



L'accès aux données pour la connaissance du territoire : le cas du Brésil [article]

 Hervé Théry  Philippe Waniez

L'Espace géographique / Année 2000 / 29-1 / pp. 53-57

L'accès aux données pour la connaissance du territoire : le cas du Brésil

Hervé Théry

École normale supérieure, 45 rue d'Ulm, 75005 Paris.

Philippe Waniez

Institut de recherche pour le développement, École normale supérieure, 45 rue d'Ulm, 75005 Paris.

RÉSUMÉ.— L'accès aux données statistiques localisées paraît, paradoxalement, plus facile au Brésil, pays en voie de développement, qu'en France, si fière de ses statistiques publiques.

BRÉSIL, RECHERCHE SCIENTIFIQUE, STATISTIQUES PUBLIQUES

ABSTRACT.— **Access to data to enhance knowledge of territory: the case of Brazil.**— Access to localised statistical data seems, paradoxically, to be easier in Brazil, a developing country, than in France, which is so proud of its public statistics.

BRAZIL, PUBLIC STATISTICS, SCIENTIFIC RESEARCH

RESUMO.— **O acesso aos dados para o contrecimento do território : a caso do Brasil.**— O acesso aos dados estatísticos localizados parece, paradoxalmente, mais fácil no Brasil, país em desenvolvimento, do que na França, tão orgulhosa das suas estatísticas públicas.

BRASIL, ESTATÍSTICAS PÚBLICAS, PESQUISA CIENTÍFICA

L'accès aux données statistiques localisées paraît, paradoxalement, plus facile au Brésil, pays en voie de développement, qu'en France, vieille terre de statistiques publiques. La tradition statistique héritée de la période coloniale portugaise a été maintenue et, malgré ses lourdeurs, l'IBGE (Institut brésilien de géographie et statistique) a pu jusqu'à présent fournir des données fines dont l'accès est libre, et facilité par le développement rapide d'Internet. Et d'autres sources, non censitaires, existent et sont largement ouvertes aux chercheurs par les organismes qui les produisent.

Le Brésil, un pays de tradition statistique

Cette tradition est sans doute un résultat positif du centralisme de l'État brésilien, hérité de la colonisation portugaise : il fallait dénombrer pour maîtriser un territoire immense et très peu densément occupé. La majeure partie des données disponibles à un niveau territorial fin (celui des communes, ou *municípios*) est produite par l'IBGE. Celui-ci a été créé en 1938, par la fusion du Conseil national de la statistique et du Conseil national de

la géographie, institués quelques années auparavant. Sa mission était de réunir et d'organiser les informations quantitatives spatialisées sur le territoire national, en utilisant les outils de la géographie moderne, notamment française. Sa tradition, liée à ses origines les plus lointaines, était fondée sur les expéditions scientifiques de terrain : délimitation des frontières litigieuses entre les États, reconnaissances dans le Mato Grosso, dans la vallée du Rio Doce, relevés sur le site de la future capitale, Brasília. À la confection des cartes de base et à l'organisation des recensements, premières missions de l'IBGE, s'ajoutèrent l'étude des systèmes agricoles, de l'urbanisation et de l'industrialisation, ce qui impliqua de grands progrès méthodologiques.

Dans le domaine de la statistique économique, l'IBGE réalise les recensements démographiques et économiques, mène les enquêtes sectorielles, élabore les indices de prix, construit les principaux indicateurs sociaux officiels. Il s'agit donc du principal producteur de statistiques du pays, statistiques de référence utilisées par tous les acteurs économiques et sociaux. On pourrait dire, par analogie avec les organismes français œuvrant dans le même domaine, que l'IBGE regroupe les missions de l'IGN et de l'INSEE, sans les études économiques destinées à la planification, lesquelles reviennent à d'autres organismes comme l'Institut de recherche économique appliquée, IPEA, ou la Fondation Getúlio Vargas.

En raison de la place qu'elles occupent à la croisée des compétences statistiques et cartographiques, les géosciences tiennent une place de choix à l'IBGE. Sous une direction spécifique, sont regroupées toutes les activités relatives au territoire : géodésie, cartographie de base, ressources naturelles, structures territoriales et géographie. La mission attribuée au département de géographie est de « construire les informations de nature géographique nécessaires à la planification du territoire en réponse aux demandes de la société et touchant à l'organisation de l'espace ainsi que les transformations de la maille politico-administrative du pays ». Les géographes de l'IBGE sont donc très impliqués dans les études touchant aux divisions du territoire, divisions officielles comme celle des États de la Fédération, ou opérationnelle comme les « micro-régions », ou les régions polarisées par les villes.

Les « micro-régions » sont un bel exemple de l'action des géographes dans l'IBGE : ils avaient délimité les « micro-régions

homogènes » (fig. 1) dans les années 1960, à une époque où l'on attachait une grande importance à la notion de paysage et d'unité naturelle. Et trente ans plus tard, c'est à nouveau aux géographes que l'on a demandé de dresser la carte des « micro-régions géographiques » (fig. 2), souvent polarisées par une ville ou relativement homogènes sur le plan économique (et non pas seulement physique), afin de publier les données du recensement démographique de 1991 sur cette base, qui n'a pourtant aucune existence administrative. Dans un pays où l'organisation territoriale est en perpétuelle mutation, on comprend l'intérêt de ces travaux : ils fournissent une base opérationnelle à tout projet devant avoir un impact sur le territoire ; une belle réussite pour les géographes de l'IBGE !

Les gisements de données disponibles à l'IBGE

Le stock de données est très fourni.

- Des recensements démographiques, menés depuis 1940 tous les dix ans, permettent de percevoir l'évolution du peuplement ; le dernier a eu lieu en 1991.
- Des recensements agricoles, fournissant des informations sur les productions, les tailles des exploitations, les modes de faire-valoir, les financements, l'équipement, etc. ; ils ont lieu tous les cinq ans ; le dernier date de 1996.
- Des enquêtes annuelles de production des cultures et de l'élevage (superficie, volume et valeur de plus de 70 productions) de 1977 à 1992, collectées par un réseau d'informateurs présents dans tout le pays.
- Les recensements économiques, quinquennaux eux aussi, couvrent les domaines de l'industrie, du commerce et des services ; plus précisément, ils donnent le nombre et la taille (en valeur de la production et en travailleurs) des établissements de chaque *município* ; ils sont malheureusement interrompus depuis 1985.
- Des enquêtes thématiques, sans périodicité particulière, éclairent certains sujets méconnus, comme les équipements de santé, les productions de l'« extraction végétale » (cueillette), l'infrastructure médicale, les moyens d'assainissement, etc.

Traditionnellement, les recensements de la population et des logements avaient lieu tous les 10 ans, aux années dont le numéro s'achève par le chiffre 0. Mais en 1990, l'IBGE, sous la présidence de Fernando Collor de Melo, n'a pu entreprendre la réalisation du recensement et l'opération a été réalisée en



Fig. 1.— Les microrégions homogènes des années 1960 au Brésil.



Fig. 2.— Les microrégions géographiques des années 1990 au Brésil.

1991, mais la publication des données statistiques habituelles a tardé, pour des raisons d'ordre technique liées à la mise en œuvre du logiciel de critique des données (vérification de la cohérence des réponses enregistrées par les agents recenseurs). Les principales données de cadrage (le nombre d'habitants selon le sexe et la localisation urbaine ou rurale, ainsi que le nombre de logements) n'ont été rendues publiques qu'en 1992. En 1994, les premiers tableaux basés sur les réponses aux questions composant le formulaire 1.01 du recensement démographique ont commencé à être disponibles. Malheureusement, un fâcheux retard portant sur les données de l'État amazonien du Pará a interdit toute analyse de l'ensemble national jusqu'au début de l'année 1995. Mais la « saga » n'était pas terminée pour autant, puisqu'une grande partie du recensement de 1991 n'avait pas encore été exploitée, celle qui provenait du formulaire 1.02, permettant d'analyser les conditions de vie. Et c'est seulement dans le courant de l'année 1998 que la publication de ces informations par l'IBGE a été achevée.

Comme on l'a compris, l'IBGE est une énorme organisation, une machine lourde face à laquelle l'utilisateur de base comme le chercheur le plus averti se trouvait, encore dans les années 1985-1995, assez dépourvu. De plus, l'accès aux données

était dépendant d'un système informatique lourd que seuls quelques spécialistes maîtrisaient et, pour obtenir un ensemble de données particulier, il fallait en quelque sorte obtenir leur faveur, ou bien connaître une personne qui connaissait quelqu'un disposant de l'indispensable sésame. L'impossibilité pour l'IBGE de répondre à la demande sociale était devenue critique avec l'élection de Fernando Collor de Melo à la présidence de la République en 1989, celui-ci n'hésitant pas à couper l'essentiel des moyens financiers de l'IBGE au nom d'un hypothétique équilibre budgétaire. Et il fallut attendre l'arrivée à la présidence de l'IBGE de Simon Schwartzman, économiste et professeur d'université (nommé par Fernando Henrique Cardoso, lui-même élu président de la République en 1994), pour que les choses changent. Il était temps ! Au cours de la Conférence nationale des usagers du système statistique brésilien (Confège-Confest) qui s'est tenue à Rio de Janeiro à la fin de l'année 1995, nombre de participants, au premier rang desquels l'Association brésilienne pour les études de population (ABEP), avaient souligné les manquements de l'IBGE, certains évoquant même son éventuel démantèlement.

Soutenu par le gouvernement, Simon Schwartzman réussit le tour de force de remettre l'IBGE sur les rails. Décision

fut prise de réaliser un comptage rapide de la population afin de pallier les insuffisances de l'exploitation du recensement démographique de 1991, autoriser l'exploitation par les chercheurs des microdonnées du recensement démographique de 1991, réaliser un recensement agropastoral, utiliser Internet comme vecteur de diffusion des données.

L'accès aux données individuelles

Quel chercheur n'a rêvé d'avoir accès aux fiches individuelles de recensement? En leur permettant d'accéder aux *microdados* (littéralement microdonnées) l'IBGE a, en partie, exaucé ce vœu. L'effort consenti par le pays et la communauté internationale pour constituer une base d'informations fiables et exhaustives (il faut rappeler que la réalisation d'un recensement est une opération très lourde et très coûteuse, en particulier dans un pays aussi vaste et diversifié que le Brésil) impose qu'aucune « miette » des données obtenues à grands frais n'échappe à la sagacité des chercheurs de toutes disciplines. La présidence de l'IBGE l'a bien compris, et a ouvert largement ses fichiers aux chercheurs, politique qui contraste avec une culture du « secret statistique » qui prévaut dans de nombreux pays.

Ces *microdados*, fournis sur 10 *cédéroms*, proviennent d'un échantillon (dont la taille est variable en fonction des caractéristiques de chaque *município*) comprenant 17 millions d'enregistrements séquentiels. Les enregistrements de type logement ou individu renferment des valeurs numériques et des codes alphanumériques correspondant à un dictionnaire fourni. Chaque enregistrement est doté d'un poids (facteur d'expansion) qui permet d'estimer des tableaux statistiques pour l'ensemble de la population, du *município* au pays entier. Les opérations de construction des tableaux peuvent être réalisées avec les principaux logiciels spécialisés (SAS, Stata, SPSS, etc.). Nous avons préféré réaliser un logiciel spécifique (dénommé *MicroDados*, micro-ordinateur Apple Power Macintosh), adapté à la cartographie, et permettant de réaliser des tableaux croisant une dimension géographique (*município*, microrégion, unité de la fédération) avec une ou deux variables relatives à la population et au logement.

La révolution Internet

Le site Web de l'IBGE constitue un exemple remarquable de l'utilité d'Internet dans le domaine de la mise

à la disposition du public d'informations statistiques publiques. Lorsqu'on se connecte à certains sites d'instituts de statistique (dont, par charité, on taira le nom), on nous propose le plus souvent une vitrine plus ou moins joliment agrémentée, comprenant un catalogue de produits qui, au mieux, sont accessibles en téléachat, ainsi que quelques échantillons de données destinés à aguicher le chaland! Le site de l'IBGE (www.ibge.gov.br) prend le contre-pied de cette parade mercantile. Conscient de son rôle citoyen, l'IBGE permet l'accès aux données les plus récentes, et autorise, de plus, leur téléchargement.

La mise à disposition du recensement agropastoral de 1996 est un modèle du genre. Tous les tableaux ont été mis en ligne au fur et à mesure de l'exploitation des questionnaires dans chacun des États. Aujourd'hui, les données couvrant l'ensemble du pays (par *município*) et la documentation y afférant sont accessibles librement (pas de mot de passe) et gratuitement (pas d'autre coût que la communication téléphonique avec le fournisseur d'accès). On pourra critiquer telle ou telle limitation du logiciel d'interrogation, la lenteur ou la fiabilité du serveur, c'est à une véritable révolution que Simon Schwartzman a amené son institut.

Qu'en sera-t-il du Recensement démographique prévu pour l'an 2000? Déjà des craintes se font jour : aura-t-il lieu effectivement en 2000, ou un peu plus tard? Son exploitation statistique sera-t-elle réalisée dans des délais raisonnables? Le plan de redressement économique lancé par Fernando Henrique Cardoso à partir de 1993 (et qui pris en 1994 le nom de Plan Real) ayant été bousculé au début de l'année 1999 en raison d'une forte dévaluation de la monnaie, il est désormais plus difficile à l'État brésilien de trouver les ressources financières nécessaires à la réalisation de grands projets. L'IBGE connaîtra-t-il une nouvelle période d'incapacité à remplir ses missions? Heureusement, d'autres sources existent.

De nombreuses sources de données non censitaires largement ouvertes

Au Brésil comme ailleurs, les chercheurs sont de plus en plus souvent amenés à commencer leur travail par la détection de sources plus « fraîches » et plus fines que les sources officielles, ce que l'on appelle parfois le *data harvesting* ou *data mining*. Il arrive même, pour reprendre une expression du géographe britannique John Shepherd, que

cette activité relève plutôt du *garbage recycling*, le recyclage des « déchets » ou au mieux des « sous-produits » que sont, pour de grandes administrations, les données qu'elles produisent inévitablement en menant leur activité principale, qui peut être de soigner la population, de percevoir l'impôt ou de distribuer le courrier.

Le Brésil possède en effet un appareil administratif très développé et couvrant l'ensemble du pays. On dispose de ce fait de bases de données constituées par les organismes en charge d'un domaine particulier (Fondation nationale de la santé, ministère de la Réforme agraire, ministère du Travail, etc.) qui, en général, sont largement ouvertes aux chercheurs désirant évaluer l'impact de la politique sociale, et cela depuis une dizaine d'années.

Ces organismes fournissent leurs données sur cédérom, et le plus souvent gratuitement. Les conditions d'accès sont réduites au minimum : lettre formelle de demande d'information, ou convention aux exigences réduites. Pour répondre au souci de rendre leurs données facilement exploitables, ils fournissent gracieusement un logiciel, vrai modèle de « convivialité » et permettent aux chercheurs en sciences sociales de construire eux-mêmes des tableaux parfois complexes avec une réelle facilité et une grande puissance de traitement. Dans tous les cas, le fournisseur demande qu'on lui retourne les résultats obtenus à l'aide de ses données. Il en est ainsi, par exemple, du ministère du Travail qui diffuse un Rapport annuel d'informations sociales (RAIS) et un Fichier général des employés (CAGED). Pour sa part, la Fondation nationale de la santé assure le suivi de trois grands ensembles de données : le fichier des autorisations d'admission hospitalières (AIH), celui des soins ambulatoires (AIA) et le système d'information sur la mortalité (SIM).

Cette dernière source mérite que l'on insiste un peu. En 1944, le Service fédéral de biostatistique publiait déjà un annuaire des causes de décès intervenus entre 1929 et 1932 dans les capitales des États de la Fédération. Mais ce n'est qu'en 1975 que furent jetées les bases d'un système de surveillance épidémiologique national (SIS) permettant une collecte régulière d'informations fiables sur la mortalité. L'adoption d'un formulaire de déclaration type, valable dans tout le pays, et destiné à remplacer les quelque 40 types d'attestations alors en vigueur, fut une

étape significative de la mise en place du SIS. Le sous-système d'information sur la mortalité (SIM) fut alors implanté dans tout le pays avec l'appui informatique de la Fondation pour les services de santé publique (FSESP) créée en 1977 par le ministère de la Santé. C'est finalement la Fondation nationale de la santé (FNS) qui gère depuis 1992 le système micro-informatique implanté dans chaque secrétariat d'État à la Santé alimentant le SIM.

La collecte des données du SIM s'appuie sur les déclarations de décès (en théorie obligatoires pour procéder à toute inhumation) faites dans les bureaux d'état civil (*Cartório de Registro Civil*) à l'aide d'un formulaire unifié en deux exemplaires. Le premier est ramassé périodiquement dans ces bureaux par les services locaux chargés des statistiques de santé. Cette source de base est complétée, lorsque l'état civil est déficient, par d'autres informations provenant des hôpitaux ou des cimetières. Les gestionnaires du SIM estiment qu'environ 80 % du total des décès sont enregistrés par ce système. Ils ajoutent que le sous-enregistrement est patent dans les régions Nord, Nord-Est et Centre-Ouest. Les fichiers fournissent, pour chaque décès de nombreuses informations, sur le décès lui-même (date, lieu, cause), et sur le décédé (sexe, âge, lieu de résidence, degré d'instruction). Pour l'année 1991, on trouve 803 835 décès pour l'ensemble du Brésil, soit un peu moins que le chiffre avancé par l'IBGE, 839 819 (4,3 % de différence). Une exploitation du fichier montre que, selon que l'on s'intéresse au sexe, à l'âge, au lieu de résidence, ou à la cause de décès, le nombre d'enregistrements utilisables décroît fortement pour atteindre, toutes lacunes cumulées, le nombre de 161 236, soit 20 % du total.

La situation d'ensemble est donc favorable et permet de suivre de façon fine les transformations territoriales que connaît encore le Brésil. Un problème se pose toutefois, la disponibilité des données anciennes : autant il est possible, voire facile, de se procurer des données récentes, autant il est souvent impossible de remettre la main sur des données d'il y a dix ou vingt ans. Le cédérom Samba 2000 réalisé par l'Orstom en coopération avec l'IBGE est désormais le seul moyen d'accéder aux données municipales du recensement agricole de 1985 à l'échelle des *municípios*, si bien que les chercheurs de l'IBGE s'en servent à usage interne. De l'intérêt, entre autres, de la coopération française avec le Brésil, même si ce n'était pas, à l'origine, le but visé...