

**D'un *Dictionnaire des communes*  
à une base de données statistiques**

par

**D. DELVAUX**

*F.R.F.C. - BELGIQUE*

**Et. HELIN**

*Université de Liège - BELGIQUE*

Il s'agit de tirer quelques leçons d'une expérience au cours de laquelle des spécialistes de trois sciences humaines en pleine mutation (l'histoire, la géographie, la sociologie) se trouvent aux prises avec les techniques informatiques elles-mêmes en rapide évolution.

Le projet initialement intitulé *Dictionnaire statistique des Communes de Wallonie*, a été subsidié par le F.R.F.C., de 1975 à 1979. Il se poursuit toujours et d'autre part, il répond à une série de besoins devenus manifestes bien avant 1975. Quels sont-ils ?

Un constat élémentaire d'abord. L'écart entre d'une part les informations que publient au compte-gouttes les *Dictionnaires* de communes et d'autre part, la masse des données tantôt enfouies dans les recueils officiels (on compte plus d'un millier de volumes depuis l'indépendance de la Belgique), tantôt éparses dans les archives du Royaume, des provinces, des communes et d'innombrables centres de documentation.

En dépit du succès de librairie dont profitent les éditeurs de Dictionnaires, les mises à jour sont tardives ou partielles et le lecteur de 1981 doit souvent se contenter d'informations périmées. Par définition, un dictionnaire adopte l'ordre alphabétique, le plus aveugle de tous puisqu'il ne se contente pas d'atomiser les données; il ignore les liens tissés par la géographie et par l'histoire entre deux villages voisins, entre une ville et sa banlieue. Il est donc incapable de reconstituer un ensemble, de suggérer une dynamique. Lorsque la tâche de documentation est perçue comme aléatoire, le chercheur qui ne met pas aussitôt la main sur le renseignement dont il a besoin, "fait comme si" on pouvait s'en passer. Combien en avons-nous vu de ces étudiants qui se résignent à ignorer la structure de la propriété foncière parce qu'ils ne soupçonnent pas les ressources des recensements agricoles de 1846 et de 1955 !

Il y a plus grave : les chiffres recueillis à grand-peine sont le plus souvent acceptés sans vérification. Il serait cependant si commode de les disposer en séries chronologiques afin de repérer les fluctuations les plus invraisemblables. Rien d'étonnant alors si l'on se contente d'une typologie impressionniste. On continue d'opposer les "pauvres" villages d'Ardenne à ceux de la plantureuse Hesbaye, les communes "désindustrialisées" du pays de Herve à celles de la vallée de la Meuse, qui passent pour encore vouées à la croissance. Aussi n'a-t-on jamais construit un échantillon stratifié ni même un *master sample* valable pour la province ou une quelconque des régions économiques ou culturelles qui la composent.

Bref, les chercheurs qui veulent tirer parti des statistiques de tout genre (population, logements, emploi du sol, ressources agricoles, industries, main d'oeuvre, écoles, équipement hospitalier, finances locales, élections) qui sont recueillies dans le cadre de la commune, se trouvaient, avant 1975, placés devant une tâche de Sisyphe : le temps passé à collecter leurs données empiétait sur tout le reste, au point qu'il compromettait le travail de contrôle, de comparaison et d'interprétation.

C'est au départ de tels constats qu'a été élaboré, en 1974, le projet intitulé *Dictionnaire Statistique des Communes de Wallonie*.

Le but ultime était de confier à des mémoires d'ordinateur toutes les données quantifiables recueillies à l'échelon communal. Même avec l'enthousiasme des néophytes, une tâche aussi démesurée ne pouvait être entreprise qu'en ménageant, dès le début, des étapes. Avant d'englober la Wallonie, il était prudent de s'en tenir à la province de Liège, comme banc d'essai : ses archives sont concentrées au chef-lieu; nous en connaissons les ressources documentaires et les témoins privilégiés. Des correspondants du Limbourg hollandais, de Namur et de Luxembourg, mis au courant de nos essais, ont d'ores et déjà été interrogés quant aux adaptations qui s'imposeraient le jour où le *Dictionnaire* serait étendu aux provinces voisines.

Plus délicat, le choix d'un cadre chronologique. En tant qu'historiens, nous estimions qu'il était essentiel de saisir l'évolution à très long terme, celle que négligent précisément, économistes et sociologues. Si les communes naissent au Moyen Age et forment depuis lors le cadre de base dans lequel s'inscrit la vie publique et sociale des Belges, leurs administrations ne produisent régulièrement des informations quantifiées qu'à partir de la réorganisation des municipalités par la Constitution de l'An VIII (25 décembre 1799).

Il ne pouvait être question de tout collectionner. Dès le départ et jusqu'à présent, on a concentré l'effort sur 4 étapes qui jalonnent l'évolution de nos populations.

- 1800-1814 : à de rares exceptions près (Verviers) la révolution industrielle n'a pas encore marqué l'économie traditionnelle;
- 1846-1866 : formation du bassin industriel liégeois, premières manifestations de la transition démographique;
- 1895-1914 : apogée de l'industrialisation;
- 1947-1970 : régime de population stable ; désindustrialisation.

Les périodes retenues coïncident avec celles de l'abondance et de l'excellence des statistiques. Elles embrassent une révolution capitale dans l'histoire de Liège et de la Wallonie : l'industrialisation, qui ne se réduit pas seulement à ses dimensions économiques mais qui a métamorphosé nos paysages et rendu mobiles nos populations tout en altérant de fond en comble leur vie quotidienne.

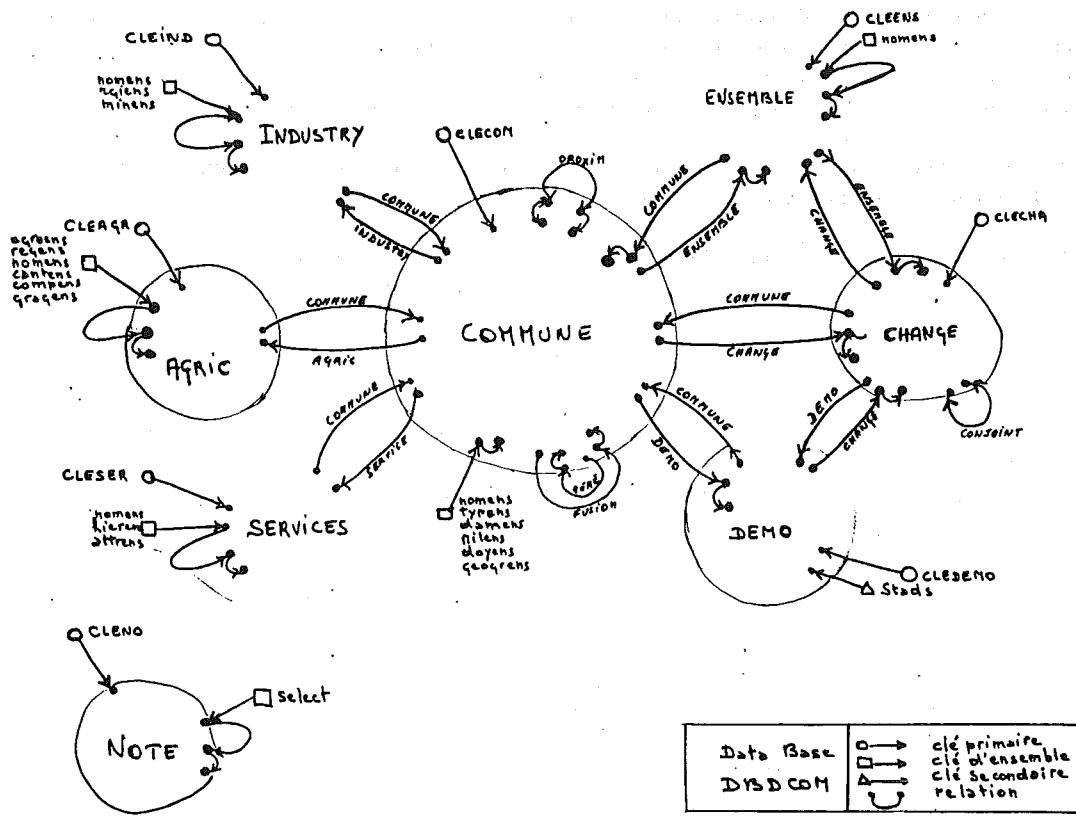
De même que l'historien se doit de retrouver la chronologie de la révolution industrielle, par vocation le sociologue s'efforcera de repérer quels sont les meilleurs, c'est-à-dire les plus sensibles indicateurs d'un changement social sans précédent, tandis que le géographe, interprétant les aires d'attraction, la hiérarchie de villages, bourgades et grands centres, fera apercevoir comment s'opère la redistribution de l'espace social.

Pluri-disciplinaire, notre entreprise l'était au départ par ses sources et la formation des trois chercheurs qui s'y sont attelés. Elle devait l'être aussi par la démarche. Il ne s'agissait plus de stocker telles quelles des statistiques; il fallait désormais les classer, "distinguer pour unir", ordonner une masse de chiffres de manière telle que leurs relations et corrélations puissent être investiguées en tous sens.

C'est ce que l'organisation des données en banque informatisée permet par la structuration du mode d'interaction entre les informations et les utilisateurs.

Une équipe du Centre de Calcul de l'Université de Liège, dirigée par Louis LEDONNE, a mis au point S.L.B.D./1, système et langage pour base de données actuellement sous PL/1. Nous l'avons adopté pour plusieurs raisons :

- 1.- il a été créé pour la gestion des données des services publics et notre application sciences humaines y entrainait donc exactement;
- 2.- les objets sont immédiatement accessibles par leur nom et leur manipulation par des non-initiés ne soulève pas d'obstacles;
- 3.- des relations multiples structurent les données, ainsi que le montre le schéma.



A l'aide de quelques macros de base du D.M.L. (Data Manipulation Language), telles que :

FIND, pour entrer dans l'entité,  
DATA pour avoir accès à une information particulière,  
JUMP pour utiliser le réseau de relations,

il est possible d'ignorer tout de l'organisation physique des fichiers de données. Grâce aux chemins ainsi balisés, on va de la structure par âge de la population de Liège en 1970 à celle des communes voisines de Liège (opérateur PROXIM) ou qui ont fusionné avec Liège (opérateur FUSION).

Un saut vers l'entité CHANGE (opérateur CHANGE) attire l'attention sur des modifications territoriales de la commune, tandis qu'un accès à l'entité NOTE fournit la référence complète et les précautions d'utilisation de la variable choisie.

Un autre saut encore par les clés d'ensemble permet de parcourir le temps (NOMENS) et de retrouver Liège en 1947, 1910, 1880, 1866 et 1846 ou encore de repartir vers les communes appartenant à la même région géographique (GEOGRENS) ou au même canton électoral (MILENS).

Cette expérience de construction d'un outil documentaire pour les chercheurs en sciences humaines pose bien sûr autant de questions qu'elle apporte de solutions. Les liens doivent encore être établis entre la banque de données et les packages statistiques ou cartographiques et la documentation doit sans cesse être enrichie des données contemporaines.

Une expérience ne doit faire état que de ce qui est acquis. Pourtant, le principal avantage de la base des données, c'est qu'elle reste ouverte. Elle peut s'enrichir. Au fur et à mesure que parviendront les résultats du recensement de 1981, ils seront incorporés à la suite des grands relevés décennaux antérieurs.

L'actuelle division du territoire en communes fusionnées n'est plus un obstacle aux comparaisons diachroniques. C'est un atout majeur aux yeux des historiens. Nous envisageons de collecter non seulement des résultats électoraux mais aussi des renseignements non quantitatifs : notices bibliographiques, événements marquants, conclusions de monographies ou de mémoires de fin d'études. Le découpage en 4 étapes (cfr ci-dessus, p. 3) n'est qu'un procédé opératoire et il ne doit pas imposer des oeillères. En cours de dépouillement, on a découvert une foule de données restées inédites et qui éclairent la période 1919-1939 que nous avons cru, à tort, traiter par interpolation.

Beaucoup reste à faire dans le domaine de la détection automatique des erreurs et tout en matière de cartographie. Nous reconnaissons d'autant plus volontiers lacunes et imperfections que nous voyons déjà le moyen d'y remédier.

La véritable inquiétude n'est pas là. Elle est dans le fait que notre Base de données reste sous-utilisée. Il serait trop facile d'en prendre son parti sous prétexte que tel est le sort de bien des recherches en sciences humaines, qu'il s'agisse de traditionnelles thèses de doctorat ou d'ingénieux systèmes informatiques. Il est plus positif de se demander si le caractère interdisciplinaire de l'entreprise ne lui a pas fait manquer son public. Insistons-y : il n'est pas question de revenir aux vieilles recettes de la spécialisation à outrance. Nous avons montré au contraire, qu'une base de données recueillies à l'échelon communal n'a de sens que si elle éclaire le présent par le passé, le local par le régional, le démographique par l'économique, le social, le culturel. A nos yeux, les relations entre les faits importent plus que les faits eux-mêmes.

Il n'en reste pas moins que la plupart des historiens sont des perfectionnistes et, à ce titre, ils se tiennent à l'écart de tout ce qui est raisonnement probabiliste et technique d'ordinateur; que bien des sociologues réprouvent l'intérêt que nous portons à tout ce qui précède 1970; que plus d'un géographe se contenterait des communes fusionnées d'après 1976. A vouloir intégrer tous les points de vue, nous finissons par n'atteindre aucun public.

Dans un premier temps, le recours à l'ordinateur n'arrange rien; au contraire, il vous isole. La technique de la carte perforée était encore largement dominante en 1974, lorsque fut conçu le projet. Lente, coûteuse, fastidieuse, elle mobilise trop d'efforts et accrédite l'illusion que l'essentiel de la tâche du chercheur consiste à collecter une masse de données. A aucun moment, les utilisateurs potentiels de la Base de données, ne sont incités à faire un essai, à proposer une option. Les étudiants, qui formeront la clientèle de demain, soupçonnent à peine les activités de recherche qui se déroulent dans une Université trop cloisonnée et dont les programmes ignorent toujours le bouleversement qu'apporte l'informatique dans tous les domaines des sciences humaines. En somme, une poignée de chercheurs isolés par leurs conditions de travail; en face, une foule d'étudiants victimes d'archaïques structures d'enseignement.

La micro-informatique pourrait offrir une alternative. Parce qu'elle est bon marché, elle brise le monopole des centres de calcul réservés aux seuls spécialistes de haut niveau. Elle dissémine aujourd'hui le savoir-faire et demain le savoir. Pour que l'étudiant se familiarise avec la machine, il faut qu'il la trouve là où il en a besoin : dans le séminaire qu'il fréquente dès la candidature (en licence, il est trop tard, les choix sont faits), dans un coin de la Salle de Lecture, dans les bureaux de l'Administration, au milieu des archives. Petit à petit, avec les précautions qui s'imposent, il apprendra les quelques instructions qui font apparaître sur écran le chiffre de population qui lui manque, la date nécessaire pour retrouver une référence. Il découvrira qu'une mémoire d'ordinateur contient infiniment plus que n'importe quel Dictionnaire des Communes.