

Russie, octobre 2002

Un recensement de transition

Le recensement de la population russe qui s'est déroulé du 9 au 16 octobre 2002 a été symbolique à plusieurs égards. Premier recensement démographique réalisé après l'éclatement de l'URSS en décembre 1991¹, il devait, entre autres objectifs, apporter une réponse à une question lancinante pour les Russes depuis fin 1991 : combien sommes-nous ? Cette interrogation fit le titre de nombreux articles de presse consacrés à la préparation du recensement. Depuis 1991, la Russie a connu de forts flux migratoires internes et externes, une baisse de la natalité et une forte augmentation de la mortalité². En 1992, l'accroissement naturel est devenu négatif, posant de manière d'autant plus aiguë la question de l'immigration, en particulier de citoyens en provenance des ex-républiques soviétiques d'Asie centrale jadis considérés comme soviétiques et donc non migrants³. Ces bouleversements ont entretenu une inquiétude démographique et politique : aujourd'hui, quelle est la puissance de la Russie seule, de ce pays dont le territoire

comptait 147,4 millions d'habitants lors du recensement précédent de 1989 ? Les chiffres de 2002 allaient-ils confirmer la catastrophe démographique annoncée par certains ? Autre inconnue, quelle a été l'ampleur réelle des flux migratoires engendrés sur le territoire de la Fédération de Russie par l'éclatement de l'URSS ? Et comment maintenant se répartit la population en Russie ? Le Président du Goskomstat, Comité d'État de la statistique⁴, Vladimir Sokoline, l'avouait lui-même dans l'attente des résultats : « *Nous ne savons pas où vit notre population.* »⁵

Les incertitudes sur la situation démographique ont donc dominé l'organisation de ce premier recensement de la nouvelle Fédération de Russie. Ses résultats étaient d'autant plus attendus que la loi du 15 novembre 1997 sur le système d'enregistrement de l'état civil a contribué à appauvrir les données chiffrées disponibles au sujet de la population⁶. Beaucoup d'informations qui jusque-là servaient de base

à la statistique courante de la population ne furent plus demandées aux personnes. De plus, la transmission aux organes de la statistique d'État du feuillet détachable appelé « deuxième exemplaire des actes d'enregistrement de l'état civil » fut supprimée en raison du caractère confidentiel des renseignements portés sur ces documents⁷. Statisticiens et démographes russes expriment régulièrement leur dépit face à cette décision de l'État russe, qui les a brutalement privés d'une source de données régulière, et facile d'accès, qui existait depuis les années 1920. Ils ont été privés en particulier d'informations qui permettaient d'établir une statistique des naissances, des mariages, des divorces, des décès et causes de décès, par catégories sociales et nationales, et de répartir les migrants selon la taille et le type de famille⁸. Le recensement est donc devenu la principale source d'obtention de ces données.

Une autre inquiétude a dominé la préparation du recensement de 2002 : quelle allait être la motivation de la population à répondre ? En effet, l'organisation de ce premier recensement de la période post-soviétique a été marquée par la prise en compte d'une nouvelle réalité en Russie, le respect des droits individuels garantis par la loi. Pour la première fois, les citoyens russes ne furent pas obligés par la loi de répondre au recensement, la participation se fit sur la base du libre choix de chaque individu⁹. Dans un pays habitué depuis 1926 à un fort enca-



ВНЕШНЯЯ ПЕРЕПИСЬ НАСЕЛЕНИЯ!

ВСЕРОССИЙСКАЯ ПЕРЕПИСЬ НАСЕЛЕНИЯ
с 9 по 16 октября 2002

Le logo et l'annonce du recensement

drement des opérations de collecte des données du recensement par les organisations du Parti, cette nouvelle disposition avait de quoi nourrir bien des inquiétudes¹⁰.

Il a fallu tenir compte de cette nouvelle situation pour la campagne d'information du recensement. Celle-ci s'est caractérisée notamment par l'effort du gouvernement pour intégrer ce recensement dans un discours renouant avec le passé russe. Ainsi, son annonce officielle mit en avant l'idée que « pour la première fois dans l'histoire de la nouvelle Russie aurait lieu, du 9 au 16 octobre 2002, le recensement général de la population de la Russie ». Elle précisait aussi qu'il s'agissait du deuxième recensement démographique effectué par l'État russe depuis 1897, date du premier recensement de la population de ce qui était pourtant à l'époque l'Empire russe et pas seulement la Russie circonscrite à son territoire actuel. Instrument de l'affirmation de la légitimité du nouvel État né en décembre 1991, le recensement de 2002 a été présenté officiellement comme un symbole de la construction de l'unité nationale et de la continuité institutionnelle avec la Russie d'avant octobre 1917, comme un des signes d'effacement de la période soviétique. Chaque citoyen a été appelé à « inscrire son nom dans l'histoire de la Russie », slogan du recensement, et un document sur l'histoire des recensements en Russie a été joint au dossier de presse. Enfin, une médaille a été émise par l'État pour récompenser une participation active au recensement, sur le modèle de celle qui avait été créée par le tsar Nicolas II « pour les travaux du premier recensement général de la population de 1897 », afin de remercier les agents recenseurs bénévoles. Celle décernée par le président Vladimir V. Poutine a récompensé les employés du Goskomstat jugés les plus méritants, à Moscou et en province, mais aussi des journalistes dont le travail d'information sur le recensement a été estimé particulièrement utile pour la réussite de celui-ci¹¹.

Si le recensement de 2002 a été ancré symboliquement dans le long terme

de l'histoire de la Russie, en revanche les principes de son organisation furent, pour la première fois, officiellement rattachés aux recommandations énoncées par la résolution 1995/7 du Conseil économique et social de l'ONU (Organisation des Nations unies), adoptée le 19 juillet 1995¹². L'État russe a affiché ainsi clairement sa volonté de retour dans l'internationalisme statistique.

En réalité, il s'est agi d'un recensement de transition, de différents points de vue.

Changements et continuité

Tout d'abord, l'histoire du recensement d'octobre 2002 a été marquée par les transformations économiques et sociales liées au passage à l'économie de marché, mais aussi par les réformes des cadres législatif et juridique de l'État russe, notamment en ce qui concerne la reconnaissance des droits des individus. Elle a porté également l'empreinte des difficultés financières du nouvel État. Programmée dans un premier temps pour l'année 1999, la réalisation du recensement fut retardée pour des raisons budgétaires. Décidée par le gouvernement deux mois après le micro-recensement effectué en février 1994 auprès de 5 % de la population, son organisation fut freinée après la crise financière d'août 1998 et la démission du gouverne-



Le tsar Nicolas II.
Détail d'un tableau
d'Ernest Karlovitch von Liphart.

ment. Enfin, un arrêté du nouveau gouvernement (22 septembre 1999) annonça que la collecte serait réalisée courant 2002¹³. Un autre suivit, le 12 avril 2000, pour fixer le déroulement des opérations¹⁴, et un recensement test fut effectué du 11 au 18 octobre 2000. La date de référence fut fixée au 9 octobre à 0 heure, les formulaires devant être remplis et collectés entre le 9 et le 16 octobre.

Premier bilan statistique à grande échelle des principales transformations démographiques et sociales de la Russie depuis l'éclatement de l'URSS, ce recensement se caractérise également par le fait qu'il a été réalisé dans le contexte de passage du système statistique soviétique à un autre, plus proche des normes internationales. Cela s'est traduit par un certain nombre de modifications méthodologiques, sur fond de continuités institutionnelles et de discussions entre statisticiens sur l'opportunité des choix effectués¹⁵.

Un dispositif institutionnel centralisé

Le dispositif institutionnel de préparation et d'organisation du recensement de 2002 rappelle celui qui était mis en œuvre pendant la période soviétique : s'il s'agit toujours d'une opération fortement centralisée, l'habitude de faire appel aux autorités et administrations locales pour organiser toute la partie opérationnelle sur le terrain a été conservée. Comme par le passé, les autorités locales ont été chargées de faciliter tout le travail de vérification des noms de rues, de numérotation des logements et des registres de la population, d'aider à l'organisation du recrutement du personnel, de fournir les locaux et moyens de transport nécessaires¹⁶.

Une commission pour la réalisation du recensement a été créée dans chaque municipalité et chaque région ou république de la Fédération. À la tête du dispositif d'ensemble, un organe fédéral, la commission d'État pour la réalisation du recensement de 2002, a été chargé d'arrêter le pro-

gramme et les questionnaires, de veiller à la coopération entre les organes de pouvoir exécutif fédéraux et leurs homologues dans les 89 « sujets de la Fédération »¹⁷, de contrôler l'exécution des mesures décidées et de décider des dépenses budgétaires.

Soumis à la tutelle administrative de la commission d'État, le Goskomstat organise et supervise le dispositif de collecte et de traitement des données sur l'ensemble du territoire de la Fédération (17 075 400 km²), subdivisé à cette fin en 400 000 circonscriptions de recensement, regroupées elles-mêmes en près de 90 000 circonscriptions d'instruction formant 11 000 districts de recensement. Après différents contrôles effectués dans les circonscriptions, les formulaires devaient remonter au comité de statistique de chacun des 89 « sujets de la Fédération », qui, après de nouvelles vérifications et corrections, devait les transmettre au Centre d'exploitation des données du recensement du Goskomstat.

Modification de l'unité d'observation

Tout en s'inspirant largement de l'expérience de la pratique des recensements soviétiques antérieurs, organisés de manière décennale¹⁸, la méthodologie de celui de 2002 se caractérise par l'introduction d'un certain nombre de modifications liées aux informations à collecter mais aussi au nouveau système d'organisation de la statistique de l'État russe, notamment aux efforts destinés à homogénéiser ses procédures d'enquête dans différents domaines, et à la volonté du Goskomstat de se conformer à un certain nombre de normes statistiques fixées par des organismes internationaux, l'ONU et le BIT (Bureau international du travail) en particulier.

Par exemple, alors que, de 1920 à 1989, l'unité d'observation des recensements soviétiques avait toujours été la famille, considérée comme unité biologique, mais aussi comme unité de budget, le ménage a



Kuzma Petrov-Vodkin. L'année 1918 à Petrograd.

été choisi comme unité de base en 2002, conformément aux usages statistiques internationaux. Cela représente une réelle rupture par rapport au choix très nettement affirmé dès 1918, et repris à partir du recensement général de 1920, de retenir la famille et non pas le ménage comme unité d'observation. Lors du recensement de la ville de Petrograd en 1918, Grigori S. Polliak, qui prit ensuite une part active dans la préparation des recensements de 1920, 1926 et 1937¹⁹, légitimait ce choix pour la Russie post-révolutionnaire de la manière suivante : « *Mettre à l'ordre du jour, à l'échelle d'un recensement de la population, l'étude de la famille comme un des facteurs les plus importants, et peu étudiés jusque-là, de la vie sociale, représente une nouveauté qu'il est impossible de ne pas saluer. Un des défauts essentiels des recensements de la population du type de ceux qui sont effectués en Europe de l'Ouest est, à nos yeux, ce principe étroit, "atomistique", qui est posé comme base de leur exploitation. L'objet des diverses classifications utilisées dans cette exploitation des données est invariablement une personne individuelle, et non cette cellule sociale naturelle à laquelle cet individu appartient par sa naissance et son*

éducation. Les liens internes qui tirent un être humain vers les rangs de tel ou tel groupe social, en vertu du fait de son appartenance à une famille déterminée, demeurent, dans ce système de traitement des données, en dehors du champ visuel du chercheur. On ne peut nier le fait, par exemple, qu'une sérieuse différence existe entre, d'une part, le fait d'être dans la position sociale d'un ouvrier ou d'un apprenti qui appartient à la famille d'un patron indépendant, et, d'autre part, le fait d'être dans la position d'un ouvrier qui vient d'un milieu ouvrier lui-même. »²⁰

L'usage de la famille comme unité d'observation demeura une caractéristique de tous les recensements soviétiques jusqu'en 1989. L'approche par ménage fut adoptée pour la première fois lors du micro-recensement de 1994, puis en 1997 lors du remaniement des enquêtes sur les budgets des familles²¹. En fait, le choix de la famille ayant été légitimé dès le début des années 1920 dans la perspective d'une analyse en termes de classes sociales, l'abandon de cet usage ne fut réellement entériné que quand le terme de classes fut lui-même abandonné. Déjà envisagé et discuté lors du recensement de 1989, il ne fut

effectif qu'en 2002²². Actuellement, la définition du ménage est la même qu'en France : ensemble des personnes occupant un même logement à titre de résidence principale.

Qui recenser ?

Une autre innovation a été introduite en octobre 2002 : seule la population permanente a été prise en compte alors que les recensements précédents enregistraient l'ensemble de la population présente le jour du recensement. L'information collectée au sujet de celle-ci était exploitée en particulier pour traiter les questions d'approvisionnement, celle relative à la population permanente étant plutôt destinée à la planification de la construction de bâtiments divers²³.

Les individus ont donc été recensés sur la base de leur logement de résidence permanente, défini, dans les instructions, comme « le lieu d'habitation, maison, appartement, où la personne interrogée passe la plus grosse partie de son temps, de manière permanente (habituelle) »²⁴. Ont été inclus dans la population permanente « les citoyens de la Fédération de Russie, les citoyens étrangers et les personnes apatrides résidant sur le territoire de la Fédération de Russie à la date du recensement de la population, mais aussi les citoyens de la Fédération de Russie vivant de manière permanente sur le territoire de la Fédération mais absents temporairement de ce territoire à la date du recensement de la population »²⁵. Les Russes résidant à l'étranger n'ont pas été recensés, même s'ils se trouvaient temporairement en Russie au moment du recensement.



Le drapeau de l'ex-URSS

Le critère de « durée minimale de résidence continue » intervenant dans la définition de la population permanente a été objet de discussions : de 6 mois en 1989, cette durée est passée à 12 mois en 2002. Cette modification a correspondu à une nouvelle définition de la notion d'immigration dans le contexte des flux importants d'arrivées de personnes en provenance des anciennes républiques de l'URSS, en particulier d'Asie centrale. D'après la nouvelle manière de comptabiliser les anciens citoyens soviétiques devenus citoyens des nouveaux États indépendants issus de l'ex-URSS, près d'un million d'immigrés non russes sont arrivés en Russie, entre 1990 et 2000, depuis les pays d'Asie centrale²⁶. Contrairement à l'usage dans les recensements de la période soviétique, les migrants de ces pays-là qui arrivent en Russie au printemps pour travailler sur les chantiers de construction ou dans des structures commerciales, et restent au-delà de six mois, mais moins de douze, ne furent pas comptés dans la population permanente en octobre 2002. Ces anciens citoyens soviétiques sont devenus des étrangers pour le nouvel État russe.

Cette traduction statistique de l'instauration des nouvelles frontières a conduit à effacer une catégorie de migrants spécifique, héritée des pratiques de migration temporaire entre les anciennes républiques soviétiques d'Asie centrale et la nation centre de l'Empire soviétique. Valeri Tichkov, directeur de l'Institut d'ethnologie et d'anthropologie de l'Académie des sciences de Russie, s'est alarmé des conséquences de la nouvelle manière de compter cette population habituelle de travailleurs temporaires qui vont et viennent entre leur pays d'origine et la Russie selon les saisons mais qui restent suffisamment de temps pour participer à la vie économique et sociale russe : « Ainsi, il faudrait revenir au principe du recensement de 1989, ce qui donnerait la possibilité d'enregistrer la résidence dans le pays de pas moins d'un million de personnes qui sont encore en Russie à la mi-

octobre, après être arrivées au printemps pour effectuer un travail dans la construction ou le commerce. Sont concernés ici les individus qui arrivent comme immigrants dans le pays, dont le délai ultime de séjour permanent en Russie est fixé à moins de 12 mois. Leurs enfants vont dans les écoles russes, eux-mêmes travaillent en Russie, y compris dans des entreprises d'État. Dans ce cas, pourquoi ne sont-ils pas comptés dans la "catégorie de base" de la population du pays ? »²⁷

Selon V. Tichkov, le fait de ne plus prendre en compte la population présente non permanente conduit à ignorer un grand nombre de personnes « qui, dans leur grande majorité, deviendront tôt ou tard citoyens de la Russie et contribueront à compenser le déficit naturel de la population russe »²⁸. Visant à relativiser le discours catastrophiste sur la situation actuelle de la population en Russie, l'argument politique rejoint ici l'argument démographique.

Comment éviter les doubles comptes ?

Suite à la forte mobilité de la population depuis 1991, sur l'ensemble du territoire de la Fédération de Russie mais aussi à l'intérieur des régions et des villes elles-mêmes, une attention particulière a été apportée au risque de double compte, notamment dans le cas de l'usage de deux logements, l'appartement où une personne est enregistrée auprès de la police et un autre où elle vit²⁹. Cette situation a été rencontrée notamment dans le cas de jeunes ménages partis vivre dans un appartement en location, indépendant de celui de leurs parents, mais qui restaient enregistrés chez ces derniers, qui les ont déclarés comme vivant avec eux. Élément de base du dispositif de contrôle, toute personne présente de manière temporaire dans un logement devait indiquer son adresse permanente afin de faciliter la vérification de son enregistrement statistique sur son lieu de résidence permanente. Trois documents ont été utilisés pour repérer les cas de double compte. Tout d'abord, une attestation

de recensement a été délivrée systématiquement à chaque personne recensée afin qu'elle puisse la produire comme preuve à tout agent recenseur qui se présenterait à son deuxième domicile. En outre, une feuille de contrôle a été établie pour chaque personne recensée dans un autre endroit que son lieu de résidence permanente. Enfin, une fiche de contrôle individuelle a été également remplie pour affermir la vérification.

Le mode d'enquête adopté a renforcé la nécessité de l'usage de l'attestation de recensement. Devant l'inquiétude exprimée par beaucoup de citoyens au sujet du fait de devoir ouvrir leur porte à une personne inconnue, susceptible d'être un agresseur potentiel, le Goskomstat décida de proposer trois méthodes au choix pour se faire recenser : au domicile, par téléphone ou au bureau de recensement de sa circonscription. Dans les trois cas, la personne était interrogée par un agent recenseur. En réalité, la crainte de faire entrer un inconnu chez soi a caractérisé principalement les très grandes villes du pays, et surtout Moscou. Tandis que 40 % de la population recensée à Moscou s'est déplacée

dans les bureaux de recensement, cela n'a été le cas que de 10 % de la population recensée à Saint-Pétersbourg, et, en moyenne, de 1 à 2 % de la population recensée dans l'ensemble du pays³⁰.

Le principe du questionnement des individus en face-à-face par un agent recenseur a été reconduit, l'agent ayant pour consigne, comme par le passé, de noter les réponses fidèlement, sans rien changer aux termes employés par la personne interrogée. Tous les recensements depuis 1920 ont été effectués selon ce principe, bien que des tentatives aient été faites lors du recensement test de 1967 pour passer à la méthode d'auto-administration du questionnaire par la personne recensée. Cette année-là, 25 % des logements retenus après les tournées de vérification préliminaires avaient été recensés selon la méthode du questionnement en face-à-face et les autres logements selon la méthode d'auto-administration du questionnaire³¹. Les résultats du test poussèrent à laisser de côté cette méthode pour le recensement de 1970. La conférence des statisticiens soviétiques réunie à cette occasion conclut : « La méthode d'auto-administration du

questionnaire non seulement n'a pas produit d'économie des forces et des moyens, mais elle a aussi compliqué l'organisation des travaux du recensement... Une majorité écrasante des directions statistiques régionales qui ont participé à la réalisation du recensement test s'est exprimée contre la méthode d'auto-administration. »³²

Comment traiter les refus de réponse et les absents ?

Outre les mesures de contrôle déjà évoquées, les instructeurs-contrôleurs ont eu recours à deux méthodes pour recueillir les informations sur les absents et les réticents à répondre, l'interrogation des voisins et l'utilisation des données administratives : « Quelques jours avant la fin du recensement, quand il fut clair qu'il y avait des difficultés pour le réussir, une directive arriva du Goskomstat au sujet de la nécessité absolue de remplir les objectifs fixés (la copie de cette lettre électronique a été envoyée à toutes les circonscriptions de recensement et d'instruction). Dans la mesure où il n'était pas possible de remplir les objectifs fixés selon les moyens habituels, des mesures complémentaires ont été mises en œuvre. La remise en jeu fondamentale fut faite sur la base des données des DEZ³³ et des listes d'électeurs. »³⁴

Pratique ancienne, les registres des DEZ, bureaux chargés de l'entretien des bâtiments et des listes d'habitants des logements de chaque immeuble, furent mis à profit pour compléter les formulaires du recensement. La loi interdit à ces bureaux de fournir d'autres informations que le sexe et la date de naissance des individus. Toutefois, dans la pratique, ils donnèrent souvent d'autres informations manquantes sur les formulaires. Ce fut le cas notamment à Moscou³⁵.

Enfin, autre mesure de contrôle, une enquête par sondage a été effectuée après la fin du recensement, entre le 17 et le 21 octobre, par chaque instructeur-contrôleur, en compagnie de



Le bureau de recensement n° 4 à Saratov

l'agent recenseur concerné, auprès de 10 % des logements de chaque circonscription de recensement pour vérifier l'exactitude du dénombrement de la population³⁶. Les corrections éventuelles furent apportées ensuite sur l'ensemble des formulaires.

Trois questionnaires principaux

Le programme du recensement a reposé principalement sur trois questionnaires. Un bulletin individuel court (formulaire K, cf. pages 17 à 19) a servi de support au recensement exhaustif de la population. Il était constitué de questions à caractère socio-démographique essentiellement : relation de la personne interrogée avec les autres membres du ménage, sexe, âge, lieu de naissance, citoyenneté, nationalité, situation matrimoniale, formation, langue maternelle, langues parlées, sources des moyens de subsistance, existence d'une occupation procurant un revenu, statut de l'emploi. Le formulaire D (cf. pages 20 et 21) était destiné à une enquête complémentaire, effectuée auprès de 25 % de la population permanente et centrée sur les questions d'emploi : type d'activité, caractère du travail effectué, lieu de travail, activité donnant lieu à des déplacements géographiques, situation des personnes sans emploi. Un troisième formulaire était consacré aux conditions de logement : surface habitable, éléments de confort, date de construction, propriétaire ou locataire.

L'échantillon de l'enquête complémentaire a été constitué selon la méthode utilisée lors des recensements de 1979 et 1989. Le logement a été pris comme unité de sondage, qu'il s'agisse d'une maison individuelle, d'un appartement, ou encore d'une pièce dans un appartement communautaire ou un foyer, par exemple³⁷. Une base de sondage a été constituée dans chaque circonscription d'instruction à partir des listes des logements vérifiées par chaque agent recenseur dans sa propre circonscription d'enquête lors

de la tournée préliminaire avant le recensement. Les logements de l'échantillon ont été tirés au sort de manière aléatoire par les instructeurs-contrôleurs à l'échelle de chaque circonscription d'instruction, en suivant la règle d'un logement sur quatre.

Dans les « sujets de la Fédération » de moins de 500 000 habitants au recensement de 1989, de même que dans les régions difficiles d'accès, le questionnaire long (formulaire K + formulaire D) a été administré à l'ensemble de la population. Dans tout le reste du territoire, deux catégories de personnes n'eurent en revanche à remplir que le questionnaire court (formulaire K) : les personnes vivant dans des logements collectifs (par exemple internats, maisons de retraite, établissements pour invalides, casernes de l'armée ou de la police) et les citoyens russes en mission à l'étranger³⁸.

Une campagne d'information multiforme

Les individus n'étant pas obligés par la loi de se faire recenser, on ne pouvait donc compter que sur leur bonne volonté. Ceci explique l'attention particulière portée à la campagne d'information sur le recensement en direction de la population. Le texte

du décret du gouvernement donna le ton en faisant appel au sens civique de celle-ci : « *Prendre part au recensement général de la population russe est un devoir social pour l'individu et le citoyen.* » Cet appel fut relayé dans chaque région par la campagne d'information locale. Toutefois, il fut toujours précisé que refuser de se faire recenser ne donnerait pas lieu à des sanctions pénales : « *La Russie ne s'est pas engagée dans la voie de l'introduction de sanctions en cas de refus de se faire recenser. La seule incitation à répondre aux questions réside dans le fait que chaque citoyen comprend qu'il est une partie de la société qui a besoin de connaissances sur elle-même.* »³⁹

Différents moyens furent mis en œuvre pour sensibiliser la population à l'importance du recensement. Même si de nouvelles formes de diffusion de l'information furent expérimentées, le caractère multiforme de l'ancienne mobilisation de la période soviétique, l'« agitatsia », est resté d'actualité⁴⁰. Aux articles de presse et affichages habituels sont venues s'ajouter les ressources offertes par Internet, par exemple la rubrique « 100 questions, 100 réponses » du site Web du Goskomstat ou des rubriques similaires sur les sites de certains comités statistiques de province, tel celui de Samara⁴¹. Le jour référence du 9 octobre, le président Poutine a été filmé par une chaîne de télévision quand il répondait aux questions de l'agent recenseur venu l'interroger. Des personnalités des arts et des spectacles furent interrogées sur la façon de répondre au recensement qu'elles avaient choisie, au domicile, par téléphone ou au bureau de leur circonscription.

À partir du mois de septembre, un jeu, « Le recensement des enfants », a été organisé par le ministère de l'Éducation dans toutes les écoles de Russie. À la suite des « leçons du recensement » consacrées à la situation démographique du pays et de la région, les enseignants distribuèrent aux écoliers des devoirs à faire à la maison afin qu'ils puissent échanger



L'affichette du recensement

avec leurs parents à ce sujet et les sensibiliser à l'idée de se faire recenser. Les plus jeunes eurent à faire des dessins sur les thèmes « Ma famille », « Ma cour d'immeuble », et à les commenter en énumérant leurs amis et les membres de leur famille, ce qui constitua une source d'information complémentaire sur les familles pour les agents recenseurs. Les écoliers les plus âgés participèrent à un jeu de rôles : jouant celui d'un agent recenseur, ils devaient poser les questions du bulletin individuel du recensement et confectionner des affiches d'appel au recensement avec un contenu patriotique. Ce « jeu des enfants » fut intégré aux cours de géographie et d'histoire locale. Dans certaines régions, il fut complété par d'autres initiatives, orientées sur l'affirmation d'une identité régionale ou ethnique, selon les cas. Par exemple, dans la région de Stavropol, à forte diversité ethnique, un concours fut organisé par les autorités politiques locales sur le thème « Souviens-toi de ta généalogie ». Dans ce cas, le recensement fut présenté comme un moment de construction d'une identité ethnique.

Des instructions pour convaincre de répondre

Les instructions données aux agents recenseurs leur présentèrent les principaux arguments à développer face aux cas les plus probables de refus de réponse. Par exemple, à la question « à quoi sert le recensement ? », il leur était conseillé de répondre : « *Le recensement est une procédure statistique traditionnelle. Les données du recensement sont indispensables pour prendre des décisions de gestion correctes. Pour tout le peuple, mais aussi pour vous-même, il sera intéressant de connaître quelle est la situation de notre société quand seront publiés les résultats du recensement. Cela nous donne l'espoir en des changements en vue d'une amélioration. On a besoin du recensement pour recueillir des données objectives sur les caractéristiques socio-démographiques, économiques et nationales de la population. Par la suite, ces*

données doivent être utilisées pour améliorer la situation dans le pays. Les données du recensement permettront au gouvernement de travailler de manière plus efficace, d'organiser correctement la politique sociale. Cette action n'est pas politique, mais économique. »⁴²

Dans les réponses envisagées, l'accent fut mis particulièrement sur l'utilité des informations recueillies en vue d'améliorer les mesures de politique économique de l'État, car il était essentiel de ne pas donner d'argument à l'expression d'un refus politique à l'égard du recensement. Or, dans certains immeubles ou quartiers urbains, des groupes d'habitants avaient déjà manifesté dans la presse leur volonté de ne pas répondre aux agents recenseurs en mettant en avant leur mécontentement vis-à-vis du gouvernement. Ce fut le cas, par exemple, à Saratov, ville de la Moyenne Volga, où quarante habitants de deux immeubles mitoyens ont, avant le recensement, envoyé une lettre collective au gouverneur de la province, au maire de la ville et aux journaux locaux pour expliquer leur intention de ne pas

participer au recensement en signe de protestation contre l'absence d'entretien et de réparations d'urgence dans leurs immeubles, aux risques et péril de la population qui y vit⁴³. D'autres situations de ce genre ont été rapportées par les journaux de différentes villes et régions. À Moscou, les habitants d'immeubles entiers ont refusé de se faire recenser pour protester contre le fait qu'ils n'avaient pas d'eau chaude. Dans tous ces cas, le refus de répondre a été utilisé par les citoyens comme un moyen de protestation contre l'État ou les autorités locales.

Face à ces refus argumentés de manière politique, les instructions conseillaient aux agents recenseurs de répondre de la manière suivante : « *Mais vous, vous voulez que cela change en mieux ? Mais si vous ne participez pas, alors qu'est-ce qui pourra s'améliorer ? Si, pour vous, la situation n'est pas bonne, alors il faut faire en sorte que l'État comprenne et prenne des mesures afin que pour vous cela aille un petit peu mieux. Mais sans informations précises cela ne pourra pas être réalisé. C'est pourquoi il faut répondre aux questions.* »



À Saratov, l'annonce du recensement sur un trolleybus

Face à un refus formulé de manière plus vindicative à l'égard du pouvoir, tel que « j'ai une piètre opinion des autorités au pouvoir, c'est pourquoi je ne veux pas participer au recensement », les instructions recommandaient de dire : « *En aucun cas il ne faut identifier pouvoir et recensement. On ne vous propose pas de voter pour quelqu'un. En donnant des informations selon les questions du recensement, vous aidez non seulement les dirigeants à prendre des mesures justes, mais aussi les entrepreneurs privés, et les organisations sociales. Le recensement n'est pas une affaire politique, mais une affaire sociale. Anton Pavlovitch Tchekhov n'était pas bienveillant envers le gouvernement tsariste, cependant il fit la propagande pour le recensement, et participa lui-même au recensement de 1897. En qualité d'agent recenseur, tout comme moi.* »⁴⁴

Cette référence à la participation active de Tchekhov aux opérations du recensement de 1897 fut souvent citée au cours de la campagne d'information.

Outre les réticences d'une partie de la population à répondre aux agents recenseurs, il fallut faire face également aux difficultés de recrutement du personnel.



Anton Pavlovitch Tchekhov, 1860-1904



10 roubles, un peu moins de 30 centimes d'euro

Des difficultés pour recruter le personnel

Les travaux de préparation et de conduite du recensement sur le terrain ont été supervisés par les comités locaux du Goskomstat en faisant appel à un personnel recruté de manière temporaire. En tout, 669 300 personnes furent employées pour effectuer le recensement⁴⁵. Les agents recenseurs ont constitué 69,6 % du personnel, et les instructeurs-contrôleurs 17,5 %. Ceci a donné en moyenne un instructeur pour quatre agents recenseurs.

La norme de travail a été fixée sur la base de l'expérience du recensement test effectué en octobre 2000. En 8 jours, chaque agent recenseur avait dû interroger 420 personnes en ville et 320 dans les zones rurales. En dépit des difficultés rencontrées pour remplir la charge de travail fixée, cette norme fut conservée pour le recensement de 2002, mais abaissée à 350 personnes en ville et 270 en zone rurale dans les régions difficiles d'accès, le Grand Nord sibérien par exemple. À titre de comparaison, en 1989, la charge de travail avait été fixée à 450 personnes par agent recenseur en ville et 350 à la campagne.

Le paiement des agents recenseurs fut un réel sujet de discussion et la somme prévue au départ dut être réajustée. Le montant de 700 roubles envisagé dans un premier temps dut être porté à 1 500 roubles en moyenne pour 19 jours de travail, ce

qui a correspondu à 78,95 roubles par jour⁴⁶. La rémunération des instructeurs-contrôleurs et des chefs de circonscription adjoints n'était guère plus élevée, 86,20 roubles par jour pour les premiers et 79,31 pour les seconds. Un chiffre permet de juger du niveau faible de ces rémunérations : à l'automne 2002, le minimum vital en Russie était évalué à 1 893 roubles par mois, ce qui revient à 63,10 roubles par jour⁴⁷.

Ces sommes furent trop faibles pour attirer suffisamment de volontaires dans les grandes villes, en particulier à Moscou où le coût de la vie est plus élevé qu'en province. Dans ce domaine aussi, les temps ont changé : contrairement à la période soviétique, il est devenu difficile de solliciter des personnes au nom seulement du devoir civique. Aussi, on fit en priorité appel aux enseignants et aux étudiants de 3^e et 4^e années. Un accord a été passé dans ce sens avec le ministère de l'Éducation et les conseils de recteurs des établissements de l'enseignement supérieur : pour les étudiants des années supérieures, le travail pour le recensement fut compté comme période de stage pratique⁴⁸. Dans de nombreux établissements, à Moscou et en province, reprenant une pratique ancienne, les enseignants de statistique, de démographie, d'économie ou de sociologie encadrèrent leurs étudiants pour faire du recensement un moment de formation. Ils exercèrent la fonction d'instructeur-contrôleur alors que la plupart des étudiants furent recrutés comme agents recenseurs, à l'exception des

étudiants de 5^e année en statistique qui renforcèrent les rangs des instructeurs-contrôleurs.

L'appel aux étudiants ne fut pas sans poser problème quand ceux-ci avaient un emploi parallèlement à leurs études. Une étudiante de 5^e année de l'Institut d'économie et de statistique de Moscou, qui assumait la fonction d'instructeur-contrôleur, en témoigne : « *Nous n'avons eu aucun problème lors de la tournée de vérification préalable. L'unique et, peut-être, principal [sic] problème dans le travail avec les agents recenseurs fut leur mauvaise volonté pour travailler. De nombreux étudiants ont participé au recensement "selon un volontariat contraint". En particulier les étudiants qui travaillent dans un autre endroit parallèlement à leurs cours. Ils ont exprimé leur mécontentement de participer au recensement. C'est justement chez ces étudiants que l'on a trouvé le plus d'imprécisions dans les questionnaires remplis.* »⁴⁹

Nouvelle société, nouvelles questions

Quelques modifications méthodologiques ont caractérisé le dispositif de questionnement du recensement de 2002. Certaines s'expliquent par la volonté d'harmoniser les procédures du recensement russe avec celles en vigueur dans d'autres pays, dans l'optique d'une comparaison internationale, d'autres par la nécessité de mieux cerner l'apparition de situations économiques et sociales nouvelles avec le passage à l'économie de marché.

Les thèmes principaux des questions posées dans le questionnaire court adressé à tout le monde (cf. pages 17 à 19) ont peu varié par rapport à ceux abordés en 1989, à l'exception de l'introduction d'une question sur l'emploi (Q.11). En revanche, la formulation a été modifiée dans certains cas, ce qui a pu donner lieu à de vives discussions au cours de la préparation, en particulier au sujet des questions liées à la nationalité des individus.

Dans la pratique, la formulation de la question 1, « votre lien de parenté avec les personnes vivant ensemble (selon le lien avec la personne inscrite en premier dans cette unité de recensement) », a posé quelques problèmes de compréhension. L'expression utilisée pour désigner le ménage, « unité de recensement », était trop peu explicite pour les personnes interrogées.

En 2002, les questions destinées à cerner la nationalité changèrent pour la première fois depuis 1920, ce qui donna lieu à beaucoup de discussions et controverses entre statisticiens, ethnologues et hommes politiques. Depuis les années 1920, étaient posées deux questions : « quelle est votre nationalité » et « quelle est votre langue maternelle ». Subsidièrement, il pouvait être demandé à la personne d'indiquer également « l'autre langue des peuples de l'URSS » qu'elle parlait couramment. Cette fois-ci, les différents débats qui entourèrent la formulation de ces questions, notamment à l'Assemblée des députés de la Russie, la Douma, aboutirent à l'abandon de celles qui figuraient dans le formulaire approuvé le 20 juin 2001 par la Commission d'État du recensement. La question sur la citoyenneté (Q.6) fut maintenue telle quelle, mais la suivante (Q.7) fut simplifiée : « à quelle ethnie (peuple) ou groupe ethnique estimez-vous appartenir ? » laissa la place à « votre appartenance nationale (selon la libre déclaration de la personne interrogée) ». Les instructions demandèrent aux agents recenseurs de respecter strictement la formulation de la question telle qu'elle était écrite sur le formulaire, sans ajouter quoi que ce soit. Il était exigé également de noter fidèlement la réponse de la personne interrogée, en respectant ses propres mots.

D'après différents témoignages d'agents recenseurs, il semble que beaucoup de personnes aient eu du mal à faire la différence entre citoyenneté et appartenance nationale. D'autres se sont volontairement déclarées de nationalité

« russe » alors qu'elles ne l'étaient pas⁵⁰. Enfin, la déclaration de la nationalité a revêtu un enjeu politique réel pour de nombreux groupes ethniques qui jugèrent essentiel d'affirmer leur poids numérique, et donc leur existence. Par exemple, le conseil des atamans du district cosaque de Stavropol appela l'ensemble des Cosaques à participer activement au recensement et à répondre « Cosaque » à la question sur l'appartenance nationale. L'objectif était de faire reconnaître comme « peuple » la communauté cosaque de cette région. Des consignes de ce type furent aussi données au sein des communautés de Tatars réparties dans différentes régions de la Volga⁵¹. Cette question prit une dimension particulièrement symbolique en Tchétchénie.

Les trois questions prévues dans un premier temps sur la connaissance des langues furent réduites à deux, et leur formulation simplifiée, appauvrie aux yeux de certains chercheurs, ethnologues et démographes en particulier. Une comparaison entre le bloc des trois questions envisagées au départ et celui des deux posées à l'arrivée éclaire les contestations que certains d'entre eux ont exprimées publiquement⁵² :



Cyrille et Méthode, apôtres des Slaves et inventeurs de l'alphabet glagolitique, à l'origine de l'actuel alphabet cyrillique

- Bloc des 3 questions approuvées en 2001 :

11. Votre langue maternelle ?

a) Si c'est « le russe », notez russe et passez à la question 13.

b) Si ce n'est pas « le russe », inscrivez quelle langue :

12. Parlez-vous le russe couramment ?

13. Quelle autre langue parlez-vous couramment ?

- Dans la version finale du questionnaire, l'enchaînement de ces trois questions laissa la place à :

9.1 Parlez-vous russe ?

9.2 Quelles autres langues parlez-vous ? (3 réponses possibles)

La suppression de l'adverbe « couramment » n'a pas été sans conséquence. En effet, la consigne donnée à ce sujet dans les instructions (« connaître une langue signifie pouvoir parler, lire et écrire ou seulement converser dans une cette langue ») a entraîné bien des interprétations de la part des agents recenseurs eux-mêmes et d'approximations dans les réponses, comme en ont témoigné différents agents recenseurs, par exemple : « En ce qui concerne la question sur la connaissance des langues étrangères, si j'en juge d'après l'ensemble des personnes recensées dans ma circonscription qui ont indiqué connaître une langue étrangère indépendamment du niveau de connaissance de cette langue, alors la quantité de personnes connaissant une langue étrangère sera chez nous bien supérieure à la réalité. »⁵³

Néanmoins, la rédaction des instructions montre nettement la volonté des statisticiens de contourner la simplification politique des questions sur les langues pour cerner au mieux la langue maternelle, faute d'une question spécifique à ce sujet. C'est ce que suggèrent les consignes données aux agents

recenseurs pour noter les réponses à la question 9.2 : « *La première des trois lignes est destinée à noter la langue maternelle de la personne interrogée. Pour cela il est indispensable de lui poser la question : Quelle est votre langue maternelle ? Si la personne interrogée nomme comme langue maternelle la langue russe, cela n'est pas à prendre en compte à la question 9.2. Mais, dans la zone « réserve » placée en bas du bulletin individuel, la première marque doit être obligatoirement cochée [...] Si le russe n'est pas la langue maternelle, alors le nom de la langue maternelle est inscrit sur la première ligne. Sur les deux autres lignes du bloc, on inscrit le nom des autres langues parlées par la personne interrogée.* »⁵⁴

Le souci de prendre en compte les langues des minorités nationales est évident également : « *Si la personne interrogée parle un nombre d'autres langues plus grand que celui prévu par le bloc, alors notez les langues telles que choisies par la personne interrogée. Si celle-ci a des difficultés pour choisir quelles langues noter, alors il faut donner la préférence aux langues des peuples de Russie et à celles de peuples autochtones des États membres de la Communauté des États indépendants (CEI). En outre, si la personne interrogée vit sur un territoire de la république (région autonome, district), alors, parmi les langues des peuples de Russie, la préférence est donnée à la langue d'État de ce territoire.* »⁵⁵

La question 10 sur les ressources des individus fit l'objet de nombreuses rumeurs, la plus étonnante étant qu'elle était posée pour la première fois. En réalité, elle figurait déjà dans tous les questionnaires des recensements depuis 1959 avec la même formulation : « Sources des moyens d'existence ». En 2002, les modifications ne concernèrent que les réponses proposées : alors que la liste de 1989 égrenait non pas des types de revenus différents, mais les catégories différentes de travail donnant lieu à un revenu, selon le type de structure d'emploi (travail dans

une entreprise d'État ou une administration, travail dans un kolkhoze, travail en coopérative, travail chez des particuliers, activité individuelle), celle de 2002 fut constituée sur la base de la distinction classique dans les économies libérales entre revenu du travail, revenu du capital et diverses formes d'allocations. La formulation de la question fut adaptée aux transformations introduites par le passage à une économie de marché.

Réelle nouveauté en 2002, une question 11 sur l'emploi a été placée à la fin du bulletin individuel, et posée selon les normes du BIT pour compter l'emploi et le chômage, les salariés et les non-salariés. À partir du recensement de 1970, les questions sur l'emploi ne furent posées que dans le cadre de l'enquête complémentaire par sondage, et centrées sur ce qui était nommé « caractéristiques économiques » : type d'occupation, branche de production (saisie sur la base du lieu de travail), statut de l'emploi, et groupe social⁵⁶. Dans une société où l'emploi pour tous était garanti par la Constitution, poser une question sur le chômage n'avait pas de sens. En octobre 2002, en revanche, la question sur le fait d'avoir un emploi ou pas, mais aussi d'être salarié ou pas, fut posée à tout le monde, et les questions plus précises à ce sujet furent réservées à l'enquête complémentaire réalisée auprès de 25 % de la population (cf. pages 20 et 21). Celle-ci comprit des questions sur les caractéristiques de l'emploi (Q.11.2 à 11.5), la situation de chômage (Q.11.6) et le repérage des migrations, notamment depuis le recensement précédent (Q.12). Comme par le passé, les instructions recommandèrent aux agents recenseurs de noter la branche de l'économie sur la base du lieu de travail des individus (Q.11.2)⁵⁷. Dans la continuité des recensements soviétiques depuis 1926, le terme occupation (*zaniatie*) fut maintenu en 2002 de préférence à celui de profession (Q.11.2, 11.5)⁵⁸. La persistance de cet usage en Russie s'explique par le maintien d'une

approche de la situation professionnelle, non pas justement selon la profession exercée, mais en croisant fonction ou métier et branche. Ainsi, les instructions du recensement précisait : « *Notez l'occupation en indiquant de la manière la plus précise possible le caractère du travail. Par exemple, il faut noter "chef du service de marketing et de l'écoulement de la production", "spécialiste dans la protection de l'information", "chercheur stagiaire en médecine", "conseiller dans le domaine juridique", "vendeur de produits industriels dans un magasin" [...]* »⁵⁹

Cette manière de présenter la catégorie occupation est très proche de la définition de l'occupation donnée par le Bureau du recensement américain pour le recensement de 2000 : « une occupation décrit le genre de travail qu'une personne fait dans son emploi. »⁶⁰ Dans la période soviétique, la classification sociale ne reposait pas sur la constitution de groupes sociaux à partir d'une classification professionnelle des individus, cela demeure vrai aujourd'hui⁶¹. Au-delà de l'habitude de la pratique statistique dans ce domaine, sans doute faut-il voir là aussi aujourd'hui un effet du flou dans la perception des situations professionnelles introduit par la période de transition.

Il est à noter qu'aucun des formulaires K et D ne contenait de questions permettant de mieux cerner deux groupes dont l'effectif a augmenté au cours des dix dernières années, celui des familles monoparentales, en particulier des femmes vivant seules avec un ou plusieurs enfants, et celui également des personnes sans domicile fixe, adultes et enfants⁶².

Premiers résultats

Les premiers résultats du recensement, annoncés officiellement au printemps 2003, confirment les prévisions du Goskomstat.

Combien sommes-nous ? Les premiers résultats établissent un chiffre de 145 287 400 personnes, dont

53,45 % de femmes⁶³. Si l'on considère la population permanente seule, soit 145 181 900 personnes, elle a diminué de 1,3 % par rapport à 1989 (moins 1 840 000). Ce chiffre confirme les projections effectuées jusque-là, ainsi que les résultats déjà établis pour les années antérieures : la diminution de l'effectif de la population est la conséquence principalement d'un déficit naturel qui n'est plus compensé par l'apport migratoire depuis 1992. Seulement trois quarts de ce déficit ont été compensés par l'immigration, dont la plus grosse partie a été constituée de flux en provenance des pays de la CEI et des États baltes.

Les refus de répondre n'ont pas mis en échec le recensement. Les méthodes utilisées par le Goskomstat pour circonscrire leur effet sur l'évaluation de l'effectif de la population ont permis à son président, Vladimir Sokoline, de déclarer à la presse, le 23 octobre : « *Nous avons couvert 100 % de la population [...] C'est une autre affaire, pour le moment, de dire quel pourcentage de la population a répondu à toutes les questions. Même les personnes qui ont refusé de répondre à nos questions, nous les avons "fixées". Dans tous les cas, elles ont été comptées dans le recensement, mais avec des lacunes.* »⁶⁴

Cette déclaration met toutefois en évidence la principale difficulté posée par les refus de réponse : s'ils n'ont pas mis en péril le dénombrement total de la population, ils ont néanmoins affecté la collecte des informations sur les caractéristiques de la population. L'enquête complémentaire par sondage servira à corriger cela.

Qui sommes-nous ? En 2002, la part de la population urbaine en Russie (73,3 %) est proche de celle de 1989 (73,4 %), bien que le nombre de citoyens ait baissé (106 470 200) par rapport à l'année 1989 (107 959 000), ce qui est le cas aussi de la population rurale (38 711 100 en 2002 et 39 062 900 en 1989). Une part dominante de la population urbaine (64 %) vit dans des villes de 100 000 habitants et plus. Toutefois, les migrations internes ont modifié la répartition de la population urbaine sur le territoire en se dirigeant principalement vers deux catégories de villes : la population n'a augmenté que dans celles qui comptent entre 100 000 et 250 000 habitants et dans celles qui ont un million d'habitants ou plus. Aujourd'hui, 40 % de la population des villes de 100 000 habitants ou plus vit dans une des 13 villes qui comptent un million d'habitants au moins : Moscou, Saint-Petersbourg,



Novossibirsk, Nijni Novgorod, Ekaterinbourg, Samara, Omsk, Kazan, Tcheliabinsk, Rostov sur le Don, Oufa, Volgograd, Perm.

La répartition de la population par régions a connu aussi de profondes modifications à la suite des courants de migration internes, qui ont représenté près de 80 % des flux globaux de migrations en Russie, mais aussi de l'arrivée de réfugiés en provenance des anciennes républiques soviétiques et des pays du Caucase. Les flux de migration intérieure se sont dirigés principalement vers les régions du Centre, de la Volga et du Sud. Dans ce dernier cas, sont venus s'ajouter les réfugiés des pays du Caucase. Des départs nombreux ont entraîné une diminution importante de la population dans les régions du Nord, de Sibérie et d'Extrême-Orient. Entre 1989 et 2002, plus d'un million de personnes ont quitté la Sibérie et les régions de l'Extrême-Orient russe pour d'autres régions de la Russie.

Aussi, aujourd'hui, plus de 60 % de la population vivent dans les régions administratives (*okrougi*) du Centre (26,2 %), de la Volga (21,5 %) et du Sud (15,8 %). Seulement 13,8 % vivent en Sibérie et 8,5 % dans l'Oural, ancienne région minière et industrielle. L'Extrême-Orient est la région la moins peuplée.

Ces pourcentages cachent toutefois des situations très différentes. Ainsi la population n'a augmenté strictement que dans les régions administratives du Sud (+ 11,6 %) et du Centre (+ 0,2 %). Encore, dans ce dernier cas, faut-il préciser que la plus grosse partie de cette croissance est due principalement à la force d'attraction de la capitale, Moscou, dont la population a augmenté de 20 % depuis 1989. Même si la région de la Volga reste une des plus peuplées, elle a néanmoins perdu 2 % de sa population totale depuis 1989.

Les transformations dans la répartition du peuplement des différentes régions russes depuis 1989 se sont accompagnées d'une diminution de la population dans 66 « sujets de la Fédération » sur 89. Il est à noter que

À Ekaterinbourg, 5^e ville de Russie



Yana

C'est à Ekaterinbourg, dans la nuit du 16 au 17 juillet 1918, que furent exécutés Nicolas II et sa famille.

beaucoup de régions autonomes peuplées de minorités nationales sont dans ce cas, notamment en Sibérie, ce qui a rendu d'autant plus sensible le traitement des questions ethniques dans ce recensement. L'afflux de réfugiés en provenance du Caucase dans les régions du Sud a contribué également à poser de manière aiguë la question ethnique sur ces territoires.

L'ensemble de ces changements démographiques explique que la question « qui sommes-nous ? », posée au début du recensement, ait revêtu une signification hautement politique dans cette période de recomposition des territoires et de l'unité nationale après l'éclatement de l'URSS. Ainsi on pouvait lire dans la rubrique « 100 questions, 100 réponses » du site Web du Goskomstat consacré au recensement :

Question : « *Quelle est la signification politique du recensement ?* »

Réponse : « *Avant tout, le recensement est un facteur de consolidation de l'ensemble de la nation, une action qui contribue à réduire les tensions sociales. Tous les individus sont*

intéressés par sa réalisation, indépendamment de leur idéologie, de leurs opinions politiques, de leur position sociale. C'est la seule forme de participation à la vie politique dans laquelle personne n'est opposé à personne. Le recensement, c'est le début du processus d'inventaire de tout le pays initié par le président Poutine. Ainsi, il constitue un des premiers pas de mise en ordre du pays : en effet l'ordre commence par un inventaire. Le recensement est une partie intégrante de la nouvelle politique de la Russie. »⁶⁵

La réalisation de ce recensement, conçu comme un instrument de la construction du nouvel État, a en effet représenté un enjeu politique réel pour un pouvoir fédéral désireux d'asseoir son autorité sur les régions et les républiques, mais aussi pour les gouverneurs des régions et les présidents des républiques, soucieux d'affirmer leur pouvoir au niveau local et face à l'État fédéral. Il l'a été également pour les minorités nationales qui aspirent à être reconnues comme « sujets » à part entière de la Fédération.

Martine MESPOULET
Université d'Angers

Notes

1. L'acte officiel qui a marqué l'éclatement de l'URSS a été cosigné le 8 décembre 1991 par les trois présidents de Russie, Ukraine et Biélorussie. Le 25 décembre, le Parlement russe vota l'indépendance de la Russie et la création d'un nouvel État, la Fédération de Russie.
2. À ce sujet, voir Alexandre Avdeev, Alain Blum, « La population russe : des raisons d'espérer ? », *Population & sociétés*, n° 351, nov. 1999.
3. Sur l'évolution des indicateurs démographiques depuis 1991, voir en particulier Alexandre Avdeev, Alain Monnier, « À la découverte de la fécondité russe contemporaine », *Population*, n° 4-5, 1994, pp. 859-902.
4. Sur le Goskomstat et la réforme de l'administration statistique d'État russe dans les années 1990, voir Martine Mespoulet, « Dix ans de réforme de la statistique d'État en Russie (1991-2001) », *Le kiosque international*, Insee, n° 12, décembre 2001.
5. Voir l'article « Sensatsia perepissi nomer odin : my ne znaem, gde jivet nache nasselenie » (Nouvelle à sensation numéro un du recensement : nous ne savons pas où vit notre population), *Rossiiskaia gazeta*, 23/10/2002, pp. 1, 9-10.
6. Loi n° 143-F3 du 15/11/1997 « Sur les actes de l'état civil ».
7. Voir A.E. Sourinov, *Ofitsial'naia statistika v Rossii : problemy reformirovaniia (La statistique officielle en Russie : les problèmes de sa réforme)*, Moscou, RETSEP, 2002, p. 114. Aleksandr Sourinov est premier adjoint du président du Goskomstat de Russie.
8. *ibid.*, p. 9.
9. Article 6, point 2, de la loi sur le recensement.
10. Sur l'encadrement des opérations de collecte des données par les organisations du Parti, voir Martine Mespoulet, *Statistique et révolution en Russie. Un compromis impossible (1880-1930)*, Rennes, Presses universitaires de Rennes, 2001, chap. 9.
11. Voir l'article « Perepis' vseobchtchaia, nagrada personal'naia » (Le recensement est général, la distinction est personnelle), *Rossiiskaia gazeta*, 29/07/2003, p. 9.
12. A.E. Sourinov, *Ofitsial'naia statistika v Rossii...*, *op. cit.*, p. 137.
13. Décret n° 1064 « O Vserossiiskoi perepissi nasselenia 2002 goda » (Au sujet du recensement de la population de la Fédération de Russie de 2002).
14. Décret n° 231 « Ob organizatsii Vserossiiskoi perepissi nasselenia v 2002 godu » (Sur l'organisation du recensement de la population de la Fédération de Russie de 2002).
15. Voir, par exemple, la rencontre organisée le 12 novembre 2002 par la section de démographie de la Maison centrale des savants de l'Académie des sciences de Russie pour faire le bilan de l'expérience des démographes-statisticiens, des enseignants et des étudiants de statistique ayant participé activement aux opérations du recensement à Moscou. Un compte rendu de cette réunion a été publié dans *Voprossy statistiki*, revue éditée par le Goskomstat : V.A. Borissov, « Kak prokhodila perepis' nasselenia 2002 goda v g. Moskve : vpetchatlenia outchastnikov » (Le déroulement du recensement de la population de 2002 dans la ville de Moscou : impressions de participants), *Voprossy statistiki*, 2/2003, pp. 54-62.
16. Décret n° 231 « Ob organizatsii Vserossiiskoi perepissi nasselenia v 2002 godu ».
17. Selon la Constitution de 1993, la Fédération de Russie regroupe 89 entités territoriales, administratives et nationales, appelées « sujets de la Fédération », à statut différent : républiques, oblasts, kraï, régions et districts autonomes, villes d'importance fédérale. À ce sujet, voir Jean Radvanyi, *La Nouvelle Russie*, Paris, Armand Colin, 2000, chap. 9.
18. L'URSS a connu officiellement 6 recensements de la population, en 1920, 1926, 1939, 1959, 1970 et 1989. Il y en eut aussi un en 1937, mais ses résultats furent annulés par Staline pour des raisons politiques. Voir Alain Blum, Martine Mespoulet, *L'anarchie bureaucratique. Statistique et pouvoir sous Staline*, Paris, La Découverte, 2003. Sur les recensements soviétiques des années 1920 et 1930, voir Martine Mespoulet, « Catégories professionnelles et classement social dans les recensements russes et soviétiques (1897-1939) », *Annales de Démographie Historique*, 1, 2003, pp. 81-109 ; voir aussi Alain Blum, *Naître, vivre et mourir en URSS. 1917-1991*, Paris, Plon, 1994.
19. G.S. Polliak élaborera la nomenclature des occupations des recensements de 1920 et 1926, et participa à l'élaboration de celle du recensement de 1937. Au sujet de ces nomenclatures, voir M. Mespoulet, « Catégories professionnelles... », *art. cit.*
20. G.S. Polliak, « Metody outcheta nasselenia v petrogradskoi gorodskoi perepissi 1918 goda » (Les méthodes de dénombrement de la population dans le recensement de Petrograd de 1918), *Vestnik statistiki*, 4-7, 1919, pp. 19-32.
21. Lors du micro-recensement de la population de 1994, le Goskomstat adopta un certain nombre de catégories statistiques recommandées par les principaux organismes internationaux, en particulier le BIT et l'ONU.
22. L.L. Kozlova, *Organizatsionno-metodologicheskie ossobennosti vsesoiuznoi perepissi nasselenia 1989 goda (Les particularités organisationnelles et méthodologiques du recensement de la population soviétique de 1989)*, Moscou, MESI, 1990, p. 6.
23. L.L. Kozlova, *ibid.*, p. 6.
24. *Roukovodstvo dlia perepistchika o poriadke provedeniia Vserossiiskoi perepissi nasselenia 2002 goda i zapolneniia perepisnykh dokumentov (Manuel pour l'agent recenseur à propos des règles à suivre pour effectuer le recensement général de la population de Russie en 2002 et pour remplir les documents du recensement)*, Goskomstat, 2002, p. 8.

25. A.E. Sourinov, *op. cit.*, pp. 137-138.
26. Valeri Tichkov, « Ne nado iskousstvenno konstrouirovat' demografitcheskii kriz » (Il ne faut pas construire artificiellement la crise démographique), site Web : *Strana.Ru*, 22/01/2002, p. 2.
27. *Ibid.*
28. *Ibid.*, pp. 1-2.
29. Ces cas se sont multipliés ces dernières années.
30. Voir « Sensatsia perepissi nomer odin : my ne znaem, gde jivet nache nasselenie », *Rossiiskaia gazeta*, 23/10/2002, *art. cit.*, p. 9.
31. L.L. Kozlova, *op. cit.*, pp. 4-5.
32. Extrait de la conférence des statisticiens soviétiques réunie à cette occasion, cité dans L.L. Kozlova, *ibid.*, p. 7. Un recensement test a été effectué par voie postale en 1997, mais cette méthode n'a pas été retenue non plus.
33. DEZ : Direktsia Eksploatatsii zdanii (Direction de l'exploitation des immeubles).
34. T.R. Kalimouline, « Seminar o perepissi nasselenia v GOU-VCHE » (Séminaire au GOU-VCHE au sujet du recensement de la population), *Voprossy statistiki*, 2, 2003, pp. 62-66. Le GOU-VCHE est l'université d'État de Haut enseignement commercial de Moscou.
35. *Ibid.*
36. Le taux de sondage avait été de 25 % en 1989 et de 50 % en 1970. Voir L.L.Kozlova, *op. cit.*, pp. 12-13.
37. *Roukovodstvo...*, *op. cit.*, p. 26.
38. Goskomstat, *Vserossiiskaia perepiss' nasselenia 2002 goda. Osnovnye polozenia (Le recensement général de la population de Russie. Principes fondamentaux)*, Moscou, 2001, p. 7.
39. Cette précision fut notamment affichée dans les pages du site Web du Goskomstat consacrées à l'information sur le recensement (<http://www.perepis2002.ru/100.html>).
40. Sur l'« agitatsia », campagne de mobilisation des recensements des années 1920, voir M. Mespoulet, *Statistique et révolution en Russie...*, *op. cit.*, chap. 9.
41. <http://www.samarastat.ru>
42. *Roukovodstvo...*, *op. cit.*, pp. 86-89.
43. Denis Lebed, « Sur la rue Sobornaïa on ne veut pas être du matériau statistique », *Saratovskie Vesti*, 09/10/2002.
44. *Roukovodstvo...*, *op. cit.*, pp. 86-89.
45. A.E. Sourinov, *op. cit.*, p. 144.
46. En juillet 2003, 1 euro était équivalent à 34 roubles.
47. Goskomstat Rossii, *Rossia v tsifrakh (La Russie en chiffres)*, Moscou, 2003, p. 107.
48. V.L. Sokoline, « O khode podgotovki k vsrossiiskoi perepissi nasselenia 2002 goda » (Sur l'état de la préparation du recensement général de la population de Russie de 2002), *Voprossy statistiki*, 7, 2002, pp. 3-6.
49. V.A. Borissov, « Kak prokhodila perepiss' nasselenia 2002 goda... », *art. cit.*, p. 55.
50. T.R. Kalimouline, « Seminar o perespissi naselenia... », *art. cit.*
51. À ce sujet, voir Valeri Tichkov (dir.), *Na puti k perepissi (Sur la route du recensement)*, Moscou, Aviaizdat, 2003.
52. *Ibid.* Voir aussi Anatoli Vichnevski, in T.R. Kalimouline, « Seminar o perepissi nasselenia... », *art. cit.*, pp. 65-66.
53. Témoignage d'un agent recenseur de Moscou, in T.R. Kalimouline, *art. cit.*, p. 63.
54. *Roukovodstvo...*, *op. cit.*, p. 56.
55. *Ibid.*
56. L.L. Kozlova, *op. cit.*, pp. 24-36.
57. *Roukovodstvo...*, *op. cit.*, pp. 61-64.
58. Sur l'histoire de l'usage de cette notion dans les recensements des années 1920 et 1930, voir M. Mespoulet, « Catégories professionnelles... », *art. cit.*
59. *Roukovodstvo...*, *op. cit.*,
60. Peter Fronczek, Patricia Johnson, « Occupations: 2000. *Census 2000 Brief* », US Census Bureau, août 2003, p. 12.
61. M. Mespoulet, « Catégories professionnelles... », *art. cit.*
62. Sur ces questions, voir Cécile Lefèvre, « Enfance, famille et institutions en Russie au cours des années 1990 : un aperçu des évolutions récentes », *Revue d'études comparatives Est-Ouest*, n° 4, 2003 (à paraître).
63. Voir le site du Goskomstat : <http://www.gks.ru/PEREPIS/predv.htm>.
64. « Sensatsia perepissi nomer... », *art. cit.*, *Rossiiskaia gazeta*, 23/10/2002, p. 9.
65. <http://www.perepis2002.ru>, rubrique « 100 questions, 100 réponses ».

Le formulaire K

Indiquer le n° d'ordre de la personne dans le logement d'après la colonne 1 de la liste AfS : □□

1. Votre lien de parenté avec les personnes vivant ensemble (selon le lien avec la personne inscrite en premier dans cette unité de recensement)

Cochez une des possibilités de réponse

personne inscrite en premier	<input type="checkbox"/>	belle-fille, gendre	<input type="checkbox"/>
femme, mari	<input type="checkbox"/>	grand-mère, grand-père	<input type="checkbox"/>
filles, fils	<input type="checkbox"/>	petite-fille, petit-fils	<input type="checkbox"/>
mère, père	<input type="checkbox"/>	autre degré de parenté	<input type="checkbox"/>
sœur, frère	<input type="checkbox"/>	par le sang, par alliance	<input type="checkbox"/>
belle-mère, beau-père	<input type="checkbox"/>	sans parenté	<input type="checkbox"/>

2. Votre sexe masculin féminin

3. Votre date de naissance □□ □□ □□□□
 jour mois année

4. Votre situation matrimoniale

Citez toutes les possibilités de réponse

marié(e) n'a jamais été marié(e)
 ↓ séparé(e) veuf, veuve

Votre mariage a-t-il été enregistré ?

oui → Si l'époux(se) de cette personne habite avec elle, indiquez le n° sous lequel il (elle) figure dans la colonne 1 de la liste AfS □□
 non

5. Votre lieu de naissance

6. Votre citoyenneté

Russie
 Autre État (indiquez lequel) :
 Apatride

7. Votre appartenance nationale

(selon la libre déclaration de la personne interrogée)

8. ÉDUCATION

Pour les personnes de 6 ans et plus

8.1 Faites-vous vos études dans un établissement d'enseignement ?

(école primaire, collège secondaire, lycée technique, collège universitaire, établissement de l'enseignement supérieur, etc.)

- oui → Indiquez le type d'établissement d'enseignement :
- | | | | |
|-------------------------------------|--------------------------|---------------------------------------|--------------------------|
| enseignement général | <input type="checkbox"/> | enseignement secondaire professionnel | <input type="checkbox"/> |
| enseignement primaire professionnel | <input type="checkbox"/> | enseignement supérieur professionnel | <input type="checkbox"/> |
- non
-

Pour les enfants de 3 à 9 ans non scolarisés

8.2 Est-ce que votre enfant fréquente un établissement préscolaire ?

oui non

Pour les personnes de 10 ans et plus

8.3 a) Votre formation

- | | | | | | |
|--|--------------------------|---|--------------------------|-----------------------------------|--------------------------|
| primaire générale | <input type="checkbox"/> | générale de base
(secondaire non terminée) | <input type="checkbox"/> | secondaire générale
(complète) | <input type="checkbox"/> |
| secondaire professionnelle
(secondaire spécialisée) | <input type="checkbox"/> | supérieure professionnelle non terminée | <input type="checkbox"/> | | |
| supérieure professionnelle | <input type="checkbox"/> | post-universitaire professionnelle | <input type="checkbox"/> | | |
-

Pour ceux qui n'ont pas reçu de formation primaire générale

b) Savez-vous lire et écrire ?

oui non

Pour les personnes de 15 ans et plus qui possèdent une formation générale (primaire, de base, ou secondaire)

c) Avez-vous obtenu un diplôme dans un établissement professionnel ou dans un établissement professionnel et technique ?

oui non

9. CONNAISSANCE DES LANGUES

9.1 Parlez-vous russe ?

oui non

9.2 Quelles autres langues parlez-vous ?

10. Indiquez toutes les ressources dont vous disposez pour vivre

Montrez la carte des ressources à la personne interrogée

- | | | | |
|--|--------------------------|---|--------------------------|
| - revenu du travail
(à l'exception du travail dans une exploitation personnelle auxiliaire) | <input type="checkbox"/> | - allocation chômage | <input type="checkbox"/> |
| - exploitation personnelle auxiliaire | <input type="checkbox"/> | - autre type d'aide de l'État | <input type="checkbox"/> |
| - bourse | <input type="checkbox"/> | - revenus de l'épargne placée | <input type="checkbox"/> |
| - pension (à l'exception d'une pension d'invalidité) | <input type="checkbox"/> | - revenus de la location de biens immobiliers et fonciers | <input type="checkbox"/> |
| - pension d'invalidité | <input type="checkbox"/> | - à la charge de quelqu'un | <input type="checkbox"/> |
| - allocation (à l'exception de l'allocation chômage) | <input type="checkbox"/> | - autre ressource | <input type="checkbox"/> |

EMPLOI

11.1 Avez-vous eu un travail qui vous a procuré un salaire ou un revenu durant la semaine précédant le début du recensement de la population ?

oui non → *fin du questionnaire*



En quelle qualité avez-vous travaillé dans votre emploi principal ?

- | | | | |
|---------|--------------------------|-------------------|--------------------------|
| salarié | <input type="checkbox"/> | non salarié | <input type="checkbox"/> |
| | | ↓ | |
| | | avec des salariés | <input type="checkbox"/> |
| | | sans salarié | <input type="checkbox"/> |
| | | autre cas | <input type="checkbox"/> |

Le formulaire D

Les questions 11.2 à 11.5 concernent l'emploi que la personne interrogée considère comme principal, si elle en occupe plusieurs.

Uniquement pour ceux qui ont répondu « oui » à la question 11.1

11.2 Dans quelle branche de l'économie êtes-vous occupé ?

- | | | | | | |
|---------------------------|--------------------------|--|--------------------------|--------------------------------|--------------------------|
| - industrie | <input type="checkbox"/> | - commerce de gros | <input type="checkbox"/> | - éducation | <input type="checkbox"/> |
| - agriculture | <input type="checkbox"/> | et de détail | <input type="checkbox"/> | - culture et arts | <input type="checkbox"/> |
| - exploitation forestière | <input type="checkbox"/> | - restauration et cafés | <input type="checkbox"/> | - recherche | <input type="checkbox"/> |
| - construction | <input type="checkbox"/> | - service municipal d'entretien des logements, services courants à la population | <input type="checkbox"/> | - finances, crédit, assurances | <input type="checkbox"/> |
| - transport | <input type="checkbox"/> | - santé publique et aide sociale | <input type="checkbox"/> | - gestion | <input type="checkbox"/> |
| - communications | <input type="checkbox"/> | | | - autre | <input type="checkbox"/> |

↓
indiquez laquelle :

11.3 Quelle production principale ou quels services produit l'entreprise, l'organisation dans laquelle vous travaillez (y compris si vous êtes entrepreneur individuel) ?

11.4 Votre emploi se trouve-t-il sur le territoire de votre ville (arrondissement) de résidence ?

oui non → sur le territoire de votre région (district autonome, kraï, république)

sur un autre territoire

↓
indiquez le nom du district autonome, de l'oblast, du kraï, de la république de Russie, ou le nom d'un autre État :

11.5 Votre occupation ou le travail que vous effectuez

Il faut noter la réponse de manière précise en indiquant la caractère du travail effectué (par exemple : vendeur, chef d'atelier)

Passez à la question 12

Uniquement pour les personnes âgées de 15 à 64 ans qui ont répondu « non » à la question 11.1

11.6 Dans le cas où vous n'avez pas d'emploi, en avez-vous cherché un au cours du dernier mois ?

oui → si l'on vous proposait un emploi qui vous convienne, pourriez-vous commencer dans les deux semaines qui viennent : oui non

non → indiquez une seule raison, la raison principale :

- | | | | |
|---|--------------------------|--|--------------------------|
| j'ai trouvé un emploi et je commence dans un délai de deux semaines | <input type="checkbox"/> | il n'y a pas de possibilité de trouver un emploi | <input type="checkbox"/> |
| j'ai trouvé un travail et attends une réponse | <input type="checkbox"/> | je m'occupe de mon foyer | <input type="checkbox"/> |
| j'attends le début de la saison | <input type="checkbox"/> | autre raison | <input type="checkbox"/> |

12. Avez-vous habité dans cette ville (bourg ou localité rurale) sans interruption depuis votre naissance ?

oui
non → **1. Indiquez l'année depuis laquelle vous habitez ici en permanence : □□□□**

Si l'année est comprise entre 1989 et 2002, passez à la partie 2 de cette question

2. Où habitez-vous en janvier 1989 ?

a) nom de la république autonome, de l'arrondissement autonome, du kraï, de l'oblast, du district autonome de la RSFSR, de la république soviétique de l'ancienne URSS ou de l'État :

b) type du lieu d'habitation où vous résidiez : urbain rural

Uniquement pour les femmes de 15 ans et plus

13. À combien d'enfants avez-vous donné naissance ?

Indiquez le nombre total d'enfants nés (en excluant les morts-nés) :

Destination Libourne

Plus de 400 stagiaires étrangers accueillis au CEFIL de 1996 à 2003

Le CEFIL, Centre de formation de l'Insee à Libourne, a été inauguré en mars 1996. Entièrement financée par l'État, sa construction s'était accompagnée de la mise en chantier par la municipalité de Libourne d'une résidence meublée, essentiellement des studios mais aussi des duplex dont soixante seraient loués à l'année par l'Insee pour y loger les stagiaires (et les intervenants).

Cette installation à Libourne répondait tout à la fois à la volonté politique de décentralisation des services de l'État telle qu'elle s'était affirmée au début des années 90 et au souci de l'Insee d'aménager un nouveau centre de formation¹ en un lieu offrant de bonnes conditions de travail et d'hébergement, à des prix compétitifs. C'était le cas de la ville de Libourne, qui présentait en outre des avantages comparatifs appréciables : gare TGV et proximité

(40 kilomètres) de l'aéroport de Bordeaux-Mérignac.

Appelé à devenir une véritable école de formation des contrôleurs recrutés sur concours, avec en particulier une période de scolarité continue de six mois², ce nouvel établissement de l'Insee avait également vocation à s'ouvrir en grand à l'accueil de stagiaires étrangers, ceci à un moment où notre Institut devait faire face à une demande accrue d'assistance technique, principalement du fait de la conversion des pays d'Europe centrale et orientale (les PECO, États baltes compris) et de la CEI³ à l'économie de marché.

C'est de cette seconde dimension de l'activité du CEFIL qu'il va être question dans les lignes qui suivent, à commencer par les séminaires de formation aux métiers de la statistique publique qui y sont organisés

chaque année en direction des principaux publics visés par la coopération française et européenne : d'une part les pays en développement francophones, d'autre part les PECO, les pays de la CEI et les pays du pourtour méditerranéen.

Été 1996 : le baptême du feu

Le premier séminaire international organisé par le CEFIL s'est tenu du 1^{er} au 12 juillet 1996, trois mois après l'inauguration des locaux. Consacré à la comptabilité nationale, il a réuni dix participants provenant des pays du Maghreb.

Suivront en 1997 deux nouveaux séminaires. Le premier, centré sur les statistiques sociales, a réuni pendant trois semaines, du 7 au 30 avril, vingt-huit participants provenant des pays en transition (Europe centrale et orientale et CEI, dix-sept pays représentés au total, et trois langues de travail : français, anglais, russe). Le second, dédié aux techniques de sondage et couplé à une réunion de l'AISE⁴, a rassemblé, du 9 au 17 juin,

Le CEFIL, au bord de l'Isle



3, rue de la Cité - 33500 Libourne
téléphone : 33 (0)5 57 55 56 00

1. En complément aux écoles nationales de statistique (l'Ensaie et l'Ensaï, qui forment respectivement les administrateurs et les attachés de l'Insee) et au Centre d'études des programmes économiques (le CEPE, unité de formation continue de la direction des enseignements supérieurs et de la recherche).

2. Antérieurement à l'installation du CEFIL, la première année d'activité des contrôleurs de l'Insee recrutés sur concours se comprenait comme une alternance de courtes périodes de formation et de travail effectif, ce qui n'était satisfaisant, ni pour les intéressés, ni pour les services dans lesquels ils avaient été affectés.

3. Communauté d'États indépendants nés de l'éclatement de l'ex-URSS.

4. Association internationale des statisticiens d'enquêtes. Le CEFIL assure le secrétariat de cette association, partie intégrante de l'Institut international de statistique.

vingt-deux participants représentant autant de pays francophones.

L'année 1998, particulièrement chargée, a notamment été marquée par l'organisation, à l'intention de neuf stagiaires provenant de six pays méditerranéens, d'un séminaire de longue durée sur la statistique publique qui s'est étalé sur près de trois mois, du 6 avril au 26 juin. Deux autres séminaires ont été organisés pendant cette même période. Le premier, centré sur les enquêtes auprès des ménages et lui aussi destiné à des statisticiens méditerranéens, s'est tenu du 6 au 24 avril. Il a réuni quinze stagiaires représentant sept pays. Le second, à destination des pays en transition, s'est tenu du 2 au 19 juin. Consacré aux statistiques d'entreprises, il était le pendant du séminaire sur les statistiques sociales qui avait été organisé l'année précédente à l'intention de ce même public. L'audience a été à peu près identique, avec à nouveau vingt-huit participants représentant cette fois seize pays. Enfin, le stage de longue durée « statistique publique » n'était pas encore terminé quand a débuté un quatrième séminaire, à nouveau centré sur les enquêtes auprès des ménages mais à l'intention cette fois d'un public africain. Ce quatrième stage de 1998, spécifiquement tourné vers l'Afrique subsaharienne, a rassemblé, du 24 juin au 10 juillet, vingt et un statisticiens provenant de quatorze pays.

5. Observatoire économique et statistique d'Afrique subsaharienne. Afristat coordonne aujourd'hui l'action statistique de dix-huit pays subsahariens ou de l'océan Indien : Bénin, Burkina Faso, Cameroun, Cap-Vert, Centrafrique, Comores, Congo, Côte d'Ivoire, Gabon, Guinée, Guinée-Bissau, Guinée équatoriale, Mali, Mauritanie, Niger, Sénégal, Tchad et Togo. Cf. notamment l'article de Lamine Diop : « Afristat, premiers jalons sur la voie de l'intégration statistique régionale », *Courrier des statistiques* n° 93, mars 2000.

6. DIAL : groupement d'intérêt scientifique spécialisé dans les études et recherches relatives aux politiques économiques à mettre en œuvre dans les pays en développement. Cf. notamment l'article de Guy de Monchy : « DIAL, développement d'investigations sur l'ajustement à long terme », *Courrier des statistiques* n° 60, décembre 1991.

7. Institut de recherche pour le développement (ex-ORSTOM) spécialisé sur les projets en zone intertropicale.

L'offre socle du CEFIL, 1999-2003		PECO
Afrique subsaharienne		
12-30 avril		17 mai-4 juin
Statistiques d'entreprises	1999	Statistiques sur l'emploi et les salaires
19 participants représentant 13 pays		19 participants représentant 12 pays
15-31 mai		13-23 juin
Emploi, revenu, secteur informel	2000	Recensements de population
21 participants représentant 16 pays		21 participants représentant 11 pays
11-22 juin		9-23 mai
Statistiques agricoles	2001	Statistiques et analyse conjoncturelles
28 participants représentant 17 pays		25 participants représentant 13 pays
10-25 juillet		18-28 mars
Enquêtes auprès des ménages	2002	Pratique des statistiques locales : pourquoi et comment ?
25 participants représentant 15 pays		25 participants représentant 10 pays
8-23 juillet		31 mars-11 avril
Cartographie et systèmes d'information géographique	2003	Utilisation des données administratives : emploi, chômage, salaires
30 participants représentant 24 pays		18 participants représentant 11 pays

Pour 2004, les sujets envisagés sont l'utilisation des données administratives pour les statistiques d'entreprises (PECO, en avril) et la stratégie de diffusion des données statistiques (Afrique subsaharienne, en juillet).

Depuis 1999, une offre socle en deux séminaires annuels

De ces premières années de fonctionnement, des différents formats de séminaires expérimentés, est née l'« offre socle » du CEFIL. Elle consiste en l'organisation de deux séminaires annuels, chacun d'une durée de deux à trois semaines, l'un monté en concertation avec Afristat⁵, l'autre avec Eurostat, et toujours en ligne avec la politique de coopération statistique de l'Insee.

Le premier s'adresse prioritairement aux pays francophones de la zone Afristat. Peuvent également y être accueillis, dès lors qu'ils maîtrisent le français, des stagiaires venant d'autres pays en développement, francophones ou non francophones.

Le second (bilingue : français/anglais) se veut (ou se voulait) prioritairement tourné vers des pays ayant vocation à rejoindre l'Union européenne à une échéance prévisible, essentiellement les PECO mais aussi certains pays du bassin méditerranéen comme Chypre et Malte. Quant aux pays de la CEI, précédemment accueillis en même temps que les PECO, ils se voient désormais offrir

des formations spécifiques (en russe), à l'instar des pays du pourtour méditerranéen.

C'est ainsi qu'à l'offre socle du CEFIL se sont ajoutés deux séminaires supplémentaires en 2000, du 3 au 21 juillet et du 29 août au 8 septembre. Le premier, consacré aux statistiques d'entreprises, a réuni vingt participants provenant de dix pays méditerranéens. Le second, centré sur les répertoires d'entreprises, a rassemblé dix-huit participants représentant dix pays de la CEI.

Le contenu et le contenant

L'identification et la bonne prise en compte des attentes des différents publics accueillis sont une préoccupation première, tant s'agissant du contenu des formations proposées que du choix des intervenants, à rechercher bien sûr au sein de l'Insee et des SSM (services statistiques des ministères), mais pas seulement. Ainsi est-il régulièrement fait appel, pour les deux séminaires formant l'offre socle du CEFIL, d'un côté à des experts d'Afristat, de DIAL⁶ ou de l'IRD⁷, de l'autre à des représentants d'Eurostat et à nos confrères statisticiens des pays de l'Union pour illustrer les meilleures pratiques.

Des programmes à connaître

PHARE (Poland and Hungary: Action for Restructuring of the Economy).

D'abord destiné à favoriser le processus d'ouverture de la Pologne et de la Hongrie, puis de l'ensemble des pays d'Europe centrale et orientale, à l'économie de marché, le programme PHARE est devenu le principal instrument du soutien financier de l'Union européenne à l'adhésion des dix PECO candidats : Bulgarie, Estonie, Hongrie, Lettonie, Lituanie, Pologne, République tchèque, Roumanie, Slovaquie et Slovénie. Les pays des Balkans occidentaux et la Turquie peuvent être concernés par des programmes similaires.

TACIS (Technical Assistance to the Commonwealth of Independent States).

Le programme TACIS est le principal instrument européen d'aide à la mise en œuvre d'accords de partenariat et de coopération avec les douze pays de la CEI : Arménie, Azerbaïdjan, Biélorussie, Géorgie, Kazakhstan, Kirghizstan, Moldavie, Ouzbékistan, Russie, Tadjikistan, Turkménistan et Ukraine. Un treizième partenaire potentiel, hors CEI, est la Mongolie.

Medstat. Medstat est le volet statistique du programme MEDA, principal instrument financier de l'Union européenne pour la mise en œuvre du partenariat euro-méditerranéen. Les pays partenaires sont au nombre de douze : Algérie, Chypre, Égypte, Israël, Jordanie, Liban, Malte, Maroc, Territoires autonomes de Palestine, Syrie, Tunisie et Turquie. L'appel d'offres pour le programme Medstat2 est actuellement suspendu.

Naturellement, il y a toujours une place pour la présentation des travaux français les plus pertinents, et, quand le sujet s'y prête, pour des exercices concrets (prise en main de logiciels cartographiques, par exemple), toujours vivement attendus. Des échanges d'expériences formalisés entre les participants sont systématiquement inscrits au programme, et très appréciés. Chaque séminaire s'achève sur une évaluation, le plus souvent précédée d'une table ronde.

Le contenant est au diapason, avec un amphithéâtre convivial et bien équipé de 80 places, une salle de 35 places (divisible en deux salles jumelles indépendantes) équipée pour l'interprétation simultanée, des salles micro-informatiques en réseau, d'autres salles encore, offrant de 10 à 25 places chacune, sans oublier un centre de documentation bien fourni en littérature économique et statistique. Rappelons en conclusion de ce rapide inventaire que, complémentirement au français, les langues de travail pratiquées à l'occasion des séminaires internationaux du CEFIL sont essentiellement l'anglais (séminaires à l'inten-

tion des pays candidats à l'Union européenne ou du pourtour méditerranéen) et le russe (pour les pays de la CEI), avec interprétation simultanée du français à l'anglais ou au russe et vice versa⁸.

Il va sans dire que la logistique du CEFIL, l'hébergement, le programme social et un accueil chaleureux s'appuient sur une équipe attentionnée et parfaitement rodée.

Convivialité

Outre les locaux et équipements pédagogiques proprement dits, et tout aussi nécessaires au bon fonctionnement du CEFIL, des espaces de détente très fréquentés lors des pauses café fournissent d'autres occasions d'échanges, entre les participants comme entre ceux-ci et les conférenciers invités.

Nous nous efforçons par ailleurs, afin de conforter l'esprit de groupe, de proposer à nos stagiaires étrangers des activités touristiques variées pendant les week-ends. S'y ajoutent, à l'initiative d'une association locale (Libourne Accueil), des

invitations à dîner, toujours très appréciées, dans des familles libournaises. Enfin, faut-il le préciser, il est fait abstraction au CEFIL des éventuels contentieux ou conflits politico-historiques pouvant opposer les pays d'origine de nos hôtes : ainsi l'ordre alphabétique des pays a-t-il pu placer côte à côte des stagiaires d'Azerbaïdjan et d'Arménie, d'Israël et du Liban.

Les bailleurs de fonds

Les séminaires internationaux organisés au CEFIL impliquent des participations financières diverses, notamment s'agissant de la prise en charge des *per diem* (dépenses quotidiennes pour l'hébergement et la restauration) et frais de voyage des stagiaires accueillis⁹.

Pour faire court, on se contentera d'indiquer ici, en s'en tenant à ces seules dépenses, que les séminaires à l'intention des pays en développement francophones sont financés pour l'essentiel par le ministère français des Affaires étrangères, les autres par l'Union européenne dans le cadre des programmes PHARE (cas des PECO), TACIS (pays de la CEI) et Medstat (pays du pourtour méditerranéen). L'intermédiation avec ces bailleurs de fonds est assurée par le CESD-Paris¹⁰, qui joue le rôle de trésorier du CEFIL, et contribue parfois sur ses ressources propres.

8. Les séminaires qui avaient été organisés en 1997 et 1998 à l'intention conjointe des PECO et des pays de la CEI s'étaient tenus en trois langues, français-anglais-russe.

9. Les *per diem* suggérés sont relativement peu élevés (50 à 60 euros) compte tenu de la modicité des coûts afférents (studio à 100 euros la semaine, restaurant administratif à proximité immédiate de la résidence). Naturellement, le CEFIL se fait un devoir d'aller chercher ses stagiaires à la gare ou à l'aéroport et de les y reconduire.

10. Le CESD-Paris est le plus ancien des établissements du Centre européen de formation des statisticiens économistes des pays en voie de développement, aujourd'hui déployé en un réseau européen. À l'origine école de formation (issue de l'Ensaie) des statisticiens africains francophones, il organise toujours le concours commun des écoles africaines et apporte un soutien à diverses actions de formation.

D'autres prestations

En complément des séminaires internationaux de formation aux métiers de la statistique publique, le CEFIL peut monter des actions plus ciblées, répondant à des demandes ou besoins spécifiques. La première opération de ce type avait été la mise sur pied, dès 1996 et en partenariat avec le rectorat de Bordeaux, d'un programme de formation et d'assistance technique à l'intention des statisticiens du ministère marocain de l'Éducation nationale. Plus près de nous, en 2001, le CEFIL a organisé un atelier de formateurs en comptabilité nationale assistée par ordinateur, autour de la nouvelle version du logiciel ERETES¹¹, et, événement marquant, un séminaire de formation aux techniques quantitatives de la macroéconomie qui a réuni vingt participants chinois. En 2002, nous avons accueilli, à nouveau dans le cadre de la formation de formateurs, une délégation ukrainienne, le domaine abordé étant cette fois celui de la statistique d'entreprises.

11. Logiciel d'aide à l'élaboration des comptes nationaux (ERE = équilibres ressources-emploi, TES = tableaux entrées-sorties).

12. Créé en novembre 1999, le consortium international PARIS21 a vocation à œuvrer au renforcement de la statistique dans les pays en développement.

13. Le nombre des contrôleurs recrutés sur concours n'a en effet cessé d'augmenter ces dernières années afin d'anticiper les départs en retraite. La gamme des formations nationales assurées au CEFIL s'est par ailleurs sensiblement enrichie, incluant désormais l'accompagnement des promotions au choix des agents nommés dans un corps de catégorie supérieure (de C en B ou de B en A).

Quand les créneaux le permettaient, le CEFIL a mis volontiers ses installations, logements compris, à la disposition de l'Institut TES (Training of European Statisticians). Dans la même veine, Libourne a abrité en 2000 la réunion du 50^e anniversaire du comité 28 (traitant des inégalités) de l'Association internationale de sociologie, en 2001 une réunion du CEIES (l'équivalent européen de notre Conseil national de l'information statistique), en 2002 une réunion de l'EEG6 (European Board for EDI Standardization Expert Group 6-Statistics).

Enfin, la notoriété venant, le conseil en formation constitue un autre volet de l'activité internationale du CEFIL. De nombreuses délégations étrangères viennent ainsi s'enquérir de nos méthodes de formation des stagiaires nationaux, pour des suites variables. Pour ce qui concerne le proche avenir, quelques opérations se dessinent : appui à l'africanisation des formations statistiques professionnelles dans la zone Afristat, aide à la rénovation des formations statistiques au Vietnam, conseil pour la mise en place d'un centre de formation statistique des pays andins.

Mieux faire connaître notre offre

Même si nos stages font le plein, une amélioration de la communication sur la nature et la diversité de notre offre ainsi que sur nos réalisations et projets paraît nécessaire, afin de faciliter l'accès à ces informations

des pays demandeurs d'assistance technique.

Des contacts ont été noués à cet effet avec le consortium PARIS21 (Partnership in Statistics for Development in the 21st Century)¹², en vue d'insertions dans sa *Newsletter* trimestrielle, éditée en plusieurs langues et consultable sur le Net (www.paris21.org). Une première concrétisation a été la publication, dans la lettre n° 2 (juillet 2002), d'un article présentant le CEFIL. Un projet de séminaire organisé en commun est par ailleurs sur les rails.

D'autres développements pourraient suivre, en particulier la mise en ligne d'informations relatives à l'activité du CEFIL sur www.insee.fr, avec possibilité d'également y accéder à partir des sites du CEPE et du CESD Paris.

Le mot de la fin sera pour signaler une prochaine extension de nos installations, essentiellement commandée par les besoins des formations nationales¹³ mais qui bien sûr profitera également aux formations internationales : des salles de l'autre côté de la rue, qui dès 2004 agrandiront le CEFIL, et recherche de nouvelles possibilités résidentielles.

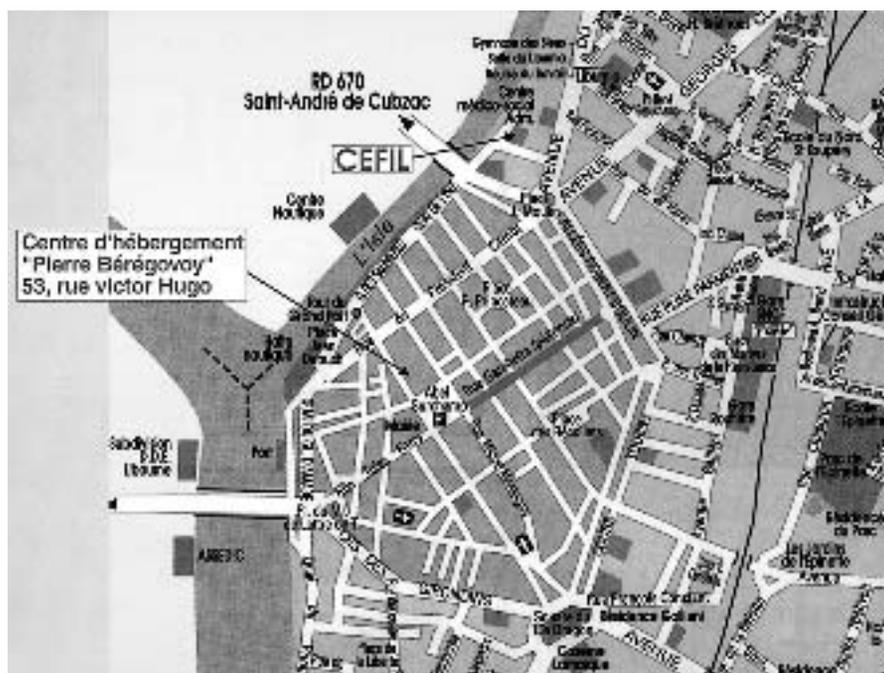
Michel BOËDA

Directeur-adjoint du CEFIL,
chargé des relations internationales

Pays ayant envoyé au moins 6 participants aux stages du CEFIL, 1996-2003

<u>PECO</u>		<u>Zone Afristat.</u>		<u>Autres pays</u>	
Bulgarie	10	Bénin	6	Albanie	9
Estonie	7	Burkina Faso	6	Angola	8
Hongrie	11	Cameroun	11	Chine	20
Lettonie	8	Cap-Vert	6	Haïti	12
Lituanie	14	Côte d'Ivoire	9	Macédoine	10
Pologne	18	Gabon	8	Madagascar	7
République tchèque	12	Mali	8	Rwanda	8
Roumanie	18	Mauritanie	10	Sao Tomé et Príncipe	6
Slovaquie	12	Sénégal	6		
Slovénie	13	Togo	6		
		<u>Pays méditerranéens</u>			
		Algérie	11		
		Égypte	6		
		Maroc	9		
		Tunisie	14		
<u>CEI</u>					
Biélorussie	6				
Russie	7				
Ukraine	6				

De 1996 à 2003, le CEFIL a accueilli au total 421 stagiaires étrangers provenant de 63 pays, sans compter les participants aux séminaires organisés dans les locaux du CEFIL par divers organismes.



Le centre d'hébergement



Une infrastructure pour les données en sciences humaines et sociales

Les observations statistiques à grande échelle sont pour les sciences humaines et sociales des ressources précieuses, indispensables à des recherches de qualité. De premiers efforts avaient été entrepris au milieu des années quatre-vingt, à l'initiative du CNRS, afin de faciliter l'accès des chercheurs à ce type de données, qui s'étaient notamment concrétisés par l'acquisition de fichiers d'enquêtes auprès de l'Insee ou d'autres services statistiques publics avec droit d'utilisation par l'ensemble des laboratoires du Centre. De nouvelles et spectaculaires avancées ont été enregistrées ces dernières années avec la création d'un comité interministériel des données en SHS, la constitution d'une unité mixte de service, le Centre Quetelet, pour l'archivage, la documentation et la diffusion des données, l'installation enfin de « plates-formes universitaires ». Le déploiement de cette infrastructure s'inscrivait dans le droit fil des conclusions d'un rapport sur les sciences sociales et leurs données qu'avait commandé début 1999 Claude Allègre, en sa qualité de ministre chargé de la recherche, à M^{me} Roxane Silberman, alors (et toujours) directrice du Lasmal-IdL.

Le comité des données...

Le « comité de concertation pour les données en sciences humaines et sociales » est régi par le décret n° 2001-139 du 12 février 2001. Créé auprès des ministres chargés de l'économie, de l'emploi, de l'éducation nationale et de la recherche¹ et placé sous la présidence du dernier cité, il propose les orientations de la politique publique de données pour la recherche en sciences humaines et sociales.

Ses principales missions sont énumérées à l'article 2 du décret :

- *s'informer des activités de production et d'utilisation des données ;*

- *déterminer les ensembles de données dont l'intérêt scientifique justifie le recueil, la préservation et la diffusion pour la recherche ;*

- *faire des propositions afin de contribuer à la promotion de la formation à l'utilisation des données, à l'accès des chercheurs et des enseignants-chercheurs aux données utiles aux sciences humaines et sociales ainsi qu'au développement des recherches utilisant les grands fichiers de données ;*

- *formuler des propositions en vue de la production de grandes enquêtes utiles à la recherche ;*

- *émettre des avis sur les conditions favorisant la mobilité des personnels entre l'enseignement supérieur, les organismes de recherche et les organismes producteurs de données.*

Le comité, qui comprend deux représentants des ministres auprès desquels il a été créé², peut associer à ses travaux des représentants de tout autre ministre intéressé. Un représentant du CNRS et un représentant de la Conférence des présidents d'université assistent de droit à ses réunions, au moins deux par an, auxquelles peuvent être invitées en tant que de besoin d'autres personnalités extérieures. Son secrétariat général est constitué de deux personnes, M^{me} Ève Caroli, professeur des universités, et l'auteur de ces lignes.

... et son conseil scientifique

Le conseil scientifique institué auprès du comité de concertation (article 6 du décret) comprend quinze membres, français ou étrangers et représentant différentes disciplines des sciences humaines et sociales, nommés pour trois ans³ par arrêté du ministre chargé de la recherche. Il élit en son sein, à la majorité des deux tiers, un président et un vice-président.

Les avis attendus de ce conseil, qui lui aussi se réunit au moins deux fois par an avec participation éventuelle de personnalités extérieures, doivent notamment éclairer le comité concernant les matières suivantes :

- *priorités à donner dans la collecte, la documentation, le stockage et le traitement des ensembles de données utiles pour les sciences humaines et sociales ;*

- *principes déontologiques applicables à la collecte, l'archivage et l'utilisation de ces ensembles de données ;*

- *domaines dans lesquels des efforts particuliers doivent être réalisés ;*

- *conditions et procédures d'accès des chercheurs et des enseignants-chercheurs aux données ;*

- *dispositifs de formation préparant les étudiants, doctorants, chercheurs et enseignants-chercheurs à l'utilisation des données.*

1. Le comité est susceptible de s'ouvrir à d'autres ministères.

2. L'un des deux représentants du ministre chargé de la recherche assure la vice-présidence.

3. Le mandat est renouvelable une fois.

Dictionnaire des sigles

CEREQ	Centre d'études et de recherches sur les qualifications
CIDSP	Centre d'information des données socio-politiques <i>Le CIDSP est une unité de recherche et de service du CNRS.</i>
CNRS	Centre national de la recherche scientifique
EHESS	École des hautes études en sciences sociales
FQP	Formation-qualification professionnelle
INED	Institut national d'études démographiques
IRESCO	Institut de recherche sur les sociétés contemporaines
Lasmas-IdL	Laboratoire d'analyse secondaire et de méthodes appliquées à la sociologie - Institut du longitudinal <i>Le Lasmas-IdL est une unité mixte de recherche CNRS, EHESS et université de Caen.</i>
MRSH	Maison de la recherche en sciences humaines, université de Caen
MSH	Maison des sciences de l'homme
SHS	Sciences humaines et sociales
UMS	Unité mixte de service

Le comité et son conseil ont travaillé à l'élaboration d'un cadre réglementaire fixant les conditions d'accès des chercheurs à différents types de données statistiques. Ils ont aussi émis des avis sur l'opportunité du financement de certaines enquêtes. La participation française à l'enquête sociale européenne (première vague en 2002-2003, deuxième en 2004) et une contribution de la Recherche à la réalisation par l'Insee de l'enquête FQP 2003 ont pu ainsi être décidées après évaluation de l'intérêt scientifique de ces opérations.

Archivage et diffusion des données : le Centre Quetelet et ses unités partenaires

Conformément à l'article 10 du décret du 12 février 2001, le CNRS est associé à la mise en œuvre de la politique publique des données en SHS par l'intermédiaire d'une unité mixte de service, le Centre Quetelet⁴.

Créée au 2^e semestre 2001 en partenariat avec l'EHESS, l'INED et l'université de Caen, cette UMS est elle-

même chargée d'appuyer, valoriser et coordonner les activités d'archivage et de diffusion d'*unités partenaires*, initialement (et toujours) au nombre de trois : le Lasmas-IdL, qui depuis sa création documente et diffuse des données issues de la statistique publique française (recensements de population, autres enquêtes de l'Insee, enquêtes du CEREQ, enquêtes des services statistiques des ministères chargés du travail, de l'emploi, de la formation professionnelle, de l'éducation nationale, de la culture...), le CIDSP, Centre d'informatisation des don-

nées socio-politiques (françaises et étrangères)⁵, enfin le service des enquêtes de l'INED.

Le Centre Quetelet comprend aujourd'hui une dizaine de personnes, ingénieurs et techniciens, réparties entre Paris, Caen et Grenoble, dans les mêmes locaux que ceux abritant le Lasmas-IdL (hébergé par l'IRESCO à Paris, et à Caen par la MRSH) et le CIDSP (hébergé par l'Institut d'études politiques de Grenoble, lui-même rattaché à l'université Pierre Mendès France).



Université de Caen, la MRSH

4. Cf. *Courrier des statistiques* n° 100, décembre 2001, pp. 55-56.

5. La diffusion de données socio-politiques est organisée à l'échelle européenne dans le cadre du CEESDA (Council of European Social Sciences Data Archives).

Son action est supervisée par un comité de pilotage rassemblant le directeur du département des SHS du CNRS, le président de l'EHESS, le directeur de l'INED, le président de l'université de Caen, le secrétaire général du comité de concertation pour les données en SHS et le président du conseil scientifique institué auprès de ce comité. Ce comité de pilotage, dont fait également et naturellement partie le propre directeur du Centre Quetelet, élit en son sein un président pour un mandat de deux ans renouvelables. Il se réunit au moins une fois par an pour examiner et approuver les projets du Centre, son budget prévisionnel, sa politique en matière de contrats, de valorisation, de publication, ainsi que les

éventuelles demandes d'adhésion d'équipes de recherche ou de service en qualité d'unités partenaires. Le rapport d'activité établi tous les deux ans par le directeur du Centre (cf. l'article suivant de Benoît Riandey) est examiné et approuvé dans ce même cadre.

Les plates-formes universitaires

Les plates-formes universitaires pour les données ont pour fonction de fournir un soutien rapproché à des travaux de recherche et éventuellement à des activités d'enseignement faisant appel à l'exploitation de fichiers statistiques diffusés par le

Centre Quetelet. À partir d'un établissement d'enseignement supérieur, éventuellement en liaison avec une MSH, et en concertation avec le Centre Quetelet, elles fournissent cette aide auprès d'un ensemble d'écoles doctorales et d'équipes de recherche appartenant à une même aire géographique et représentant plusieurs disciplines des SHS. La première a été créée dans le Nord.

Alain CHENU

Professeur des universités
Secrétaire général du comité
de concertation
pour les données en sciences
humaines et sociales

Quelques adresses Web

Rapport Silberman

www.education.gouv.fr/rapport/silberman/

Comité de concertation

www.recherche.gouv.fr/comite/shumaines.htm

Centre Quetelet

www.centre.quetelet.cnrs.fr

Lasmas-IdL

www.iresco.fr/labos/lasmas/enquetes.htm

CIDSP

www-bdsp.upmf-grenoble.fr/contentfr.htm

INED

www-enquetes.ined.fr/enquete.php

Enquête sociale européenne

www.europeansocialsurvey.org/

Centre Quetelet, deuxième anniversaire

Créé le 7 décembre 2001 avec mission de soutenir la politique publique de valorisation des données en sciences humaines et sociales (cf. l'article d'Alain Chenu, pages 29 à 31), le Centre Quetelet a rapidement pris son envol : en 2002, près de 1 500 fichiers d'enquêtes ont ainsi été mis à la disposition de la recherche, plus de 160 chercheurs au total. Nos premiers efforts avaient naturellement porté sur l'amélioration des conditions et procédures d'accès des chercheurs aux données. De nombreux projets ont par ailleurs été initiés, visant tout à la fois à accroître notre offre et à stimuler la demande. Bonne route, donc, à l'équipe du Centre Quetelet, aujourd'hui composée de dix ingénieurs et techniciens répartis sur trois sites (Paris, Caen et Grenoble) et pilotée par un nouveau directeur, Pascal Arduin.



Adolphe Quetelet, 1796-1874

Centre Quetelet — IRESCO
59, rue Pouchet - 75489 Paris Cedex 17
tél. : 33 (0)1 40 25 11 61
CQFD, Centre Quetelet Fournisseur de Données
www.centre.quetelet.cnrs.fr

Trois grands fonds d'enquêtes

Le Centre Quetelet valorise et coordonne les activités d'archivage et de diffusion de trois unités partenaires : la BDSP (Banque de données socio-politiques), le service des enquêtes de l'INED (Institut national d'études démographiques) et le Lasmas-IdL (Laboratoire d'analyse secondaire et des méthodes appliquées à la sociologie - Institut du longitudinal).

Abritée par le CIDSP (Centre d'informatisation des données socio-politiques), dans les locaux de l'Institut d'études politiques de Grenoble, la BDSP archive aujourd'hui quelque 600 enquêtes, nationales ou internationales, couvrant de nombreux aspects de la vie politique et sociale, en France et en Europe : enquêtes Eurobaromètres, ISSP (International Social Survey Program), World Values Survey, enquêtes post-

électorales du CEVIPOF (Centre de recherches politiques de Sciences Po), baromètres BVA ou encore OIP (Observatoire interrégional du politique). On retrouve bien sûr dans tout cet ensemble nombre de sujets correspondant à des problématiques de recherche très actuelles : élections, comportements et attitudes politiques, études sur les médias, valeurs et questions sociales, religion, élites et leadership, éducation, structure et stratification sociale, agriculture ou encore immigration.

Côté INED, on citera notamment ces grands thèmes d'étude que sont la fécondité, la contraception, la sexualité, l'immigration, les situations et histoires familiales, sans oublier une série longue d'enquêtes d'opinion sur la conjoncture démographique et la politique familiale. Les données et la documentation de plus de 50 enquêtes de l'INED ont été validées par les Archives nationales, à

l'intention certes des générations futures mais aussi pour le plus grand profit des chercheurs d'aujourd'hui.

Enfin, le Lasmas-IdL documente et diffuse des données issues de la statistique publique française : recensements de population et autres enquêtes de l'Insee, enquêtes des services statistiques des ministères chargés du travail, de l'emploi, de la formation professionnelle, de l'éducation nationale, de la culture, enquêtes du CEREQ (Centre d'études et de recherches sur les qualifications).

À ces trois grands fonds « institutionnels » est venu s'ajouter l'accès aux enquêtes du CREDES (Centre de recherche, d'étude et de documentation en économie de la santé), de l'INPES (Institut national de prévention et d'éducation pour la santé) et de l'OFDT (Observatoire français des drogues et des toxicomanies).

Les outils de la documentation

DDI, la norme internationale de documentation des données, a été traduite en français à l'instigation du Centre Quetelet. De la diffusion de ce nouveau document de référence est bien sûr attendue une homogénéisation de la qualité des données statistiques.

DDS, le logiciel de documentation des données statistiques développé par l'Insee, est déjà mis à profit dans bon nombre de services statistiques ministériels, notamment celui du ministère du Travail. Un objectif fort est de systématiser l'utilisation de ce logiciel dans l'ensemble de la statistique publique.

Suivront bientôt les enquêtes de transport urbain du CERTU (Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques).

De tout cet ensemble ressort une grande cohérence - et une forte complémentarité - des thématiques s'offrant aux chercheurs.

La dynamique est créée

Le dispositif Quetelet prend pleinement son sens dans une utilisation accrue des données. De fait, en 2002, ce sont 1 448 fichiers qui ont ainsi été mis à disposition de 167 chercheurs, soit, par rapport à 2001, trois fois plus de fichiers à deux fois plus de chercheurs...

Ce résultat est l'aboutissement d'une activité soutenue en matière de documentation, de veille et d'archivage, d'aide aux utilisateurs. S'y sont ajoutées l'élaboration et la mise à disposition sur notre site Web d'un moteur de recherche d'enquêtes (baptisé ORDS, outil de recherche de données en sciences sociales), grâce auquel le demandeur peut directement accéder à de la documentation sur toute enquête référencée par le Centre (www.centre.quetelet.cnrs.fr/ords).

Trois égales priorités

C'est fondamentalement pour favoriser et développer l'accès des chercheurs en sciences humaines et sociales à des observations statistiques à grande échelle qu'a été créé le Centre Quetelet. De cette raison d'être s'ensuivent trois égales priori-

tés : *le renforcement de notre vocation pluridisciplinaire*, qui déborde le cadre de ces trois disciplines fondatrices que sont la démographie, la science politique et la sociologie, *l'encouragement à la documentation des données* d'enquêtes, au fil de leur production, enfin *la stimulation de la demande*.

Renforcer notre vocation pluridisciplinaire. De nouvelles compétences, et sans doute de nouveaux partenariats, doivent être introduits dans le dispositif Quetelet. L'économie, la géographie, l'histoire sont de bons exemples, même si ces disciplines traitent abondamment de données d'une nature pour nous inhabituelle : des agrégats liés à un instant ou à un territoire. Des avancées ont été enregistrées dans le domaine de l'économie, avec l'ouverture aux chercheurs d'un accès gratuit à la BDM (banque de données macroéconomiques de l'Insee). Le Centre leur offre également son soutien s'agissant des démarches à effectuer auprès du comité du secret du Conseil national de l'information statistique (le Cnis) en vue de pouvoir accéder aux enquêtes d'entreprises. Déjà, les économistes sont les premiers utilisateurs des « données Quetelet ».

Encourager la documentation des données. Les services producteurs attendent du Centre Quetelet un soutien dans la documentation de leurs données. En la matière, beaucoup a déjà été fait, ainsi que l'atteste l'évolution rapide de l'offre d'enquêtes dûment documentées émanant des fonds BDSF et Lasmas-IdL. Mais le véritable objectif se situe bien sûr en amont : les données d'enquêtes doivent être documentées au fil de leur

production, de sorte à pouvoir être immédiatement exploitables à des fins de recherche. Les Archives de France et le Centre Quetelet invitent producteurs, archivistes, documentalistes et utilisateurs à venir participer le 29 janvier 2004 à Paris à une journée d'étude « Archivage historique et mise à disposition des données des sciences sociales », qui traitera de cet objectif.

Stimuler la demande. Trop de chercheurs ignorent encore la richesse voire même l'existence de l'offre du Centre Quetelet et de ses unités partenaires. Des actions d'information sont donc nécessaires, concernant la localisation et les conditions d'utilisation de « nos » fichiers d'enquêtes. Les ingénieurs du Centre ont également vocation à former les utilisateurs en matière de traitement et d'analyse des données. Ont ainsi été organisés, en octobre 2003 un atelier thématique sur les sondages, avec l'appui de la délégation de Paris A du CNRS (Centre national de la recherche scientifique), puis fin novembre une journée d'étude sur les enquêtes formation/éducation, en collaboration avec le PIREF (Programme incitatif de recherche en éducation et formation). Suivra le 23 février 2004 une journée de formation aux techniques quantitatives, organisée avec la participation du groupe « méthodologie » de l'Association française de sociologie. Enfin, à l'image des utilisations pédagogiques, le Centre Quetelet souhaite développer des clubs de chercheurs utilisateurs d'une même enquête, et ceci dans un double objectif : réduire les coûts d'accès aux données, et favoriser une recherche synergétique.

Les fichiers pédagogiques

Le Centre Quetelet propose aux universitaires des fichiers pédagogiques issus d'enquêtes emblématiques de leur discipline (<http://centre.quetelet.cnrs.fr/pedagogie.htm>). Destiné à servir de support pratique à des cours de méthodes et/ou de statistiques¹, ce matériel se présente sous la forme d'un « kit-cédérom » comprenant une extraction de données du fichier initial, la documentation relative à l'enquête dont est issu ce fichier ainsi qu'un ou deux diaporamas de présentation.

Sept fichiers sont disponibles depuis la rentrée universitaire 2002-2003 :

- *démographie* : enquête Fécondité et famille de l'INED
- *économie* : enquête Budgets des familles 1995, enquête Emploi 2001 (Insee)
- *géographie* : fichier historique de l'urbanisation en France
- *histoire* : enquête de Louis Henry sur la population de la France (1670-1829)
- *sciences politiques* : enquête post-électorale de 1995 du CEVIPOF
- *sociologie* : enquête Choix du conjoint, d'Alain Girard (INED)

La collection devrait prochainement s'enrichir de ces six autres fichiers :

- *démographie* : sondage au 1/1000 du recensement de 1999, extrait de l'enquête sur l'histoire familiale associée à ce recensement (Insee)
- *économie* : enquête sur les petites entreprises industrielles (Insee, 1997), « mini-fichiers » de salaires, de revenus fiscaux, de montants de patrimoine
- *épidémiologie* : enquête ACSF sur les comportements sexuels en France
- *sciences de l'éducation* : enquête sur le niveau intellectuel des enfants d'âge scolaire (INED/INOP², 1965)
- *socio-linguistique* : extraits du fichier des prénoms en France et de l'enquête sur l'histoire familiale (Insee)
- *sociologie* : enquête Formation des couples (INED), enquête sur les pratiques culturelles des Français (ministère de la Culture).

Destinés à une très large diffusion³, ces fichiers pédagogiques doivent être rigoureusement anonymes en sorte de ne présenter aucun risque de divulgation d'informations personnelles. Ils doivent respecter la propriété intellectuelle du producteur. Ils ne doivent permettre aucun détournement de finalité, aucune utilisation commerciale au détriment de la propriété commerciale des producteurs et de l'égalité de traitement des entreprises face à la concurrence. Leur production doit être menée en accord avec les producteurs et en concertation avec les universitaires.

Indiquons enfin que la prise en compte des demandes de fichiers pédagogiques s'effectue de façon centralisée et automatisée au moyen d'un logiciel développé à cet effet par le Centre Quetelet : le logiciel PALMIPED, en clair « petite application logicielle pour mission pédagogique ».

1. Ces fichiers pédagogiques apportent ainsi une première réponse à une recommandation insistante de l'Académie des sciences qui, dans son rapport sur la statistique publié en décembre 2000 (rapport RST n° 8, Éditions Tec & Doc), préconisait un recours accru aux grandes bases de données dans l'enseignement de la statistique ainsi qu'une mise à niveau des laboratoires universitaires en informatique statistique.

2. Institut national d'étude du travail et d'orientation professionnelle.

3. Notamment dans le cadre des travaux pratiques, pour lesquels est proscrite l'utilisation des fichiers sources dont sont issus les fichiers pédagogiques. L'utilisation de ces fichiers sources pour des cours magistraux est en revanche conforme à leur finalité. Sous réserve de la signature d'un engagement de l'enseignant, et sous la responsabilité de ce dernier, ils sont également accessibles aux étudiants préparant une thèse, un DEA, un DESS ou une maîtrise.

Ouverture sur l'Europe

Les centres européens d'archives vivantes se sont regroupés en un réseau très actif, le CESSDA (Council of European Social Sciences Data Archives), dont la BDSP est le membre français. Le Centre Quetelet entretient déjà des relations étroites avec trois de ces centres, le Data Archives de l'université d'Essex, le SIDOS de Neuchâtel (Service suisse d'information et d'archivage de données pour les sciences sociales) et le ZUMA de Mannheim (Zentrum für Umfragen, Methoden und Analysen).

Par ailleurs, et naturellement, nous participerons activement, chaque fois que possible, au congrès annuel de l'IASSIST (International Association of Social Science Information & Technology).

Le dispositif Quetelet et l'Université

Avec les chercheurs des grands instituts de recherche, les universitaires

sont nos principaux utilisateurs, en même temps que des partenaires. Favorisées par l'implantation géographique du Centre Quetelet (réparti, rappelons-le, sur trois sites : Paris, Caen et Grenoble), des relations privilégiées se sont en effet nouées avec l'université de Paris 1, avec celle de Caen, plus particulièrement avec sa maison de la recherche en sciences humaines et son UFR (unité de formation et de recherche) d'économie, avec l'université Pierre Mendès France de Grenoble et plus particulièrement son IEP (Institut d'études politiques) et sa MSH (maison des sciences de l'homme). Au-delà de ces trois « fiefs », le Centre a développé d'étroites relations avec les universités d'Aix-Marseille, Dijon, Lille, Strasbourg et Toulouse. Une concrétisation particulièrement attendue a été la création, à Lille, d'une première plate-forme universitaire de données, interface d'un genre nouveau entre le Centre Quetelet et les jeunes chercheurs et doctorants. Cette plate-forme joue déjà un rôle moteur dans la formation

aux recherches quantitatives et l'enseignement des méthodes quantitatives.

Un autre aspect remarquable de l'activité du Centre Quetelet en direction des universités est le développement au sein de ces dernières, dans le cadre d'une dynamique initiée à l'été 2003, de clubs d'utilisateurs de « fichiers pédagogiques » (cf. page 35) spécifiquement élaborés par nos soins à des fins d'apprentissage de la science statistique. Ces clubs seront bientôt dotés d'un forum sur notre site Internet. Une bourse aux fichiers pédagogiques devait par ailleurs se tenir à l'automne 2003 à la maison de l'économie de l'université de Paris 1.

Benoît RIANDEY
Premier directeur
du Centre Quetelet

Données sociales

Une publication unique en son genre

Depuis déjà trente ans, *Données sociales* propose tous les trois ans un panorama rigoureux et attendu de la société française et de son évolution. Avec une telle ancienneté, il va de soi que ce titre, créé en 1973 par Alain Darbel, porte en lui une doctrine, ou, pour parler en termes plus chargés de professionnalisme, une « ligne éditoriale ». En quoi consiste exactement cette fameuse ligne éditoriale, sur laquelle, évidemment, rédacteurs et auteurs vont avoir les yeux fixés ? Il s'agit simplement d'un certain nombre de choix, compromis entre des objectifs et les résistances de la réalité : on se propose ici de présenter les principaux d'entre eux et les raisons qui y ont présidé.

Une histoire pas si linéaire

La longue lignée de *Données sociales* pourrait donner à penser que la vie éditoriale de cet ouvrage de référence s'est déroulée comme un fleuve tranquille. En fait, il n'en est rien. Il a fallu, au fil des éditions successives, faire des arbitrages délicats et les assumer, lever des doutes et des préventions : finalement, la réalisation à l'Insee d'un recueil périodique sur les données sociales était peut-être une gageure. Et le résultat n'est pas banal : comme on le verra, la ligne éditoriale de *Données sociales* est assez unique en son genre.

Social ou social ?

Le mot « social » se partage entre deux significations : dans la première (sens « collectif »), il se réfère à la collectivité considérée dans son ensemble, dans la seconde (sens « individuel ») aux différentes situations personnelles.

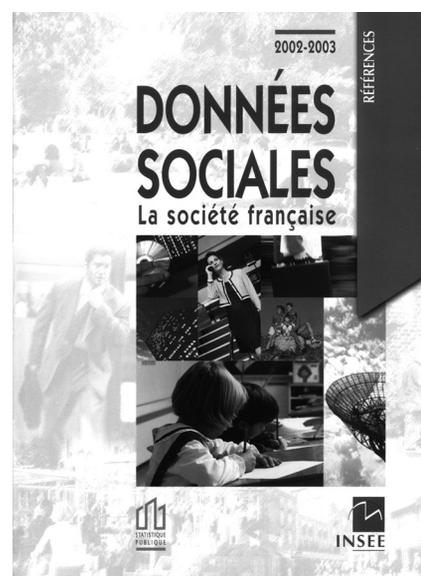
Au sens collectif, le social, c'est la description de l'ensemble de la société dans ses disparités et ses différences entre groupes sociaux. Au sens individuel, c'est l'attention portée aux problèmes quotidiens des personnes et aux politiques sociales afférentes : politiques de l'emploi et d'indemnisation du chômage, politique familiale, politiques de santé, d'aide sociale, etc.

Données sociales a d'abord privilégié la première approche. La tendance s'est inversée à partir de la cinquième édition (1984).

Le tournant des années quatre-vingt

Les premières éditions (1973, 1974, 1978 et 1981) entendaient réaliser une synthèse sur l'ensemble des données des domaines sociaux, selon des idées à l'honneur dans les années soixante-dix. Cette préoccupation s'est traduite par la recherche d'une présentation systématique et de rapprochements entre domaines. Il s'agissait de transposer dans le champ social le modèle de la comptabilité nationale, qui connaissait alors ses heures de gloire dans le domaine économique. De cette approche initiale il est resté le plan général de l'ouvrage, de quelque dix chapitres, qui part des « ressources » (démographie, formation) pour aboutir aux fruits de l'activité et à leur répartition : niveau de vie, redistribution.

Au tournant des années quatre-vingt, prenant acte de la résistance des données sociales à se fondre dans une présentation modélisée, les responsables de la publication optent



pour une approche « éclatée » : désormais, l'éclairage sera individuellement porté sur les différentes facettes du domaine social et les politiques afférentes, sans chercher à y superposer une vision d'ensemble. Au plan de la forme, *Données sociales* devient un recueil de plusieurs dizaines d'articles, rédigés par autant d'auteurs provenant de divers horizons. Le rôle de l'équipe rédactionnelle s'en trouve transformé, de rédacteur unique à coordonnateur d'un travail collectif.

Le social et la quantification

D'un ouvrage traitant des domaines du social et engageant la responsabilité de l'Insee, le lecteur attend forcément des études « rigoureuses » et « objectives », c'est-à-dire appuyées sur des données statistiques dûment établies et référencées. Tout cela suppose, 1) que les domaines étudiés soient effectivement couverts

par des données chiffrées, 2) que les auteurs construisent des raisonnements rigoureux, en étayant leurs arguments à partir des données en question, 3) que la validité de ce genre d'études soit assez largement reconnue.

En ce qui concerne le point 1, il faut évoquer le considérable développement des informations sur les domaines du social depuis les années soixante : enquêtes auprès des ménages et mobilisation des sources administratives.

Le point 3 renvoie aux débats sur la sociologie quantitative. Jusque vers les années soixante, la sociologie était en France une discipline uniquement littéraire : il était assez généralement convenu, parmi les sociologues français, qu'une analyse sociale était chose trop subtile pour qu'un tableau de chiffres puisse lui apporter quelque chose. Cependant, la sociologie quantitative a connu une rapide avancée, par imitation des travaux américains et grâce au soutien de personnalités comme Pierre Bourdieu. *Données sociales* a accompagné cet essor.

Quant au point 2, il incombe évidemment à l'équipe rédactionnelle de veiller à la qualité des articles. Cependant, il existe le risque de la facilité : réduire ses ambitions et se limiter à une simple présentation et description de résultats chiffrés. Jusqu'à présent, *Données sociales* a su conserver son caractère de recueil d'articles et non d'annuaire de chiffres.

Une œuvre collective

Privilégier une approche « individualisée », comme c'est le cas depuis la cinquième édition de *Données sociales*, conduit à couvrir largement le champ étudié sous l'étiquette de « social ». Un ensemble d'auteurs recrutés à l'Insee est alors insuffisant ; il faut faire appel aux compétences des services statistiques des ministères (les SSM), dans les domaines de l'éducation, de l'em-

ploi, de la santé, de la culture, etc. Ces services ont aussi une expérience particulière en matière de suivi et d'évaluation des mesures de politique sociale.

Mais, bien sûr, il serait dommage de se limiter aux seuls aspects concernés par la gestion économique ou administrative. La nécessité de compléter, en introduisant par exemple des études de nature sociologique, a été reconnue depuis déjà bien longtemps. Certes, il ne peut être question, dans un volume limité, et avec une information statistique qui ne se renouvelle pas toujours d'une édition à l'autre, de prendre au pied de la lettre l'exigence d'exhaustivité. Mais certains thèmes particuliers peuvent être analysés à tour de rôle, toutes les deux ou trois éditions. Pour ce type de contributions, les collaborations sont essentiellement à rechercher dans les milieux universitaires.

La vulgarisation à l'honneur

Le public de *Données sociales* est surtout composé de non-spécialistes, qui cherchent à s'informer de façon générale sur le domaine social. L'exemple typique est celui des enseignants.

Naturellement, rendre accessible à un auditoire de non-initiés un article nourri d'information statistique exige des efforts de rédaction spécifiques : il faut mettre en évidence l'intérêt du sujet traité, suivre un fil directeur clair et naturel, simplifier le plus possible l'expression, bannir ellipses et sous-entendus, qui vont de soi entre spécialistes mais font pour les autres obstacle à la compréhension. Tableaux et graphiques doivent être assortis de légendes et notes de lecture parfaitement explicites. Enfin et à l'intention des lecteurs désireux d'en savoir davantage, on n'omettra bien sûr pas de mentionner la source des données et les références bibliographiques essentielles.

La longue histoire de *Données sociales* et les efforts accomplis par

les équipes de rédaction successives pour améliorer la lisibilité de cet ouvrage en rendent maintenant la réalisation plus aisée : la ligne rédactionnelle et les règles auxquelles doivent se conformer les auteurs sont de mieux en mieux connues, et acceptées.

L'édition 2002-2003

Le plan général de l'édition 2002-2003 de *Données sociales*, la onzième du nom, reprend la ligne directrice des parutions antérieures : état de la population sur le plan démographique et celui de la formation, activité professionnelle, conditions de travail, santé, revenus et patrimoine, modes de vie, enfin solidarités collectives en distinguant entre liens sociaux et protection sociale. Un chapitre nouveau, intitulé « Dimension spatiale », accueille des études réalisées à la suite du recensement de la population de 1999. Signalons aussi, dans le chapitre « Liens sociaux », les articles sur les valeurs sociologiques et la religion, sujets qui n'avaient pas été abordés depuis de nombreuses éditions.

Autant que possible, les chapitres ont été organisés selon un schéma commun : d'abord un ou deux articles de cadrage, puis l'analyse de la politique collective et de ses moyens, ensuite celle des besoins et des risques, enfin l'étude de cas particuliers.

Au total, sont proposés près de 70 articles, largement étayés de données chiffrées présentées sous forme de tableaux, de graphiques ou de cartes. Le nombre des auteurs dépasse allègrement la centaine. Signe de l'ouverture croissante de *Données sociales* en direction des SSM et du monde universitaire, plus de la moitié de ces auteurs n'appartenaient pas à l'Insee.

Un chantier de trois ans

L'élaboration de *Données sociales* est une entreprise de longue haleine, qui mobilise nombre d'acteurs différents.

L'équipe rédactionnelle est celle qui ouvre le bal, lequel va s'étaler sur trois ans. Schématiquement, on peut distinguer trois étapes : d'abord la validation des grandes lignes de la nouvelle édition, ensuite la construction du sommaire, enfin la mise au point des articles.

Dans le premier temps, il s'agit de consulter les grands services contributeurs pour voir s'ils sont toujours disponibles ; s'ils ont des demandes à faire ou des modifications à proposer ; etc. Dans le cas de l'édition 2002-2003, c'est à ce stade qu'avait été décidée la création du chapitre « Dimension spatiale ». Une fois les grandes lignes ainsi arrêtées, il faut prendre contact avec les futurs auteurs pour avoir une idée des articles qu'ils proposent, et des échéances des premières versions. Un sommaire provisoire de la nouvelle édition est alors établi, et l'équipe de rédaction commence à s'inquiéter de la tenue des échéances. Naturellement, cette préoccupation ne fera qu'aller *crescendo*... La troisième et dernière étape recouvre un travail de rédaction conforme aux idées courantes : aménagement des articles en liaison avec leurs auteurs et d'autres intervenants.

Tout cela représente beaucoup de travail pour une équipe de rédaction de taille réduite (trois personnes pour l'édition 2002-2003 : Agnès Betton, Ravi Baktavatsalou et l'auteur de ces lignes), même si celle-ci ne manque pas d'alliés. En particulier, les hiérarchies des services auteurs participent à plusieurs titres à la mise au point des articles : relecture, proposition d'aménagements et validation (visa).

La phase technique inclut une mise en forme harmonisée des graphiques, la composition d'ensemble, la réalisation de la couverture et bien sûr l'im-

pression. Pour l'édition 2002-2003, ces travaux ont été respectivement assurés par la société AFDEC (M^{me} Marmouget), le secrétariat technique aux éditions de l'Insee (M^{me} Quesada et son équipe), la société Inéaiki et enfin la société Jouve.

Une exception culturelle française

Une référence importante pour *Données sociales* a été le *Social Trends* anglais, apparu au début des années soixante-dix, précisément au moment où le projet français prenait corps. Cependant, ces deux ouvrages, outre que l'un est annuel et l'autre (plus ou moins) triennal, présentent finalement des différences importantes qui renvoient au contraste entre recueil de chiffres et recueil d'analyses.

En effet, *Socials Trends* consiste pour l'essentiel en une mise à jour annuelle de séries statistiques, organisées en quinze grands domaines, à quoi s'ajoutent en tout et pour tout deux articles analysant des points particuliers, variables d'une livraison à l'autre. Dans *Données sociales*, rappelons-le, sont proposés plus de 50 articles (69 pour l'édition 2002-2003) : la plupart des thèmes étudiés sont certes repris d'une édition à l'autre, mais certains peuvent apparaître ou disparaître et tous les articles sont nouveaux à chaque édition. La nécessité de reprendre à zéro la rédaction de chaque édition explique le rythme, plus lent, de *Données sociales*. Une autre explication réside dans l'inertie du social : un article d'analyse (interprétation des tendances, recherche des facteurs, évaluation de mesures politiques) n'est pas obsolète sur un an.

De fait, dans le domaine du social, les ouvrages du type « recueil annuel de chiffres » sont les plus répandus, en Europe comme aux États-Unis. Chez nous, deux publications bien connues relevant de cette formule, mais qui ne se limitent pas au champ social, sont le *TEF* (Tableaux de l'économie française) et l'*Annuaire statistique de la France*, le premier à vocation grand public, le second tourné vers une clientèle plus professionnelle. On citera aussi *France portrait social*, où la part des tableaux est très importante.

Données sociales, dans sa formule actuelle façonnée par trente ans d'expérience, paraît bien une exception. C'est peut-être l'organisation de la statistique française qui explique cette originalité. Notre système statistique public, le SSP, est en effet moins centralisé qu'on ne le pense souvent, incluant, à côté de l'Insee, près de vingt services statistiques ministériels. En même temps, la quasi-totalité des personnels d'encadrement de ces services ont eux aussi été formés dans les écoles nationales de statistique rattachées à l'Insee (l'Ensaef, École nationale de la statistique et de l'administration économique, et l'Ensa, École nationale de la statistique et de l'analyse de l'information). D'où la richesse possible d'études réalisées dans une « langue commune » sur des sujets très variés et abordés sous des points de vue différents.

Philippe L'HARDY

Philippe L'Hardy a assuré pendant dix ans la rédaction en chef de Données sociales avant de prendre sa retraite en avril 2003.

Références

Dominique Deruelle : « En chantier : *Données sociales* 1980 », *Courrier des statistiques*, n° 9, janvier 1979.

Jean Bégué : « La sortie de *Données sociales* numéro 4 », *Courrier des statistiques*, n° 20, octobre 1981.

Annie Fouquet : « *Données sociales*, édition 1984 », *Courrier des statistiques*, n° 30, avril 1984.

Patrice Cazier : « Une comparaison entre *Données sociales* et quelques publications étrangères analogues », *Courrier des statistiques*, n° 36, octobre 1985.

Annie Fouquet, Maryse Marpsat : « *Données sociales*, édition 1987 », *Courrier des statistiques*, n° 42, avril 1987.

Insee, division des études sociales : Construire les *Données sociales*, *Les collections de l'Insee*, série M, n° 128, 1988 (fiche de lecture dans le n° 46, avril 1998, du *Courrier des statistiques*).

2002-2003

DONNÉES SOCIALES

La société française

INSEE
INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES
ÉCONOMIQUES

**STATISTIQUE
PUBLIQUE**

Tous les trois ans, depuis trente ans, « *Données sociales - La société française* » dresse un panorama rigoureux et attendu de la société française et de son évolution. Sous la plume d'une centaine de spécialistes des sciences sociales, de l'Insee et d'autres organismes, sont réunis les chiffres, les informations et les arguments qui permettent d'appréhender la réalité sociale d'aujourd'hui. L'impact de la réduction du temps de travail sur l'emploi, le bilan de l'enseignement secondaire des années quatre-vingt, les pratiques et croyances religieuses actuelles, les bénéficiaires de la couverture maladie universelle, l'évolution de la place des femmes dans le monde du travail, l'effet des innovations techniques sur l'emploi... sont quelques-uns des nombreux thèmes traités dans cet ouvrage. Chaque nouvelle édition de « *Données sociales - La société française* » est saluée par l'ensemble de la presse qui souligne que la richesse de son contenu, la fiabilité de ses données et la rigueur de ses analyses en font une référence unique.

ISSN 0758-6531
ISBN 2 11 06 8100 4
REF DONSOC02

Novembre 2002 - Prix 37,30 €

9 782110 681003

RÉFÉRENCES

Données sociales

Quel usage des catégories socioprofessionnelles ?

On choisit ici un angle d'attaque particulier : la part faite au classement socioprofessionnel¹ dans la dernière édition de *Données sociales* (2002/2003), riche de quelque 70 contributions organisées en 9 grands chapitres.

Pour faire rapide, nous nous sommes focalisés sur les seuls tableaux et graphiques, à l'exclusion de toute analyse textuelle.

Le simple dénombrement, chapitre par chapitre, des pages comportant au moins un tableau ou graphique se référant au classement socioprofessionnel nous a dès lors offert notre indicateur, exprimé ci-dessous en pourcentage de pages (tableau 1).

Le « maximum » est atteint pour le chapitre « emploi », ce qui a priori n'a

rien de surprenant. Que l'« éducation/formation » arrive juste après, devant les « salaires/conditions de travail », peut en revanche étonner : faut-il y voir une postérité de l'école de Bourdieu dans les milieux de la recherche sur l'éducation ? Enfin, nos deux zéros donnent bien sûr à réfléchir, tout particulièrement le premier : quid des inégalités sociales face à la santé, que l'on sait pourtant bien réelles ? Au total, tous chapitres confondus, on est à un peu moins de 6 % (37 pages sur 640).

Mais qu'en était-il « autrefois » ? La même méthode, appliquée à l'une des premières éditions de *Données sociales*, celle de 1978, remontant donc à un quart de siècle en arrière, conduit à des résultats sensiblement différents, en tout cas en valeur absolue² (tableau 2).

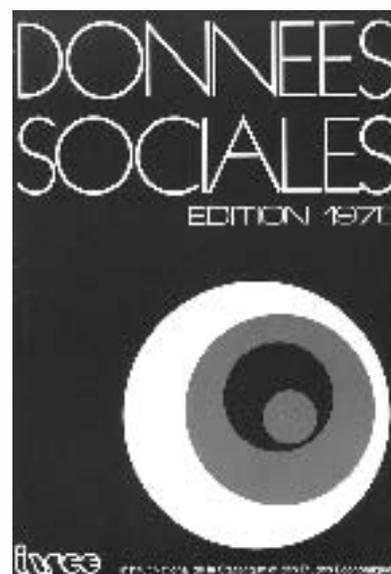


Tableau 1

1,7	Population, famille	1 page sur 60
9,8	Éducation, formation	10/102
12	Emploi	12/100
7,4	Salaires, conditions de travail	5/68
0	Santé	0/60
5,7	Revenus, patrimoine, conditions de vie	5/88
2,9	Dimension spatiale	2/70
4,5	Liens sociaux	2/44
0	Protection sociale	0/48
5,8	<i>Total</i>	37/640

Tableau 2

5,3	Population, santé, formation	2 pages sur 38
27,4	Activité professionnelle, conditions de travail	20/73
37,5	Prise en charge individuelle des besoins	30/80
4,4	Prise en charge collective des besoins	4/90
57,9	Socialisation et pratiques culturelles	22/38
39,3	Activité féminine et famille : aspects démographiques	11/27
33,3	Activité féminine et famille : aspects socioéconomiques	8/34
25,5	<i>Total</i>	97/380

1. La nomenclature de référence a d'abord été celle de 1954 avant l'entrée en vigueur, en 1982, de la PCS (professions et catégories socioprofessionnelles).

2. Encore faut-il préciser que nous n'avons pas tenu compte d'un dossier de 20 pages spécifiquement consacré à la condition sociale des petits paysans, dont on aurait très bien pu considérer qu'il se référait tout entier au classement socioprofessionnel.

Une approche par auteurs

Comment les auteurs de *Données sociales* 2002/2003, 117 au total, dont 47 étaient en fonction à l'Insee et 55 dans un SSM¹, les 15 autres provenant pour l'essentiel du monde universitaire, ont-ils manifesté leur sensibilité à la catégorie socioprofessionnelle en tant que variable descriptive de la société française ?

Le simple dénombrement des signataires des contributions comportant au moins un tableau ou graphique se rapportant à cette variable apporte une première réponse : 32 de nos auteurs, soit un peu plus du quart de l'effectif total, se sont peu ou prou référés au classement socioprofessionnel dans le fil de leur analyse.

Curieusement, cette proportion d'un peu plus d'un quart apparaît quasiment constante, quelle que soit l'origine des auteurs. Et, bien sûr, les résultats corroborent ceux observés au niveau thématique. La moitié des auteurs de l'Insee qui ont recouru à la catégorie socioprofessionnelle sont des contributeurs du chapitre relatif à l'emploi. De même, les deux SSM qui s'appuient le plus volontiers sur cette variable sont ceux des ministères du Travail et de l'Éducation nationale.

1. Les SSM sont les services statistiques des ministères.

De 26 % en 1978 à 6 % pour 2002/2003, la chute apparaît bien sévère. Quand on se rappelle que la nouvelle nomenclature PCS avait eu droit à un chapitre tout entier³ dans l'édition de 1984...

Nous voudrions ici suggérer qu'un groupe de recherche approfondisse le sujet, *via* une analyse détaillée couvrant l'ensemble des onze édi-

tions de *Données sociales* parues depuis 1973.

Naturellement, il ne s'agira pas seulement de retracer, thème par thème, l'évolution de la part faite au classement socioprofessionnel dans ces éditions successives. Il faudra aussi s'employer à « expliquer » les courbes afférentes, à la lumière des tournants éditoriaux qu'a connus la publication (cf. l'article précédent de Philippe L'Hardy) et des événements ou réflexions, externes ou internes, qui ont présidé à ces réorientations. Ainsi montrera-t-on à la communauté professionnelle en quoi et pourquoi

l'acuité de sa vision diminue régulièrement et quelle perte d'information et de compréhension en résulte pour la société française tout entière.

Disparition des oppositions de classe, entend-on à tout bout de champ... Et si la réalité était qu'on ne voulait plus voir, ni regarder, ni montrer ?

Bernard SUJOBERT

Insee, direction générale,
mission « conduite de l'action
régionale »

3. Intitulé « De nouveaux outils pour l'analyse socioprofessionnelle », ce chapitre avait été mis en musique par Alain Desrosières, Jean-Paul Girard et Laurent Thévenot.

HID-prisons

Leçons d'une enquête particulière

En 1997, Maud Guillonnet et Annie Kensey, toutes deux démographes à la Direction de l'administration pénitentiaire, constatent la pénurie d'informations sur l'état de santé des détenus, et soulignent en particulier cet oubli : « *Bien que les handicapés constituent une évidente catégorie qui a des besoins spécifiques, les données quantitatives les concernant ne sont pas disponibles.* » C'est de ce constat qu'est né le projet de conduire une enquête HID (handicaps-incapacités-dépendance) en milieu carcéral.

La 3^e enquête HID

L'enquête HID-prisons, dont la conception a été confiée à l'INED (Institut national d'études démographiques), s'inscrit dans le prolongement des enquêtes de même nature précédemment conduites par l'Insee auprès des personnes vivant en institutions socio-sanitaires (en 1998) ou en ménages ordinaires (1999).

L'objectif premier était de mesurer la prévalence des incapacités (d'origine physique ou psychique, mais aussi culturelle : difficultés de lecture, d'écriture, problèmes de langue...) en milieu carcéral, dont certains éléments indiquaient qu'elle pouvait être particulièrement élevée. D'abord, la population carcérale a connu un vieillissement important. Cette tendance tient, d'une part à une modification de la législation sur les crimes et délits sexuels, qui a fait entrer en prison des personnes d'âge mûr, et, d'autre part, à l'allongement de la durée des peines. Ensuite, les personnes incarcérées sont souvent issues de milieux populaires [Cassan, Kensey et Toulemon, 2000]. Or, comme l'avait bien démontré

Le musée national des prisons,
installé dans l'ancienne maison d'arrêt de Fontainebleau



Visite virtuelle : <http://www.justice.gouv.fr/musee/indexation>

l'enquête HID-ménages [Mormiche et groupe de projet HID, 2000], les inégalités sociales sont fortes en matière d'incapacités.

La présence en prison de personnes souffrant d'incapacités pose bien entendu la question de leur réinsertion, en même temps qu'elle interroge sur leurs conditions de vie. L'enquête, grâce aux modules du questionnaire consacrés à l'environnement familial, à la scolarité, à l'emploi et aux ressources, devait également être l'occasion de recueillir des éléments d'information sur ces différents aspects. *In fine*, l'apport de cette nouvelle enquête, relativement aux études portant spécifiquement sur la psychopathologie en milieu carcéral ou sur d'autres pathologies ou comportements très prévalents chez les détenus (SIDA, hépatite C, toxicomanie, consommation d'alcool), réside tout à la fois dans son

caractère généraliste et les rapprochements qu'elle allait permettre avec les précédentes enquêtes HID auprès des institutions socio-sanitaires et des ménages ordinaires.

Une entreprise collective

Le groupe de projet constitué par l'INED en vue de la mise au point de l'enquête HID-prisons réunissait des représentants de la DAP (Direction de l'administration pénitentiaire), de la DGS (Direction générale de la santé), de la DREES¹ et de plusieurs organismes de recherche, dont

1. Créée en 1998 au sein du ministère de l'Emploi et de la Solidarité, la DREES, Direction de la recherche, des études, de l'évaluation et des statistiques, est le « SSM » (service statistique ministériel) chargé de la santé et des affaires sociales. Elle est aujourd'hui rattachée à deux ministères : celui des Affaires sociales, du Travail et de la Solidarité et celui de la Santé, de la Famille et des Personnes handicapées.

l'Inserm (Institut national de la santé et de la recherche médicale) et le CNRS (Centre national de la recherche scientifique). L'Insee, riche de l'expérience acquise à l'occasion des enquêtes HID-institutions et HID-ménages mais aussi de l'édition 1999 de l'enquête Famille [Cassan, Héran et Toulemon, 2000], dont le champ avait été étendu à la population carcérale, a apporté son appui logistique et notamment son réseau d'enquêteurs. Outre la part de l'Insee, le montage financier rassemblait des contributions de la DAP, de la DREES et bien sûr de l'INED.

La mise au point du protocole d'enquête

Initialement, il était prévu de soumettre directement le questionnaire HID-prisons à un échantillon d'environ 700 détenus. Mais, compte tenu de l'incertitude sur la fréquence des situations d'incapacité en prison, on prenait le risque de ne toucher qu'un trop petit nombre de personnes souffrant effectivement d'une incapacité. Fallait-il donc faire précéder la « passation » du questionnaire HID d'une opération de filtrage visant à repérer les personnes souffrant d'incapacités, à l'image de ce qui avait été fait pour l'enquête HID-ménages² ? Et si oui, quand et comment administrer ce questionnaire de filtrage ?

Il a été décidé de procéder à un premier test, dont les résultats permettraient de se forger une opinion quant

2. Lors du recensement de la population 1999, un questionnaire spécifique « vie quotidienne et santé », en abrégé VQS, avait été distribué par un échantillon d'agents recenseurs à quelque 400 000 personnes. Voir Mormiche 2000.

3. Compte tenu de la grande mobilité de la population carcérale (taux de renouvellement mensuel de l'ordre de 15 %), il a été très vite admis que le laps de temps entre les deux opérations ne devait pas dépasser un mois.

4. Dans le cas de l'enquête Famille en milieu carcéral, les enquêteurs étaient très majoritairement des personnes ayant une expérience de la prison (visiteurs de prison, anciens avocats, chercheurs travaillant sur la population carcérale...).

5. Administration du questionnaire HID à toute personne répondant positivement à au moins une question de VQS-prisons.

La prison, une institution « presque » comme une autre

Afin d'assurer la comparabilité des résultats obtenus en milieu carcéral et en population générale, on a veillé à ce que le questionnaire HID-prisons soit aussi semblable que possible à ceux qui avaient été utilisés pour les précédentes enquêtes « handicaps-incapacités-dépendance ». De fait, il ne diffère guère du questionnaire HID-institutions. La prison, en effet, est « presque » une institution comme une autre, à ceci près, c'est une litote, que les personnes qui s'y trouvent n'ont pas la liberté d'aller et venir. La formulation des questions relatives aux déplacements a donc dû être adaptée ; d'autres comme celles relatives aux achats ou aux vacances ont bien sûr été supprimées.

à l'utilité et la faisabilité d'un filtrage préalable par le biais d'un questionnaire auto-administré, baptisé VQS-prisons (VQS signifiant « vie quotidienne et santé »). L'opération a été organisée en juin 1999 dans deux établissements pénitentiaires (la maison d'arrêt d'Osny et le centre de détention de Nantes). Les principaux enseignements ont été, d'une part que cette procédure n'était pas adaptée (les taux de réponse ont été très décevants, 26 % à Nantes et 32 % à Osny), d'autre part que la prévalence des incapacités était beaucoup plus élevée parmi les détenus que dans la population générale.

Dès lors, l'administration en face-à-face des questionnaires nous est apparue incontournable. Trois options restaient toutefois envisageables :

- passation directe du questionnaire HID : si la prévalence des incapacités est très élevée, l'opération de filtrage n'est plus pertinente ;
- administration d'un questionnaire de filtrage en face-à-face suivie trois semaines plus tard³ de la passation du questionnaire HID ;
- administration d'un questionnaire de filtrage en face-à-face et passation dans la foulée du questionnaire HID.

Décision a été prise de tester cette troisième option, qui, comparativement à la deuxième, présentait certes un inconvénient, celui de devoir choisir un critère de filtrage immédiatement discernable par les enquêteurs et donc très simple,

mais aussi et surtout de très nombreux avantages pratiques : mobilisation du personnel des établissements pénitentiaires pendant une période plus courte, pas de « pertes » entre le filtrage et la passation du questionnaire HID, moindre coût.

L'opération a été menée en janvier 2000 dans trois établissements pénitentiaires (les maisons d'arrêt de Rouen et d'Amiens et le centre de détention de Villenauxe-la-Grande). Au-delà de la question du protocole de l'enquête, elle devait bien sûr permettre de tester le questionnaire HID-prisons ainsi que la possibilité de confier à des enquêteurs de l'Insee la réalisation d'une enquête en milieu carcéral⁴. Les taux de réponse ont été plutôt satisfaisants (76 % à Amiens, 75 % à Villenauxe et 69 % à Rouen). Une nouvelle fois, il est apparu, certes au moyen d'un critère de filtrage assez peu sélectif⁵, que la prévalence des incapacités en milieu carcéral était extrêmement élevée (51 % à Amiens, 60 % à Villenauxe et 62 % à Rouen). On aura remarqué que c'est là où le taux de réponse est le plus élevé que la prévalence des incapacités apparaît la moins forte : ce constat nous a amenés à nous interroger sur la possible existence d'un biais de sélection. La confrontation des réponses données successivement aux deux questionnaires, VQS-prisons puis HID-prisons, a par ailleurs révélé un certain nombre de discordances, confirmant la nécessité de mettre en place un « échantillon témoin » de taille conséquente. Comme pour l'enquête HID-ménages, cet échantillon témoin devait répondre à un double objectif : comparer la situa-

tion des personnes souffrant d'incapacités avec celle des autres personnes et « récupérer » des personnes souffrant effectivement d'incapacités bien que non repérées comme telles par l'opération de filtrage (faux « négatifs »).

Ce test a également montré que les enquêteurs de l'Insee seraient parfaitement à la hauteur de la tâche et que le temps de passation du questionnaire HID-prisons, 32 minutes en moyenne⁶, était tout à fait acceptable.

Protocole final

En dépit de la forte prévalence des incapacités en milieu carcéral, nous avons décidé de maintenir l'opération de filtrage, dont le coût marginal apparaissait relativement peu élevé. La procédure sans filtrage aurait en effet réclamé d'administrer beaucoup plus de questionnaires témoins pour obtenir au final un même nombre de questionnaires « positifs ». Une telle approche aurait assurément rendu le travail des enquêteurs plus fastidieux, sans compter qu'elle aurait probablement affecté les taux de réponse.

Dans la pratique, l'ensemble des détenus formant l'échantillon de l'enquête se sont donc d'abord vu administrer le questionnaire VQS-prisons. Ceux ayant répondu positivement à l'une au moins des questions portant sur l'existence d'incapacités, de limitations d'activité, d'un handicap ou d'un besoin d'aide lié à un problème de santé (questions numérotées 3 à 17, cf. fac-similé du questionnaire pages 51 à 54) ont été invités à répondre au questionnaire HID, de même qu'un détenu sur trois, indépendamment des réponses données au questionnaire VQS, et ce aux fins de constitution de l'échantillon témoin.

Enfin, on l'a dit, compte-tenu du taux d'échec escompté, le risque qu'un effet de sélection vienne biaiser significativement les résultats était important. On pouvait craindre que les détenus en bonne santé, ne se sentant pas concernés par l'en-

quête, choisissent de ne pas y participer, et que ne soient pas non plus touchés ceux qui en raison de leur mauvais état général auraient des difficultés à se rendre au parloir. Pour contrôler ce risque de biais, un recueil d'informations complémentaires a été mis en place. Les médecins des établissements enquêtés ont accepté de procéder à une évaluation individuelle de l'état général de tous les détenus initialement sélectionnés, qu'ils aient ou non répondu par la suite au questionnaire VQS. Cette évaluation selon une échelle à quatre niveaux (excellent/bon/mauvais/très mauvais) a le plus souvent été effectuée sur la base de la consultation des dossiers médicaux. L'information ainsi recueillie, croisée avec celle par ailleurs disponible sur la participation à l'enquête, devait permettre de tester l'hypothèse d'un biais de sélection de l'échantillon.

Un échantillon constitué en deux temps

Le tirage des établissements a été effectué à partir de la statistique de la population pénitentiaire au 1^{er} avril 2000. La France métropolitaine comptait alors 174 établissements pénitentiaires, dont le plus petit ne comptait que 15 détenus alors que le plus grand en hébergeait plus de 3 000. Afin d'éviter un travail trop important aux médecins des UCSA⁷, il a été convenu qu'on ne devrait pas enquêter plus d'une centaine de détenus par établissement visité. Inversement, l'affectation souhaitable d'au moins deux enquêteurs à chacun de ces établissements impliquait un seuil minimal d'entretiens à réaliser par établissement. Nous avons *in fine* opté pour le tirage de 50 ou 100 détenus par établissement. Les établissements de moins de 50 détenus (ils étaient 9 au total) ont donc été exclus du champ de l'enquête. En revanche, nous y avons inclus d'office les 4 établissements comptant plus de 1 000 détenus, trois dans la région parisienne (Fresnes, La Santé et Fleury-Mérogis) et le quatrième à Marseille (Les

Baumettes), qui accueillent une population sensiblement différente du reste de la population carcérale en bien des aspects : structure par âge, par statut socioéconomique, proportion d'étrangers, nature des infractions commises, etc. Les autres établissements, dans la mesure où ils étaient situés dans les zones couvertes par les 9 directions régionales de l'Insee mobilisées pour la réalisation de l'enquête⁸, ont été répartis en 6 strates, par croisement de deux variables : d'une part le nombre de détenus présents dans l'établissement (trois modalités : petits établissements de 50 à 149 détenus, établissements moyens de 150 à 349 détenus et grands établissements de 350 à 999 détenus), d'autre part la catégorie d'établissement (deux modalités : maison d'arrêt, centre de détention ou maison centrale). On a ensuite procédé au tirage au sort avec des probabilités égales au sein de chacune des 6 strates ainsi constituées. Ont ainsi été sélectionnés 28 établissements, s'ajoutant aux 4 très grands établissements précités, tous 4 des maisons d'arrêt. Ces 28 établissements tirés au sort se répartissaient comme suit : 21 maisons d'arrêt, dont 7 petites, 8 moyennes et 6 grandes, et 7 centres de détention ou maisons centrales, dont 2 petits établissements, 1 moyen et 4 grands.

La sélection des détenus à enquêter dans chacun de ces 32 établissements, 2 800 personnes au total⁹ (soit 550 en petit établissement, 850 en établissement moyen et 1 400 en

6. L'enquête devait être entièrement réalisée au moyen de questionnaires papier, hors le dispositif CAPI (collecte assistée par informatique). Il n'était en effet pas envisageable d'autoriser l'introduction de micro-ordinateurs dans les établissements.

7. Unités de consultations et de soins ambulatoires.

8. Directions régionales d'Alsace, Champagne-Ardenne, Haute-Normandie, Île-de-France, Limousin, Nord-Pas-de-Calais, Pays de la Loire, Poitou-Charentes et Provence-Alpes-Côte d'Azur.

9. Soit un taux de sondage d'un peu plus de 6 %. Au 1^{er} mai 2001, la population carcérale entrant dans le champ de l'enquête s'élevait en effet à près de 44 000 personnes.

Il y a prison et prison

Les peines les plus courtes et les détentions provisoires dans l'attente d'un jugement définitif sont purgées dans les **maisons d'arrêt**. Les **centres de détention** reçoivent les condamnés à des peines d'une durée supérieure à 1 an et dont les perspectives de réinsertion apparaissent les meilleures. Enfin, les condamnés les plus difficiles sont pris en charge dans les **maisons centrales**, dotées à cette fin d'un régime sécuritaire renforcé. Ces différents régimes de détention peuvent coexister, sous forme de « quartiers », au sein d'un même établissement pénitentiaire.

SICOM/Ministère de la Justice



La prison de la Santé à Paris, construite en 1867 par l'architecte Vaudremer

grand ou très grand établissement), a quant à elle été assurée par les enquêteurs, le vendredi précédant la semaine de collecte, par tirage aléatoire sur la liste des détenus alors présents dans l'établissement¹⁰. Devaient être exclues du tirage au sort trois catégories de personnes. D'abord les mineurs, qui n'auraient pu être interrogés sans autorisation

parentale. Ensuite les détenus bénéficiant d'un régime de semi-liberté, qui n'auraient pu être joints aux heures autorisées de présence des enquêteurs. Enfin les détenus hospitalisés, car l'objectif était de repérer les incapacités chroniques. Or, à l'hôpital, les incapacités sont le plus souvent temporaires. Ce choix était en outre conforme à celui fait pour

l'enquête en population générale, puisque seules les unités de soin de long séjour des hôpitaux avaient été prises en compte dans l'enquête HID-institutions.

Les résultats de la collecte

La collecte s'est étalée du lundi 14 mai 2001 au vendredi 18 mai 2001. Elle a mobilisé 64 enquêteurs de l'Insee, 2 par établissement visité.

Sur les 2 800 personnes qui formaient l'échantillon de départ, 2 031 (1 951 hommes et 80 femmes) ont effectivement répondu au questionnaire VQS. Le taux de participation, 72,5 % en moyenne, s'est échelonné de 47 à 96 % selon les établissements. De façon générale, la participation a été plus élevée dans les petits établissements, où il est plus aisé de mobiliser le personnel et les détenus (82 %, contre 70 % dans les établissements de plus de 149 détenus). Le taux de refus s'est établi à 20,9 %, le taux d'échec pour autre motif¹¹ à 6,6 %.

Sur ces 2 031 répondants au questionnaire VQS, 950 ont répondu positivement à l'une au moins des questions 3 à 17 et ont ainsi été invités à répondre au questionnaire HID. Ont par ailleurs été sélectionnées 364 personnes supplémentaires, qui allaient constituer notre échantillon témoin. Au total, 1 314 personnes devaient donc répondre au questionnaire HID. À cette étape, on a enregistré 30 échecs, dont 13 refus de réponse et 17 cas d'inaptitude à répondre (il s'agissait le plus souvent de personnes qui ne maîtrisaient pas suffisamment bien le français).

Nombres d'établissements sélectionnés et de détenus enquêtés par taille d'établissement

	Établissements sélectionnés	Détenus enquêtés
Petits établissements	9	550
Moyens	9	850
Grands	10	1 000
Très grands	4	400
Total	32	2 800

10. Naturellement, chacune des personnes ainsi tirées au sort a ensuite reçu une lettre l'informant qu'elle avait été sélectionnée pour participer à l'enquête.

11. Le plus souvent parce que la personne venait d'être libérée ou transférée dans un autre établissement (61 cas) ou parce qu'elle avait été écartée a priori car jugée dangereuse (46 cas).

Participation	État de santé général					Total
	Excellent	Bon	Mauvais	Très mauvais	Inconnu	
Non	30,2 %	41,5 %	17,7 %	1,4 %	9,3 %	100,0 %
Oui	33,3 %	44,4 %	17,5 %	2,2 %	2,6 %	100,0 %
Total	32,5 %	43,6 %	17,5 %	2,0 %	4,4 %	100,0 %

Participation	État de santé général					Total
	Excellent	Bon	Mauvais	Très mauvais	Inconnu	
Non	24,9 %	25,5 %	27,1 %	18,9 %	56,3 %	27,5 %
Oui	75,1 %	74,5 %	72,9 %	81,1 %	43,7 %	72,5 %
Total	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %

État de santé général : pas de différence significative entre répondants et non-répondants...

Nous avons naturellement examiné comment se répartissaient les répondants et les non-répondants au questionnaire VQS selon leur état de santé général tel qu'évalué par les médecins des UCSA et, inversement, comment se répartissaient les personnes dont l'état de santé avait été jugé excellent, bon, mauvais ou très mauvais selon qu'elles avaient ou non participé à l'enquête. Cet examen (cf. tableaux ci-dessus) n'a révélé aucune différence significative entre la population des répondants et celle des non-répondants.

... mais de fortes variations entre les établissements

Sur le graphique 1, on a porté, pour chaque établissement, en abscisses le taux de participation à l'enquête et en ordonnées la proportion de personnes dont l'état général de santé a été jugé mauvais ou très mauvais par les médecins des UCSA. L'absence de lien entre ces deux grandeurs apparaît clairement, de même que la très forte variation d'un établissement à l'autre (de 3 à 58 %) de la proportion de détenus en mauvais ou très mauvais état de santé.

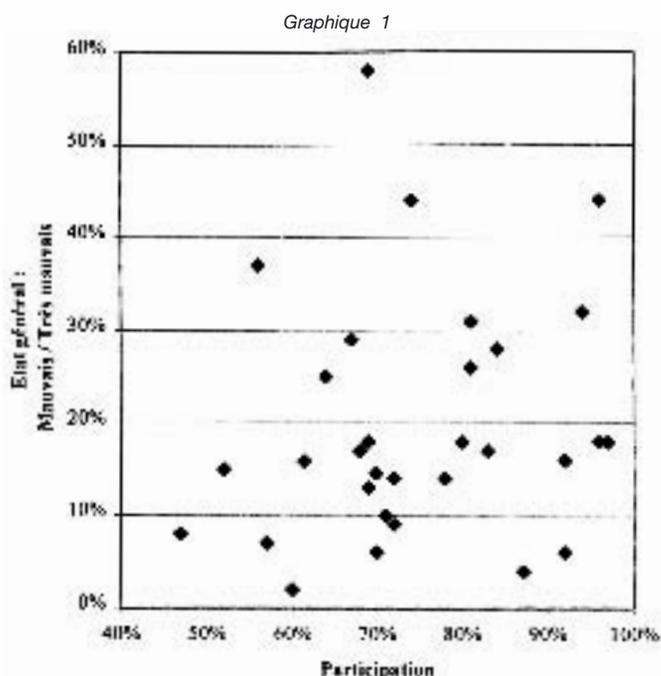
Une précédente étude sur la santé des entrants en prison [Mouquet, 1999] a mis en évidence de fortes disparités inter-établissements. Dans le cas présent, il est très vraisemblable que la dispersion observée ne

correspond que pour partie à des différences réelles et que la façon d'évaluer l'état général des personnes, plus ou moins optimiste selon les médecins, participe de cette variation, en l'atténuant ou en la renforçant. Les 32 établissements pénitentiaires enquêtés peuvent être schématiquement partagés en trois groupes : un groupe « moyen », formé de 11 établissements où plus des trois quarts des détenus ont été classés dans les 2 catégories centrales (bon/mauvais), et deux groupes « extrêmes », où l'échelle d'évaluation semble s'être déplacée, pour l'un vers l'excellent et pour l'autre vers le très mauvais. On conçoit sans peine que l'évaluation des médecins soit toute relative. De fait, dans un établissement où les détenus sont plutôt en bonne santé,

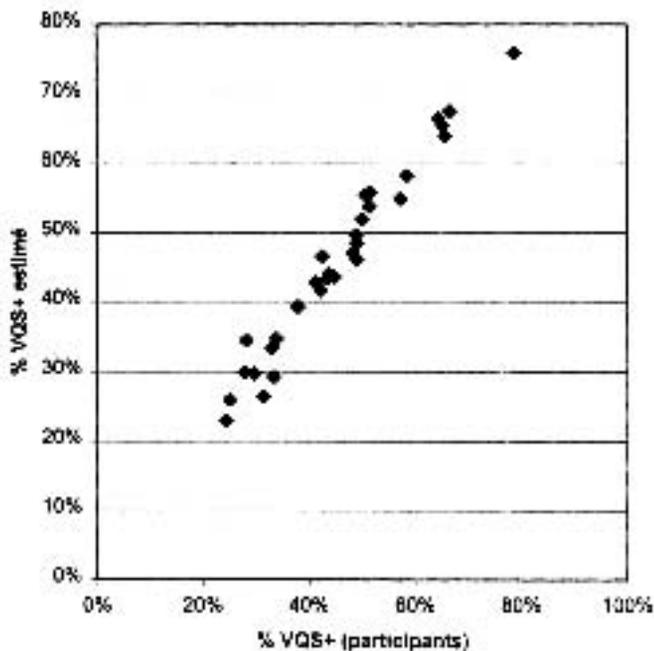
les critères retenus par le médecin pour considérer que l'un ou l'autre est très bonne santé sont sans doute plus exigeants. Ces considérations nous conduisent à la conclusion suivante : pour tester l'hypothèse d'un effet de sélection, on ne peut se contenter de comparer globalement l'état général des participants et des non-participants à l'enquête. Cette comparaison doit être effectuée au niveau de chaque établissement.

L'échantillon à la loupe

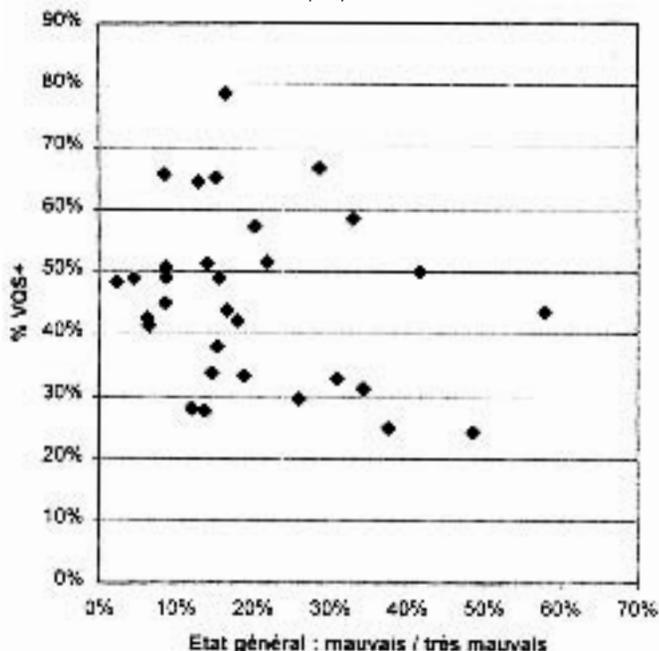
L'absence de différence significative entre l'état de santé général des répondants et celui des non-répondants, constatée au niveau de l'échantillon tout entier, est également vérifiée établissement par établissement.



Graphique 2



Graphique 3



Le graphique 2 en donne une illustration assez saisissante. On a porté en abscisses la proportion de personnes qui ont été « filtrées » par le questionnaire VQS, et, en ordonnées, ce qu'aurait été cette proportion si tous les détenus initialement tirés au sort avaient effectivement participé à l'enquête. Cette dernière grandeur a été estimée comme suit : supposons que, dans un établissement donné, la proportion de répondants filtrés par le questionnaire VQS soit comprise

entre la proportion de détenus jugés par le médecin en très mauvais état général et la proportion de détenus jugés en mauvais ou très mauvais état général. On peut alors déterminer les coordonnées de la première proportion par rapport aux deux autres (coordonnées barycentriques). En faisant l'hypothèse que ces coordonnées restent valides chez les non-répondants, on en déduit d'abord une estimation de la proportion de non-répondants qui auraient

été « filtrés », puis le résultat général incluant répondants et non-répondants. On voit sur le graphique que les points obtenus s'alignent presque parfaitement selon la première diagonale. Dans chacun des 32 établissements enquêtés, l'estimation de la proportion de personnes filtrées n'est donc quasiment pas affectée par la prise en compte des non-participants, et le résultat du filtrage VQS n'est pas biaisé par la non-réponse.

L'évaluation de l'évaluation

Pour être complet, il faut encore s'assurer que l'évaluation de l'état de santé général des détenus telle que l'ont opérée les médecins des UCSA constitue un bon « proxy » de la présence d'incapacités. Bien sûr, la lettre d'information qui avait été envoyée aux médecins pénitentiaires donnait des consignes en ce sens, mais il n'est pas sûr qu'il ait été aisé de les suivre. Le graphique 3 croise, établissement par établissement, le pourcentage de répondants qui ont été filtrés par le questionnaire VQS (ordonnées) et le pourcentage de ceux qui ont été jugés par le médecin en mauvais ou très mauvais état général de santé (abscisses). La belle relation linéaire espérée n'est malheureusement pas perceptible. Mais l'absence de corrélation entre les deux indicateurs observés n'est pas surprenante si l'on admet que l'évaluation des médecins a été opérée selon des modalités différentes d'un établissement à l'autre.

La spécificité des enquêtes en milieu carcéral

Le point crucial est sans doute celui de la participation des détenus, qui, on l'a vu, peut varier très fortement selon le protocole d'enquête retenu et selon les établissements. Si l'on a certes affaire à une population captive, elle n'en est pas pour autant docile. En outre, cette population se renouvelle très rapidement (cf. note de bas de page n° 3). Enfin, compte tenu des diverses activités auxquelles participent les détenus

(travail, sport, promenade...), il faut parfois faire preuve de persévérance pour « obtenir » un entretien. Dans ce contexte, il apparaît absolument indispensable de s'assurer une coopération optimale du personnel des établissements pénitentiaires. De ce point de vue, l'organisation de l'enquête HID-prisons, même si elle s'est révélée globalement très satisfaisante, aurait sans doute pu être meilleure encore si les contacts avec les établissements sélectionnés avaient été plus précoces et les enjeux de l'opération mieux explicités.

Une autre question importante est bien sûr celle de la sécurité : sécurité des enquêteurs, mais aussi difficultés d'organisation liées aux contraintes sécuritaires des établissements. En la matière, tout s'est passé pour le mieux, y compris dans la maison centrale que, grâce au très bon rapport de coopération établi avec la DAP, nous avons pu inclure dans l'échantillon de l'enquête. Notre regret, qui est aussi une leçon pour de prochaines enquêtes en milieu carcéral, est de ne pas avoir prévu davantage d'entretiens dans ce type d'établissements, suffisamment au moins pour pouvoir engager après

Garantir l'anonymat des informations recueillies

Dans chaque établissement, la gestion de la liste des personnes sélectionnées pour participer à l'enquête (cf. fac-similé page 50) avait été confiée à un surveillant. C'est ce dernier qui appelait les détenus au parloir des avocats, là où devait se dérouler l'entretien, et cochait, sur la partie de la liste prévue à cet effet (colonnes de droite), les cas de refus (ou d'impossibilité) de participation. De ce fait, lors de l'entretien, l'enquêteur ignorait l'identité de la personne interviewée (sauf bien sûr si cette dernière avait fait le choix de se présenter). Ce n'est qu'à la fin du dernier entretien que l'enquêteur rentrait en possession de la liste, afin de s'assurer de sa complétude et de l'exhaustivité de la collecte. Une fois ces vérifications achevées, il devait détacher la partie droite de la liste, relative à la participation à l'enquête, et l'envoyer à l'INED. La partie gauche a quant à elle été remise au médecin-chef de l'UCSA qui, une fois son évaluation effectuée, a détruit les données nominatives et envoyé le reste du formulaire à l'INED. L'appariement entre l'information sur l'état général et celle sur la participation à l'enquête a été effectué grâce au code-barre préimprimé sur le formulaire.

coup des analyses spécifiques. Les détenus en maison centrale forment en effet une population aux caractéristiques très particulières, tant sur le plan pénal que sur le plan socio-démographique.

Le mot de la fin sera pour les enquêteurs de l'Insee. Après ces deux expériences elles aussi particulièrement délicates (sujet « dur », milieu inhabituel, méthodes de collecte inédites) qu'avaient été l'enquête HID-institutions et l'enquête auprès des usagers des services d'aide aux

personnes en difficulté¹², l'enquête HID-prisons a confirmé de façon éclatante l'étendue de leur motivation et de leur talent.

Aline DÉSESQUELLES
Institut national
d'études démographiques

12. Cf. l'article de Cécile Brousse, Bernadette de la Rochère et Emmanuel Massé : « L'enquête de l'Insee auprès des usagers des services d'hébergement et de distribution de repas chauds : une méthodologie inédite pour étudier la population des sans-domicile », *Courrier des statistiques* n° 104, décembre 2002.

Références bibliographiques

- Cassan F., Héran F., Toulemon L.**, 2000, « Étude de l'histoire familiale : l'édition 1999 de l'enquête Famille », *Courrier des statistiques*, n° 93, mars 2000, pp. 25-37.
- Cassan F., Kensey A., Toulemon L.**, 2000, « L'histoire familiale des détenus », *Insee Première*, n° 706, avril 2000.
- Désesquelles A. et groupe de projet HID-prisons**, 2002, « Le handicap est plus fréquent en prison qu'à l'extérieur », *Insee Première*, n° 854, juin 2002.
- Guillonnet M., Kensey A.**, 1997, « La santé en milieu carcéral - Éléments d'analyse démographique », *Revue française des affaires sociales*, n° 1, janvier-mars 1997, pp. 41-60.
- Mormiche P.**, 1998, « L'enquête HID de l'Insee : objectifs et schéma organisationnel », *Courrier des statistiques*, n° 87-88, décembre 1998, pp. 7-18.
- Mormiche P. et groupe de projet HID**, 2000, « Le handicap se conjugue au pluriel », *Insee Première*, n° 742, octobre 2000.
- Mouquet M.-C.**, 1999, « La santé à l'entrée en prison en 1997 : un cumul des facteurs de risque », *DREES Études et résultats*, n° 4, janvier 1999.

Enquête VIE QUOTIDIENNE ET SANTÉ

N°ENQ :	Questionnaire destiné à l'Institut national de la statistique et des études économiques	N°VQS :	RÉPUBLIQUE FRANÇAISE
----------------	--	----------------	---------------------------------

Jour : **Heure de début de l'entretien VQS :**h **Heure de fin de l'entretien VQS :**h

Nom de l'établissement : **S'agit-il d'un individu témoin ?** Oui 1 Non 2

Cochez la case correspondant à la situation de la personne

1. Sexe	Homme <input type="checkbox"/> 1 Femme <input type="checkbox"/> 2
2. Année de naissance en →	1 2 _ _
3. Avez-vous des difficultés pour lire les caractères ordinaires d'un article de journal (avec des lunettes ou des verres de contact, si vous en portez habituellement) ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
4. Avez-vous des difficultés pour reconnaître le visage de quelqu'un dans la cour de promenade (avec des lunettes ou des verres de contact, si vous en portez habituellement) ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
5. Avez-vous des difficultés pour remplir seul (e) un formulaire simple (bon de cantine...) ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
6. Au cours d'une conversation avec des personnes parlant votre langue, avez-vous des difficultés pour parler et vous faire comprendre ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
7. Avez-vous des difficultés pour entendre ce qui se dit au cours d'une conversation avec plusieurs personnes ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2

Vu l'avis favorable du Conseil National de L'Information Statistique, cette enquête est reconnue d'intérêt général et n'a pas de caractère obligatoire.
 Label n° 2001 X 701 EC du Conseil National de L'Information Statistique, valable pour l'année 2001. Questionnaire confidentiel destiné à l'INSEE.
 La loi n° 78-17 du 6 janvier 1978, relative à l'informatique, aux fichiers et aux libertés, s'applique aux réponses faites à la présente enquête. Elle garantit aux personnes concernées un droit d'accès et de rectification pour les données les concernant. Ce droit peut être exercé auprès de la Direction Régionale de l'INSEE.

8. Lorsque vous êtes debout, avez-vous des difficultés pour vous pencher et ramasser un objet sur le sol ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
9. Avez-vous des difficultés pour mettre ou enlever vos vêtements ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
10. En raison de problèmes de santé, auriez-vous besoin de l'aide d'une autre personne dans la vie quotidienne ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
11. En raison de problèmes de santé, auriez-vous besoin d'aménagements particuliers dans votre cellule ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
12. En raison de problèmes de santé, utilisez-vous de façon régulière ou auriez-vous besoin d'une prothèse, d'un appareillage ou d'un dispositif technique (cane, béquilles, membre artificiel, fauteuil roulant, prothèse auditive...) ? <small>(Attention : ne pas tenir compte des lunettes, dentiers et autres prothèses dentaires)</small>	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
13. En raison de problèmes de santé, êtes-vous limité(e) dans les activités que vous pouvez faire? <small>Si non : passez directement à la question 16.</small>	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
14. Cette difficulté dure-t-elle depuis au moins six mois (ou pensez-vous qu'elle peut durer au moins six mois) ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2
15. Cette difficulté est-elle due à l'une des raisons de santé suivante ? (cochez la ou les bonnes réponses) : 1. Etat nerveux (anxiété, troubles du comportement) 2. Etat dépressif 3. Vieillesse 4. Douleurs importantes 5. Maladies chroniques 6. Troubles de la vision ou de l'audition 7. Malformation de naissance 8. Maladie génétique 9. Séquelles de maladie 10. Accident de la circulation routière 11. Autre accident 12. Autre raison de santé (préciser)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/> 11 <input type="checkbox"/> 12
16. Considérez-vous que vous avez un handicap ?	Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2

<p>17. Avez-vous fait ou a-t-on fait pour vous une demande de reconnaissance d'un handicap ou d'une invalidité (carte d'invalidité, pension, allocation, rente, capital) ?</p>	<p>Oui <input type="checkbox"/> 1 Non <input type="checkbox"/> 2 Ne sait pas <input type="checkbox"/> 3</p>
<p>18. Si oui, cette demande a-t-elle été acceptée ou refusée ?</p>	<p>Acceptée <input type="checkbox"/> 1 Refusée <input type="checkbox"/> 2 Pas encore <input type="checkbox"/> 3 de réponse/ ne sait pas</p>

Si la personne a répondu positivement à au moins une des questions précédentes ou si elle fait partie des témoins :

IAPTE. La personne est-elle apte à répondre au questionnaire ?

- 1. Oui 1 **→ ACCEPI**
- 2. Non 2

Si non, Pour quelle(s) raison(s) ?

- 1. Problème de santé, infirmité ou handicap 1 ; Préciser en clair :
- 2. Illégitimité 2
- 3. Problème de langue 3
- 4. Autre 4 ; Préciser en clair :

ACCEPI. Acceptez-vous de répondre à un questionnaire plus détaillé portant sur votre état de santé et ses conséquences sur votre vie quotidienne ?

- 1. Oui 1
- 2. Non 2

II.ANG. La personne parle-t-elle le français...

- | | | |
|---------------------------------------|--------------------------|---|
| 1. sans objet : personne muette | <input type="checkbox"/> | 1 |
| 2. sans difficulté | <input type="checkbox"/> | 2 |
| 3. avec quelques difficultés..... | <input type="checkbox"/> | 3 |
| 4. avec beaucoup de difficultés..... | <input type="checkbox"/> | 4 |
| 5. pas du tout..... | <input type="checkbox"/> | 5 |
| 6. Autre | <input type="checkbox"/> | 6 |

AUTRES OBSERVATIONS :

Une posture bayésienne pour le statisticien régional

Le premier, statisticien « fréquentiste », prétend ne rien savoir qui n'ait été issu d'une inférence statistique classique¹. Le second, qui fait profession d'expert, sait d'avance : nous l'appellerons *l'homme de l'art*. Nos deux personnages se heurtent bien souvent. Le second s'agace de l'ignorance affectée du premier, s'échauffe quand celui-ci, après moult investigations et calculs, lui exhibe une vérité qu'il croit avoir toujours sue. De son côté, le statisticien s'irrite au spectacle des querelles d'experts, explose quand est présenté comme donnée de référence ce qui ne lui paraît être qu'un pré-supposé, qui plus est infondé. Le terrain d'entente est pourtant tout trouvé. Les deux protagonistes, en effet, peuvent mettre à profit leurs compétences respectives dans le cadre d'une approche commune : l'approche bayésienne.

D'où viennent les connaissances *a priori* ?

Supposons que je veuille évaluer le nombre de personnes effectivement dotées d'un emploi dans une certaine commune. Je sais que cette commune comptait il y a trois ans un millier d'habitants, qu'elle s'insère dans une région où le taux d'activité s'élève à 40 %, et que le taux de chômage a été mesuré à 20 % dans une commune voisine. Comment, en l'absence d'autres informations, ne pas estimer *autour* de 320, soit $1\ 000 \times 0,4 \times (1 - 0,2)$, le nombre recherché ? Naturellement, je puis aussi m'employer à estimer les bornes inférieure et supérieure fixant l'ampleur de cet *autour*, et ainsi déterminer une « fourchette ».

Ce n'est pas autrement que procède l'homme de l'art : ses connaissances *a priori*, les « ordres de grandeur » à partir desquels va se forger son opinion d'expert, sont le plus souvent le résultat d'une extrapolation basée sur des données statistiques. Une différence majeure, toutefois, est qu'il ne dispose pas nécessairement de l'ensemble des données auxquelles j'ai accès et que, en principe, « sa » fourchette sera l'expression d'une intuition plus ou moins raisonnée et non le résultat d'un calcul statistique conduit selon les règles du genre. C'est pourquoi, selon nous, le statisticien, même s'il est bien sûr dans son rôle quand il rappelle la nécessité absolue de l'expérience (l'enquête), ne doit pas s'interdire de prêter son concours à l'élaboration des informations *a priori* sur lesquelles s'appuie l'homme de l'art. Or, trop souvent, il a de sa mission une conception restrictive, se refusant à procéder à des estimations qu'il est pourtant, sinon le seul à pouvoir effectuer, du moins « le mieux placé pour les faire le moins mal ».

La posture bayésienne

Pour l'homme de l'art, l'enquête, l'inférence statistique, n'ont d'utilité que si elles doivent améliorer ses connaissances *a priori*, si donc elles vont permettre de préciser la fourchette. Ce sera le cas, par définition, d'une inférence bayésienne. Si en effet l'inférence classique s'appuie sur les seules données observées, l'inférence bayésienne, quant à elle, s'appuie également sur les informations *a priori* disponibles.

La « posture » bayésienne que nous tentons de dessiner ici suppose naturellement une grande proximité,



Thomas Bayes, 1702-1761

non synonyme de dépendance, avec l'homme de l'art. Toute la difficulté, en effet, réside dans l'évaluation de la pertinence des informations *a priori* sur lesquelles on entend s'appuyer. Et si le résultat de l'enquête vient semer le trouble, parce que très peu probable si les informations *a priori* sont « vraies » (voir page 57), devra alors s'engager un dialogue complémentaire, pas toujours commode à cadrer. D'autre part, le statut des estimations bayésiennes (qui résultent de la combinaison d'informations *a priori* détenues au moins pour partie par l'homme de l'art, de l'enquête effectuée par le statisticien, et de la procédure d'estimation bayésienne proprement dite) n'est pas simple : il est

1. L'inférence statistique consiste, en substance, à élaborer des conclusions applicables à une population tout entière à partir d'observations recueillies auprès d'une fraction seulement de cette population. Elle est dite « classique » si, à la différence de l'inférence bayésienne, elle s'appuie uniquement sur les observations en question, hors tout *a priori*.

celui d'une propriété conjointe, ce qui n'est pas dans les habitudes de la statistique publique.

Un cas qui n'est pas de pure école : la statistique régionale et locale

Le statisticien régional se trouve souvent dans une conjoncture favorable pour « expérimenter » la posture bayésienne. D'une part, la rareté des données statistiques de niveau régio-

nal ou local impose certes d'en produire de nouvelles, mais incite avant tout à ne négliger aucune information déjà existante. D'où l'intérêt d'un dialogue avec les acteurs locaux en vue de déterminer conjointement l'information *a priori* disponible. Il sera d'autant plus fécond que la trame spatiale permet de structurer l'information en question, pour ainsi dire de l'incarner. Ainsi, nombre d'énoncés géographiques en forme de slogan (« le carrefour avignonnais », ou bien « les percées alpines », ou encore

« Annemasse, banlieue de Genève ») recèlent en réalité toute une batterie de présomptions (idée voisine de celle de probabilité *a priori*) englobant des domaines très variés, de la délinquance à la sociologie urbaine en passant par les complémentarités économiques ou sociales, la logistique, l'environnement... Le phénomène d'*autocorrélation spatiale*² vient renforcer l'efficacité de cette « fixation » géographique de l'information *a priori*.

En fait, le statisticien, habitué à fréquenter des espaces abstraits à *n* dimensions, sous-estime les facilités qu'autorise notre bon vieil espace géographique³. Par exemple, le succès de la cartographie statistique tient en grande partie au rapprochement, bayésien dans son principe, qu'elle permet d'opérer entre l'information nouvelle et la riche information *a priori* que détient déjà l'utilisateur éclairé, tirant parti de l'autocorrélation aussi bien que des corrélations entre variables distinctes en un même lieu.

Ainsi, au-delà de la stricte méthode bayésienne, un vaste domaine est à occuper et structurer par une *posture bayésienne* en un sens large, qui confère au statisticien un rôle bien défini dans un processus cognitif collectif. Le champ de la statistique régionale et locale apparaît un de ceux où cette démarche est la plus riche de promesses.

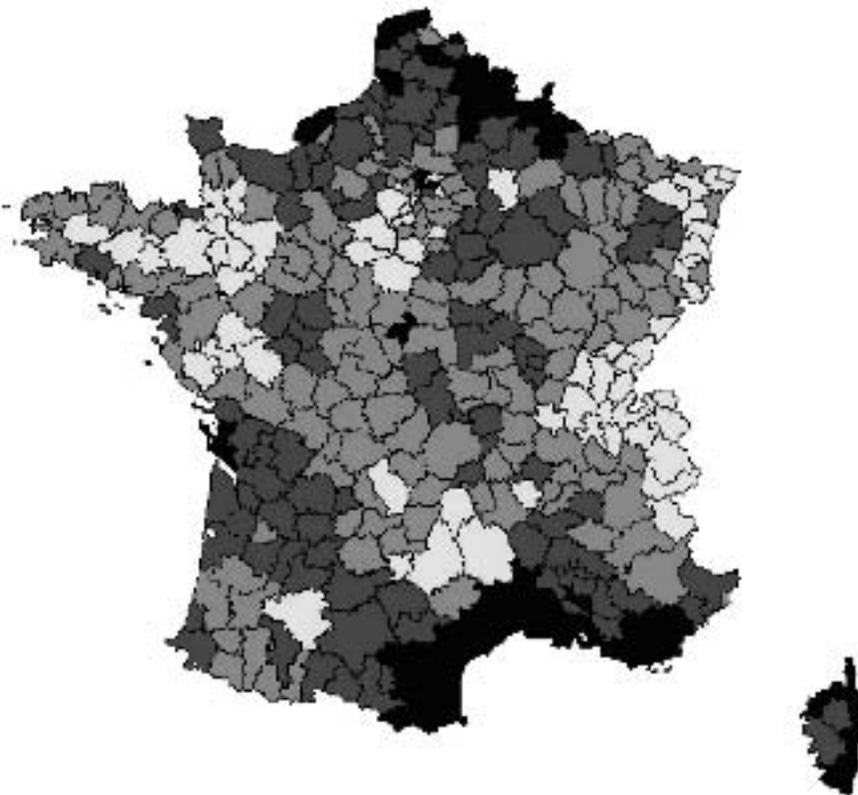
Jean LAGANIER

Insee
Direction régionale
de Provence-Alpes-Côte d'Azur

Une illustration de la notion d'autocorrélation spatiale

La carte ci-dessous représente les taux de chômage par zone d'emploi. On y constate visuellement le phénomène d'autocorrélation spatiale des taux de chômage : chaque zone d'emploi tend à présenter un taux voisin de celui des zones proches. La carte présente ainsi un nombre limité de « taches » de même couleur.

Les mécanismes sous-jacents à l'autocorrélation spatiale, notion purement statistique, peuvent être très divers. Dans le cas présenté, il s'agit notamment d'un mécanisme de mobilité des travailleurs. Si une très forte différence de taux de chômage existe entre deux zones voisines A et B, A présentant le taux supérieur, cela induit souvent la recherche par les résidents de A d'emplois dans la zone B : A « exporte son chômage » vers B, ce qui tend à rééquilibrer les taux.



Taux de chômage par zone d'emploi
Source : RP 99

Tranches de valeurs, du plus clair au plus foncé : 5,4-9,8 ; 9,8-12,6 ; 12,6-16,1 ; 16,1-23,2

2. Le niveau atteint par une variable en un certain lieu est souvent dépendant de la façon dont évolue cette variable en un lieu voisin. Sont concernées par ce phénomène d'*autocorrélation spatiale* de nombreuses variables, démographiques, sociales et économiques.

3. L'obstination qui a été mise à découper l'espace géographique selon des « zonages » ou « nomenclatures spatiales » est selon nous significative d'un certain refus de la naturelle « simplicité » de cet espace, d'une volonté de le « traiter » à l'instar des espaces abstraits de l'économie et de la société.

**Présentation d'un cas :
opportunité d'une enquête en présence de connaissances *a priori***

Dans cette région, on sait de façon certaine qu'il y a 10 000 entreprises. L'agence pour l'innovation estime que, sur cet effectif, le nombre θ d'entreprises innovantes se situe entre 400 et 600. Elle voudrait confirmer et si possible préciser cette estimation.

Un premier bureau d'études propose une enquête, et une méthode d'inférence classique. Le budget disponible limite l'échantillon à 200 entreprises. La proposition donne l'estimation et l'intervalle de confiance à 90 % au moins (arrondi) correspondant au nombre x d'entreprises innovantes qui seront identifiées dans l'échantillon (tableau 1). Au vu du tableau, l'agence ne donne pas suite. Elle constate en effet qu'en aucun cas l'enquête ainsi exploitée ne pourra même confirmer son évaluation *a priori* !

Un autre bureau d'études lui propose alors une méthode bayésienne. L'enquête est identique et pratiquée également auprès de 200 entreprises. D'un commun accord, la probabilité *a priori* est prise de type « biweight », soit de la forme $\Pr(\theta = t) = 0,15 (t - 400)^2 (600 - t)^2 / 40000^2$ sur l'intervalle [400, 600]. La mise en œuvre de la méthode bayésienne permet alors d'estimer θ par la médiane de sa distribution *a posteriori* et de donner un intervalle de confiance à 90 % au moins selon les valeurs trouvées de x (tableau 2). Cette proposition est jugée plus intéressante par l'agence. Le bureau d'études précise cependant qu'en cas de résultat x peu probable (voir les probabilités marginales de x dans la dernière colonne), il ne livrerait les résultats que sous toutes réserves, car cela jetterait un doute sérieux sur la loi *a priori* retenue.

Tableau 1

x	θ estimé	intervalle de confiance	
0	0	0	0
1	50	0	100
2	100	0	200
3	150	0	300
4	200	0	350
5	250	50	450
6	300	50	500
7	350	100	550
8	400	150	650
9	450	200	700
10	500	200	750
11	550	250	800
12	600	300	900
13	650	350	950
14	700	350	1 000
15	750	400	1 050
16	800	450	1 100
17	850	500	1 150
18	900	550	1 250
19	950	600	1 300
20	1 000	600	1 350
...
200	10 000	10 000	10 000

Tableau 2

x	θ estimé	intervalle de confiance			Pr (x)
0	468	422	536	0,00	
1	470	423	539	0,00	
2	473	424	542	0,00	
3	476	425	545	0,01	
4	480	427	548	0,02	
5	483	428	550	0,04	
6	486	430	553	0,06	
7	490	432	555	0,09	
8	493	434	558	0,11	
9	496	436	560	0,12	
10	500	438	562	0,12	
11	503	440	564	0,11	
12	506	443	566	0,09	
13	509	445	568	0,07	
14	513	448	569	0,05	
15	516	451	571	0,03	
16	519	454	572	0,02	
17	521	457	574	0,01	
18	524	459	575	0,01	
19	527	462	576	0,00	
20	529	465	577	0,00	
...	
200	593	583	599	0,00	

Au sujet du tableau 2

Connaissant la loi *a priori* $\Pr(\theta = t)$ et la loi (binomiale) $\Pr(x/\theta)$ du tirage aléatoire, on en déduit, par application du théorème de Bayes (voir *Courrier des statistiques* n° 100, décembre 2001), la probabilité *a posteriori* $\Pr(\theta = t / x)$, ce qui permet de déterminer l'estimation centrale de θ par la médiane de cette probabilité *a posteriori*, ainsi que l'intervalle de confiance. On en déduit aussi la probabilité marginale de x , dans l'hypothèse bien sûr où la loi *a priori* est valide.

À lire

« On peut tout montrer avec des statistiques... sauf bien sûr l'essentiel. » Il est vrai qu'il est facile d'ironiser sur les chiffres, sur leur propension à mentir et à se contredire. Pourtant, nous le savons bien, la production d'une information chiffrée de qualité, qui viendra éclairer la prise de décision, est une nécessité absolue dans une société démocratique évoluée¹. Aussi l'effort pédagogique doit-il être permanent afin de mieux faire comprendre la méthode statistique, véritable boîte noire pour beaucoup de nos concitoyens. C'est dans cet esprit que notre collègue Bernard Aubry, de la direction régionale de l'Insee-Alsace, a rédigé ce petit livre, publié en novembre 2002 aux éditions Ellipses.

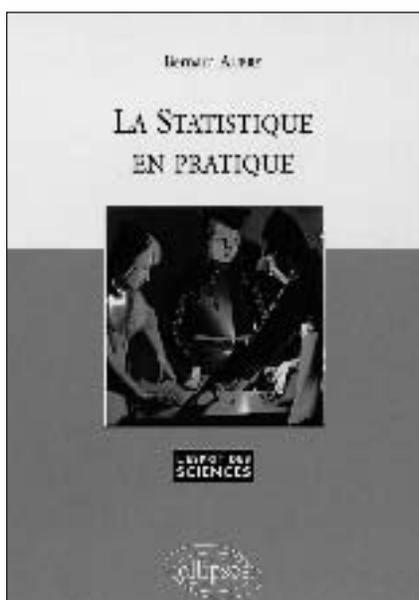


Illustration de couverture :
Georges de La Tour,
Les joueurs de dés (détail)

Le dé pour apprivoiser le hasard

Le point de départ de *La statistique en pratique* est l'observation des résultats d'une expérience très simple : on lance un dé 120 fois, on examine la distribution des résultats obtenus et on la compare à la distribution théorique, fondée sur la propriété de symétrie du dé. Ce premier chapitre est l'occasion de préciser les concepts et le vocabulaire de base (hypothèse, épreuve, loi de probabilité...) et de présenter les indicateurs statistiques classiques. L'accent est mis sur la *répétition* des épreuves qui, d'une loi quelconque, nous conduit à la loi normale, et une bonne place est donnée à la réflexion sur le hasard, omniprésent dans notre vie quotidienne comme dans l'approche scientifique. Quand on dit qu'« *il ne faut rien laisser au hasard* », c'est bien sûr qu'on redoute que le *sort* ne soit pas favorable. Mais une autre locution, tout aussi courante, est qu'« *il faut faire confiance au hasard* » : quand il ne sait pas décider, l'homme s'en remet au hasard. Ainsi la théorie des sondages a-t-elle reconnu, depuis les débats qui avaient agité la communauté statistique au début du XX^e siècle, la suprématie du tirage aléatoire sur les méthodes de sélection empiriques, du moins pour les échantillons de grande taille.

1. L'une des conditions requises pour l'entrée dans l'Union européenne n'est-elle pas la capacité du pays candidat à mettre en œuvre un appareil statistique performant ?



Montesquieu, 1689-1755
« Le suffrage par le sort est de la nature de la démocratie. »

Aux confins de la science

L'esprit peut-il agir sur la matière ? Sommes-nous influencés par les astres ? L'auteur, sans se prononcer sur la validité des expériences en question, rappelle que dans les années 1930, Joseph B. Rhine, de l'université de Durham (Caroline du Nord, USA), puis d'autres dans son sillage, auraient scientifiquement mis en évidence la capacité de certains sujets à « tromper le hasard » (cas par exemple d'un sujet obtenant des résultats sortant largement de l'intervalle de confiance à la recherche du 6 au lancer de dé). Lui-même s'emploie, à partir du recensement de la population, à rechercher des traces éventuelles de la réalité de l'astrologie dans la distribution de deux variables, la profession exercée et le signe zodiacal de la personne avec qui l'on partage sa vie. Ainsi la diagonale apparaît-elle sur-représentée dans un tableau donnant la répartition des couples selon le signe zodiacal des deux partenaires.

Serait-on davantage prédisposé à s'attacher à une personne de même signe ? En poussant un peu plus l'investigation, on observe que les « jumeaux astraux » (nés le même jour du même mois) ont tendance à s'unir de façon privilégiée : forcément, le statisticien aguerri subodorera une explication prosaïque, liée à des questions de collecte, de saisie ou de redressement... La leçon est que d'apparentes connexions statistiques ne sauraient démontrer la réalité de tel ou tel phénomène : elles invitent seulement à poursuivre plus loin l'analyse.

De la distribution à la loi

Dans le chapitre intitulé « L'éventail des salaires », Bernard Aubry va droit à l'essentiel, offrant en une petite quinzaine de pages une présentation lumineuse des principaux instruments de l'analyse statistique : histogramme, fonctions de distribution et de répartition, indicateurs de valeur centrale, de dispersion et de concentration... Il s'attarde sur les difficultés de mesure et les erreurs à ne pas commettre dans l'interprétation des résultats obtenus, et, naturellement, nous invite à passer *de la distribution à la loi*. Dans le cas du lancer de dé, l'exercice n'était pas bien difficile... S'agissant d'une distribution de salaires, il faut découvrir l'« *ordre qui se cache* » derrière les chiffres, l'ordre donc la loi, formalisable d'une façon plus ou moins compliquée.

2. Rappelons à ce sujet que beaucoup de joueurs choisissent des numéros correspondant à des quantités ou mois de naissance.

SOMMAIRE	
Introduction	3
Avant-propos	7
I. La loi binomiale et l'hypergéométrique	11
II. La loi de Poisson	19
III. La loi normale	25
IV. La loi de Laplace	25
V. La loi de Cauchy	25
VI. La loi de Student	25
VII. La loi de Fisher	25
VIII. La loi de Rayleigh	25
IX. La loi de Weibull	25
X. La loi de Gumbel	25
XI. La loi de Pareto	25
XII. La loi de Zipf	25
XIII. La loi de Benford	25
XIV. La loi de Benford	25
XV. La loi de Benford	25
XVI. La loi de Benford	25
XVII. La loi de Benford	25
XVIII. La loi de Benford	25
XIX. La loi de Benford	25
Annexes	21
Bibliographie	25

Le Loto

L'avantage, avec le Loto, est qu'il permet de présenter l'essentiel des lois de probabilité élémentaires : la loi binomiale, la loi hypergéométrique, la loi de Poisson, celle de Pascal, la loi normale et celle du khi-deux. L'auteur s'attache d'abord à nous montrer, sur la base de l'ensemble des tirages opérés depuis l'origine du jeu, que les résultats sont bien conformes à l'hypothèse d'équiprobabilité de sortie de chacun des 49 numéros. Vient alors le moment de se pencher sur les distributions de gains, d'où il ressort que les numéros, s'ils sont égaux devant le hasard, ne sont pas égaux en termes de gain procuré. En effet, partage du pactole oblige, les numéros les plus joués² engendrent un moindre gain. Enfin, Bernard Aubry nous enseigne comment, à l'aide de l'informatique, simuler en vraie grandeur une édition du jeu de Loto avec détermination des gains de chaque niveau (de 3 numéros sans le complémentaire

jusqu'aux heureux gagnants du premier rang, ceux qui ont rempli une grille comportant les 6 bons numéros). La méthode présentée est voisine de celle mise en œuvre pour la détermination des intervalles de confiance dans le cas d'un plan de sondage complexe.

La météo

Les statistiques collectées par les services de Météo France offrent un corpus d'informations considérable et qui pourrait aussi servir à des usages pédagogiques, notamment autour de la notion de corrélation. Premier aspect du problème, qui renvoie aux distributions de salaires : quels sont les indicateurs pertinents ? Et l'on remarque au passage que la distribution des températures moyennes est parfaitement « normale » alors que l'on s'accorde volontiers à dire qu'il n'y a plus de saisons.

Un livre grand public

Dans un ouvrage de taille aussi réduite (128 pages), il est évident que certains aspects importants de la méthode statistique ne peuvent qu'avoir été esquissés. De fait, on retient surtout le souci de l'auteur de montrer ce qui peut réunir des sujets a priori aussi différents que le Loto, la météo, l'astrologie, la distribution des salaires et la théorie des sondages. C'est précisément pourquoi ce bien joli petit livre, précieux outil d'initiation à notre discipline, s'adresse plus largement encore à tous ceux que fascinent les chiffres et la loi des grands nombres.

Le Courier