

LES MICRODADOS DU RECENSEMENT DÉMOGRAPHIQUE DE 1991

Philippe WANIEZ*

Quel chercheur n'a rêvé d'avoir accès aux fiches individuelles de recensement ? En nous permettant d'accéder aux microdados (littéralement micro-données) l'IBGE a, en partie, exaucé ce vœu ! Pour comprendre précisément de quoi il s'agit, un petit rappel de l'histoire du recensement de 1991 s'avère nécessaire.

Les vicissitudes du recensement démographique de 1991 au Brésil

De manière traditionnelle, les recensements de la population et des logements ont lieu, au Brésil, tous les 10 ans, aux années dont le numéro s'achève par le chiffre 0 (1960, 1970, 1980, etc.). En 1990, le Brésil, sous la présidence de Fernando Collor de Mello n'a pu entreprendre la réalisation du recensement. Finalement, l'opération a été réalisée en 1991, mais les données statistiques habituellement publiées par l'IBGE tardent à être produites 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). Finalement, nos relations scientifiques établies depuis plusieurs années avec le Département de Géographie de l'IBGE, nous ont permis d'obtenir le dénombrement en 1992 ; il s'agissait seulement de données de cadrage indiquant le nombre d'habitants selon le sexe et la localisation urbaine ou rurale, ainsi que le nombre de logements : des données utiles, mais rien qui permettent une étude solide de la dynamique du territoire brésilien.

En 1994, les premiers tableaux basés sur les réponses aux questions composant le formulaire 1.01 du recensement démographique commencèrent à être disponibles. Malheureusement, un fâcheux retard portant sur les données de l'État du Pará a interdit toute analyse de l'ensemble national jusqu'au début de l'année 1995. Et c'est seulement dans le courant de la même année que ces données furent intégrées par nos soins au Système pour l'Analyse des *Municípios* Brésiliens pour l'An 2000 (Samba 2000) et

* Institut de Recherche pour le Développement, Paris.

utilisées dans différents projets de recherche, notamment à la Maison de la Géographie de Montpellier (GIP RECLUS) et au CREDAL.

Mais la “saga” n’était pas terminée pour autant puisqu’une partie importante du recensement de 1991 n’avait pas encore été exploitée, celle composant le formulaire 1.02 permettant d’analyser les conditions de vie. À la fin de l’année 1996, dans le cadre d’une convention passée entre l’IBGE et l’Orstom, j’ai eu la possibilité d’accéder aux fiches individuelles du formulaire 1.02 communément appelées *microdados*. Malheureusement, la réalisation de tableaux statistiques à partir de ces fiches imposait le choix et la mise au point d’une méthodologie de traitement informatique, et l’ajout de ces données à Samba 2000 reste à faire.

Ainsi, de manière paradoxale, les résultats présentés ici s’appuient sur des données élaborées très récemment, mais portant sur un recensement qui a déjà sept ans d’ancienneté à l’heure où sont écrites ces lignes. On pourrait dire “à quoi bon analyser des données si anciennes sachant qu’au Brésil, les transformations économiques et sociales sont extrêmement rapides et massives, et s’accompagnent de bouleversements territoriaux parfois impressionnants” ? Plusieurs arguments peuvent être présentés pour justifier l’intérêt scientifique d’entreprendre et approfondir l’étude de ces données.

En premier lieu, elles offrent une meilleure “résolution” géographique que les Enquêtes Nationales par Échantillon de Domiciles (PNAD) habituellement utilisées pour évaluer les conditions de vie, ces enquêtes ne donnant malheureusement pas une vue détaillée du territoire puisque les niveaux habituellement utilisés par les géographes (*município* et *microrégion*) n’y sont pas disponibles, en raison même de la méthode d’échantillonnage utilisée par l’IBGE.

Le second argument en faveur de la démarche présentée ici réside dans le fait que, même si le recensement démographique de l’An 2000 a bien lieu cette année là (c’est-à-dire si les coupes budgétaires que ne manqueront pas de subir les organismes fédéraux ne touchent pas de plein fouet cette opération), les tableaux portant sur les conditions de vie seront au mieux disponibles en 2003-2004 (à en juger par notre expérience en la matière). Il nous a semblé préférable de ne pas attendre trois à quatre années de plus pour entreprendre ces recherches. Dans la même perspective, on peut considérer que les résultats présentés ici constituent une marque dans le temps, un point de repère spatio-temporel qui facilitera, au début du XXI^e siècle, l’évaluation des transformations au cours de la décennie qui s’achève.

Enfin, l'un des objectifs du présent dossier est de montrer le parti que l'on peut tirer du croisement des réponses à différentes questions. La recherche d'Hervé Théry qui met en œuvre cette méthode en croisant la couleur de peau avec les niveaux de revenus ouvre cette piste de recherche et apporte quelques éléments nouveaux à un vieux débat.

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. L'effort consenti par le pays et la communauté internationale pour offrir au public une base d'informations fiables et exhaustives impose qu'aucune "miette" des données obtenues à grand frais n'échappe à la sagacité des chercheurs de toutes disciplines. La Présidence de l'IBGE l'a bien compris en ouvrant largement ses fichiers aux chercheurs, politique qui contraste avec une culture du "secret statistique" qui prévaut dans de nombreux pays.

Des questionnaires aux tableaux statistiques régionaux

Les institutions brésiliennes qui diffusent des fichiers de données résultant de leur fonctionnement les fournissent en général sur CD-Rom. C'est le cas, par exemple du Ministère du Travail (fichiers Rais et Caged) ou de la Fondation Nationale de la Santé (fichiers Aih Aia). Ces deux organismes ont en commun le souci de rendre leurs données exploitables et fournissent gracieusement un logiciel destiné à réaliser des tableaux statistiques. Ainsi, les logiciels SGT et TabWin sont des modèles de "convivialité" qui 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 le cas des *microdados* du recensement démographique de 1991, l'IBGE s'est contenté de copier les fichiers sur 10 CD-Rom, sans y joindre un logiciel de tabulation, laissant au chercheur le soin de se procurer les moyens de traitement adéquats, ou de sous-traiter cette opération auprès de sociétés de service. Ce travail est d'autant moins facile à réaliser par un non-spécialiste que les *microdados* proviennent d'un échantillon (dont la taille est variable en fonction des caractéristiques de chaque *município*) et qu'il nécessite quelques connaissances —élémentaires mais non nulles— de l'estimation statistique.

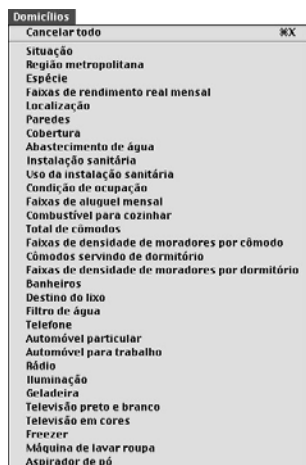


Figure nº1

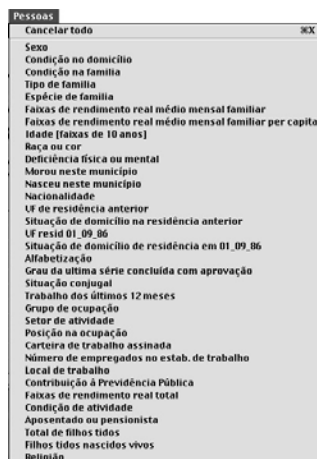


Figure nº2



Figure nº4



Figure nº3



Figure nº5



Figure nº6

Pour combler cette lacune, un logiciel de tabulation simple, permettant de réaliser des tris croisés de profondeur 2 ou 3 a été conçu afin d'élaborer des tableaux statistiques croisant une dimension géographique (*município*, micro-région, unité de la fédération) avec une ou deux variables relatives à la population et au logement. C'est ce logiciel, dénommé simplement MicroDados, fonctionnant exclusivement sur micro-ordinateur Apple Macintosh® ou Power Macintosh®, qui a permis d'élaborer l'ensemble des statistiques exploitées par les articles qui composent le présent dossier (sauf l'Indice de Développement Humain). Pour fonctionner, ce logiciel exige que les données soient stockées sur un disque magnétique d'une capacité de 5 gigaoctets, et doit disposer de 70 mégaoctets de mémoire vive.

L'interface du logiciel est simplifiée à l'extrême. Elle se compose des 6 menus suivants :

- *Arquivo* permet de demander l'exécution d'une tabulation ou de quitter le logiciel ;
- *Domicílios* (figure n°1) et *Pessoas* (figure n°2) renferment la liste des questions du recensement sur lesquelles des tableaux peuvent être réalisés (il s'agit d'un sous-ensemble de ces questions choisi en fonction de nos domaines d'intérêt et du nombre de réponses possible, ce nombre déterminant pour une grande part l'espace mémoire requis) ; pour un niveau géographique donné, il est possible de croiser une ou deux variables relatives aux logement, une ou deux variables relatives aux personnes, ou bien encore une variable portant sur le logement et une autre sur les personnes ;
- les deux menus suivants, *Regiões* (figure n°3) et *UF* (figure n°4) permettent de choisir l'espace géographique, régions ou États, pour lequel un tableau doit être produit. Il est naturellement possible de choisir plusieurs régions, plusieurs États (même s'il ne sont pas dans la même région) ou le Brésil tout entier ;
- enfin, le menu *Níveis* (figure n°5) permet d'indiquer à quel niveau géographique (État, micro-région ou *município*) le tableau doit être calculé ; par défaut, les tableaux correspondant à ces trois niveaux sont réalisés.

Les résultats sont enregistrés dans un fichier texte dont les champs sont séparés par des tabulations : chaque case du tableau croisé forme une variable identifiée par les modalités de réponse à laquelle elle correspond. Le temps de traitement dépend du nombre d'observations, et va de quelques minutes pour un petit État de la Région Nord (figure n°6), à plusieurs heures pour l'ensemble du Brésil.