

INTERPRETATION STATISTIQUE D'UN RECENSEMENT NOMINATIF : TESTS, CODIFICATION, PREMIERS TRAITEMENTS

par E. HELIN

Professeur ordinaire à l'U.Lg.

Afin de me conformer aux trois objectifs assignés à ce Colloque, il me sera permis de faire état non seulement de ce qui est déjà réalisé, mais de projets en cours d'exécution. Grâce à M. Cl. Desama qui, dès 1968, m'a apporté son dévoué concours et a pris une part active aux recherches exposées ici même par d'autres historiens liégeois, les perspectives qu'offrent les procédés d'*output* seront exposés ci-dessous. Enfin, le présent Colloque insiste à juste titre sur la nécessité d'organiser la collaboration. Nos conclusions abonderont dans ce sens; mais qu'il soit bien établi dès à présent qu'aucune tentative n'aurait vu le jour à Liège sans l'activité et la compétence de l'équipe réunie au L.A.S.L.A. sous la direction du Professeur L. Delatte. C'est dans son Laboratoire que fut menée à bien, pour la première fois en Belgique, l'indexation automatique de textes d'Histoire moderne. A présent encore, il n'est pas de semaine qui s'écoule sans que nous bénéficions de l'enseignement dispensé par MM. Evrard et Denooz.

Par ailleurs, M. l'Archiviste Général du Royaume, M. le Conservateur des Archives de l'Etat à Liège et son personnel ont fait montre d'une infatigable obligeance à l'égard des lecteurs souvent insatiables que sont les historiens démographes.

En effet, c'est d'abord et avant tout de démographie qu'il s'agit. On sait l'essor pris par cette discipline dans les pays occidentaux depuis une quinzaine d'années (1). Le défrichage n'a pu être mené de front dans tous les secteurs. Les registres paroissiaux ont été mis à contribution bien plus que les recensements; les villages et petites bourgades sont étudiées plus minutieusement que les villes; on en sait davantage sur les sujets du Roi Soleil que sur les bourgeois ou les ouvriers que Karl Marx croisait dans la rue lorsqu'il habitait Bruxelles. Ce n'est un secret pour aucun historien professionnel que plus une documentation est copieuse, moins elle attire les chercheurs. Naguère encore, l'inverse aussi était vrai (2). Pareilles options ne peuvent guère se justifier à l'aide d'arguments scientifiques. Elles sont devenues indéfendables pour quiconque est au courant de la mémoire, de la rapidité et de l'endurance à toute épreuve des ordinateurs.

Une fois au travail, on s'aperçoit que l'ordinateur est bien davantage qu'un appareil qu'il faut gaver d'informations. Il est capable d'apercevoir les moindres caractères des masses de données qu'il traite. Il

importe donc de lui faire restituer les plus subtiles nuances. En revanche, dépourvu du flair qui permet d'improviser en cours de route, l'ordinateur nous contraint à justifier logiquement les démarches consacrées par la routine. Itinéraire éprouvant certes, mais récompensé d'une riche moisson d'observations.

Ces deux objectifs —

- 1) recueillir les variables disponibles pour toute une population, quelle que soit l'abondance statistique (3),
- 2) reconsidérer à la lumière d'une technique nouvelle les diverses opérations traditionnelles en démographie rétrospective —

ont télé-commandé les solutions apportées aux problèmes d'*input*. Les particularités inhérentes aux documents de base ont suggéré le reste.

Les documents de base, ce sont cinq gros in-folio, cotés 7 à 12, dans une série qui en compte 158 et qui, presque chaque année entre 1800 et 1846, recense rue par rue, maison par maison, les habitants de la ville de Liège.

Nous avons là une multitude de données susceptibles de montrer en détail l'impact de la croissance économique sur une population qui va être une des premières à suivre l'exemple anglais et à s'engager dans la voie d'une industrialisation qui se poursuit encore sous nos yeux. Cette approche est surtout familière à M. Cl. Desama. Pour ma part, dans la démographie du XIXe siècle, je vois se perpétuer une foule de traits qui caractérisaient la société d'Ancien Régime. C'est l'occasion ou jamais de prolonger les enquêtes déjà entreprises sur la période de l'indépendance liégeoise. Le recensement de l'an IX satisfait à ces desiderata. Il n'est que de sept ans postérieur à la conquête française; plus complet que les 4 registres de l'an VIII, il est aussi mieux soigné que ceux qui furent dressés sous l'Empire. Comme tous ceux de cette série, il a pour auteurs les commissaires de police de chaque quartier, aidés dans cette besogne par les membres du Bureau municipal de bienfaisance.

m

Pour chaque adulte, les recenseurs ont inscrit les renseignements suivants :

- numéro d'ordre, rue, numéro de la maison,
- noms, prénom, lien de parenté avec le chef de ménage ou état matrimonial,
- âge,
- "condition" et "état" ou profession, c'est-à-dire ce que l'on appellerait de nos jours le statut et le métier (4),
- provenance et date d'entrée à Liège, pour ceux qui n'en sont pas natifs,
- nombre d'enfants de moins de 12 ans, au domicile du chef de ménage, en distinguant les garçons des filles,
- nom et parfois adresse du propriétaire de l'immeuble,
- indigence, infirmité. De telles remarques sont rares et laconiques : elles précisent qu'un enfant est orphelin, qu'une veuve "vit de ses parents", qu'un cordonnier "n'a guère d'ouvrage".

Chacune de ces variables, ainsi que les manières dont elles ont été interprétées et parfois, hélas, omises, seront l'objet de commentaires détaillés. On se bornera ici à faire état des vérifications préalables et systématiques. A cet effet, on a mis à contribution :

m

- 1) les recensements les plus contemporains, à savoir ceux des ans VIII, X et XI. Ils n'ont servi à interpoler que dans le cas de minimes lacunes (2 ff arrachés) ou de lecture incertaine (5). Pour le reste, on a évité de créer un hybride en mélangeant les données de dates différentes; on s'est borné à compter, sur des échantillons, les concordances et les discordances;
- 2) les matrices primitives du Cadastre. Celles-ci datent de 1810-1813 et ont été tenues à jour jusqu'en 1834. Durant cette période, les Liégeois s'étaient remis à construire dans les quartiers périphériques. Le contrôle n'est donc praticable que dans les rues entièrement bordées de maisons riveraines dès avant 1800. A cette réserve près, aucune omission n'est à reprocher aux agents recenseurs;
- 3) les annonces de la *Gazette de Liège*, déjà utilisées par les commerçants et les personnes qui exercent des professions libérales. Bien sûr, on ne retrouve pas trace, dans le recensement, des marchands ambulants ni des musiciens en tournée; en revanche, que de précisions sur la nature des commerces et les activités des hommes de loi (6)!
- 4) enfin et surtout, les actes de naissance, mariage et décès (enregistrés par l'officier de l'état civil) procurent les adresses à l'aide desquelles nous avons vérifié les noms, âges et qualités des adultes recensés. Les résultats de ces confrontations valent d'être publiés en détail. Une surprise cependant : le fort pourcentage de gens qui ont changé de domicile (7). Ce taux s'amenuise si l'on considère les actes les plus contemporains du recensement; il est d'ailleurs plus élevé pour les locataires et les gens sans métier que pour les propriétaires et ceux qui ont pignon sur rue. En dernière analyse, la discordance entre ces deux séries de documents donne davantage une mesure de la mobilité sociale différentielle que de l'inexactitude des déclarations.

C'est que nos vérifications ne sont pas exhaustives : un banal déménagement suffit à dérouter nos investigations. De là, l'intérêt qu'il y aurait à confier à l'ordinateur l'exploration d'un catalogue des habitants (8).

Au cours des étapes suivantes, en effet, nous n'avons pas retenu d'emblée toutes les données du recensement. Les noms des habitants et ceux du propriétaire de l'immeuble n'ont pas encore été perforés. Les noms de rues non plus. Cette dernière omission est sans gravité puisque les maisons sont numérotées de manière continue pour l'ensemble d'un quartier; un bref indicateur suffit pour reconstituer au complet n'importe quelle adresse. Une table, beaucoup plus longue à vrai dire, établit la concordance avec les parcelles cadastrales, ce qui facilitera les représentations cartographiques : indices de l'indigence ou de l'aisance, industries et commerces, propriétés nationales, immeubles dévastés par le bombardement de 1794, etc.

Puisque c'était la première fois que nous nous aventurons sur le terrain de l'informatique appliquée à

la démographie, nous n'avons pas cru pouvoir nous passer d'un intermédiaire entre le recensement original et la carte perforée. En l'occurrence, six registres dont chaque page fut divisée en 14 colonnes. Neuf d'entre elles pouvaient être remplies aussitôt (n^o d'ordre de chaque adulte, n^o de maison, n^o d'ordre de chaque ménage à l'intérieur d'une même maison, âge, date d'entrée à Liège pour ceux qui sont étrangers à la ville, nombre de garçons et nombre de filles de moins de 12 ans) ou bien en procédant sur le champ à un codage facile à mémoriser : 10 chiffres pour désigner chaque état matrimonial en fonction du sexe, 10 autres pour désigner la catégorie à laquelle appartient le propriétaire de l'immeuble. Trois autres colonnes sont réservées d'abord au métier et au statut, ensuite à la provenance des étrangers. Afin de traiter les deux premiers cas, il a fallu prévoir un codage matriciel et combinatoire qui, jusqu'à présent, requiert 356 nombres de 3 chiffres (pour déterminer la nature de l'occupation) et 87 nombres de 2 chiffres (afin de situer le niveau occupé dans la stratification des statuts socio-professionnels et de signaler infirmes ou indigents).

En ce qui concerne les provenances, on a exploité la constante selon laquelle *grosso modo* les contingents d'immigrés sont inversement proportionnels à la distance qui sépare lieu d'origine et lieu d'implantation. On a alors codé les communes limitrophes de Liège, puis les cantons du département, ensuite les sous-préfectures des départements voisins, les départements eux-mêmes et enfin les pays étrangers. Bien que les capitales et les chefs-lieux aient fait l'objet d'un codage distinct, ce système s'est avéré trop sommaire. Il sera remplacé par un autre, mis au point par M.M. Desama et Poussou. Au prix d'un travail plus long (nombres de 8 chiffres au lieu de 3), le nouveau code indiquera les distances et l'orientation par rapport au pôle d'attraction.

La répartition des zones à perforer sur chaque carte s'est modelée sur la succession des colonnes, utilisée dans les registres ayant servi au codage.

La perforation de 33 colonnes dans plus de 31.000 cartes a pu être accélérée en pré-perforant la référence, le n^o du quartier et le n^o d'ordre des adultes. La vérification ne s'est pas faite automatiquement, mais en collationnant les chiffres imprimés au sommet de chaque carte avec ceux du registre codé. Avant les principales mises en oeuvre du fichier, l'I.B.M. 1620 du L.A.S.L.A. a effectué des contrôles de séquence. C'est ce même ordinateur qui a perforé un des fichiers dérivés : celui des étrangers (plus de 4.375 cartes). Un autre fichier dérivé — celui des ménages — compte 11.496 cartes qui récapitulent, ménage par ménage et maison par maison, les totaux d'adultes, d'enfants et d'habitants. Il a été perforé en 1968, par une calculatrice I.B.M. 602.

L'exploitation proprement dite soulève des problèmes d'*output* dont traitera M. Cl. Desama. En attendant, il suffira ici de citer les têtes de chapitre et de souligner que nous recueillons nos résultats quartier par quartier. Des études antérieures ont fait ressortir, en effet, les contrastes démographiques qui continuaient à caractériser citadins d'une part, faubouriens de l'autre. Des oppositions tout aussi marquées

distinguaient quartiers riches (l'île, par exemple) et quartiers pauvres : Outre Meuse et la frange extérieure des faubourgs. En fait, nous traitons successivement 9 fichiers : 5 pour les 26.824 habitants des quartiers *intra muros*, 4 pour les 15.349 habitants des faubourgs.

Très graduellement, nous avons commencé par des dénombrements simples (une seule variable) et ce n'est qu'ensuite que nous avons abordé le calcul des fréquences de combinaisons entre plusieurs variables.

Comme premier type d'opérations, citons :

- la répartition des maisons vides et des 7.875 maisons occupées,
- le nombre de ménages par maison,
- le nombre d'habitants par ménage,
- celui des adultes occupés dans les grands secteurs de la production économique et dans chaque quartier,
- celui des indigents,
- les contingents d'étrangers, compte tenu de leur pays ou région d'origine.

Les traitements qui combinent plusieurs variables se comptent par dizaines et nous sommes loin de les avoir tous essayés. On dispose déjà des :

- pyramides des âges par quartiers : elles combinent l'âge, le sexe, l'état matrimonial. Il serait possible d'en dessiner, pour les indigents comme pour des professions caractéristiques : par exemple, les fonctionnaires républicains, les membres du clergé. Le calcul a été fait pour les chefs de ménage propriétaires de la maison qu'ils occupent et pour les étrangers;
- distribution des enfants par ménage, compte tenu de la composition des ménages et de l'âge de la mère; idem pour les domestiques, selon le métier du chef de ménage;
- contingents de 380 Français installés à Liège, ventilés selon la profession et la date d'arrivée, etc...

Sans doute aurions-nous pu allonger cette énumération si un changement d'ordinateur (mise en marche d'un I.B.M. de la série 360) et l'adoption du langage PL/1 ne nous avaient incités à retarder l'essai de nouveaux programmes : délai passager que nous espérons bien rattraper grâce à la souplesse du PL/1 et à la rapidité du traitement des données dès qu'elles sont enregistrées sur disques.

Les perspectives immédiates ? Assurément continuer la publication des résultats déjà acquis.

A moyen terme, il s'agira de réparer une omission initiale, en perforant les patronymes et les prénoms des adultes recensés en l'an IX. Il importe, en effet, d'enrichir systématiquement la documentation déjà disponible en assurant le couplage entre le recensement de l'an IX et les séries contemporaines. Outre celles qui ont servi aux contrôles préalables (cfr. ci-dessus), nous préparons le traitement des collections

de passeports, des rôles fiscaux, des listes d'indigents secourus à domicile. Dans chaque catégorie, les anthroponymes fournissent le support sur lequel reposent les identifications individuelles.

Les difficultés sont loin d'être résolues : elles proviennent tant de la fréquence des homonymies que de l'instabilité des patronymes. L'abandon de la particule nobiliaire entre 1794 et 1804 n'est pas seul en cause.

Le fichier de l'an IX doit encore se prêter à de nombreuses recherches prévisibles dans un avenir plus lointain. Nous pensons surtout à la détermination de cohortes qui, à partir de 1801, devraient servir à mesurer la fécondité, la mortalité et la mobilité professionnelle durant la première moitié du XIXe siècle. Un tel fichier permettra aussi de franchir le cap des années 1789-1794, période révolutionnaire qui clôture si arbitrairement tant de recherches sur la société traditionnelle. Ce n'est pas parce que l'Ancien Régime est aboli que disparaissent du jour au lendemain les hommes qui l'ont vécu et qui, bon gré mal gré, tendent à maintenir en place ses structures économiques, sociales et mentales.

Efforçons-nous de conclure.

Faire connaître les réalisations : nous n'y sommes parvenus que superficiellement du fait que les articles jusqu'à présent publiés se sont limités à des résultats partiels.

Faire connaître les projets à l'étude : c'est quelque peu présomptueux au moment où l'adoption du PL/1 ouvre des perspectives aussi nombreuses que prometteuses.

Faciliter la collaboration : ce ne peut être un vœu pieux, c'est un impératif. En se groupant depuis deux ans et demi, quelques historiens de Liège, de Verviers et de Namur, de clients qu'ils étaient chez les informaticiens, sont devenus apprentis. C'est aux philologues du L.A.S.L.A. qu'ils le doivent et aussi aux encouragements de l'Administration des Archives. Inutile de dire que nous avons encore presque tout à expérimenter et que nous souhaitons consulter des spécialistes de la toponymie, de l'anthroponymie, du droit ancien. A plus longue échéance, nous voudrions que les jeunes générations d'historiens reçoivent une formation qui les mette à même de tirer parti des ressources d'un ordinateur.

Les techniques d'*input* vont se perfectionnant; les appareils sont surclassés au bout de quelques années. Ce qui restera toujours indispensable, c'est la matière première, c'est-à-dire la documentation d'archives. Au moment où les ordinateurs centuplent leurs performances, il serait paradoxal d'envoyer au pilon des données démographiques de premier ordre parce que, faute de crédits, l'Administration des Archives serait dépourvue des locaux, des fichiers ou des disques susceptibles d'assurer dès aujourd'hui la conservation et, demain, l'exploitation scientifique d'irremplaçables richesses documentaires.

- (1) Il suffira de se référer aux classiques que sont devenus les livres du Rd P. Roger MILLS, de MM. Louis HENRY et Peter LASLETT. Des revues telles que *Population* (publiées par l'I.N.E.D. à Paris), *Population Studies* (London School of Economics) et *Annales de démographie historique* (Paris, Sirey) mettront le lecteur au courant des résultats les plus récents. Une synthèse est présentée par J.P. POUSSOU et Georges DUPEUX, *Démographie historique*, 415 pp., Paris, A. Colin, 1970. Pour l'usage de l'ordinateur par les historiens, se référer aux titres réunis ci-dessous, à la suite de l'exposé de M. Claude DESAMA, *Réflexions sur les programmes appliqués au recensement de la population de Liège*.
- (2) La littérature récente témoigne du prestige que conserve le mystère de chaque origine, comme des controverses qui se perpétuent autour de questions insolubles.
- (3) L'expression ne peut être entendue ici au sens que vient de lui attribuer J. STENGERS, *L'historien devant l'abondance statistique*, dans Revue de l'Institut de Sociologie, n° 3, pp. 427-458, Bruxelles, 1970. Celui-ci souligne l'in vraisemblance de statistiques d'abord publiées par des fonctionnaires dépourvus de sens critique, prises ensuite au sérieux par des vulgarisateurs trop pressés. Mais la plupart des historiens démographes se servent de sources authentiques, exposent en détail leurs résultats et s'astreignent à de longues enquêtes, précisément parce qu'ils n'ajoutent pas foi aux chiffres officiels. Aujourd'hui, l'ordinateur multiplie les contrôles, permet de vérifier chaque opération et incite à recourir aux tests statistiques : autant de précautions qu'ignorait sereinement l'historiographie traditionnelle.
- (4) Lire à ce sujet *Les codes socio-professionnels comme instruments d'analyse des populations antérieures à la Révolution Industrielle*, à paraître dans les Actes du XL^e Congrès de la Fédération archéologique, Malines, 1970.
- (5) Quelques toponymes : "Scée" pour Seilles, "Néoze" pour Néau [Eupen]. - D'autres ont pu être identifiés : un domestique natif de la Hesbaye, selon la déclaration de l'an IX est recensé, en l'an X, comme provenant de Momalle.
- (6) Henri Delloye, recensé en l'an IX au n° 873 de la place Saint-Pierre, comme "homme de loy", énumère les multiples activités auxquelles il s'adonne dans l'annonce qu'il insère dans la *Gazette de Liège* du 17 vendémiaire an IX.
- (7) Pour les hommes mariés dans l'année antérieure au recensement, il atteint 66 %.
- (8) Cfr. ci-dessous et l'exposé du couplage des données, par M. Cl. DESAMA.