

C. ROCHE

**Exemple de classification hiérarchique avec
contraintes de contiguïté : le partage d'Aix-en-
Provence en quartiers homogènes**

Les cahiers de l'analyse des données, tome 3, n° 3 (1978),
p. 289-305

http://www.numdam.org/item?id=CAD_1978__3_3_289_0

© Les cahiers de l'analyse des données, Dunod, 1978, 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/>

EXEMPLE DE CLASSIFICATION HIÉRARCHIQUE
AVEC CONTRAINTES DE CONTIGUITÉ :
LE PARTAGE D'AIX-EN-PROVENCE
EN QUARTIERS HOMOGENES
[ZONES AIX]

par C. Roche ⁽¹⁾

1 Origine du problème et plan de l'étude

La méthode de prévision de demande téléphonique adoptée par la DGT présuppose un découpage en quartiers homogènes des principales agglomérations. Le problème n'est pas nouveau. Plusieurs découpages ont été proposés par divers organismes ; mais on verra sur les exemples ci-dessous qu'aucun de ces découpages ne nous satisfait.

A) *Le Découpage de l'INSEE*. Réalisés en vue du recensement de 1968, les quartiers ont l'avantage de représenter une surface acceptable (de l'ordre de 4.000 habitants). Cependant, ne tenant pas suffisamment compte des profondes transformations qu'ont subies les agglomérations françaises, ces découpages ne présentent pas une homogénéité suffisante.

B) *Les Découpages des S.D.A.U.* (Schéma directeur d'Aménagement et d'Urbanisme). Ces découpages ont pour nous deux inconvénients :
- ils sont trop larges (couvrant environ 10.000 habitants chacun)
- réalisés en fonction de l'aménagement futur, ils ne tiennent pas compte de l'homogénéité des quartiers existants.

C) *Les Découpages des P.O.S.* (Plan d'occupation des sols). S'ils sont utiles pour la réalisation d'un zonage, ces découpages sont généralement trop fins pour constituer des zones utiles pour notre propre étude.

On a donc entrepris sur l'exemple de la ville d'Aix, de mettre au point un découpage semi-automatique des agglomérations en quartiers. La méthode statistique qui joue ici un rôle central est la classification ascendante hiérarchique avec contrainte de contiguïté, méthode dont le programme est présenté dans ce cahier (cf [C.A.H. CONTIGUITE] pp 275-287.

Le présent article vise à illustrer ce programme, tout en donnant des aperçus sur la sociologie urbaine. Au § 2 on décrit les données ; dont une analyse de correspondance objet du § 3 permet d'embrasser la structure. Le § 4 est consacré aux classifications automatiques proprement dites ; on voit sur quelques cas comment d'après leur population et leur parc de logements les flots s'agrègent en quartiers. Le partage définitif de la ville est proposé en corrigeant et élaborant d'après de telles réflexions les résultats des calculs (§ 5).

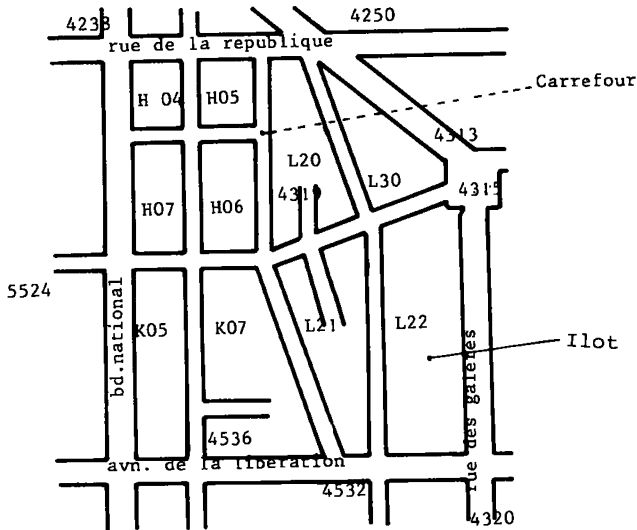
2 Les données analysées

Pour partager la ville d'Aix en quartiers, on est parti de son découpage naturel en 504 ilots ; et l'on a retenu de ceux-ci d'une part (§ 2.1) leurs relations de contiguïté, d'autre part (§ 2.2) leur description par quelques variables issues du recensement.

(1) Direction Générale des Télécommunications (S.P.A.F., Département C1)

Cette étude a été réalisée en commun avec Monsieur Sogno et Mademoiselle Coromines (D R T - Marseille)

2.1 La matrice de contiguïté : La notion de contiguïté que nous avons utilisée est assez large, ceci pour limiter au minimum l'hétérogénéité induite par la prise en compte de la contiguïté pour l'analyse. Un îlot est limité par 3, 4 ou 5 carrefours. Nous considérons comme îlot contigu à cet îlot, tout îlot limité par un au moins de ces carrefours. Sur l'exemple ci-dessous l'îlot HO6 est contigu aux îlots HO4, HO5, L20, K07, K05, HO7, L21.



La matrice donnant pour chaque îlot les numéros des îlots qui lui sont contigus, a été réalisée suivant une numérotation décimale des îlots, différente de la numérotation INSEE. La numérotation nouvelle est continue de l'îlot 001 à l'îlot de numéro maximum ; une table de passage relie les deux numérotations qu'on utilisera simultanément par la suite.

2.2 Les données sociologiques : Nos données proviennent du fichier îlot lourd, sondage au 1/5, du recensement 1975 : ce sont :

- La catégorie socioprofessionnelle C.S. (sur 8 modalités)
- Le tableau de l'âge et du sexe de la population (sur 6 modalités)
- Le statut des occupants des logements (sur 3 modalités)
- La taille des immeubles (sur 3 modalités)
- La date de construction des logements (sur 4 modalités)

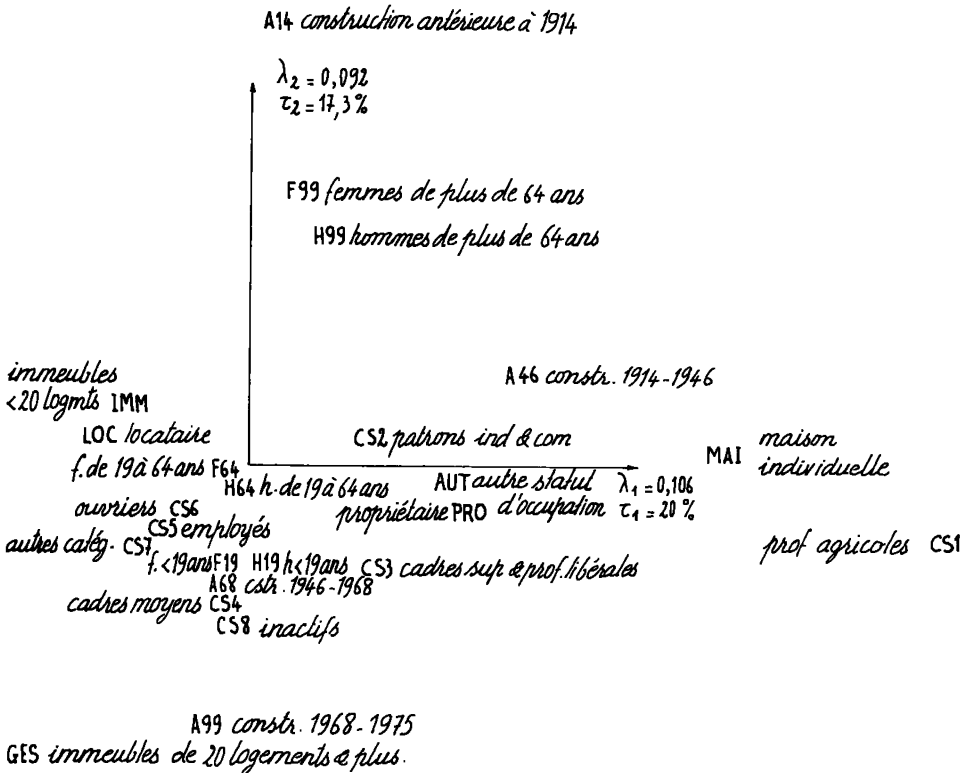
Le détail de 24 modalités de ces 5 variables est donné au § 3 avec le tableau des résultats de l'analyse des correspondances. En vue de l'analyse on a cherché à multiplier chaque groupe de modalités par un coefficient différent, afin que dans l'inertie du nuage chaque variable apporte une part égale, excepté la C.S. qui devait avoir part double (soit 1/3 pour la C.S. ; et 1/6 pour chacune des quatre autres variables). Par application d'un programme de pondération (pour un exemple de tel programme, cf Cahiers Vol II, n° 3, pp 353-359, A. W. Hamrouni) on a abouti à un système de coefficients correcteurs, comme l'explique le tableau ci-joint.

Variable	Catégories Socio-Prof.	Age du log.t	Taille des immeubles	Statut d'occupation	Age et sexe population
Inertie sur données brutes	0.140	0.350	0.0788	0.228	0.0413
Coef. Correcteurs	1.27	0.449	1.13	0.378	2.16

Tableau de recherche des coefficients correcteurs

N.B. : On remarquera que sur les données brutes, les variables âge et statut d'occupation des logements sont très prépondérantes ; au contraire le profil des CS et celui des âges varient avec une amplitude bien moindre ; on en a toutefois amplifié l'importance parce que du point de vue du comportement des abonnés au téléphone (objet de l'étude) leur rôle est essentiel !

Nous n'avons pas introduit de variables décrivant les activités professionnelles des non-résidents car cela nécessite un travail important de redressement de fichiers administratifs. Mais il est dans notre projet de le faire ultérieurement.



Graphique réduit des résultats de l'analyse des correspondances (cf § 3), les demi-axes positifs ont longueur 1.

3 Résultats de l'analyse de correspondance

On trouve sur une même page les modalités des variables avec leurs sigles, et les résultats (coordonnées factorielles et contributions) relatives aux facteurs 1 à 6 ; on a encadré les principales contributions aux facteurs 1 à 5. Ceux-ci comme l'atteste le tableau ci-dessous totalisent 54,3% de l'inertie totale du nuage ; les axes 1 et 2 à eux seuls en expliquent plus du tiers. Ces données numériques suffisent à l'interprétation de l'analyse.

	Valeur-propre	Pourcentage	Poucentage cumulé
1	.10576511	19.94	19.94
2	.09171506	17.29	37.22
3	.04970972	9.37	46.59
4	.04048148	7.63	54.22
5	.03293591	6.21	60.43
6	.02953983	5.57	66.00
7	.02440748	4.60	70.60
8	.02257857	4.26	74.85
9	.01976828	3.73	78.58
10	.01750764	3.30	81.88

Sur l'axe 1, les propriétaires PRO associés aux maisons individuelles (ou abritant un petit nombre de foyers) MAI, s'opposent aux locataires LOC ; on notera que les agriculteurs CS1, sont avec PRO et MAI.

Sur l'axe 2, se détachent nettement les points F99 (ainsi que H99), femmes et hommes âgés, associés à CS8 (inactifs) et A14 (logements anciens : construits avant 1914). A l'opposé, mais moins écartés, les cadres moyens CS4, les logements récents A99, etc.

Sur l'axe 3, les ouvriers, CS6 s'opposent aux cadres supérieurs CS3, ceux-ci associés aux immeubles de 20 logements et plus GES. Qui connaît les communes d'ortoirs de la région parisienne s'étonnera de voir les cadres supérieurs associés aux grands ensembles! Mais à Aix, les H. L. M. sont des immeubles de taille modérée ; au contraire on a construit récemment de luxueuses tours, notamment la tour d'Aygozi où logent des cadres supérieurs.

Sur l'axe 4, deux catégories d'immeubles IMM et GES s'opposent entre eux, ainsi que deux statuts d'occupation du logement : PRO et AUT.

L'axe 5 est dominé par l'opposition entre les zones bâties de 1946 à 1968 (A68) et celles comportant les logements les plus récents (A99).

En projetant les points figurant les flots sur les axes factoriels, avec les points figurant les modalités des variables, on a une première idée du zonage : on peut corriger la proximité purement topographique des flots par une proximité fondée sur les problèmes statistiques. Après avoir appliqué le programme de classification, on suivra sur les axes factoriels l'arborescence des classes projetées en éléments supplémentaires (cf 5.1).

4 Utilisation de la classification ascendante hiérarchique

Vu le but poursuivi, il s'impose comme on l'a dit dès l'introduction, d'effectuer l'agrégation des flots en imposant une contrainte de contiguïté (i.e. n'agréger entre eux que des noeuds auxquels correspondent sur le terrain des zones limitrophes). Mais de plus il faut que les classes définitives retenues définissent des quartiers à peu près tous égaux en population : cette contrainte aussi peut être introduite dans le programme. Toutefois comme l'utilisateur doit de toute façon retoucher les résultats produits par l'ordinateur, on a intérêt à utiliser simultanément les deux versions du programme avec ou sans contrainte de taille, que nous traitons ici successivement.

Professions agricoles		Logements constr. avant 1914		Maisons individ. (sens INSEE)	
Patrons de l'ind. et du comm.		Logements constr. de 1914 à 1946		Immeubles de moins de 20 logements	
Cadres sup. et prof. lib.		Logements constr. de 1946 à 1968		Hommes de moins de 19 ans	
Cadres moyens		Logements constr. de 1968 à 1975		Hommes de 19 à 64 ans	
Employés		Propriétaires		Hommes de plus de 64 ans	
Ouvriers		Locataires		Femmes de moins de 19 ans	
Autres catégories		Autre statut d'occupation		Femmes de 19 à 64 ans	
Inactifs				Femmes de plus de 64 ans	

Ci-dessus, détail des sigles ; ci-dessous, résultats de l'analyse factorielle ; le plan des axes 1x2 est au \$5.1

NOMS	MASSES	DISTO	COORDONNEES						CONTR.RELAT. (COR,COS2)						
			CTR des PTS aux FACT en 1/100						des FACT aux PTS						
			F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	
CS1	8.82	.003	1.82	-.18	.98	-.96	.36	.17	9.9	.1	6.1	7.2	1.2	1.2	.3
CS2	1.29	.014	.33	.08	.26	.06	-.00	.14	1.4	.1	1.9	.1	.0	.9	.02
CS3	.89	.032	.33	-.27	-.43	.28	.18	-.41	3.4	2.6	12.0	6.2	3.0	17.9	.02
CS4	.54	.034	-.05	-.36	-.13	.06	-.13	-.05	.1	4.9	1.2	.3	1.7	.3	.19
CS5	.22	.022	-.22	-.17	.14	.04	-.01	.06	1.0	.7	.9	.1	.0	.3	.01
CS6	.49	.044	-.18	-.12	.46	-.08	-.05	.26	1.3	.7	19.2	.7	.3	9.9	.01
CS7	1.57	.015	-.24	-.18	.31	-.04	-.01	-.51	.8	.5	2.9	.0	.0	13.2	.13
CS8	.29	.076	-.04	.41	-.22	.06	-.01	.10	.1	14.1	7.5	.7	.0	2.5	.16
A14	2.15	.015	.03	1.15	.39	-.16	.46	-.28	.0	21.6	4.5	1.0	9.8	4.1	.03
A46	3.98	.004	.70	.22	-.34	-.34	-.27	.28	1.7	.2	.9	1.1	.8	1.0	.04
A68	.92	.021	-.07	-.34	.17	.02	-.79	.09	.1	2.6	1.2	.0	39.7	.5	.02
A99	1.58	.014	-.09	-.69	-.07	.16	.80	.06	.1	7.2	.1	.9	27.7	.2	.01
AUT	1.40	.021	.50	-.04	.08	-.57	-.31	-.67	5.0	.0	.3	16.9	6.3	31.7	.00
LOC	.49	.071	.55	-.12	-.12	.28	.04	.16	19.9	1.1	2.1	13.9	.4	5.7	.00
LOC	.20	.124	-.40	.07	.06	-.06	.02	.03	18.6	.6	.8	1.1	.2	.3	.05
MAI	1.92	.019	1.20	.02	.27	-.31	.04	.17	26.6	.0	2.8	4.6	.1	2.0	.00
IMM	.28	.053	-.33	.15	.09	.28	-.04	-.14	5.4	1.3	.9	10.3	3.3	3.3	.02
GES	4.13	.009	-.61	-.75	1.14	1.18	.29	.35	3.1	5.4	23.0	30.3	2.3	3.6	.07
H19	.23	.055	.01	-.27	.19	.07	-.06	.01	.0	4.3	4.0	.7	.5	.0	.00
F19	.25	.054	-.01	-.27	.15	.06	-.05	.01	.0	4.4	2.5	.5	.3	.0	.03
H64	.07	.118	-.02	-.04	-.00	-.09	.08	-.01	.1	.3	.0	2.5	2.1	.0	.00
F64	.129	.05	-.04	-.02	-.06	.01	.02	-.03	.2	.1	.8	.0	.1	.3	.00
H99	.95	.020	.20	.60	-.20	.09	-.20	.15	.7	7.8	1.7	.4	2.3	1.6	.02
F99	.83	.034	.12	.72	-.21	.07	-.09	.03	.5	19.2	2.9	.4	.8	.1	.00

4.1 Agrégation avec contrainte de contiguïté, sans contrainte de taille :

Comme avec tout programme de C.A.H., on a pour chaque noeud fourni une indication de niveau (ou mesure de l'hétérogénéité), et les numéros des noeuds (ou individus : ici des flots) par agrégation desquels il a été formé. De plus les centres de gravité des noeuds peuvent être placés en éléments supplémentaires sur les cartes (non géographiques!) issues de l'analyse factorielle. Toutes ces informations méritent d'être confrontées, non sans recourir aussi aux données primaires et à toute information disponible. On ne pourra donner ici que quelques exemples.

Ici, le programme ne tenant pas compte de la contrainte de taille, il appartient à l'utilisateur de remonter l'arbre hiérarchique, jusqu'à parvenir à des noeuds dont la taille le satisfasse. En fait le programme peut sur simple indication de niveau réaliser une troncature donnant les branches inférieures de l'arbre. Toutefois comme toutes les branches sont coupées au même niveau, pour mieux que ce niveau soit choisi, il ne convient pas à toutes : telle branche définit un quartier acceptable tel quel ; telle autre est trop vaste ; telle autre encore n'agrège qu'un petit nombre d'îlots typiques ou isolés en banlieue. C'est pourquoi le niveau de troncature de l'arbre est choisi par tâtonnements en se guidant à la fois sur le nombre des classes à définir (légèrement supérieur au nombre des quartiers) et l'hétérogénéité jugée tolérable (l'histogramme des niveaux des noeuds en fonction de leurs rangs aidant à ce choix).

Dans l'exemple traité, nous avons choisi de procéder à la coupure au niveau 970 : comme il y a 504 flots à classer, le n° du sommet de l'arbre est $1007 = (2 \times 504) - 1$; et cette coupure fournit une partition en $38 = (1008 - 970)$ classes. Sur la carte * n° 1, présentant le découpage qui en résulte on peut mettre en évidence les faits prévus *a priori*.

Certaines zones sont de taille comparable à celle du quartier que l'on désire (1500 ménages en moyenne). Il s'agit des zones 963, 959, 945, 903. Plus ces zones ont été constituées tôt dans l'analyse, plus leur homogénéité est forte. Ainsi la zone 903 a été regroupée le plus tôt. Elle forme donc un quartier très homogène. Généralement ce type de zone forme la base du découpage de la ville. Elles dessinent des "taches" très homogènes. Quant aux zones moins homogènes, leurs frontières peuvent être éventuellement retouchées en fonction d'appréciations qualitatives.

D'autres zones : 851, 946 sont de tailles largement inférieures à celles d'un quartier ; certaines enfin sont trop étendues : telles les zones 956, 969, 960. Pour disséquer les plus grandes zones et éventuellement agréger aux fragments retenus les zones trop petites, il faut revenir au détail de la classification arborescente.

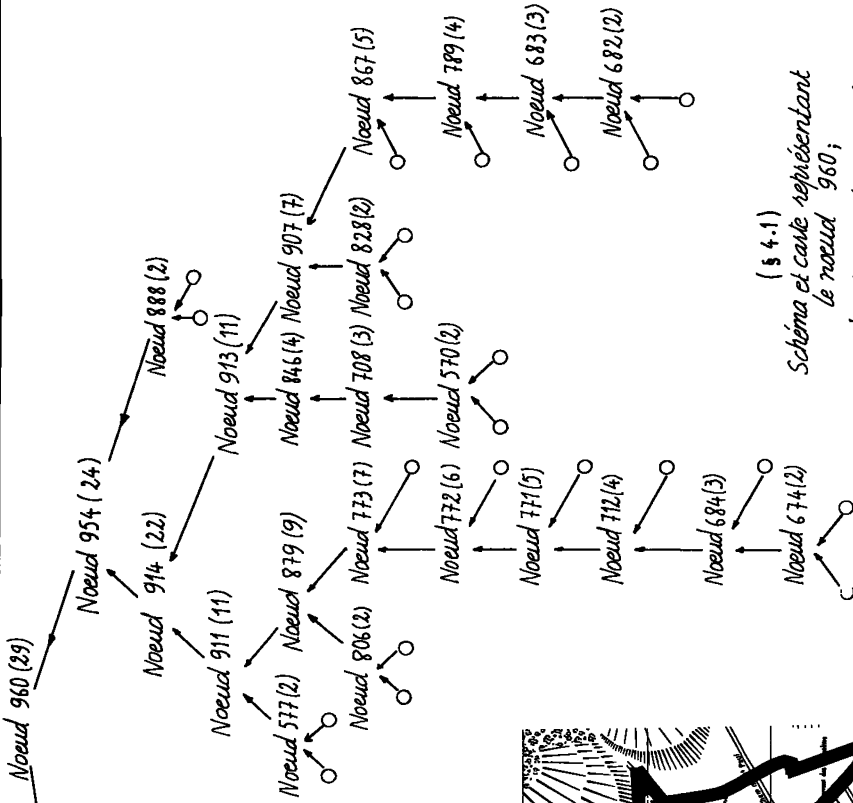
Considérons par exemple le noeud 960 en nous aidant de la carte partielle et du schéma figurés sur une même page. On lit sur le schéma que 960 est formé par agrégation de 853 et 954, noeuds dont les tailles (nombre d'îlots) respectives sont 5 et 24 ; le 954 était formé de 914 (taille 22) et 838 (taille 2) : etc.... La carte ne montre que les noeuds les plus importants :

- Le noeud 853 (traits obliques vers la droite)
- Le noeud 911 (traits obliques vers la gauche)
- Le noeud 846 (traits horizontaux)
- Le noeud 907 (traits verticaux)

Ce travail d'analyse consiste à regrouper les petits noeuds pour construire des quartiers de taille appréciable. La connaissance de la ville est certes indispensable pour opérer de tels choix. Cependant plusieurs principes simples peuvent nous guider :

Deux (ou plusieurs) grands noeuds forment des grandes zones qui sont hétérogènes entre elles. Généralement ces noeuds divisent la ville en grands secteurs socio-économiques.

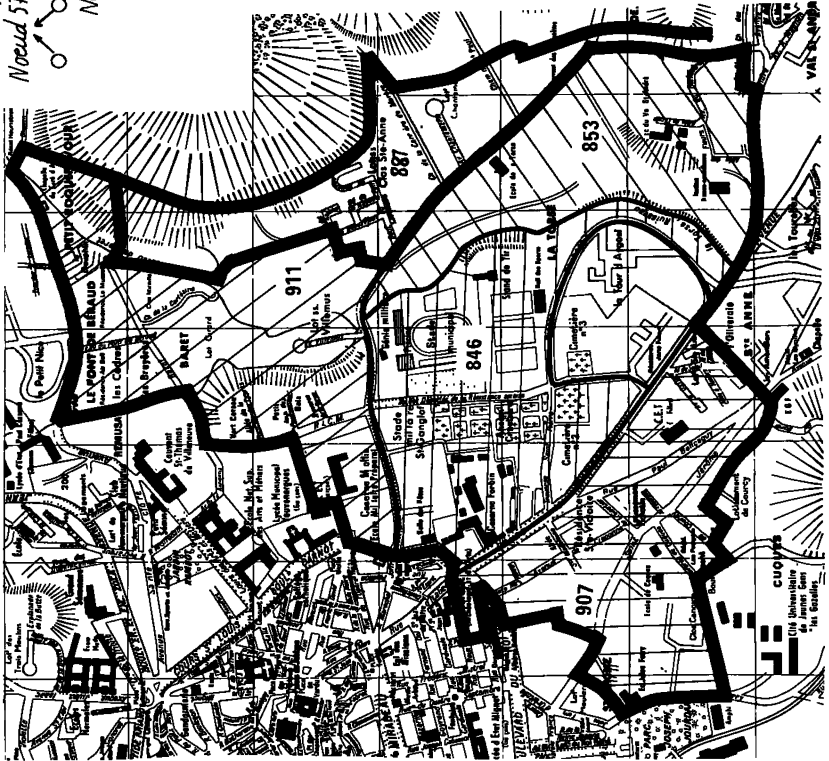
* Cette carte, comme les cartes 2 et 4, n'est qu'un schéma montrant des contours de zone : il nous a paru suffisant de donner une seule carte détaillée, la carte 3 qui d'ailleurs porte les limites des zones définitivement adoptées.



(5 4.1)

Schema et carte representant le noeud 960;

Dans le schema ci-dessus figurent enchevêtrés le nombre d'lots après le numéro du noeud ; les lots eux-mêmes sont figurés par un \circ . La carte ne compare que quelques noeuds.



Deux petites classes contiguës et formées tôt dans l'analyse (plus exactement avec un faible indice d'hétérogénéité) qui ne sont regroupées entre elles que tard (avec un indice d'hétérogénéité fort) ont toutes chances d'être fortement dissemblables puisque le programme a "attendu" longtemps avant de les regrouper.

Plus un noeud est important plus il a de chances d'absorber les îlots isolés.

Finalement on a abouti à des décisions telles que les suivantes :

Le noeud 846 est apparu suffisamment homogène et important pour former un quartier. On lui a adjoint un îlot au Nord, îlot comportant principalement un CET.

Le noeud 853 a par contre été annexé au noeud 887, noeud qui figure sur la carte n° 1 essentiellement en fonction de critères dus à l'aménagement de la cité.

4.2 Agrégation avec contrainte de taille

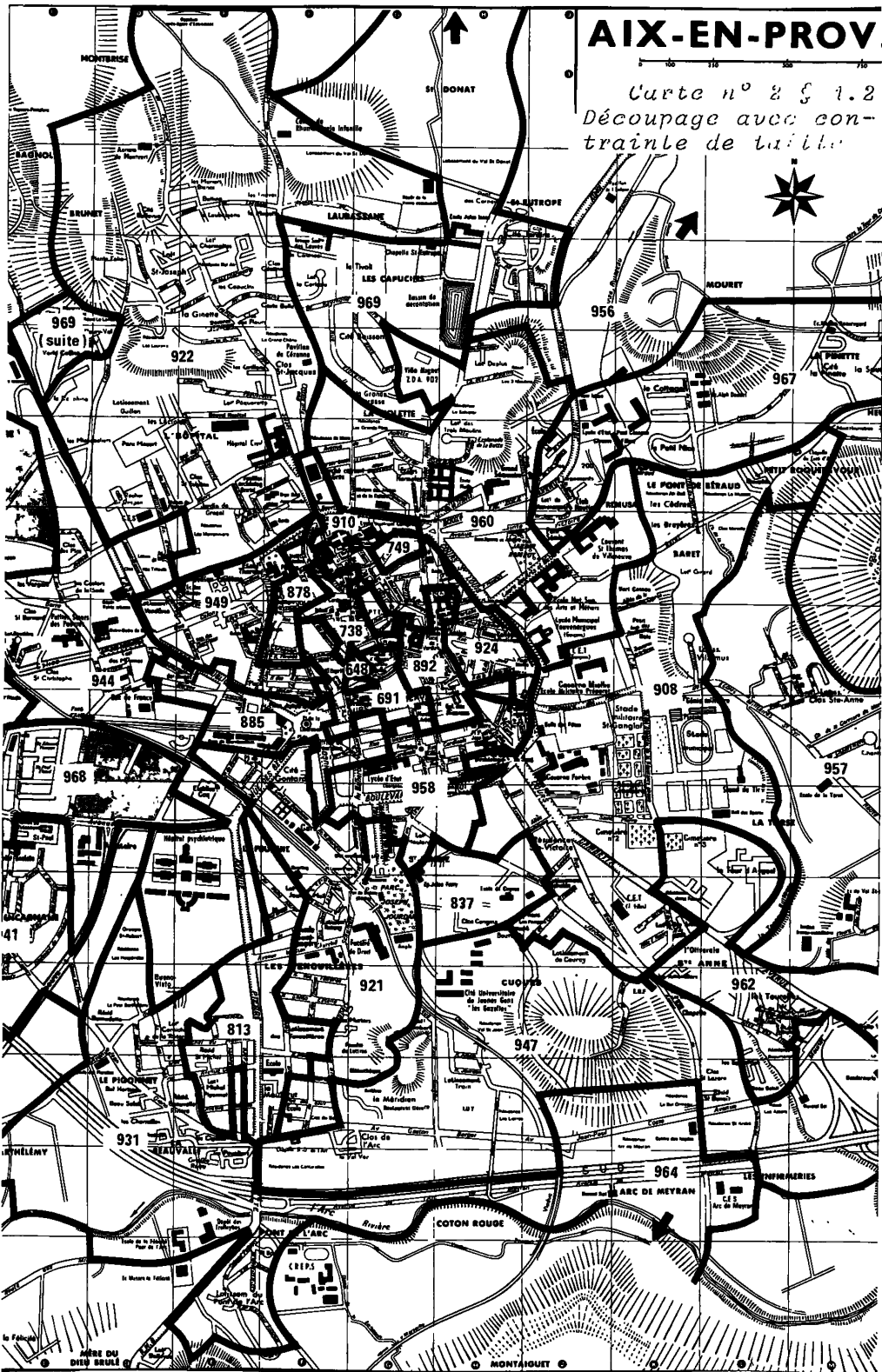
Les remarques précédentes sur la formation de "grands" noeuds justifient l'analyse avec contrainte de taille : le programme peut éliminer de l'étude tout noeud dépassant une certaine taille. Le choix du seuil de contrainte est approximatif. Si l'on désire des quartiers de n îlots on prendra comme contrainte un nombre entier d'îlots égal environ à $2/3n$ l'algorithme s'arrête au moment où aucune agrégation ne peut plus être réalisée sans créer de noeuds dépassant la limite de taille, ni unir des noeuds non contigus.

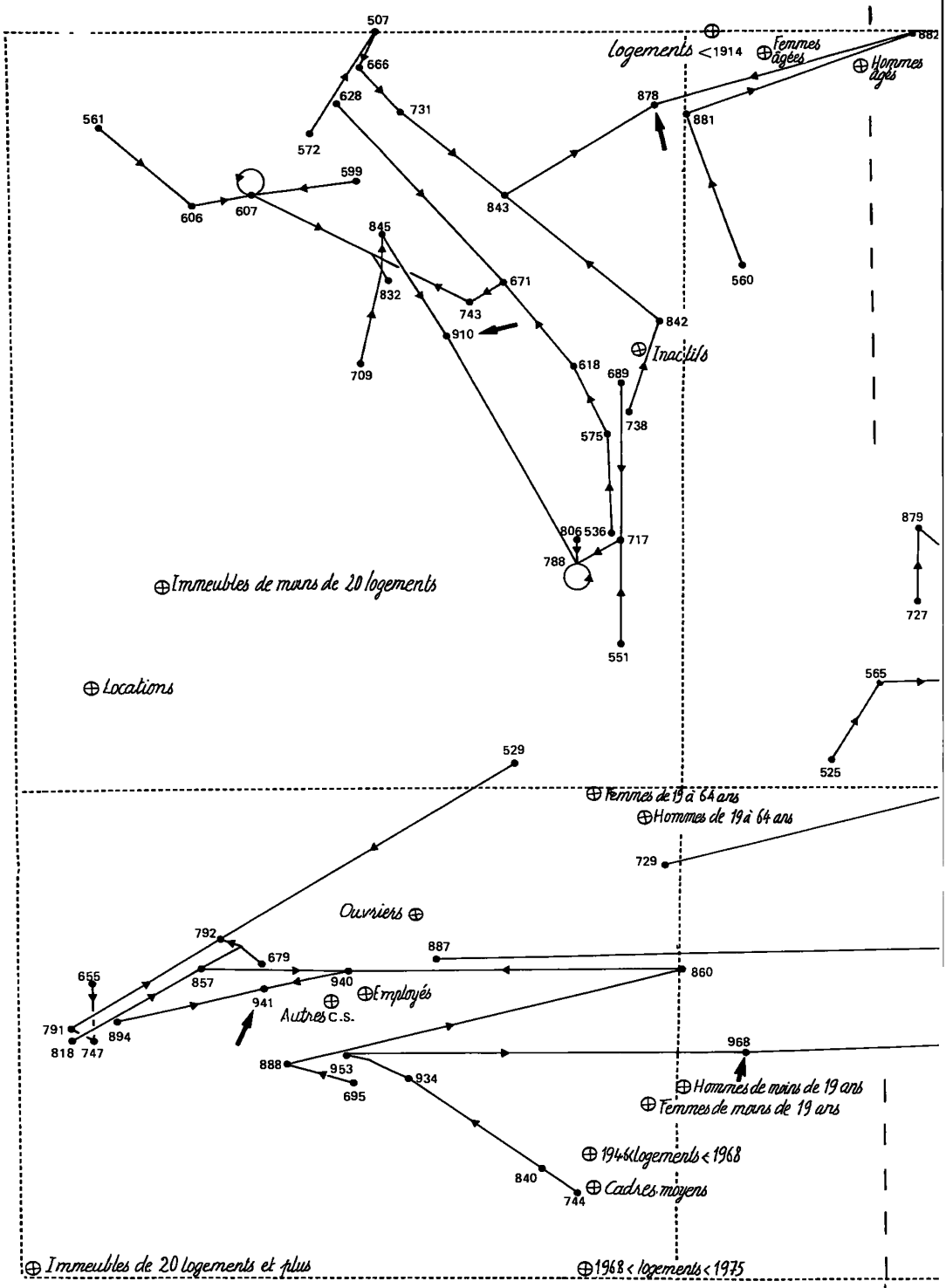
A la vérité le résultat de l'agrégation avec contrainte de taille n'est pas très différent de ce que pourrait fournir un dépouillement de l'arbre construit au § 4.1 sans cette contrainte : beaucoup des noeuds inférieurs très homogènes obtenus au § 4.1 se retrouvent ici à peu près tels quels, pourvu que leur effectif soit précisément celui requis. (De façon précise, on retrouve exactement dans l'arbre du § 4.2 la partie de l'arbre du § 4.1 inférieure au niveau du premier noeud dépassant la contrainte de taille). Mais d'une part le découpage de l'arbre du § 4.1 est en quelque sorte fait ici non à un niveau inférieur, mais sur chaque branche à un niveau adapté à la contrainte de taille. C'est ainsi que le découpage opéré dans le centre ville est nettement plus fin que le découpage issu de l'analyse sans contrainte de taille. Ceci est logique. Le centre possédant (notamment au niveau de la construction des logements) une nette spécificité l'analyse sans contrainte le distingue quasiment en bloc de la périphérie, après l'avoir agrégé à un niveau peu élevé ; tandis que la contrainte de taille conserve les subdivisions encore inférieures entre lesquelles on peut reconnaître des disparités réelles fréquemment ignorées des urbanistes plus sensibles à l'unité historique des quartiers. D'autre part la contrainte de taille est cause que s'opèrent certaines agrégations, qui ne se seraient pas faites sans elles ; en cela elle fournit des résultats différents de ceux du § 4.1 ; parfois il s'agit de perturbations regrettables ; mais parfois au contraire on a des solutions heureuses à notre problème fondamental : constituer des quartiers d'une taille prescrite.

Même si elle demande révision (§ 5.1), la classification avec double contrainte (taille et contiguïté) a le mérite de fournir un projet de découpage issu automatiquement de l'analyse des données : c'est la carte 2.

5 Elaboration des résultats de l'analyse des données

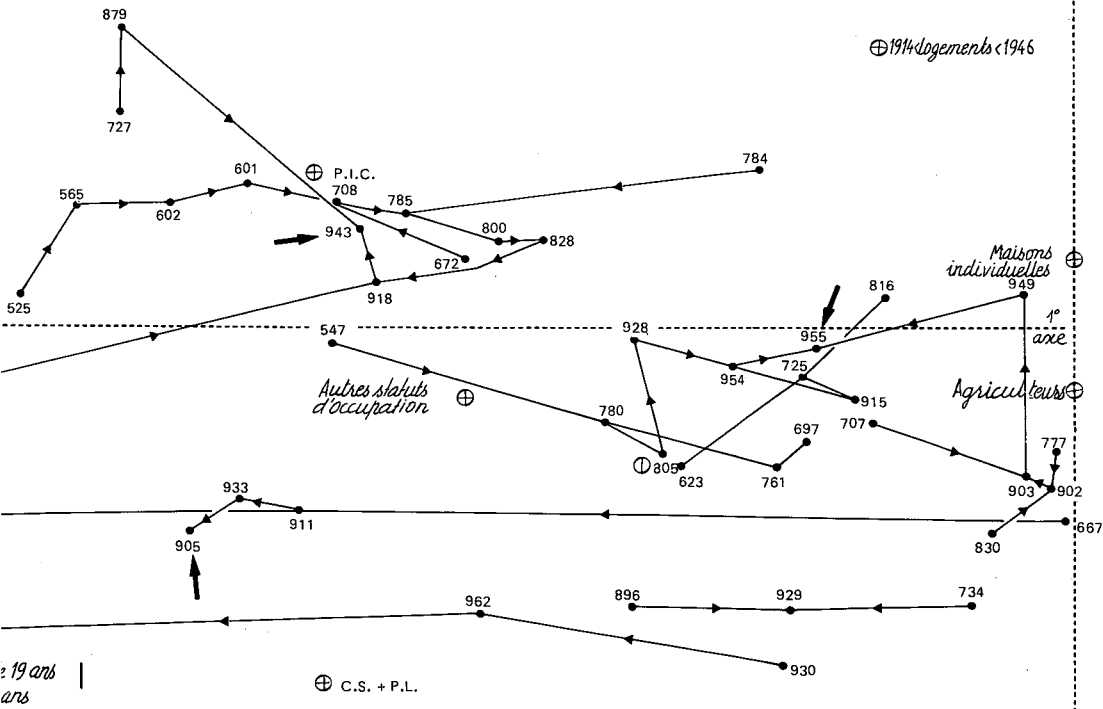
L'analyse avec contrainte permet ainsi d'obtenir un découpage de la ville de façon très rapide. La comparaison de ce découpage avec le découpage obtenu après recombinaison des noeuds issus de l'analyse sans contrainte permet de mettre en évidence les noeuds stables et aussi les petites zones qui sont rattachées tantôt à un secteur tantôt à un autre. Il reste cependant beaucoup à faire pour aboutir à une solution sinon parfaite, du moins très satisfaisante. L'élaboration des résultats de la





882
⊕ Hommes
âgés

Zonage de l'agglomération d'Aix : plan 1x2 issu de l'analyse du tableau
504 x 24 : 504 îlots x 24 modalités descriptives ; on n'a pas figuré les
îlots mais quelques classes construites en agrégeant ceux-ci (34.2.)
et mises en éléments supplémentaires.



classification se poursuit donc, d'une part en les confrontant à l'analyse factorielle, d'autre part en recourant à toutes les données connues des urbanistes.

5.1 Critique des classifications d'après l'analyse factorielle

En plaçant les noeuds en éléments supplémentaires sur les cartes issues de l'analyse factorielle (chaque noeud étant au barycentre des flots qui le constituent) on voit immédiatement par quelles similitudes (autre que les proximités et contiguités géographiques) s'explique l'agrégation progressive des flots en noeuds. (Pour plus de précision on peut remonter aux données primaires en calculant pour chaque noeud la somme des données décrivant les flots qui le composent : e.g. nombre de logements antérieurs à 1914 ; etc).

Pour illustrer ce § on a représenté le plan 1x2 issu de l'analyse factorielle (cf § 3) avec une représentation partielle de la classification faite au § 4.2. (Evidemment, notre étude a porté sur l'ensemble de la classification ; et considéré le plan 3x4 après le plan 1x2 ; mais même pour le spécialiste le recours à des représentations partielles est indispensable : si l'on veut voir clair dans un grand ensemble de points). Cette représentation plane va nous permettre d'interpréter les regroupements effectués et de suivre pas à pas la "génération d'un noeud" :

Prenons par exemple le noeud 910 (flèche en haut à gauche). Il est formé du noeud 845 et du noeud 788. Nous avons relié ces noeuds par deux flèches (l'une partant du noeud 788 vers le noeud 910, l'autre partant du noeud 845 vers le noeud 910).

De l'interprétation des axes faite au § 3 nous pouvons déduire les caractéristiques des noeuds : ainsi 845 est nettement composé de personnes âgées et marqué par les logements anciens ; il en est de même, à un moindre degré du noeud 788.

Plusieurs principes vont guider notre interprétation de la génération des noeuds :

Un noeud qui est formé de noeuds proches de lui sur la carte factorielle, comme par exemple le noeud 955, (à droite) sera forcément un noeud très homogène.

Au contraire, un noeud formé de noeuds dispersés sur la carte (par dispersés nous entendons que les noeuds n'appartiennent pas au même quadrant par exemple) sera moins homogène ; dans ce cas, la contrainte de contiguité "force" le regroupement. Répétons ici que la projection sur les axes 1 et 2 ne résume pas toutes les données. L'homogénéité des quartiers ne peut être confirmée qu'en tenant compte des proximités entre ilots sur les autres axes.

Prenons l'exemple des noeuds 941 et 968, (à gauche, en bas), qui représentent une partie de la ZUP d'Encagne. Le noeud 941 est formé de noeuds très proches, très nettement marqués par les secteurs locatifs des immeubles (axe 1), la population jeune (axe 2) : il s'agit du secteur locatif HLM. (Habitations dites à Loyers Modérés). Sur les axes 3 et 4 on note de plus la proximité des flots du noeud 941 avec la CS ouvrier. Le noeud 968 est formé de points se trouvant au même niveau sur l'axe 2, et de points très dispersés sur l'axe 1. Le noeud 968 comprend : d'une part, le noeud 962 (et ceux qui l'ont généré) fortement marqué par les valeurs positives de l'axe 1, et donc à forte proportion de propriétaires ; d'autre part, les noeuds 912, 934 qui sont proches du 941 ; à cette différence près qu'ils sont plus marqués par les CS cadres moyens.

Ainsi le noeud 968 regroupe le secteur en accession à la propriété de la ZUP (962) et les immeubles proches. L'examen des axes 3 et 4 confirme cette interprétation. Les ilots du noeud 968 sont assez proches des points CS cadres moyens et employés ; et compte-tenu du faible rôle du statut de l'occupant sur les axes 3 et 4, la grande hétérogénéité notée sur l'axe 1 n'apparaît plus.

Cependant, l'analyse peut introduire certains résultats discutables dans les zones suburbaines. Ainsi le noeud 945 (analyse avec contrainte

de taille) est le regroupement de deux lotissements :

S^{te} Anne (lotissement de standing).

Val S^t André (lotissement HLM).

La raison en est simple : les deux lotissements ont été agrégés "par défaut" ; les zones alentour étant encore marquées par l'économie rurale, ces deux zones, bâties à une même époque, ont été regroupées pour cette raison. On prendra garde à détecter ce cas particulier qui ne se retrouve qu'à de rares occasions dans les agglomérations de taille moyenne.

La carte des noeuds projetés sur les premiers axes de l'analyse de correspondance permet donc de distinguer les noeuds très homogènes (noeuds très concentrés sur les cartes). Ces noeuds seront en tout état de cause conservés dans le découpage final. Elle permet l'interprétation de regroupements *a priori* hétérogènes, comme le montre l'exemple du noeud 968. Elle signale aussi (cf noeud 945) des noeuds regroupant des flots très disparates à la réunion desquels il n'est pas possible d'apporter une explication convaincante.

5.2 Collaboration du statisticien et de l'urbaniste

A ce niveau de l'analyse nous possédons donc un zonage provisoire, issu de l'analyse de correspondance et des classifications hiérarchiques avec contrainte de contiguïté, et avec ou sans contrainte de taille. Il est évident que ce zonage, qui, du point de vue des données de départ, est un des meilleurs possible, n'intègre pas des données qualitatives comme par exemple le développement de l'agglomération. Le travail de discussion consistera à revoir, retoucher ce zonage. Ce travail est d'une importance égale au travail statistique proprement dit. Il doit être réalisé par un géographe urbaniste ayant une connaissance réelle de l'agglomération, en collaboration avec l'économiste de la cellule de prévision.

Nous nous bornerons à noter ici que les retouches, modifications, (voire très exceptionnellement les bouleversements) apportés au zonage le sont en fonction des considérations suivantes (par ordre d'importance) :

L'avenir des quartiers. Les P.O.S. permettent de connaître assez précisément l'avenir de certains quartiers. Notre projet étant d'obtenir un découpage pour la prévision de lignes il est évident que des quartiers peuvent (et doivent) être redécoupés en fonction des P.O.S.; généralement on isolera Z.A.D., Z.A.C. et Z.I..

L'activité, prise au sens large, du quartier. Ce critère n'est pas pris en compte dans l'analyse statistique proposée*. Il faudra donc en tenir compte largement. A ce propos il est recommandé de réaliser plusieurs cartes d'emplois au lieu de travail.

Emploi industriel

Emploi administratif notamment

La "réputation" des quartiers. Cette réputation peut tenir aux facilités d'accès aux équipements collectifs comme à la beauté du site et des constructions.

Cette révision du zonage est certes éminemment qualitative et possède une part d'arbitraire. Elle n'en est pas moins indispensable.

On lira les deux dernières cartes d'Aix en Provence.

La carte n° 3 fournit la carte finale retenue. Les numéros y figurant sont les numéros des noeuds issus de l'analyse avec contrainte de taille qui correspondent aux quartiers retenus.

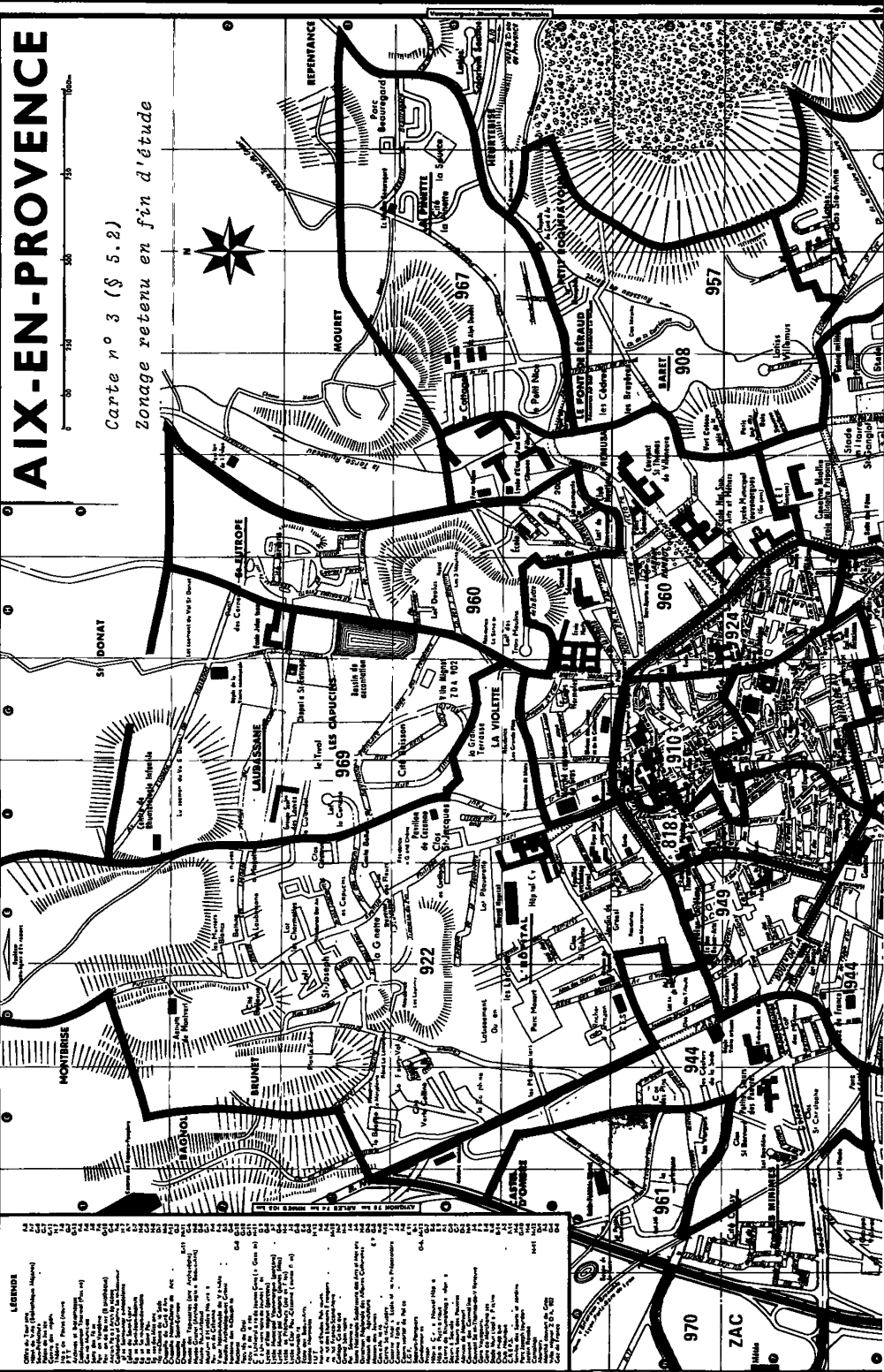
La carte n° 4 fournit le découpage réalisé "à la main" avant l'étude informatique.

* On pourrait peut être y remédier en introduisant une variable issue d'un traitement sur le fichier "SIRENE" en flots.

AIX-EN-PROVENCE

0 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000

Carte n° 3 (§ 5.2)
Zonage retenu en fin d'étude



LÉGENDE

020 : Aire de jeu
021 : Aire de jeu (à l'échelle réduite)
022 : Aire de jeu (à l'échelle réduite)
023 : Aire de jeu (à l'échelle réduite)
024 : Aire de jeu (à l'échelle réduite)
025 : Aire de jeu (à l'échelle réduite)
026 : Aire de jeu (à l'échelle réduite)
027 : Aire de jeu (à l'échelle réduite)
028 : Aire de jeu (à l'échelle réduite)
029 : Aire de jeu (à l'échelle réduite)
030 : Aire de jeu (à l'échelle réduite)
031 : Aire de jeu (à l'échelle réduite)
032 : Aire de jeu (à l'échelle réduite)
033 : Aire de jeu (à l'échelle réduite)
034 : Aire de jeu (à l'échelle réduite)
035 : Aire de jeu (à l'échelle réduite)
036 : Aire de jeu (à l'échelle réduite)
037 : Aire de jeu (à l'échelle réduite)
038 : Aire de jeu (à l'échelle réduite)
039 : Aire de jeu (à l'échelle réduite)
040 : Aire de jeu (à l'échelle réduite)
041 : Aire de jeu (à l'échelle réduite)
042 : Aire de jeu (à l'échelle réduite)
043 : Aire de jeu (à l'échelle réduite)
044 : Aire de jeu (à l'échelle réduite)
045 : Aire de jeu (à l'échelle réduite)
046 : Aire de jeu (à l'échelle réduite)
047 : Aire de jeu (à l'échelle réduite)
048 : Aire de jeu (à l'échelle réduite)
049 : Aire de jeu (à l'échelle réduite)
050 : Aire de jeu (à l'échelle réduite)
051 : Aire de jeu (à l'échelle réduite)
052 : Aire de jeu (à l'échelle réduite)
053 : Aire de jeu (à l'échelle réduite)
054 : Aire de jeu (à l'échelle réduite)
055 : Aire de jeu (à l'échelle réduite)
056 : Aire de jeu (à l'échelle réduite)
057 : Aire de jeu (à l'échelle réduite)
058 : Aire de jeu (à l'échelle réduite)
059 : Aire de jeu (à l'échelle réduite)
060 : Aire de jeu (à l'échelle réduite)
061 : Aire de jeu (à l'échelle réduite)
062 : Aire de jeu (à l'échelle réduite)
063 : Aire de jeu (à l'échelle réduite)
064 : Aire de jeu (à l'échelle réduite)
065 : Aire de jeu (à l'échelle réduite)
066 : Aire de jeu (à l'échelle réduite)
067 : Aire de jeu (à l'échelle réduite)
068 : Aire de jeu (à l'échelle réduite)
069 : Aire de jeu (à l'échelle réduite)
070 : Aire de jeu (à l'échelle réduite)
071 : Aire de jeu (à l'échelle réduite)
072 : Aire de jeu (à l'échelle réduite)
073 : Aire de jeu (à l'échelle réduite)
074 : Aire de jeu (à l'échelle réduite)
075 : Aire de jeu (à l'échelle réduite)
076 : Aire de jeu (à l'échelle réduite)
077 : Aire de jeu (à l'échelle réduite)
078 : Aire de jeu (à l'échelle réduite)
079 : Aire de jeu (à l'échelle réduite)
080 : Aire de jeu (à l'échelle réduite)
081 : Aire de jeu (à l'échelle réduite)
082 : Aire de jeu (à l'échelle réduite)
083 : Aire de jeu (à l'échelle réduite)
084 : Aire de jeu (à l'échelle réduite)
085 : Aire de jeu (à l'échelle réduite)
086 : Aire de jeu (à l'échelle réduite)
087 : Aire de jeu (à l'échelle réduite)
088 : Aire de jeu (à l'échelle réduite)
089 : Aire de jeu (à l'échelle réduite)
090 : Aire de jeu (à l'échelle réduite)
091 : Aire de jeu (à l'échelle réduite)
092 : Aire de jeu (à l'échelle réduite)
093 : Aire de jeu (à l'échelle réduite)
094 : Aire de jeu (à l'échelle réduite)
095 : Aire de jeu (à l'échelle réduite)
096 : Aire de jeu (à l'échelle réduite)
097 : Aire de jeu (à l'échelle réduite)
098 : Aire de jeu (à l'échelle réduite)
099 : Aire de jeu (à l'échelle réduite)
100 : Aire de jeu (à l'échelle réduite)

970
ZAC
961
944
949
910
960
908
957
967
969
922
966
965
964
963
962
961
960
959
958
957
956
955
954
953
952
951
950
949
948
947
946
945
944
943
942
941
940
939
938
937
936
935
934
933
932
931
930
929
928
927
926
925
924
923
922
921
920
919
918
917
916
915
914
913
912
911
910
909
908
907
906
905
904
903
902
901
900
899
898
897
896
895
894
893
892
891
890
889
888
887
886
885
884
883
882
881
880
879
878
877
876
875
874
873
872
871
870
869
868
867
866
865
864
863
862
861
860
859
858
857
856
855
854
853
852
851
850
849
848
847
846
845
844
843
842
841
840
839
838
837
836
835
834
833
832
831
830
829
828
827
826
825
824
823
822
821
820
819
818
817
816
815
814
813
812
811
810
809
808
807
806
805
804
803
802
801
800
799
798
797
796
795
794
793
792
791
790
789
788
787
786
785
784
783
782
781
780
779
778
777
776
775
774
773
772
771
770
769
768
767
766
765
764
763
762
761
760
759
758
757
756
755
754
753
752
751
750
749
748
747
746
745
744
743
742
741
740
739
738
737
736
735
734
733
732
731
730
729
728
727
726
725
724
723
722
721
720
719
718
717
716
715
714
713
712
711
710
709
708
707
706
705
704
703
702
701
700
699
698
697
696
695
694
693
692
691
690
689
688
687
686
685
684
683
682
681
680
679
678
677
676
675
674
673
672
671
670
669
668
667
666
665
664
663
662
661
660
659
658
657
656
655
654
653
652
651
650
649
648
647
646
645
644
643
642
641
640
639
638
637
636
635
634
633
632
631
630
629
628
627
626
625
624
623
622
621
620
619
618
617
616
615
614
613
612
611
610
609
608
607
606
605
604
603
602
601
600
599
598
597
596
595
594
593
592
591
590
589
588
587
586
585
584
583
582
581
580
579
578
577
576
575
574
573
572
571
570
569
568
567
566
565
564
563
562
561
560
559
558
557
556
555
554
553
552
551
550
549
548
547
546
545
544
543
542
541
540
539
538
537
536
535
534
533
532
531
530
529
528
527
526
525
524
523
522
521
520
519
518
517
516
515
514
513
512
511
510
509
508
507
506
505
504
503
502
501
500
499
498
497
496
495
494
493
492
491
490
489
488
487
486
485
484
483
482
481
480
479
478
477
476
475
474
473
472
471
470
469
468
467
466
465
464
463
462
461
460
459
458
457
456
455
454
453
452
451
450
449
448
447
446
445
444
443
442
441
440
439
438
437
436
435
434
433
432
431
430
429
428
427
426
425
424
423
422
421
420
419
418
417
416
415
414
413
412
411
410
409
408
407
406
405
404
403
402
401
400
399
398
397
396
395
394
393
392
391
390
389
388
387
386
385
384
383
382
381
380
379
378
377
376
375
374
373
372
371
370
369
368
367
366
365
364
363
362
361
360
359
358
357
356
355
354
353
352
351
350
349
348
347
346
345
344
343
342
341
340
339
338
337
336
335
334
333
332
331
330
329
328
327
326
325
324
323
322
321
320
319
318
317
316
315
314
313
312
311
310
309
308
307
306
305
304
303
302
301
300
299
298
297
296
295
294
293
292
291
290
289
288
287
286
285
284
283
282
281
280
279
278
277
276
275
274
273
272
271
270
269
268
267
266
265
264
263
262
261
260
259
258
257
256
255
254
253
252
251
250
249
248
247
246
245
244
243
242
241
240
239
238
237
236
235
234
233
232
231
230
229
228
227
226
225
224
223
222
221
220
219
218
217
216
215
214
213
212
211
210
209
208
207
206
205
204
203
202
201
200
199
198
197
196
195
194
193
192
191
190
189
188
187
186
185
184
183
182
181
180
179
178
177
176
175
174
173
172
171
170
169
168
167
166
165
164
163
162
161
160
159
158
157
156
155
154
153
152
151
150
149
148
147
146
145
144
143
142
141
140
139
138
137
136
135
134
133
132
131
130
129
128
127
126
125
124
123
122
121
120
119
118
117
116
115
114
113
112
111
110
109
108
107
106
105
104
103
102
101
100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
42
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0

On notera :

Que les quartiers du Nord-Est ont sensiblement été modifiés par l'étude qualitative ce, en fonction essentiellement des études des P.O.S.

Que les quartiers ont vu leurs contours "lissés".

Les zonages découpent globalement la ville de façon similaire ; cependant des disparités importantes subsistent qui proviennent de la difficulté de l'urbaniste à distinguer les disparités socio-économiques (l'exemple de la ZUP d'Encagnane est clair) et de l'impossibilité pour l'analyse informatique d'apprécier le développement ultérieur de l'agglomération.

Les deux méthodes, étude informatique et étude qualitative sont donc complémentaires.

BIBLIOGRAPHIE

- B. BURTSCHY : Analyse de Données et Analyse de la Consommation téléphonique de la région de Nantes T.PE/2/1/B.B (stage réalisé au SPAF sous la direction de R. CHEMAMA).
- Y. GRELET et A. W. HAMROUNI : Programme de calcul des pondérations relatives pour plusieurs groupes homogènes de variables [Pond. Pr.] in *Les Cahiers de l'Analyse des Données*. Vol II n° 3 pp 353-359 (1977).
- L. LEBART, A. MORINEAU, N. TABARD : "Techniques de la description Statistique". Méthodes et logiciels pour l'étude des grands tableaux. DUNOD. 1977.
- C. SAUTEREAU : Note de cours "Méthodes géométriques d'analyse de données". CIRO, 24 Av. Prieur de la Côte d'Or. 94110 Arcueil.

Les cartes qui illustrent cet article ont été préparées en reportant les résultats des classifications sur une réduction du Plan-Guide Blay.

Nous remercions l'éditeur qui nous a autorisé à utiliser cette carte dont il a la propriété exclusive.

AIX-EN-PROV

Carte n° 4 (§ 5.2)

Zonage manuel

