concernés et codage des questions par groupe dans une enquête : cas des créateurs qui développent leur entreprise

Les cahiers de l'analyse des données, tome 11, n° 1 (1986), p. 57-66

http://www.numdam.org/item?id=CAD_1986__11_1_57_0

© Les cahiers de l'analyse des données, Dunod, 1986, 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/>

CHOIX DES INDIVIDUS CONCERNÉS ET CODAGE DES QUESTIONS PAR GROUPE DANS UNE ENQUÊTE : CAS DES CRÉATEURS QUI DÉVELOPPENT LEUR ENTREPRISE

[DEVELOP. ENTREPRISE]

par A. Koutsandréas*, F. Dossou**

0 Rappel sur l'enquête : problème et méthode : Dans la conjoncture économique difficile des années 80, la création d'entreprises dynamiques apparaît à tous grandement souhaitable. Les pouvoirs publics encouragent des modes nouveaux de formation à la création d'entreprises. Au ministère du travail le Centre d'études de l'emploi a, au cours de l'été 1983, lancé une vaste enquête par correspondance sur les "Effets des initiatives locales et des formations à la création d'activités sur le développement des entreprises et sur l'emploi". Proposant à une population très hétérogène de créateurs potentiels, plus de 150 questions réparties en 11 thèmes, le centre a reçu quelque 800 bordereaux de réponses exploitables. Les tendances générales révélées par l'enquête sont exposées dans un premier article auquel nous renvoyons (cf. [CREATEURS D'ENTREPRISE] in CAD Vol X n° 4, 1985). La structure complexe des thèmes et questions emboîtés, a occasionné de nombreuses non-réponses, étudiées pour elles-mêmes ailleurs (cf. [NON REPONSE] in CAD Vol X n° 4, 1985). Ici nous concentrons notre attention sur une classe d'enquêtés particulièrement intéressante : les créateurs qui développent leur entreprise. Au § 1 nous exposons en détail la méthode suivie pour trier les individus et coder l'information pertinente. Au § 2 nous donnons une vue d'ensemble des résultats, sans entrer dans le détail des modalités de réponse au questionnaire.

1 Choix et codage de l'information pertinente : Nous considérerons successivement le choix des individus (§ 1.1) et le codage des groupes de questions qui offrent successivement plusieurs modalités de réponse non exclusives à un même problème (§ 1.2). Dans le cadre du présent article, nous ne pouvons donner le détail des questions retenues, mais seulement la structure du tableau analysé avec l'enchaînement des analyses faites ou à faire (§ 1.3).

1.1 Choix des individus : Pour définir une classe d'individus d'après leurs réponses à une enquête, on se fixe généralement une variable "clé", qui sert de critère : un seul passage du fichier permet d'extraire la sous-population recherchée. Une voie prometteuse pour améliorer le tri des individus nous semble être de faire préalablement une typologie d'après les non-réponses que présentent leurs bordereaux. Mais ici, non seulement la structure du questionnaire est complexe, mais la saisie des réponses n'a pas été assez fine : dans bien des cas, la réponse NON et l'absence de réponse, ou non-réponse NR sont confondues dans une seule modalité, que nous noterons en bref NNR. Dès lors on connaît mal les mobiles réels d'un sujet qui a omis de fournir à une question une réponse positive explicite : est-il non concerné ? Dissimule-t-il par prudence une information qui pourrait être utilisée contre lui par les

(*) Docteur de l'université P. & M. Curie en statistique mathématique.

(**) Chargé de recherche au Centres d'études de l'emploi - Ministère du travail, Emploi et Formation professionnelle.

inspecteurs des impôts ? (cf. *infra* § 2 : le plan (2,3)) ; ou a-t-il répondu NON ? D'autre part, fait sur lequel nous reviendrons au § 1.2 pour réformer le codage, plusieurs blocs de questions auxquelles on demande des réponses distinctes successives, n'offrent en fait que des modalités de réponse non exclusives une seule question : par exemple, si l'on s'enquiert des "obstacles rencontrés" on demande successivement A est-il un obstacle ; B est-il un obstacle ; F est-il un obstacle ; alors que le nombre des obstacles peut varier de 0 à 6.

Dans ces conditions, l'algorithme de choix des individus, nous paraît devoir traiter des informations multiples sous forme de disjonctions ("ou") et de conjonction ("et").

Dans le cas présent, nous partons de l'hypothèse que les créateurs d'entreprise qui envisagent un développement se signalent par des réponses fournies aux questions du Thème 8, intitulé "Votre entreprise demain", et comprenant 21 questions numérotées de 86 à 104. Outre une question générale 94 les questions forment trois blocs.

(84-88) : créations d'emplois au cours des deux ans à venir ; pour ces questions on distingue NON de NR.

(89-93) : moyens de recrutement envisagés pour les emplois en CSP ; pour ces questions, on ne distingue pas NON de NR, mais il y a une seule modalité NNR.

94 : objectifs pour l'entreprise : développer ? Consolider ? Réduire ?

(95-104) : divers obstacles à surmonter (trouver un marché, etc.) ; pour ces questions on distingue NON et NR.

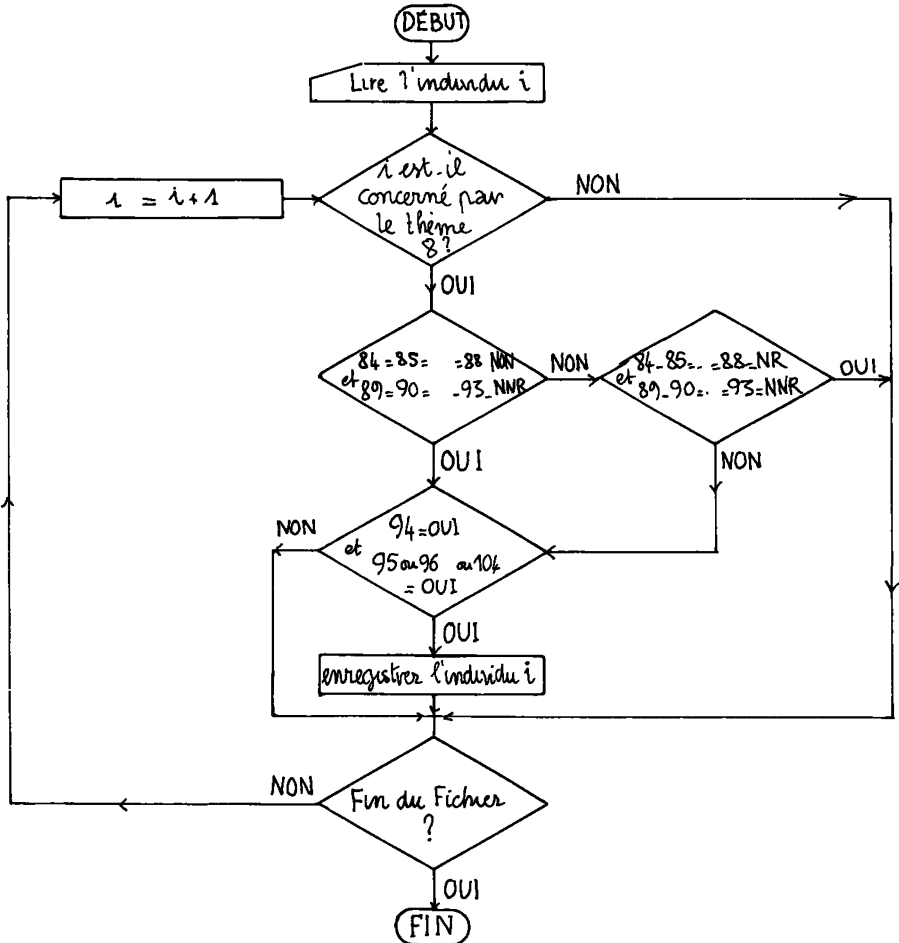
Selon nous, un individu qui, à la fois, a NON à toutes les questions du bloc (84-88) et NNR à toutes celles du bloc (89-93) est un créateur qui ne prévoit pas de nouveaux emplois. Par contre un créateur qui a NR à toutes les questions du bloc (84-88) et NNR à toutes celles du bloc (89-93), n'est pas concerné par le thème 8. Enfin, l'individu qui a répondu à la question 94 peut s'atteler au développement s'il a répondu OUI à l'une au moins des questions du bloc (95-104), signalant ainsi un ou plusieurs obstacles dont il est effectivement préoccupé.

Sur ces principes on a conçu un algorithme, qui a retenu 300 individus considérés comme des créateurs qui développent leur entreprise. Les analyses dont rend compte le § 2 semblent montrer qu'il en est bien ainsi ; à ceci près que quelques individus dont l'entreprise est en cours de création présentent des projets dont on peut mettre en doute la solidité (cf. cl 587 dans le quadrant (F1 + F2-) du plan (1,2)).

1.2 Codage conjoint des réponses à un groupe de questions : Le questionnaire offre plusieurs exemples de groupes de questions, auxquelles on demande des réponses distinctes successives, mais qui ne sont en fait que l'énumération des modalités de réponse non exclusive à une seule question. Nous prendrons ici pour exemple le groupe G3 du thème 10 : "obstacles à surmonter pour réaliser de nouveaux projets de création". Le tableau ci-dessous (dont le contenu sera complètement expliqué dans la suite) montre qu'il s'agit de six questions, numérotées de 135 à 140, ayant chacune deux modalités de réponse OUI et NNR (le cas de trois modalités : OUI, NON, NR, se traiterai de la même manière ; car notre recodage en une seule question prend seulement en compte les réponses OUI).

LES OBSTACLES A SURMONTER (THEME 10)

ANCIEN CODE				NOUVEAU CODE	
QUEST.	SIGLE	MODALITE	EFFECT.	SIGLE	TOTCOL
135	135A	Trouver un nouveau marché :	OUI 28	135o	25
	135B	"	NNR 786		
136	136A	Trouver du personnel qualifié:	OUI 21	136o	19
	136B	"	NNR 793		
137	137A	Réaliser de nouv. investiss. ^t :	OUI 62	137o	56
	137B	"	NNR 752		
138	138A	Lancer un nouveau produit :	OUI 29	138o	26
	138B	"	NNR 785		
139	139A	Trouver d'autres partenaires :	OUI 16	139o	14
	139B	"	NNR 798		
140	140A	Autres obstacles	OUI 40	140o	36
	140B	"	NNR 768		
-	-	1 item au groupe 3	-	G3<1	756
-	-	NON ou N.R. au groupe 3	-	G3N	696



Une question unique complétant toutes les informations apportées par le groupe G3 devra d'abord comprendre 6 modalités, une pour chaque obstacle prévu, héritière des réponses OUI à chacune des 6 questions. Il faut encore une modalité G3N qui sera propre aux sujets ayant fourni la réponse NNR à chacune des 6 questions. Le cas embarrassant est celui des sujets ayant adopté plusieurs modalités d'obstacles qui leur étaient proposées. Mettre plusieurs fois 1 revient à donner à ces sujets un poids supérieur à celui des autres ; ce qui est peu justifié dans la mesure où les multiples obstacles ne sont peut être pas pris au sérieux autant qu'un obstacle unique. Diviser la masse 1 entre tous les obstacles cités nous a paru aller trop loin dans l'émiettement, d'autant plus que les sujets qui citent deux obstacles méritent d'être crus à l'égal de ceux qui en citent un seul. Finalement, compte tenu de ces desiderata contradictoires, on a créé une modalité $G3 \leq 1$ qui vaut 0 pour les sujets ayant donné plusieurs obstacles, et 1 pour ceux qui en signalent 0 ou 1. Le tableau ci-dessous, où les obstacles sont désignés par les six lettres A, B, ..., F indique clairement comment l'on a procédé.

A	B	C	D	E	F	N=0	$N \leq 1$	TOTAL
					1		1	2
	1				1			2
2/3	2/3	2/3						2
2/4	2/4	2/4	2/4					2

La modalite F a été adoptée

Les modalités B et F ont été adoptées

Les modalités A, B et C ont été adoptées

Les modalités A,B,C et D ont été adoptées

⋮

						1	1	2
--	--	--	--	--	--	---	---	---

N.R. ou NON à la question

On remarque que pour tout individu le total des valeurs écrites dans l'ensemble des colonnes est 2. En revenant au tableau placé en tête de ce §, on voit, (e.g.), qu'à l'obstacle "réaliser de nouveaux investissements" cité par 62 sujets sur 814, il correspond une colonne dont le total est seulement 56 ; parce que, quand un sujet cite cet obstacle avec deux autres ou davantage, le nombre inscrit dans la colonne n'est pas 1 mais 2/3 ou 1/2 etc. .

Il va sans dire que le codage adopté ici n'est pas le seul possible : en particulier si on accepte un poids variable, on peut supprimer la colonne $G \leq 1$ (où seuls ont zéro les individus ayant donné strictement plus d'un obstacle). L'essentiel est d'avoir groupé en une seule question toute l'information ; en faisant disparaître cette série de six modalités NNR (NON ou Non Réponse) qui n'a pas de sens réel, dans la mesure où un sujet qui a répondu OUI à une au moins des 6 questions initiales, n'a pas adopté une attitude négative.

1.3 Le tableau de description et son analyse : Sans entrer dans le détail, disons que l'expérience acquise au cours de multiples analyses, nous a permis de diminuer le volume des données retenues. Le tableau principal 300×173 croise les 300 individus (choisis comme on l'a dit au § 1.1), avec 173 modalités qui concernent plus particulièrement l'entreprise et son développement (thèmes 5, 7, 8, 9) ; à quoi on a adjoint 228 colonnes supplémentaires. Ce tableau de description, du fait du codage des choix multiples expliqué au § 1.2, comporte outre des zéros et des uns, quelques valeurs fractionnaires. Toutefois le total est le même pour toute ligne. Ceci permettrait de créer un tableau de BURT kb par

la formule usuelle :

$$kb(j, j') = \Sigma \{k(i, j) k(i, j') \mid i \in I\},$$

l'analyse de ce tableau, donnant les mêmes facteurs sur J que celle du tableau de description, mais avec une normalisation différente (les v. propres étant élevées au carré) ; et les mêmes facteurs sur I (adjoint en ligne suppl. au t. de Burt), sans changement de normalisation. (D'ailleurs même si les lignes sont inégales en poids, on peut considérer un Burt généralisé, kbg :

$$kgb(j, j') = \Sigma \{k(i, j) k(i, j') / k(i) \mid i \in I\}$$

dont l'analyse se rapporte à celle du tableau k initial, comme l'analyse d'un t. de Burt usuel à celle du tableau disjonctif complet dont il est issu). L'intérêt de l'analyse du t. de Burt étant qu'il donne une appréciation plus juste des taux d'inertie afférents aux facteurs, ainsi que des contributions COR de ceux-ci aux modalités j. Toutefois les résultats publiés ici (§ 2) sont issus de l'analyse du tableau de description k lui-même.

On a de plus effectué une classification ascendante hiérarchique sur chacun des ensembles I et J. Sur les graphiques on a mis les centres de gravité des 11 classes de la partition définie par les 10 noeuds supérieurs de la hiérarchie construite sur I (classes d'individus). Nous ne dirons rien de la hiérarchie sur J (classes de modalités) parce que son interprétation met en jeu le détail du questionnaire. Il importe de rappeler ici qu'en croisant deux partitions C et D choisies respectivement sur I et J on obtient un tableau cumulé C x D, avec :

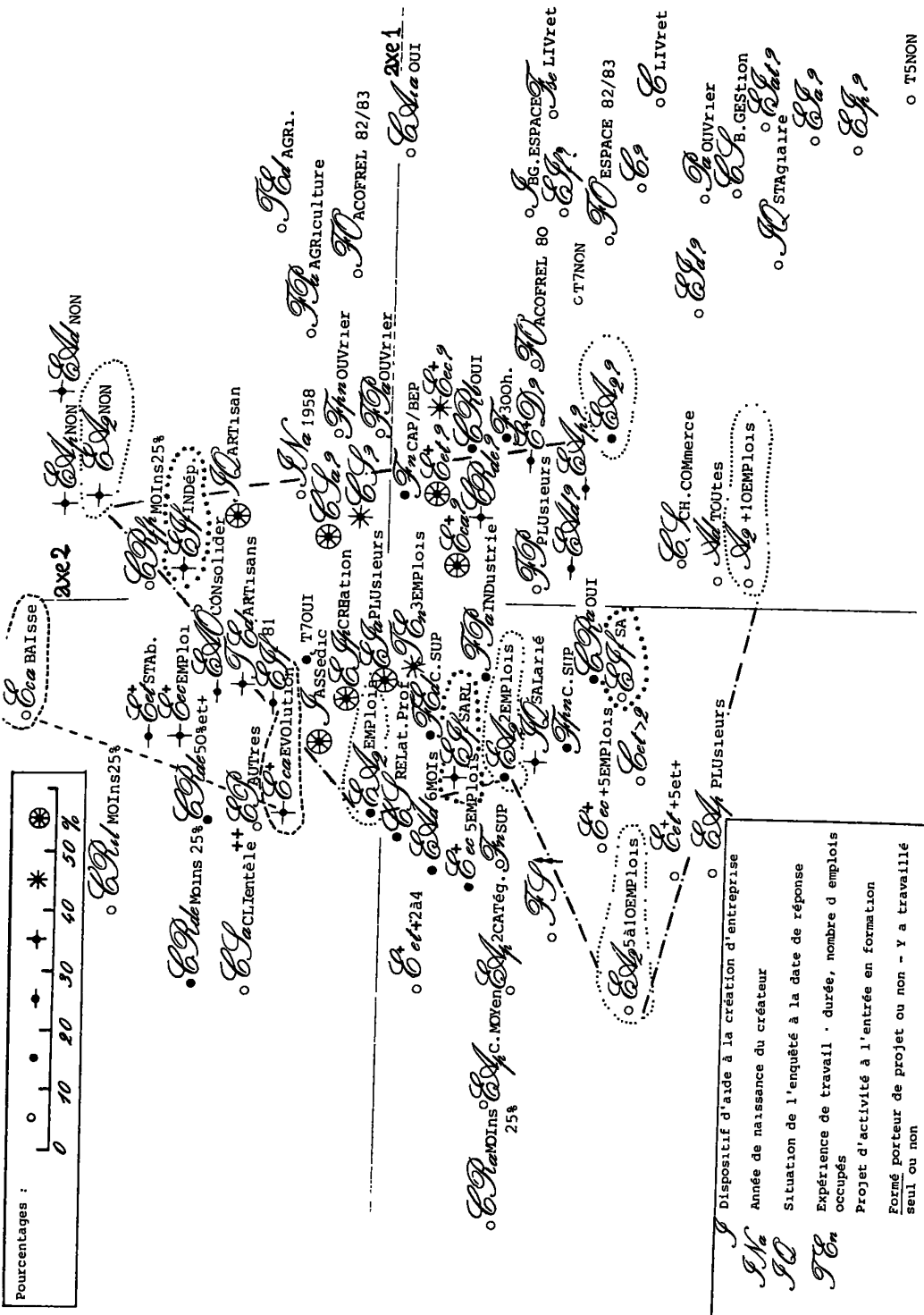
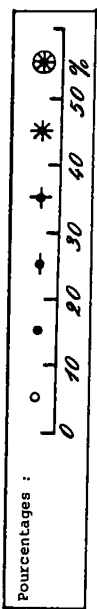
$$k(c, d) = \Sigma \{k(i, j) \mid i \in c ; j \in d\};$$

tableau qui aide grandement à interpréter les deux partitions l'une par l'autre (cf. e.g., Amalberti et coll. : [QUEST. RECH.] in CAD Vol IX n° 4 1984).

2 Résultats d'analyse : Comme on l'a expliqué au § 1.3, nous nous bornerons à rendre compte de l'analyse factorielle du tableau de description I x J ; avec sur les graphiques la partie supérieure d'une hiérarchie construite par CAH sur l'ensemble I des 300 individus retenus (cf. § 1.1).

2.1 Le premier axe : On a du côté (F1+), des sujets jeunes, (nés après 1958), issus de milieu ouvrier agricole, dépourvus d'instruction au-delà du niveau élémentaire, bénéficiant d'un Livret d'Epargne Manuel (LEM) pour créer une entreprise de caractère artisanal. A l'opposé (F1-), on trouve des créateurs plus âgés pourvu d'une instruction secondaire ou supérieure, ayant derrière eux une vie active longue (et éventuellement mouvementée), ayant souvent pour situation, à la date de réponse au questionnaire, d'être salarié (gérant) dans leur entreprise dont la forme juridique était le plus souvent celle d'une société (SARL).

2.2 Le deuxième axe : L'interprétation de cet axe apparaît si l'on considère la suite ordonnée des réponses à la question des projets d'avenir quant aux créations d'emplois sur les deux ans à venir. La réponse Non est du côté (F2+) ; la réponse maxima (10 ou davantage) est du côté (F2-). Mais il est frappant que la réponse dubitative (?) se trouve aussi du côté (F2-). Pour préciser l'interprétation de l'axe 2, il est indispensable de tenir compte de l'axe 1 ; c'est-à-dire de considérer l'interprétation propre aux divers quadrants du plan (1,2).



Dispositif d'aide à la création d'entreprise

Année de naissance du créateur

Situation de l'enquête à la date de réponse

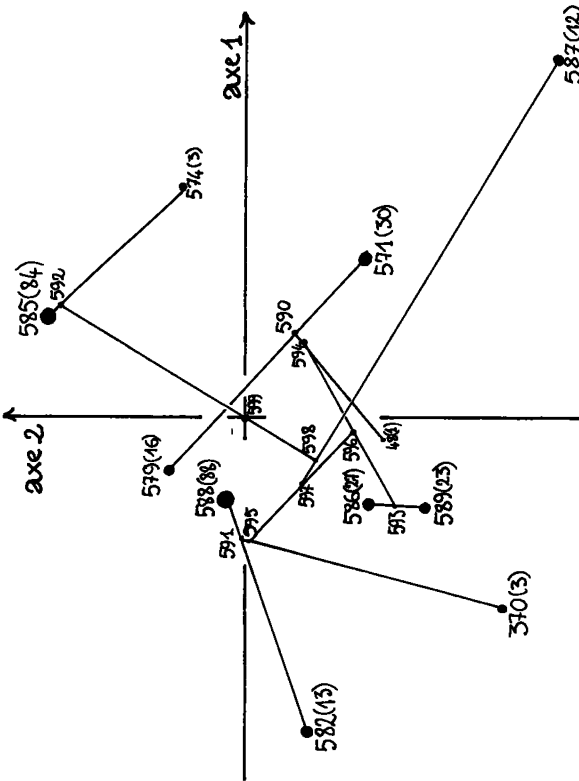
Expérience de travail · durée, nombre d emplois occupés

Projet d'activité à l'entrée en formation

Formé porteur de projet ou non - y a travaillé seul ou non

o T5NON

Formation spéciale suivie : organismes et promotions	585(84) 592
Niveau d'études à l'entrée en formation	574(15)
Niveau professionnel à l'entrée en formation	571(30)
Activité de l'entreprise créée	588(86)
Date de démarrage	579(16)
Forme juridique	591
Provenance	575
Atouts au départ	598
Nombre d'emplois créés au départ	590
Evolution de l'effectif total	594
Evolution du chiffre d'affaires	586(87)
Difficultés de l'entreprise en 1983	595
Créations d'emplois sur les deux années à venir	486
Prévisions d'emplois en C.S.P.	589(23)
Prévisions de dates d'embauche	340(3)
Objectifs pour l'entreprise	570(13)
Conditions à la création : situation antérieure du créateur	587(12)
Part de l'aide aux demandeurs d'emploi ... dans les ressources financières initiales	
Part du fonds personnel ou familial	
Part de l'indemnité de licenciement	
Part de l'apport d'un ou des associés	
Part du prêt bancaire	
Aide publique reçue : primes d'installation artisanale	
Soutiens à la création	
Nature des soutiens	
Type de projet avant création	
Pour ceux qui veulent se lancer dans un autre projet : activité envisagée	

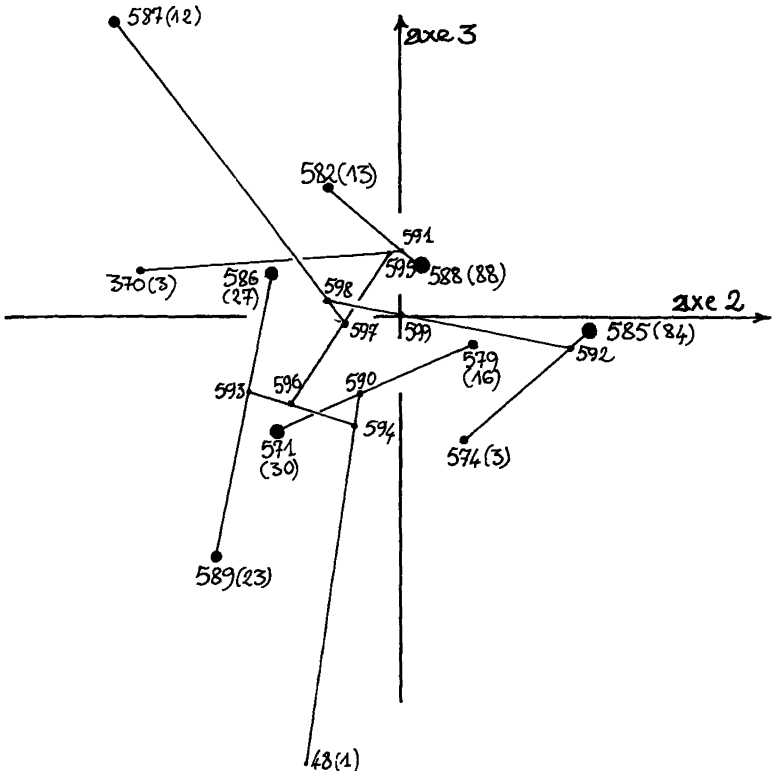


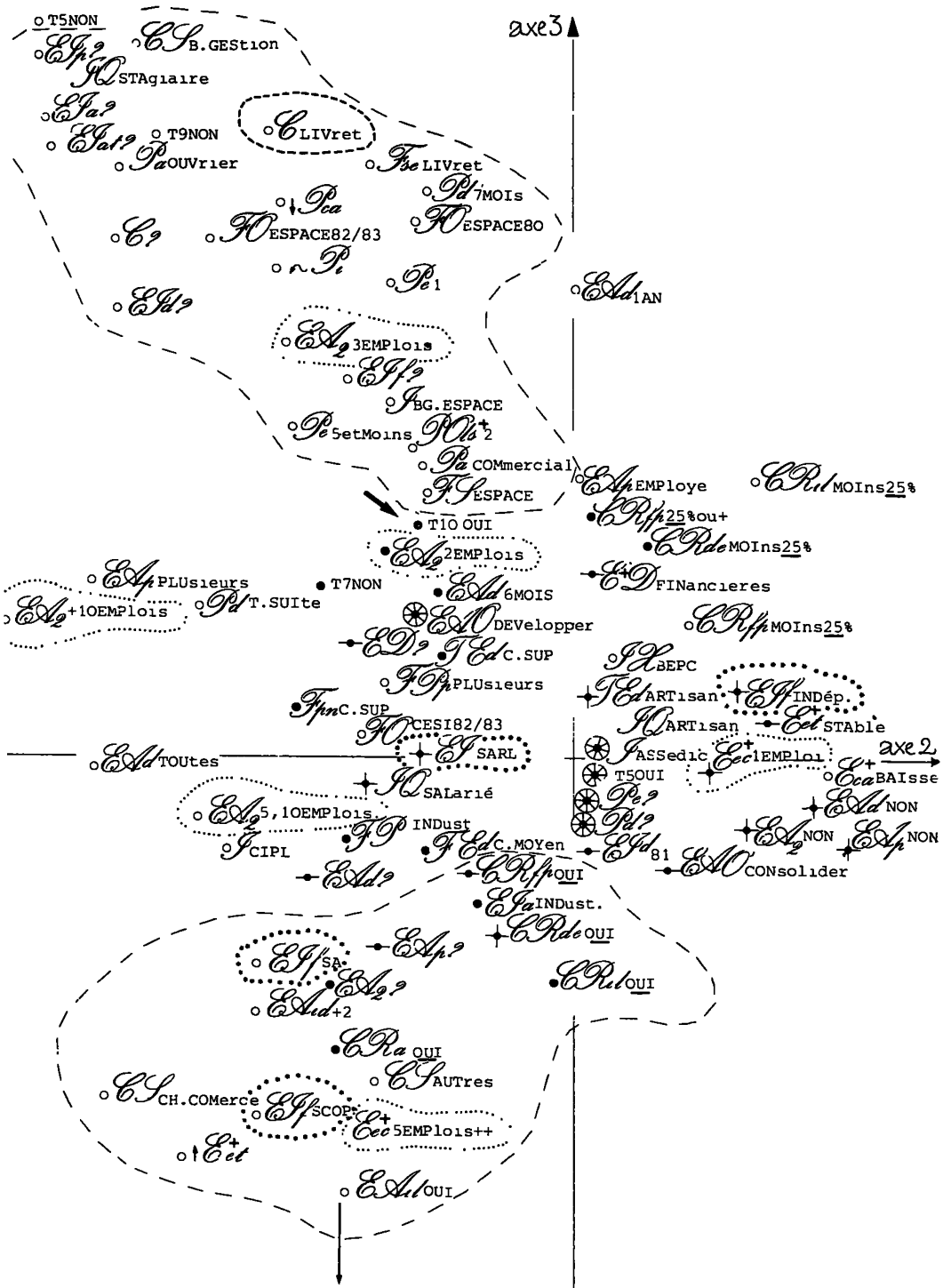
2.2' Le plan (1,2) : A l'écart dans le quadrant (F1+, F2-), attire notre attention la modalité "Stagiaire" en réponse à la question "situation" ; et avec cette modalité plusieurs points d'interrogation quant à l'état présent de l'entreprise (ainsi que les modalités supplémentaires T5 Non, T7 Non qui signalent des non-réponses aux thèmes (5) "création de l'entreprise" et (7) "votre entreprise existe encore"). Manifestement il s'agit d'entreprises en cours de création ; et quant aux projets d'avenir les créateurs encore stagiaires ont le choix entre une réponse évasive (e.g. "nombre d'emplois à créer : ?") et des perspectives plus flatteuses qu'assurées : "créer 10 emplois ou davantage..." ; modalité qui est attirée vers ce quadrant à la différence de celle, plus modeste de "5 à 10".

Dans la quadrant (F1- , F2-) on trouve des sociétés en activité dont le succès semble attesté par des projets d'embauche réaliste ("2 à 5").

Au contraire, dans le demi-plan (F2+), c'est au mieux la stabilité, voire la baisse du chiffre d'affaire ; en particulier dans le quadrant (F1+ , F2+) on a des entreprises artisanales dont la prospérité laisse à désirer. Quant au quadrant (F1- ; F2+), il est relativement moins chargé en individus que les trois autres, ou plus précisément, il compte relativement peu d'individus écartés de l'origine, ainsi qu'on en peut juger d'après la charpente de la classification dessinée dans le plan (1,2) ; les points de ce quadrant représentent des entreprises véritablement déjà créées, mais dont l'activité laisse à désirer.

2.3 Le plan (2,3) : Certains types de créations reconnus dans le plan (1,2) se montrent dans tout leur relief dans le plan (2,3). Du côté (F2+), il y a peu d'étalement suivant l'axe 3, et nous n'apprenons rien de plus quant aux entreprises qui n'annoncent pas de projets d'embauche et d'expansion. Mais du côté (F2-) les différentes formes que revêtent ces projets apparaissent bien différenciés. Dans le quadrant (F2- , F3+), on retrouve la classe des stagiaires créant leur entreprise grâce au LEM, jeunes ouvriers ayant bénéficié des services de la "boutique ESPACE" (Nord et Pas-de-Calais) : ils ont peu à dire quant à la création effective (thème 9 : modalité T9 Non), mais sont fertiles en nouveaux projets (T10 Oui) : des points d'interrogations pour le présent, des assertions pour l'avenir.





Tout au contraire dans le quadrant (F2- , F3-), on a des sociétés (notamment des SCOP) qui sans pouvoir dissimuler le présent, évitent de révéler l'avenir vraisemblablement par méfiance envers le fisc qu'ils imaginent cachés derrière les enquêteurs.

Sans entrer dans l'examen détaillé des classes de la CAH, nous soulignons que leurs centres de gravité offrent un schéma utile de la densité du nuage des "créateurs" ; et engageons le lecteur à en comparer la disposition dans les plans (1,2) et (2,3). En particulier les deux classes 586 et 589 très proches dans le plan (1,2), se séparent suivant l'axe 3 ; le syndrome de méfiance (F3- ; nombreuses omissions quant aux projets d'avenir, indices de prospérité présente) étant particulièrement caractéristique de la classe 589.

2.4 Le quatrième axe : Comme on peut le voir sur le nuage des entreprises, projeté par exemple sur le plan (2,4) (non figuré ici), la distribution des individus sur l'axe 4 est nettement dissymétrique : la majorité est concentrée du côté (F4-) ; mais lui font équilibre du côté (F4+), des individus moins nombreux mais plus dispersés, constituant les classes 586 et 574. Quant à l'interprétation nous dirons que ces individus ont bénéficié d'une formation spéciale.

Il se peut que la dette contractée envers le formateur rende ceux de la classe 586 plus loquaces, moins circonspects, que leurs homologues de la classe 589 (cf. *supra* plan (2,3)).

2.5 Le cinquième axe : Sur cet axe, l'opposition concerne principalement les moyens personnels ou familiaux dont les créateurs ont pu disposer. Du côté (F5+) ces moyens sont importants, pouvant aller jusqu'à fournir la totalité du capital requis ; corrélativement, les créateurs signalent comme difficulté qui les préoccupe le plus les rapports avec l'administration ; leurs incertitudes sont relatives à la difficulté de trouver du personnel qualifié, et aussi des marchés. Du côté (F5-), les créateurs, qui ont parfois reçu la totalité du capital nécessaire d'un service d'aide aux demandeurs d'emplois, ou d'une banque, ont du fait de cette dépendance toutes autres préoccupations.

3 Conclusion

La création et le développement de petites entreprises recèlent, on le sait, beaucoup de possibilités d'emploi, dont la mise en valeur est engagée.

La C.A.H. subdivise les créations en un certain nombre de groupes distingués par les caractéristiques et conditions du démarrage, l'évolution de la main d'oeuvre et celle du chiffre d'affaires, les objectifs de développement et les obstacles à surmonter. Sous une même forme juridique (e.g. dans les sociétés SA, SARL, etc.) on constate souvent une diversité qui rend bien compte des conditions de vie des entreprises naissantes. A côté des petites entreprises et coopératives de type traditionnel, on trouve des unités ayant un potentiel de croissance et qui pourraient accélérer leur développement à des conditions qui sont, selon les cas, techniques, commerciales, ou de financement. Mais il reste beaucoup à faire pour en assurer le succès.