

A. KOUTSANDRÉAS

Analyse de la distribution des non-réponses dans une enquête sur les créateurs d'entreprises

Les cahiers de l'analyse des données, tome 10, n° 4 (1985),
p. 437-451

http://www.numdam.org/item?id=CAD_1985__10_4_437_0

© Les cahiers de l'analyse des données, Dunod, 1985, 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/>

ANALYSE DE LA DISTRIBUTION DES NON-RÉPONSES DANS UNE ENQUÊTE SUR LES CRÉATEURS D'ENTREPRISES

[NON-RÉPONSE]

par A. Koutsandréas *

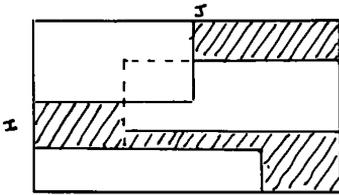
L'exploitation d'une enquête sur les créations d'entreprise (cf. [CREATEURS D'ENTREPRISE] in CAD Vol X n° 4 1985) a révélé un taux de non-réponse d'au moins (1/3) chez tous les sujets. L'ampleur et l'intérêt de ce phénomène (§ 1) nous ont conduit à chercher une typologie des sujets d'abord selon la distribution par thèmes des questions auxquelles ils omettent de répondre (§ 2) ; et ensuite selon le détail des questions laissées sans réponse (§ 3).

1 Non-réponses aléatoires et non-réponses systématiques : Toute enquête de quelque importance est en butte aux non-réponses. Dès les débuts de l'application de l'analyse des correspondances à des données issues d'enquêtes, il est apparu que les non-réponses pouvaient être distribuées non au hasard comme si elles avaient pour cause une inattention momentanée, mais systématiquement, suivant une structure. Prenons, dans le traité sur l'Analyse des Données (TIIC n° 6), l'*Etude sur les conditions du développement de la Colombie*. Quarante cinq unités territoriales ou *Municipes* sont notées, en principe, suivant 306 critères ; mais il y a, dans les tableaux des notes, de nombreux vides, dont le caractère systématique perturbe la première analyse effectuée sur les notes. Une étude propre consacrée aux omissions révèle alors la distribution géographique de celles-ci, précisée par une présentation cartographique des résultats de la CAH. L'analyse des notes est alors reprise en se bornant aux 172 critères ne comportant pas d'omission. Le directeur de l'enquête (A. Bison : d'Economie et Humanisme) s'explique ainsi " Sur le terrain même, nous savions bien que les deux enquêteurs du Sud étaient plus faibles que les autres ; et c'est nous qui avons exigé qu'ils ne donnent pas de réponse plutôt que de donner une réponse arbitraire et fantaisiste".

Dans l'enquête [COLOMBIE], plus de la moitié des questions recevaient une réponse effective pour la totalité des individus (en l'occurrence, seulement 45 municipes). Dans une enquête par questionnaire, auprès de plusieurs centaines de personnes, il est fréquent de trouver des non-réponses partout ; aucun sujet ni aucune question n'étant indemne. Or il est préférable de ne soumettre à l'analyse multidimensionnelle que des blocs rectangulaires (sous-ensemble de sujets x ensemble des modalités de réponse à un sous-ensemble de questions) où des manques d'information soient rares et non systématiquement distribués (par exemple, tels que pourraient les produire l'effacement du contenu de quelques cases touchées au hasard). Il se peut que la majorité des manques soit le fait de quelques sujets ; ou n'affecte que quelques questions ; ou encore que l'essentiel de l'information effectivement

(*) Docteur de l'U. P. et M. Curie. L'auteur remercie le Centre d'Etude de l'Emploi qui lui a confié l'analyse de l'enquête considérée ici.

recueillie soit recouverte par deux ou trois blocs rectangulaires, dont chacun est presque parfaitement rempli.



exemple : l'essentiel des manques est renfermé dans les zones hachurées ; le reste est recouvert par 3 rectangles.

Le taux des non-réponses est ordinairement plus élevé quand le sujet est laissé à lui-même que quand il est aidé ou aiguillonné par un enquêteur. Il n'en faut pas conclure que l'enquêteur soit indispensable ou même en tout cas utile : car si sa présence rend plus régulier le format des informations recueillies, elle peut en altérer la qualité en mêlant aux opinions propres du sujet des suggestions, trop facilement acceptées par celui-ci.

L'enquête sur les créations d'entreprise s'est déroulée par correspondance. Beaucoup des enquêtés, travailleurs manuels, n'avaient pas la pratique de la rédaction écrite que donne la préparation des examens! Adressé à une population inconnue, donc impossible à segmenter *a priori* et pour cette raison hétérogène, le questionnaire avait été conçu assez ample pour recueillir toutes les informations intéressantes que les sujets pourraient désirer apporter. En contre-partie, il fallait que chaque sujet fût autorisé à ne répondre qu'aux questions qui le concernaient ; ce qu'on faisait par cette mention :

".... certaines questions peuvent ne pas être appropriées à votre cas : omettez-les".

La présentation ordonnée des questions réparties en 11 thèmes ou chapitres précédés chacun d'un titre, pouvait même encourager l'enquêté à sauter des blocs entiers du questionnaire. On ne s'étonnera pas, dans ces conditions que les omissions aient été très nombreuses.

[NON RÉPONSE] § 1	N de quest. C9	N de mod. C39	Moyen: C78/ C9	NR9 (129I)	NR9 (744H)	...	NR9 (12I)
: N° et contenu du thème 9							
1. Identification	25	195	7,8	20	13	...	2
2. Expérience du travail	19	123	6,3	19	19	...	3
3. Formation à la création	18	150	8,0	18	18	...	0
4. Projets abandonnés	2	18	9,0	2	2	...	2
5. Création d' entreprise	6	40	6,6	6	3	...	0
6. Entreprise n' existe plus	4	20	5,0	4	4	...	4
7. Entreprise existe encore	8	52	6,5	8	5	...	4
8. L' entreprise demain	21	114	5,3	21	20	...	12
9. Conditions à la création	25	113	4,5	25	25	...	12
10. Nouveaux projets de créat.	12	44	3,6	12	12	...	7
11. Ce qui compte le plus	7	56	8,0	4	7	...	3
TOTAL Q	147	925	6,2	139	138	...	49

Le tableau ci-joint offre une première esquisse de la structure du questionnaire et de la distribution des non-réponses. Il comprend, outre une marge inférieure, 11 lignes ; chacune de celle-ci correspond à un thème, dont la 1-ère colonne donne le numéro et le titre : e.g. 7 = *Entreprise existe encore*. Les trois colonnes suivantes donnent pour chaque thème θ le nombre des questions qu'il comporte C_θ (e.g. 8 pour le thème 7) ; le nombre total CJ_θ des modalités de réponses afférentes à ces questions (52 modalités pour les 8 questions du thème 7) ; le nombre moyen (CJ_θ/C_θ) de modalités par question du thème (52/8 = 6,5 pour le thème 7). Les colonnes suivantes devraient être au nombre de 814, une par individu : on n'en a mis que 3 ; les individus étant rangés de gauche à droite dans l'ordre décroissant du nombre de leurs omissions sur l'ensemble du questionnaire. Le premier sujet, dont l'identification est 129I, a laissé 139 questions sans réponse (le nombre 139 se lit à la marge inférieure) ; plus précisément, il a omis 20 des 25 questions du thème 1, toutes les 19 questions du thème 2... etc. . Le sujet 12I qui est, de tous, celui qui a répondu au plus grand nombre de questions, n'en a pas moins omis 49 sur 147, soit exactement 1/3.

2 Analyse de la distribution des non-réponses des individus par thème

2.0 Le tableau analysé : Il s'agit de recenser pour chaque individu i , dans chaque thème θ , le nombre $k(i, \theta^+)$ des questions auxquelles il a répondu et celui $k(i, \theta^-)$ de celles auxquelles il a omis de répondre. On a donc un tableau à 814 lignes et 22 colonnes :

$I \times U$: 814 individus \times (11 thèmes dédoublés)

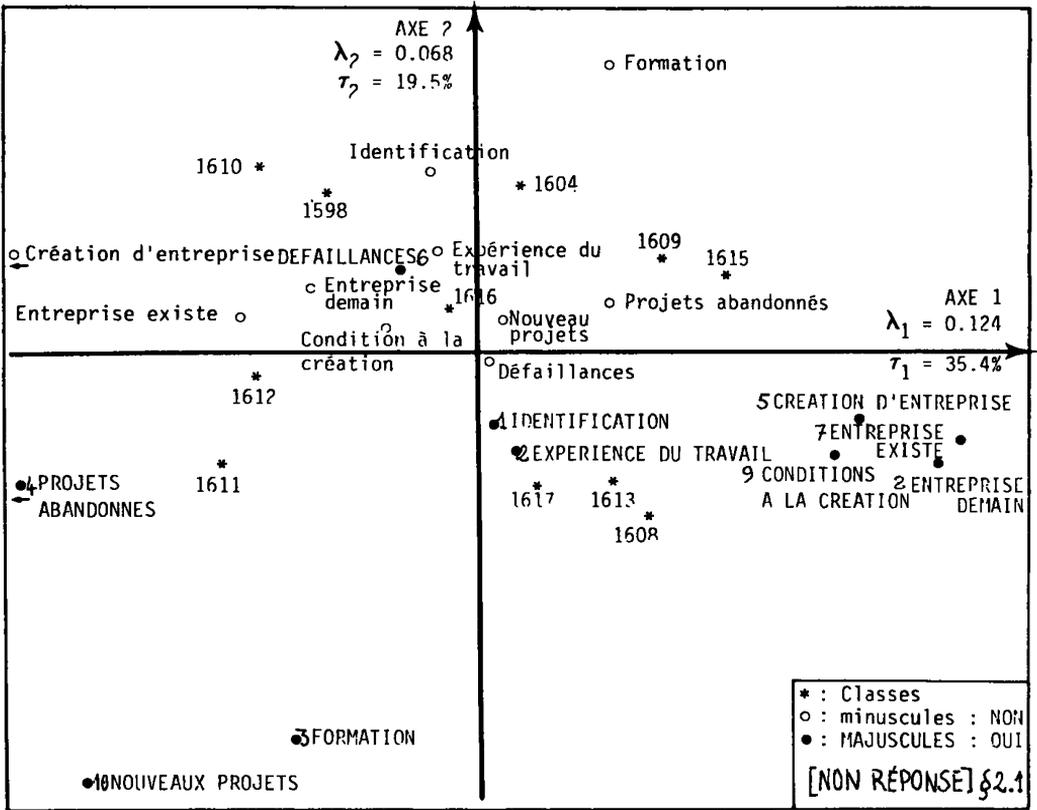
(où on a noté $U = O^+ \cup O^-$ l'ensemble de thèmes dédoublés). Sur les graphiques et tableaux, les sigles utilisés pour les thèmes sont composés d'une lettre T ou θ , initiale du mot thème, suivie d'un nombre (de 1 à 11) et de l'un des deux lettres O = Oui (a répondu) ou N = Non (n'a pas répondu).

A ce tableau, on peut adjoindre en ligne supplémentaire, toute modalité j de réponse ou de signalement concernant l'ensemble i des individus. Cette ligne supplémentaire est tout simplement définie en faisant la somme de celles afférentes aux individus i rentrant dans la modalité j .

Il importe de voir clairement que du fait de la complémentarité des colonnes (de façon précise) : $\forall i \in I : k(i, \theta^+) + k(i, \theta^-) = \text{Card } \theta$, nombre des questions du thème θ , on a dans l'espace des profils sur I , un nuage $N(U)$ constitué par un système de 11 doublets (θ^+, θ^-) ayant tous même centre de gravité, qui n'est autre que celui de $N(U)$. L'analyse du tableau $I \times U$ fournit donc exactement 11 facteurs non triviaux.

L'intérêt du présent tableau relativement à celui recensant les non-réponses par questions, (objet du § 3), est que d'une part, avec 22 colonnes seulement, l'analyse est peu coûteuse ; et d'autre part l'interprétation des résultats est facile, ne requérant pas la connaissance du contenu des questions, mais seulement celle des titres des thèmes. A condition toutefois que les thèmes aient été bien découpés et que les titres en soient évocateurs, ce qui est ici le cas.

2.1 L'analyse factorielle : L'interprétation de l'axe 1 est claire, aussi bien sur le nuage de l'ensemble U des thèmes dédoublés, que sur les extraits publiés du tableau des $COR1$ et $CTR1$. Du côté $F1 > 0$ se trouvent les réponses effectives aux thèmes 5, 7, 8, 9, concernant une entreprise que l'individu a créée, qui existe, et continuera d'exister.

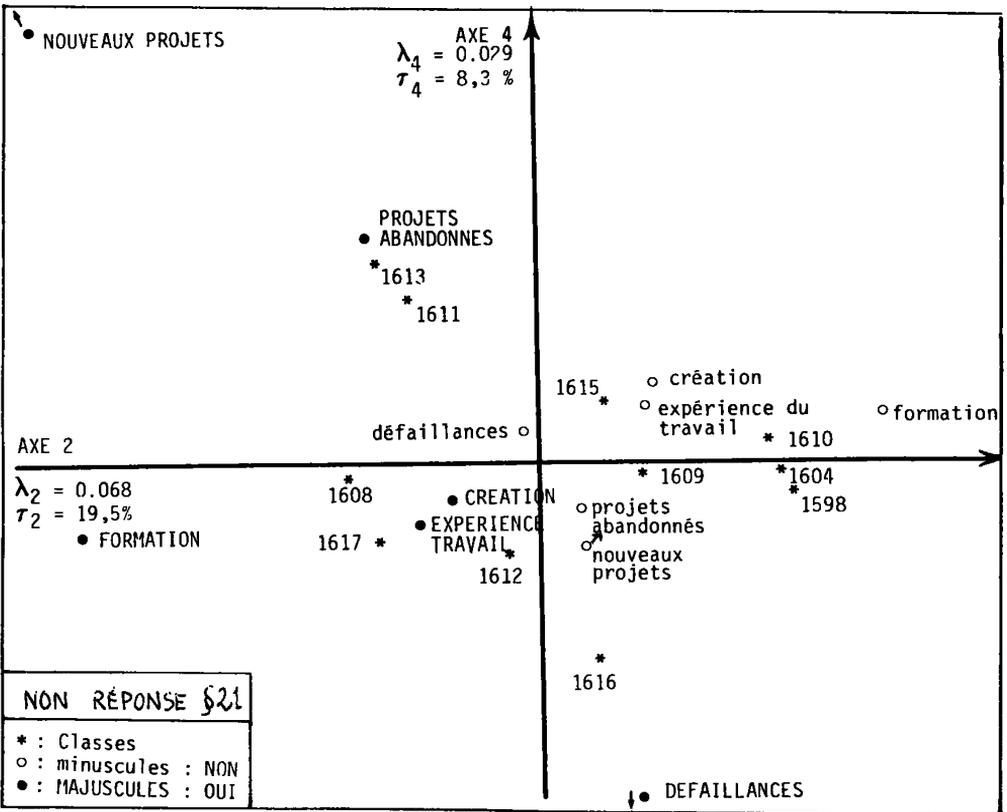


	F1<0			F1>0		
		COR	CTR		COR	CTR
5 Création d'entreprise	T5N	783	138	T5O	783	96
8 Entreprise demain	T8N	694	83	T8O	694	224
7 Entreprise existe	T7N	671	53	T7O	671	120
9 Conditions à la création	T9N	659	27	T9O	659	114
4 Projets abandonnés	T4O	313	26	T4N	313	5
3 Formation	T3O	152	45	T3N	152	34

	F2<0			F2>0		
		COR	CTR		COR	CTR
3 Formation	T3O	731	398	T3N	731	295
1 Identification	T1O	317	24	T1N	317	69
10 Nouveaux projets	T10O	133	61	T10N	133	5
2 Expérience du travail	T2O	106	30	T2N	106	27

Du côté $F_i < 0$ on ne trouve parmi beaucoup de mentions θ^- de non-réponse, que le mélancolique thème T4 Oui : PROJETS ABANDONNES. Quant à la lecture des tableaux de contributions, on remarque que sur tout axe α et pour tout θ $COR\alpha(\theta^+) = COR\alpha(\theta^-)$, car, ainsi qu'on l'a dit, (θ^+, θ^-) constitue un dipôle centré à l'origine et $COR\alpha(\theta^+)$ n'est autre que le carré du cos de l'angle que forme avec l'axe α le segment $\theta^+ \theta^-$ (on vérifie sur les graphiques, à l'imprécision du dessin près, que tout segment $\theta^+ \theta^-$, joignant un point noir à un point blanc, passe par l'origine). D'autre part pour apprécier la contribution CTR du thème θ à l'axe α , il faut additionner $(CR\alpha(\theta^+)$ et $CTR\alpha(\theta^-)$.

Les axes 2, 3, 4 sont dominés respectivement par les thèmes 3 (formation à la création), 2 (expérience du travail), 10 (nouveaux projets). Ce dernier thème se signale par un taux de non-réponses particulièrement élevé le poids de T10 Non est environ 12 fois celui de T10 Oui : ce qui se manifeste sur les graphiques-plans par une proximité toujours grande du point T10 Non (le plus lourd) avec l'origine, l'autre extrémité (T10 Oui) du dipôle étant 12 fois plus éloignée. Corrélativement, la contribution $CTR\alpha(T10 \text{ Oui})$ est 12 fois supérieure à $CTR\alpha(T10 \text{ Non})$.



[NON RÉPONSE] § 2.1

 $\lambda = 0,143$; $\tau = 11,8\%$

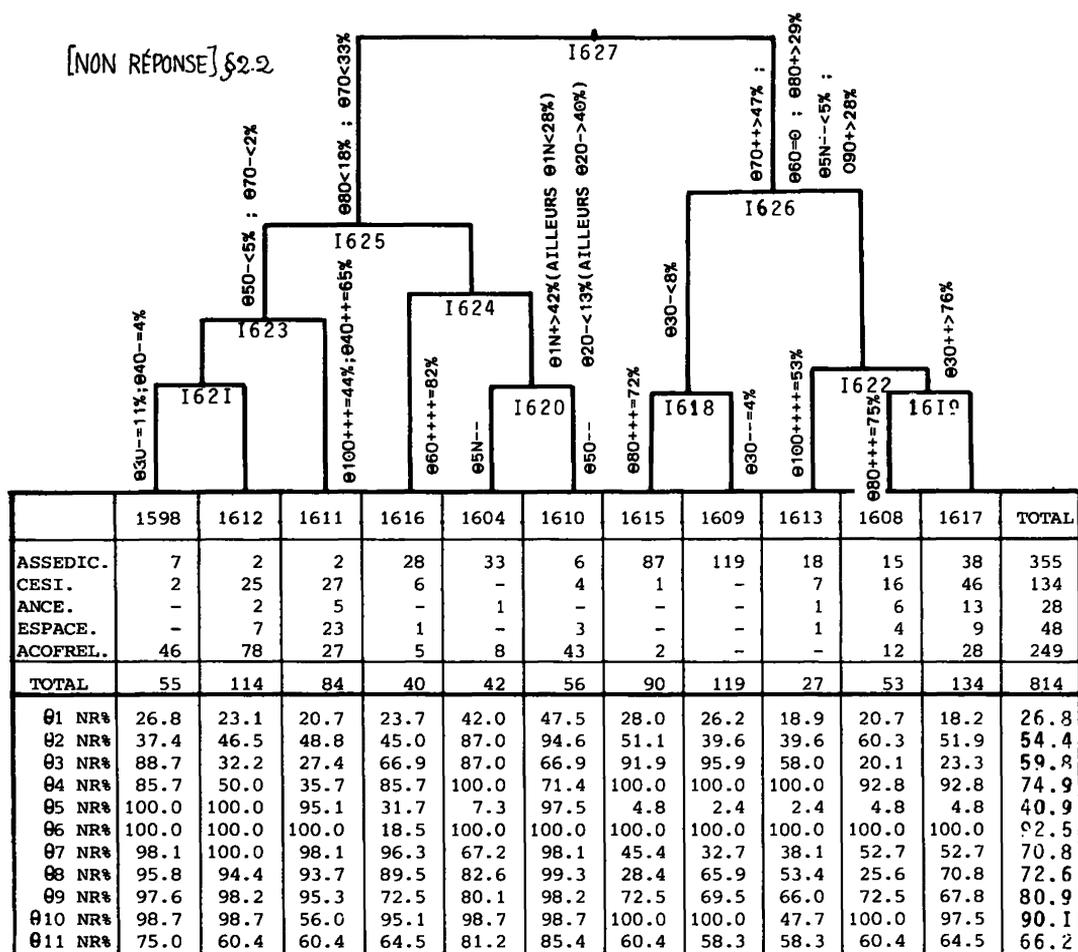
	F3<0			F3>0		
		COR	CTR		COR	CTR
2Expérience du travail	T2N	741	310	T2Q	741	346
1Identification	T1N	175	63	T1Q	175	22
3Formation	T3Q	89	80	T3N	89	59

2.2 La classification ascendante hiérarchique de l'ensemble des indi-

vidus : Une description précise des classes de partition retenue, avec son interprétation verbale proposée se trouve sur deux tableaux se faisant face. Il nous suffira d'expliquer comment lire ces tableaux, en complétant cette lecture par quelques réflexions générales.

Avec l'arbre on a noté sous le numéro de chacune des 11 classes terminales retenues sa composition par organisme d'origine ; et les taux de non-réponses aux 11 thèmes. Par exemple la classe 1609 comprend 119 individus tous issus de l'ASSEDIC ; le taux moyen de non-réponse au thème 1 (identification) est dans cette classe de 26,2% ; d'où, par complémentarité, un taux de réponse de 73,8% (non indiqué).

Il importe de préciser que ce sont ces pourcentages qui sont éventuellement notés sur les branches de l'arbre ; et non les composantes elles-mêmes des profils. Dans l'exposé [INTERPRET CAH] (cf. CAD Vol X n° 3 ; 1985), il s'agit de classer les profils annuels des importations de phosphates sur l'ensemble des principaux exportateurs : une indication telle que MAR⁺⁺⁺ = 60% signifie que la classe ainsi étiquetée reçoit du MARoc une part de ses importations en phosphates très élevée : soit 60%. Mais dans le cas présent, il n'est pas opportun de porter une indication analogue : telle que : dans le profil de la classe 1609 sur U (ensemble des thèmes dédoublés) la composante T1N (non-réponses au thème 1) représente 3,9% du total ; tandis que la composante T1Q (réponse à 01) vaut 12,5% (la somme des deux étant 16,4% ; part que représentent les 25 questions du thème 1, rapportées à l'ensemble des 147 questions des 11 thèmes). L'information directement interprétable est : alors le nombre maximum de non-réponses au thème 1



est pour chaque individu de 25, la moyenne au sein de la classe 1609 représente 26,2% de ce nombre (soit 6,55). En revanche, les signes + et - sont utilisés comme dans [INTERPRET CAH] : par exemple la mention e30-- = 4% sur la branche aboutissant à 1609, signifie que d'une part (cf. supra) les individus de cette classe ont en moyenne répondu seulement à 4% des 18 questions du thème 3 (formation à la création) ; et que d'autre part ce taux anormalement faible (relativement au taux moyen qui est de 40,2% ; complémentaire de 59,8% valeur du taux de n. r. au thème 3 qu'on lit dans la colonne TOTAL) apporte à l'écart de la classe 1609 au centre de gravité du nuage une contribution relative comprise entre 200 et 350 millièmes (exactement 258/1000 ; comme on le lit sur le listage VACOR).

L'ensemble des informations qui accompagnent l'arbre permet d'écrire une interprétation des deux classes à la fois précise et claire. L'opposition majeure est entre les deux classes 1626 et 1625 : d'une part créateurs d'une entreprise qui existe encore ; d'autre part individus qui n'ont pas actuellement d'entreprise. Comme on peut le vérifier sur le listage VACOR (COD1(1627) = 944/1000) cette opposition correspond exactement à l'axe 1 commenté au § 2.1.

- I626:05N⁻<5%;070⁺,080⁺,090⁺:Créateurs d' une entreprise qui existe encore.
 I618:030⁻<8%:N' a pas bénéficié d' une formation à la création.
 I615:080⁺⁺⁺=72%:a des projets de développement.
 I609:030⁻=4%:absence totale de formation.
 I622:mêmes caractères généraux que I626.
 I613:0I00⁺⁺⁺:nouveaux projets de création.
 I619:030⁺⁺>76%: a bénéficié d' une formation à la création.
 I608:080⁺⁺⁺>74%: a des projets d'avenir pour l'entreprise.
 I617: mêmes caractères généraux de I619.
- I625:070<35%;080<18%;n'ont pas actuellement d'entreprise.
 I623:050⁻<5%;070⁻<2%: n'ont jamais créé d'entreprise.
 I611:0I00⁺⁺⁺=44%: ont des projets de création;
 040⁺⁺=65%: ont abandonné des projets.
 I621: mêmes caractères généraux que I623;
 I598:030⁻=11%;040⁻=4%; ni formation ni projets abandonnés
 (répondent un peu à 01 et 02).
 I612:030=68%;040=50%:(formation et projets abandonnés;
 ne répondent qu'à 01,02,03,04,011).
- I624:mêmes caractères généraux que I625.
 I616:060⁺⁺⁺=82%;05N=32%: a créé et a l'expérience d'un échec.
 I620:0IN⁺>42%;020⁻<12%: ne précise ni son identité
 ni son expérience du travail.
 I604:05N⁻: avait créé; ?
 I610:050⁻: n'a jamais créé et ne répond un peu qu' à 01,03,04.

[NON RÉPONSE] §2.2 : Interprétation de la CAH sur les individus

Parmi les créateurs (1626) se signale la classe 1618 de ceux qui n'ont pas bénéficié d'une formation à la création : leur origine est presque exclusivement l'ASSEDIC, et au sein de 1618 on a 1615 : 080oui+++ = 72% : répond au thème "l'entreprise demain".

Remarquons encore les classes 1613 et 1611, qui ont la plus forte valeur de F4 (positif) ; elles s'associent toutes deux, conformément à l'interprétation de l'axe 4, avec un taux élevé au thème 10 "nouveaux projets de création". Cependant les individus de 1611 (inclus dans 1625) semblent ne pas avoir actuellement d'entreprise ; tandis que ceux de 1613 (inclus dans 1626) en ont une.

Nous terminerons ce commentaire en soulignant que la CAH comme l'analyse factorielle démontrent le caractère systématique du phénomène de non-réponse ; et en suggérant d'intéressantes interprétations ; qui sont toutefois fragiles, parce que, d'une part, les sujets peuvent déformer la réalité ou la dissimuler et d'autre part ainsi qu'on le verra, le codage a parfois confondu réponse négative et absence de réponse.

Comme il est naturel, chaque couple q^+, q^- définit un segment passant par l'origine : mais l'une des deux modalités est très lourde, elle se place à proximité de l'origine dans une zone dense que nous avons pu seulement figurer par des hachures. Au contraire, réponses et non-réponses au thème 3 (formation) étant de poids équivalent sont nettement visibles les unes et les autres. Presque sans exception, les modalités de rép. sont dans le demi-plan $F_2 < 0$; et les modalités de n.-rép. dans le demi-plan $F_2 > 0$. Mais la discrimination entre rép. et n.-rép. se voit au mieux dans le plan (2,3).

3.1.3 L'axe 4 : Dans l'analyse par thème du § 2.1, l'axe 4 est dominé par le thème 10 : Nouveaux Projets. Dans la présente analyse par questions, on a sur l'axe 4 un étalement des modalités de réponse effective (q^+) aux questions du thème 8 : "l'Entreprise Demain". Les questions de ce thème sont groupées sous quatre titres : Créations d'emplois ; Moyens (en vue de ces créations) ; Objectifs (énumérés en une seule question : e.g. maintenir son acquis... et Obstacles à surmonter.

Sur l'axe 4 les modalités q^+ afférentes aux créations d'emplois et aux objectifs sont proches de l'origine avec un facteur F_4 généralement négatif. Enfin les modalités q^+ des obstacles sont très écartées du côté $F_4 > 0$. Or, alors que dans les questions relatives aux obstacles on a prévu un item Non (distinct de la "non-réponse", ou du "non concerné") les questions rangées sous le titre "moyens" ont été codées avec une seule modalité ambiguë "non-réponse ou non à la question". Or manifestement ne pas recourir à un contrat d'apprentissage quand on recrute du personnel, n'est pas la même chose que de ne pas envisager de recruter, de quelque manière que ce soit. Cependant, les créateurs d'entreprise qui envisagent l'avenir, tout en ne redoutant pas tous les obstacles énumérés par le questionnaire, sont concernés par ceux-ci, et le montrent en répondant effectivement "Oui" ou "Non" ; laissant la "non-réponse" à ceux qui n'ont pas de projets de développement. Ceci donne à la mention "réponse effective q^+ " à telle ou telle question relative aux obstacles un sens clair, convenant à tous les créateurs face à l'avenir. Alors que la réponse effective (q^+) à une question relative à un moyen n'appartient qu'à la fraction des individus concernés qui recourent à ce moyen.

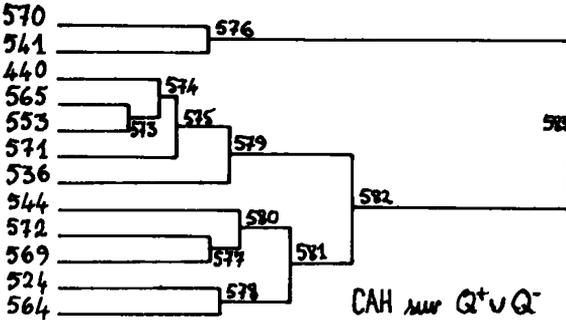
Dans un autre article (DEVELOP. ENTREPRISE] in CAD Vol XI n° 1 nous reprendrons le codage de ces blocs de questions. En effet, bien que le codage initial ait été différent selon le cas, il est toujours possible, en comparant les réponses d'un individu i à toutes les questions d'un bloc de savoir s'il est concerné par ce bloc : car alors il fournit la réponse "Oui" à l'une au moins des questions du bloc.

3.2 Classification sur l'ensemble des questions dédoublées : Sans publier la liste complète des questions, on peut rendre compte de la CAH effectuée sur $D = Q^+ \cup Q^-$, (d'après le tableau $I \times D$), en donnant par thème l'inventaire des classes d'une partition issue de cette CAH. Ici on s'est borné à la partition en 12 classes définie par les 11 noeuds les plus hauts. Dans l'ensemble cette partition s'accorde avec le schéma des 11 thèmes et valide donc la composition du questionnaire. Les exceptions elles-mêmes sont instructives. Prenons quelques exemples.

La classe 564 comprend les q^+ de toutes les questions du thème 3 (formation) ; les accompagnent 3 q^+ du thème 1 (identification) qui, en fait, concernent la formation (cf. *supra* § 3.1.1).

La plupart des classes, (même au niveau d'une partition en 12 grandes classes) sont composées exclusivement soit de modalités q^+ (rép. effective) soit de q^- (non-réponse), font exceptions 572 et 571. Dans 571, les q^+ afférentes aux questions les plus banales du thème 1 (identité : âge sexe...) s'associent aux q^- des questions de faux thèmes auxquels fort peu de sujets ont fourni une réponse effective : cette classe 571 n'est autre que la zone dense hachurée qui sur nos graphiques entoure l'origine.

Dans 572 avec les non-réponses q- aux questions des thèmes 5, 7, 8,9 caractéristiques d'une entreprise qui existe, vont les réponses q+



[NON RÉPONSE] §3.1		N° de quest. C8
: N° et contenu du thème θ		
1. Identification		25
2. Expérience du travail		19
3. Formation à la création		18
4. Projets abandonnés		2
5. Création d'entreprise		6
6. Entreprise n' existe plus		4
7. Entreprise existe encore		8
8. L'entreprise demain		21
9. Conditions à la création		25
10. Nouveaux projets de créat.		12
11. Ce qui compte le plus		7
TOTAL Q		147

Cl.	Thème θ ⁺ (oui)											Thème θ ⁻ (non)										
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10	11
570		2			6		8	11	25		1											
541								10														
440						4																
573	9	17									6											
571	12											3	10	8	2		4		5	19	12	6
536												3		10								
544												10										
572				2											6		8	16	6		1	
569												8	9									
524										12												
564	3		18																			

aux deux questions du thème 4 des projets abandonnés.

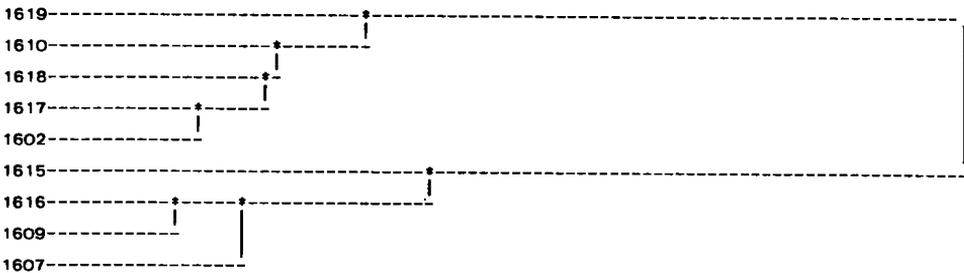
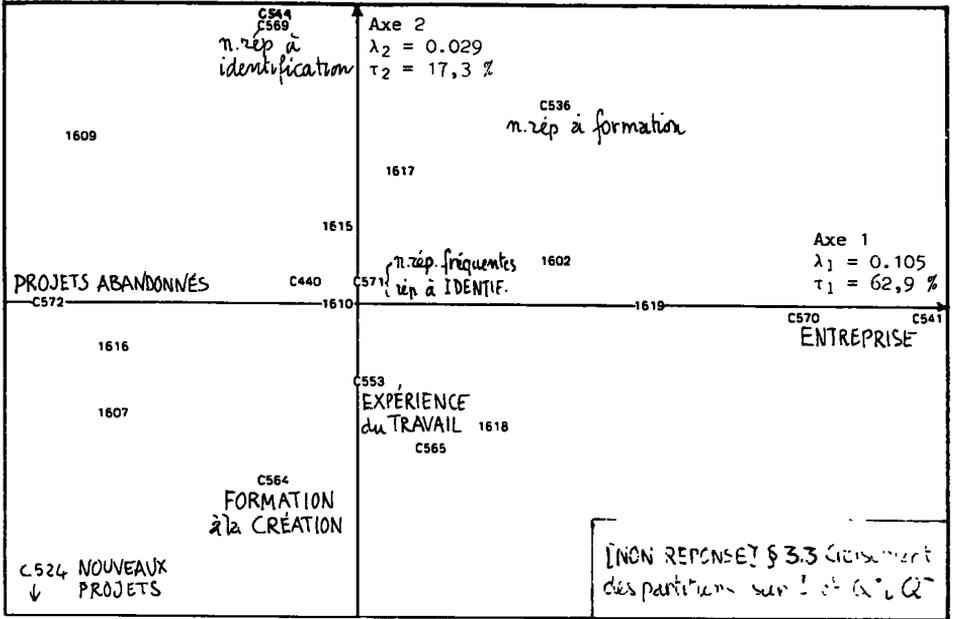
Quant au thème 8, les réponses effectives q+ aux questions relatives aux moyens (cf. supra § 3.1.3) forment une classe à elles seules (cl. 541). Les autres q+ du th. 8 sont dans la cl. 570 avec tout ce qui concerne une entreprise existante ; 541 et 570 s'unissent d'ailleurs pour constituer la cl. 576.

3.3 Classification sur l'ensemble des individus : Nous n'interprétons pas ici en détail cette classification, mais indiquerons seulement les méthodes employées pour cette interprétation.

La classification considérée ici a été effectuée d'après le tableau $I \times (Q^+ \cup Q^-) = (814 \cdot 294)$ (cf. § 3.0). Il est impossible de fonder l'interprétation sur les contributions apportées individuellement par les 294 variables. Une première voie est de recourir aux facteurs

qui ont déjà été interprétés : ce qu'on fera directement, en considérant la place des centres des classes sur les plans de l'analyse factorielle reproduits au § 3.1 : ou, pour plus de précision, sur le listage FACOR (non publié ici).

Une deuxième voie est d'utiliser non les variables, éléments de Q^+ ou Q^- (rép. ou non-rép. à chaque question), mais les variables cumulées par classes, suivant la partition C elle-même construite par la CAH (cf. § 3.2). On peut alors calculer des contributions aux classes et dipôles d'après le tableau cumulé $I \times C$, les c jouant le rôle de variables.



	1619	1618	1617	1616	1615	1610	1609	1607	1602
C572	936	2022	1539	5330	1105	940	2514	2568	1456
C571	9350	8075	5539	10192	2789	2265	5027	4355	7158
C570	3166	3167	1560	176	519	488	52	173	2640
C569	777	400	977	680	385	185	990	272	327
C565	1893	2106	509	2251	450	521	283	1095	1697
C564	1145	1919	619	1685	300	270	546	1071	148
C553	347	406	138	532	113	132	131	224	385
C541	1346	234	91	138	99	55	0	7	183
C535	853	264	505	553	308	242	362	77	1168
C524	6	204	14	36	30	28	5	355	5
C440	3	7	8	4	3	120	0	0	0
C544	34	30	35	31	323	10	18	23	17

Ici nous considérons seulement un petit tableau 9×12 croisant la partition C de $(Q^+ \cup Q^-)$ en 12 classes (arrêtée au § 3.2) et une partition CL de I en 9 classes (issue de la CAH sur I faite d'après le tableau $I \times (Q^+ \cup Q^-)$) ; tableau défini par la formule de cumul usuelle :

$$\forall c \subset Q^+ \cup Q^- = D ; \forall c\ell \subset I :$$

$$k(c, c\ell) = \sum \{k(i, j) \mid i \in c\ell ; j \in c\}.$$

L'analyse factorielle de ce petit tableau fournit un premier facteur très semblable à celui obtenu au § 3.1 ; ensuite on a un facteur 2 qui est une combinaison des facteurs 2 et 3 obtenus précédemment ; etc. . La CAH sur CL (ensemble de classes de I) effectuée d'après ce même petit tableau donne une organisation hiérarchique des 9 classes qui coïncide avec celle obtenue dans la CAH sur I d'après le tableau $I \times (Q^+ \cup Q^-)$. Ceci montre l'intérêt du tableau cumulé qui (en utilisant éventuellement des partitions plus fines que celles en 9 et 12 classes adoptées ici) fournit sous une forme facile à lire et à analyser l'essentiel de l'information contenue dans le tableau initial $I \times (Q^+ \cup Q^-)$: 814×294).

Enfin pour comparer la présente CAH à celle interprétée plus en détail au § 2.2, on a croisé les deux partitions correspondantes sur I. Dans le tableau ci-joint ces classes sont désignées par un nombre de 3 chiffres précédé d'une lettre : ainsi la classe 1609 du § 2.2 est notée T609 (T \approx Thèmes) ; et la classe 1609 du présent § est notée Q609 (Q \approx Questions). Sans coïncider exactement les classes vont souvent par paires : Q602 avec T609 ; Q610 avec T616 ; Q609 avec T610. Ou encore Q618 est à peu près la réunion de T617 et T 613.

[NON REPONSE]		CAH sur I d'après $I \times (Q^+ \cup Q^-)$								
§ 3.3		Q619	Q618	Q617	Q616	Q615	Q610	Q609	Q607	Q602
CAH sur I d'après $I \times (Q^+ \cup Q^-)$	T617	15.	57.	19.	13.	3.	3.	8.	11.	5.
	T616	1.	6.	1.	4.	1.	19.	3.	4.	1.
	T615	56.	5.	4.	11.	10.	0.	2.	1.	1.
	T613	1.	18.	1.	3.	0.	1.	0.	0.	3.
	T612	8.	11.	8.	60.	8.	1.	7.	7.	4.
	T611	6.	11.	5.	14.	3.	7.	6.	28.	4.
	T610	3.	7.	5.	1.	4.	0.	29.	3.	4.
	T609	9.	4.	9.	12.	3.	0.	7.	4.	71.
	T608	30.	3.	5.	2.	2.	1.	1.	4.	5.
	T604	2.	4.	18.	4.	4.	2.	3.	4.	1.
	T598	5.	3.	4.	24.	6.	2.	2.	4.	5.

4 Conclusion

L'analyse des non-réponses soit par thème (§2) soit par questions (§ 3) a reconnu des classes d'individus se distinguant d'abord par le fait de posséder ou non une entreprise en activité ; ensuite par la formation, l'expérience du travail, les projets d'avenir. Mais l'analyse a également appelé notre attention sur les imperfections du codage, particulièrement pour les blocs de questions qui offrent une série d'éventualités non exclusives entre elles mais au contraire susceptibles de servir ensemble à une même fin. Dans l'étude ultérieure de la sous-population des créateurs en développement, nous refonderons donc le codage de ces blocs de questions.