

## **Titre: IPUMS et AICMD apportent une valeur significative aux microdonnées de recensements africains.**

### **Auteurs:**

**Robert McCaa** (rmccaa@umn.edu), Professeur Emérite d'Histoire de la Population, Ambassadeur, IPUMS-International, Centre de Population du Minnesota

**Patricia Kelly-Hall** (pkelly@umn.edu), Coordinatrice, IPUMS-International, Centre de Population du Minnesota

### **Sites Web:**

[www.ipums.org/international](http://www.ipums.org/international)

<http://ecastats.uneca.org/aicmd/Home/tabid/40/language/en-US/Default.aspx>

### **«Diffusion [signifie] ouverture de la valeur inhérente de non données»**

Séminaire sur les Tendances Emergentes en Transmission de Données et en Statistique, New York le 19 février 2010

Walter Radermacher (Président, Eurostat) et Pieter Everaers (Directeur, Eurostat)

### **Introduction.**

Les chercheurs et les décideurs africains doivent avoir accès aux microdonnées en vue d'analyser les défis qu'impliquent les transformations démographiques et économiques en cours sur toute l'étendue du continent (McCaa, Esteve, Ruggles et Sobek 2006). IPUMS-International ([www.ipums.org/international](http://www.ipums.org/international)) propose aux agences statistiques de par le monde entier une solution unique pour la diffusion d'échantillons rendus anonymes de microdonnées. Le projet assume les responsabilités, les risques et les coûts liés à la récupération, l'archivage, la préservation de l'anonymat, l'intégration et la diffusion de microdonnées à l'échelle mondiale (McCaa et Thomas 2009). Les chercheurs et les décideurs accèdent aux données sans aucun frais quel que soit leur pays de naissance, de résidence ou de nationalité, grâce à un financement soutenu par la Fondation Nationale pour la Science et l'Institut National de la Santé (États-Unis). Commencée en 1999, IPUMS-International bénéficie de la satisfaction de plus de 5000 chercheurs dans une centaine (100) de pays et à travers plus de 90 bureaux nationaux de statistique partenaires.

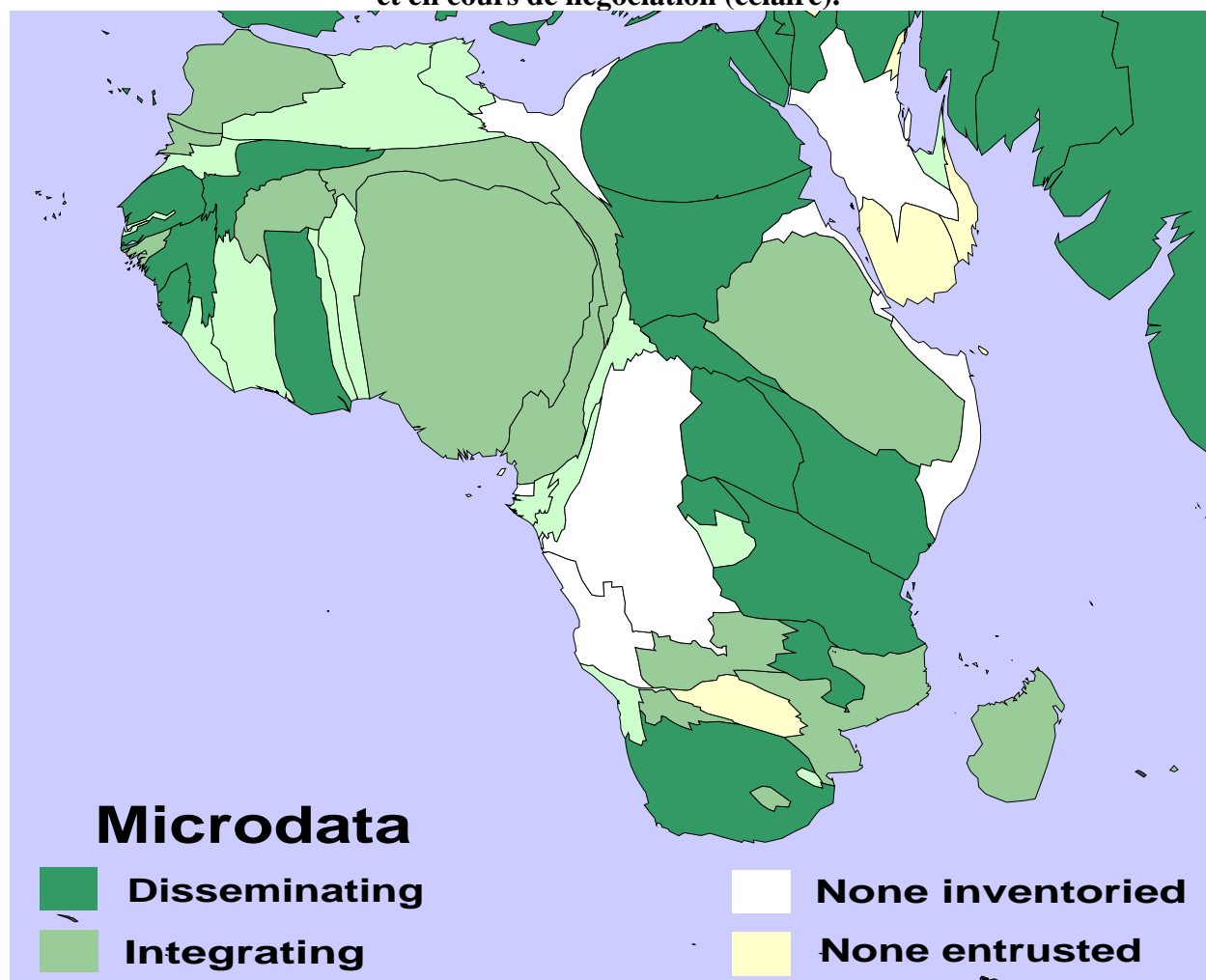
A la date de Juin 2011, le projet international a achevé l'anonymisation et l'intégration des microdonnées de population pour 62 pays dans le monde, totalisant 185 échantillons et 397 millions d'enregistrements individuels. Grâce à la coopération extrêmement généreuse des ONS, la base de données IPUMS-international est en pleine expansion au taux de 5-10 pays supplémentaires par an. En 2015, il est probable que la couverture s'étendra à 85 pays ou plus, englobant plus des quatre-cinquième de la population mondiale.

Le projet est dirigé par le Centre de Population du Minnesota (MPC). En Afrique, le Centre Africain de Statistique (AICMD) a établi un partenariat avec le MPC dans une initiative collaborative panafricaine regroupant des Offices Nationaux de Statistique, des universités et les principaux centres de recherches (<http://ecastats.uneca.org/aicmd/Home/tabid/40/language/en-US/Default.aspx>). Vingt-neuf Offices Nationaux de Statistique de pays africains, englobant 85% de la population du continent, ont confié des microdonnées de recensement et d'enquêtes nationales à IPUMS sous un protocole uniforme d'accord (annexe A, figure 1 et tableau 1). L'accès est régi par un accord de licence rigoureux et strict (Annexe B).

**Éviter la marginalisation; confier les microdonnées du cycle de recensement de 2010 sans délai.** Le Journal de la Statistique Officielle, dans un numéro spécial sur l'accès et la confidentialité des microdonnées, a averti qu'"il existe un risque réel que les ONS peuvent être marginalisés si elles sont incapables de répondre aux besoins des chercheurs dans ce domaine de plus en plus important de fourniture de [micro]données"(Editions, 2010:55). Les agences africaines de statistique qui ne coopèrent pas actuellement avec les initiatives IPUMS/ AICMD sont invitées à contacter les auteurs de ce papier. Les bureaux de statistique participants sont quant à eux encouragés à confier les microdonnées du cycle des recensements de 2010 sans plus tarder en vue de faciliter l'accès en temps opportun aux échantillons intégrés rendus anonymes. Ce faisant, cela va augmenter les parties prenantes et, selon les propos du Président d'Eurostat (voir citation plus haut), "élargir la valeur inhérente" aux microdonnées de recensement.

**Figure 1. Cartogramme pondéré par la taille de la population des pays africains participant dans le collaboratoire IPUMS-International, Novembre 2011:**

**3 nuances de vert - intégré (sombre),  
en cours d'intégration (moyen),  
et en cours de négociation (éclairé).**



Les cinquante pays africains possédant un million d'habitants ou plus peuvent être divisés en quatre groupes, selon la participation à l'initiative IPUMS/AICMD :

**a. 14 avec microdonnées en cours de diffusion** (nombre de bases de microdonnées entre parenthèses) : Egypte (2), Ghana (2), Guinée (2), Kenya (2), Malawi (3), Mali (2), Rwanda (2), Sénégal (3), Sierra Leone (1), Afrique du Sud (3), Sud-Soudan (1), Soudan (4), Tanzanie (2), et Ouganda (2).

**b. 15 avec microdonnées en cours d'intégration** : Botswana (3), Burkina Faso (3), Cameroun (3), Cap-Vert (2), Tchad (2), Ethiopie (3), Lesotho (2), Madagascar (1), Maurice (2), Maroc (3), Mozambique (2), Niger (2), Nigeria (2 enquêtes générales auprès des ménages), Togo (2) et Zambie (2).

**c. 15 ne participent pas**, en dépit d'une campagne persistante qui remonte à 2005: Algérie, Bénin, Burundi, République centrafricaine, Comores, République du Congo, Côte d'Ivoire, Gabon, Gambie, Libéria, Mauritanie, Namibie, Swaziland, Tunisie, et Zimbabwe.

**d. 6 n'ont pas, à l'heure actuelle, de microdonnées de recensement** : Angola, République démocratique du Congo, Djibouti, Erythrée, Libye et Somalie.

Le lancement IPUMS de 2012 prévoit intégrer des échantillons pour le Maroc (1982, 1994 et 2004) ainsi que ceux de six nations hors de l'Afrique (Salvador, Indonésie, Mexique, Nicaragua, Turquie et Uruguay). Plus de pays africains seront inclus dans les futurs lancements, selon la célérité avec laquelle les microdonnées du cycle de recensements de 2010 seront rendues disponibles. Malheureusement, jusqu'à la fin de 2011, sept pays n'avaient pas encore confié de microdonnées concernant les énumérations réalisées il y a de cela 2, 3, 4 et même 5 ans : il s'agit du Cameroun (2005), du Tchad (2009), de l'Éthiopie (2007), du Kenya (2009), du Lesotho (2006), du Mozambique (2007), et du Nigeria (2006).

**IPUMS-I apporte de la valeur ajoutée aux microdonnées de recensement dans 4 grandes catégories.** IPUMS-International cherche à ouvrir la valeur inhérente aux microdonnées de recensement pour les chercheurs universitaires, les décideurs politiques et les fonctionnaires. IPUMS-I *n'est pas* simplement un saut-secours pour faire passer des échantillons de recensement des ONS vers les chercheurs. Au lieu de cela, généralement, deux ou plusieurs années de travail sont investies par un groupe important de personnel expérimenté du MPC pour préparer les microdonnées anonymisées, intégrées et les métadonnées pour la diffusion. Au vingt et unième siècle, avoir à portée de main une copie des microdonnées originales et un dictionnaire n'est pas suffisant pour faire de la recherche de haute qualité. Ce document traite 16 valeurs ajoutées offertes par IPUMS-International. Celles-ci peuvent être classées en quatre groupes à savoir : le secret statistique, l'intégration, la diffusion et l'éthique.

## **I. Secret statistique et sécurité.**

**1. Sécurité des microdonnées et secret statistique.** Ni le MPC, ni ses employés ne diffusent jamais les microdonnées de source originale. Ces données sont extrêmement sensibles et pour cette raison, seuls les chercheurs chevronnés, spécialement formés, et à plein temps sont autorisés à travailler sur les données jusqu'à ce qu'elles soient rendues anonymes. Les employés du MPC sont soumis à des amendes civiles (jusqu'à 250 000 USD), des poursuites pénales et à la perte d'emploi en cas de violation des procédures de sécurité de l'Université du Minnesota. L'autorité légale de l'Université assume la responsabilité de protéger la sécurité totale et la confidentialité des données confiées au MPC.

IPUMS-I est l'une des quelques rares organisations académiques diffusant les microdonnées de recensements qui a été examinée par un auditeur externe. L'examineur, M. Dennis Trewin, le très respecté statisticien australien et président du groupe de travail de l'UNECE « Gérer la Confidentialité Statistique et l'Accès aux Microdonnées », a cité l'environnement informatique du MPC comme «meilleure pratique», «la norme des meilleurs bureaux de statistique», etc. M. Trewin a effectué une inspection d'une semaine, sur le site du MPC et a conclu son rapport comme suit (Trewin 2007) :

Sans aucun doute IPUMS-International répond aux quatre principes fondamentaux énoncés dans la CES [Conférence des Statisticiens Européens] (2007). Il est cité dans CES (2007) comme étude de cas de bonnes pratiques. Cet examen confirme son statut de bonne pratique pour les entrepôts de données (sockage). En effet, il est susceptible de fournir la meilleure pratique pour un Entrepôt des données statistiques internationales [soulignement ajouté].

Lors de la réunion du groupe d'experts de l'UNECE sur la confidentialité des données statistiques, en Novembre 2005, nous avons expliqué les procédures de sécurité sur la diffusion des données IPUMS-international comme suit (McCaa et Esteve 2005) :

Lorsque l'extrait est prêt (généralement en quelques minutes), le chercheur est avisé par courriel que les données doivent être récupérées dans les 72 heures. Un lien est fourni à un site protégé par un mot de passe pour télécharger l'extrait concerné. Les données sont cryptées pendant leur transmission à l'aide d'une procédure standard de cryptage de 128 bits de type SSL (Secure Sockets Layer), correspondant au niveau retenu par le système bancaire et d'autres industries où la sécurité et la confidentialité sont essentielles. Le chercheur peut ensuite télécharger en toute sécurité le fichier, le décompresser et procéder à l'analyse en utilisant les métadonnées intégrées fournies, composées de noms de variables et de libellés.

Cette méthode de diffusion continue de résister à l'épreuve du temps, et alors que l'utilisation croît, l'accélération rapide des vitesses de transmission sur Internet a validé les protocoles de sécurité de IPUMS-I.

**2. Protections de contrôle de la divulgation statistique.** Les microdonnées sont soumises à règles fortes de contrôle de divulgation, juridiques, administratives, techniques et statistiques uniformes offrant une plus grande protection tant pour le groupe dans son ensemble que pour n'importe quel bureau de statistique pris individuellement qui choisit de «faire cavalier seul» (McCaa, Ruggles et Sobek 2010). Le plus important contrôle technique de divulgation statistique est la suppression des enregistrements par l'utilisation de sous-échantillonnage. Toutes les valeurs dans les enregistrements à l'extérieur de l'échantillon sont supprimées. Vient ensuite la suppression des noms et des détails des niveaux géographiques inférieurs.

Chaque autorité statistique équilibre le compromis confidentialité /utilité en instruisant le projet IPUMS-I sur le seuil minimum à retenir pour les unités géographiques identifiables. Pour de nombreux pays, le seuil est couramment fixé à 20 000 habitants. D'autres le placent aussi haut que 100 000 (Etats-Unis) ou dans le cas le plus extrême (Pays-Bas) toutes les unités géographiques administratives sont supprimées. Nous sommes heureux de constater que dans certains cas, non organismes statistiques partenaires ont reconsidéré les décisions antérieures, en offrant des échantillons de précision plus élevée (Mexique 1990 est passé de un à dix pour cent)

et plus de détails. Dans le cas de la Colombie, le seuil géographique, initialement fixé à 100 000, a été réduit à 20 000 après que les géographes colombiens aient vigoureusement protesté.

L'agence colombienne de statistique a non seulement réduit le seuil, mais a aussi harmonisé les identifiants géo-statistiques, de sorte que tous les échantillons de microdonnées de recensement de la Colombie pourraient être diffusés avec un ensemble unique de codes géographiques.

D'autres mesures de protection de la divulgation des statistiques sont fournies en ordonnant de façon aléatoire les enregistrements et en échangeant les identifiants géographiques d'un nombre non divulgué des ménages. Cela signifie que personne ne peut affirmer avec certitude que l'individu ou le ménage a été identifié.

En consultation avec le bureau national de statistique, certaines variables peuvent être surcodées, d'autres peuvent être soumises à recodage global, la suppression de chiffres pour les variables hiérarchiques (profession, industrie, géographie), ou la suppression d'une variable entièrement. Les décisions sont prises en consultation avec l'autorité correspondante de statistiques nationales. Les variables de poids et les facteurs d'expansion ne sont généralement pas un problème parce que la plupart des échantillons sont implicitement stratifiés de telle sorte que tous les enregistrements portent un poids identique.

**3. Accès restreint.** L'accès aux microdonnées d'IPUMS-International est limité, malgré le «P» dans IPUMS. Les utilisateurs potentiels doivent soumettre une [demande électronique détaillée](#) pour à la fois réaliser des recherches de bonne foi et pour expliquer les besoins d'accès. Une partie essentielle du processus est d'accepter dix restrictions strictes sur les conditions d'utilisation interdisant la redistribution, limitant à un usage scientifique, interdisant une utilisation commerciale, protégeant la confidentialité, assurant la sécurité, appliquant les règles strictes de confidentialité, permettant la publication scientifique, citant correctement, traitant des mesures disciplinaires en cas de violation, et la déclaration d'erreurs. En d'autres termes, IPUMS-I est un système d'accès à "l'utilisateur de confiance".

Convenir des conditions d'utilisation lie à la fois le chercheur et l'institution du chercheur. Le Conseiller juridique de l'Université du Minnesota est sur le point de frapper à la première indication d'abus. Tant le chercheur individuel et l'institution du chercheur sont responsables de maintenir la sécurité et faire respecter le contrat de licence. Les violations sont susceptibles de conduire à des sanctions contre l'individu et l'institution du chercheur fautif.

IPUMS-I résout l'énigme de la gestion de l'accès le plus large possible de microdonnées sensibles tout en protégeant la confidentialité des statistiques. Beaucoup d'agences de statistique ont longtemps voulu rendre les microdonnées de recensement disponibles pour les chercheurs, mais manquent de matériel substantiels et de ressources humaines nécessaires pour mettre en œuvre et gérer des systèmes sécurisés. IPUMS-I est la seule organisation académique de diffusion de microdonnées du recensement qui est citée comme une bonne pratique par le groupe de travail de la Conférence des statisticiens européens sur la Gestion de la Confidentialité des statistiques et de l'Accès aux Microdonnées (2007).

Bien que l'accès aux microdonnées IPUMS-International soit gratuit, l'utilisation est limitée aux chercheurs de bonne foi qui acceptent de se conformer à des conditions strictes d'utilisation (annexe B). IPUMS diffuse des extraits, taillés sur mesure pour les besoins précis de recherche de chaque utilisateur. L'extrait moyen de IPUMS se compose de quelques 10 variables.

Cela contraste avec les pratiques de la plupart des bureaux de statistiques, où les microdonnées de recensement sont diffusées comme des ensembles complets, comprenant un dictionnaire de données et un échantillon complet contenant toutes les variables et tous les enregistrements individuels. Généralement, dans l'approche traditionnelle, lorsque les demandes sont remplies, chaque chercheur reçoit exactement le même ensemble de données et la documentation. Compte tenu de la taille colossale de la base de données IPUMS-International, la diffusion de la série complète des variables et d'une taille invariable d'échantillons est impraticable.

Malgré les conditions strictes de l'utilisation et les restrictions, plus de cinq mille chercheurs-qui représentent plus de 100 pays et 900 institutions-sont accrédités pour l'accès à la base de données IPUMS-I. Plus d'un tiers des utilisateurs autorisés de IPUMS-I demandent l'accès aux microdonnées pour un seul pays. Une fraction importante d'entre eux sont domiciliés à l'étranger et demandent l'accès aux données relatives à leur propre pays d'identité (d'origine).

## **II. Intégration**

**4. Métadonnées de source complète.** Les chercheurs doivent avoir facilement accès à la documentation originale de recensement dans la langue officielle. Au minimum, les questionnaires de recensement, les instructions aux agents énumérateurs ou les manuels de formation et les livres de codes sont nécessaires. Les métadonnées supplémentaires concernant l'organisation, le travail de préparation, et la phase pratique du recensement sont également précieuses pour le projet IPUMS-International et sont cataloguées et archivées avec tous les autres documents reçus. Une copie originale des documents papiers ou une version électronique PDF sont préférées pour les matériaux de métadonnées publiées. Notre objectif est de fournir une collection archivée de haute qualité des fichiers PDF pour toutes les formes de métadonnées relatives aux microdonnées de recensement. Les documents originaux sont mis à la disposition des chercheurs, en anglais et dans la langue officielle de l'original.

**5. Métadonnées Intégrées, compatibles DDI** -transnationales et à travers le temps. Afin de faciliter le processus de recherche des microdonnées sont intégrées pour tous les recensements et tous les pays, y compris des descriptions détaillées de chaque recensement, échantillon et variable.

**a. Les recensements et les échantillons.** Les Métadonnées IPUMS offrent des descriptions détaillées de chaque recensement dans la base de données, répertoriant le titre, l'année, l'univers, de jure/de fait, l'unité de dénombrement, le jour du recensement officiel, les questionnaires, la période du travail sur le terrain et le type, le répondant et les estimations de sous-dénombrement, le cas échéant. Des copies numérisées des questionnaires de dénombrement et des manuels d'instructions sont disponibles dans la langue officielle et le texte en traduction anglaise. Chaque échantillon est décrit en tenant compte de la source, de la conception de l'échantillon, de l'unité d'échantillonnage, de la fraction d'échantillon, du nombre d'enregistrements individuels, du poids de l'échantillon, des unités d'habitation ou de logement, des logements vacants, des ménages, des groupes de quarts et des populations spéciales.

**b. Descriptions des variables, textes sources, et codes.** Les Métadonnées IPUMS définissent chaque variable intégrée et décrivent les caractéristiques de base : la disponibilité par recensement, l'univers de la variable ou de la question, les codes, les textes source (énumération), et les variables non harmonisées utilisées pour l'intégration. L'accès à cette information est obtenu grâce à un simple clic sur un lien hypertexte sur le site IPUMS. Une discussion générale de comparabilité est fournie pour chaque variable, avec des discussions spécifiques pays ou

recensement en se concentrant sur les éloignements de la pratique standard. Le but de ces discussions est de mettre en évidence des contrastes importants. Le clic sur "texte de dénombrement" conduit à une page des questions de source et les instructions correspondantes en anglais pour chaque recensement. Des clics supplémentaires permettent de voir la documentation d'origine sous forme numérisée afin que les chercheurs puissent étudier la présentation et le libellé dans la langue officielle.

Les chercheurs parcourent les métadonnées intégrées pour examiner rapidement le libellé même des questions du recensement et des instructions aux agents recenseurs pour toute combinaison de pays et des années de recensement. L'équipe d'intégration du MPC applique des balises XML pour les documents de recensement, en associant les variables dans les microdonnées avec les concepts dans le texte. Le matériel marqué est ensuite importé dans une base de données. Une fois cette étape complétée, les métadonnées peuvent être récupérées de façon dynamique pour toute combinaison de pays et d'années de recensement, variable par variable. Initialement cet outil a été développé pour accélérer le travail de l'équipe d'intégration. Une fois que son utilité est devenue évidente, nous avons exploité le système de métadonnées dynamiques pour le site web, afin de permettre le libre accès aux métadonnées. Les métadonnées sont sur-mesure selon la demande de chaque chercheur et disponibles en un seul fichier consultable en format DDI ou en format texte.

**6. Microdonnées Intégrées.** Le principal avantage de IPUMS pour les chercheurs et aussi bien pour les ONS est l'intégration de multiples échantillons de microdonnées pour chaque pays, généralement en commençant par le premier recensement pour lequel des microdonnées existent ou sont récupérables. Pendant des décennies, plusieurs ONS ont fourni un échantillon pour le recensement le plus récent, mais peu parmi eux ont réexaminé les recensements antérieurs en vue d'harmoniser les ensembles de données successifs ou de rédiger de nouveaux documents pour faciliter l'analyse comparative de deux ou plusieurs recensements. Les offices de statistiques modernes construisent un échantillon du recensement et un dictionnaire de données pour les chercheurs nationaux. Cinq ou dix ans plus tard, avec le recensement qui a suivi, le processus est répété avec peu d'indications sur l'amélioration de la comparabilité des ensembles de données de recensement successifs.

Nous devons réaffirmer que le projet IPUMS *ne* diffuse *pas* les fichiers de recensement confiés par les offices nationaux de statistique. En lieu et place, des échantillons de recensement de haute précision sont anonymisés (McCaa et al. 2006) et intégrés, variable par variable, en utilisant un système composite de codage (Esteve et Sobek, 2003). Les échantillons sont intégrés à la fois chronologiquement et trans-nationalement. Les métadonnées intégrées sont construites à l'aide d'une étude méticuleuse de l'ensemble des documents de source originale et après une analyse approfondie des microdonnées. Des milliers d'heures sont consacrées à analyser, discuter, débattre, faire des avant-projets, tester et re-tester jusqu'à ce que l'intégration des microdonnées est validée pour la diffusion aux chercheurs. Le processus est répété à chaque lancement annuel d'échantillons de recensement supplémentaires dans la base de données IPUMS.

L'objectif de base de l'effort d'harmonisation IPUMS-I est de simplifier l'utilisation des données tout en ne perdant aucune information significative. C'est un défi parce que pour rendre les données simples pour l'analyse comparative dans le temps et l'espace, il est nécessaire de développer des schémas de codage comparables. Les microdonnées sont intégrées de sorte que des concepts identiques (variables, catégories) aient des codes identiques. Pour éviter la perte

d'informations importantes pour les échantillons qui ont encore plus de détails, IPUMS-I utilise une stratégie de codage composite en vue de conserver tous les détails d'origine, et dans le même temps fournir des codes comparables entre les échantillons. Avec des codes composites, les chercheurs peuvent comparer facilement à travers le temps et l'espace, mais des nuances dans les changements de sens sont facilement discernables. Le premier chiffre, que nous appelons le "code général", fournit des informations qui sont disponibles dans tous les échantillons (le plus bas dénominateur commun de données). Le ou les deux chiffres suivant(s) fournit (fournissent) des informations supplémentaires disponibles dans un sous-ensemble important des échantillons. Les tous derniers chiffres fournissent des détails qui ne sont que rarement disponibles. Lorsque l'information n'est pas disponible pour un échantillon particulier, un zéro par défaut est affecté à ce chiffre.

Comme exemple de la méthode IPUMS d'intégration des variables, considérons le concept «niveau d'éducation», la variable la plus largement utilisée dans la base de données IPUMS-International. La plupart des microdonnées de recensement fournissant de l'information sur cette mesure indiquent si le répondant a achevé sa scolarité primaire, secondaire ou supérieure ou s'il n'a pas été scolarisé du tout. Ainsi, le premier chiffre du code IPUMS-International composite se compose de quatre catégories (1-4), plus un code de données manquantes (9) et non dans le code-univers (0, pour les enfants trop jeunes pour aller à l'école ou d'autres pour lesquels la question n'a pas été posée). Certains échantillons de recensement contiennent de plus amples informations indiquant, par exemple, ceux qui ont fréquenté, le primaire, le secondaire ou même le supérieur, mais n'ont pas terminé le cours des études. Le deuxième chiffre rend compte de cette information.

L'intégration internationale réussie doit documenter ces distinctions afin que les chercheurs puissent facilement être informés de ces dernières et de milliers d'autres détails. L'Annexe C illustre les schémas détaillés et généraux de codage pour la variable niveau d'éducation des pays africains (représentée par ses deux chiffres code ISO 3166) et recensements (représentée par un code de l'année à deux chiffres avec le siècle omis).

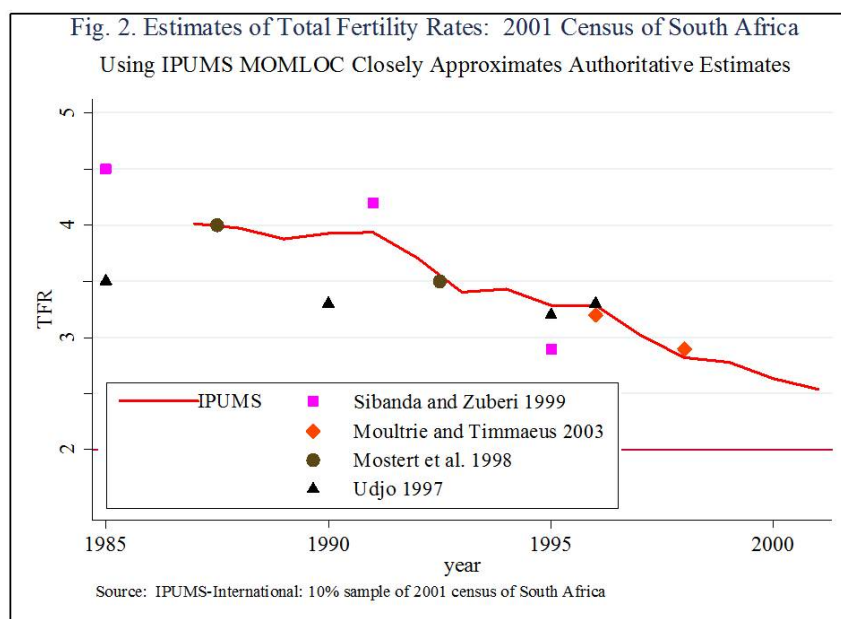
**7. Valeur ajoutée des variables IPUMS-I.** Les Annexes D et E listent les variables intégrées des ménages et des personnes pour 14 recensements africains. Il ya trois types variables d'IPUMS-I à valeur ajoutée :

**a. Variables techniques:** Type d'enregistrement, Pays, Année, Identifiant de l'échantillon IPUMS, Identifiant du ménage, Numéro d'identification de l'individu dans le ménage, Poids du ménage, Nombre de sous-échantillons, Statut de groupe quarts, Continent, Pays, Région, Résidence au premier niveau administratif, Facteurs d'expansion, (Poids de l'échantillon- pour les ménages et les personnes), etc.

**b. Résumé des variables ménages et familiales:** Classification des ménages, Nombre de familles dans le ménage, Nombre de couples mariés dans le ménage, Nombre de mères, Nombre de pères, Localisation du Chef dans le ménage, Nombre de personnes non apparentées, Appartenance à une unité familiale, Nombre des membres de famille propre dans le ménage, Nombre d'enfants dans le ménage, Nombre d'enfants de moins de 5 ans dans le ménage, Age de d'enfant aîné propre du ménage, Age du plus jeune enfant propre ménage, etc.

**c. Variables Pointeuses ("LOC")** afin d'identifier le ou les conjoints co-résident(s), les enfants et leurs parents: la localisation de la Mère, du Père et du conjoint dans le ménage, la Règle pour relier parent (s) et conjoint (s), Belle-mère probable, Beau-père probable, l'homme

avec 2 épouses ou plus, deuxième femme ou plus, etc. (Sobek et Kennedy, 2009). MOMLOC (variable pointeuse de la mère) est particulièrement puissant pour estimer la fécondité en utilisant la méthode de survie inverse. Pour restreindre MOMLOC aux mères biologiques, quand STEPMOM est supérieur à zéro, alors MOMLOC devrait être remis à zéro. La figure 2 démontre que l'utilisation MOMLOC pour faire correspondre les mères aux enfants pour le calcul de leur propre enfant estime le taux de fécondité total, produit une série de taux de fécondité totale qui est incroyablement semblable aux autres estimations des autorités, malgré le fait que les ménages de l'Afrique du Sud sont anormalement complexes (Kennedy, McCaa, Sobek, et Cleveland, 2011).



**8. Fichiers de frontières intégrés.** IPUMS-International fournit les fichiers de limites ou frontières (shapefiles) pour faciliter des travaux de cartographie des données nationales et internationales. Le chercheur télécharge les fichiers des limites et les manipule avec des logiciels comme il le désire. Deux niveaux de fichiers des limites de SIG sont disponibles: un fichier avec les frontières du monde au niveau des pays, et des fichiers au niveau des pays avec des limites supérieures niveau administratif (États, provinces, etc.) Les utilisateurs peuvent créer des cartes avec les données de IPUMS-international en utilisant un programme de logiciel de statistique et ArcMap (un logiciel de cartographie de SIG). Pour plus d'informations sur la liaison des données IPUMS-international au monde et la sélection des cartes de pays, voir : <https://international.ipums.org/international/gis.shtml>.

### III. Diffusion

**9. Accès transfrontière:** l'accès transfrontalier aux microdonnées est essentiel dans le monde global d'aujourd'hui, où les chercheurs sont très mobiles. Considérons, par exemple, le domaine de la démographie, où un cinquième des membres de la société mondiale professionnelle, l'Union Internationale pour l'Etude Scientifique de la Population (UIESP), réside en dehors de leur pays de naissance. Pour les 506 membres de l'UIESP résidant aux Etats-Unis, trente pour cent n'y sont pas nés. Des démographes chinois de naissance, près d'un tiers résident

en dehors de Chine. Pour l'allemand et le néerlandais de naissance membres de l'UIESP la fraction s'élève à deux cinquièmes<sup>1</sup>. Pour beaucoup de professionnels et de nombreux démographes, chercheurs en sciences sociales en général - l'accès transfrontalier est essentiel si l'on veut mener une recherche en utilisant les microdonnées du recensement de leur pays de naissance, que ce soit comparatif ou non. Le portail IPUMS-International offre une expérience homogène pour accéder à des échantillons pour tous les pays (pour l'AICMS tous les pays africains) et les années de recensement, indépendamment du pays de naissance, de résidence ou de nationalité du chercheur.

**10. Extraits sur mesure regroupés en un seul fichier.** Avec IPUMS-I deux extraits ne sont jamais semblables. Chaque extrait est taillé sur mesure par le chercheur par le biais d'une série de d'écrans de sélection pointer-et -cliquer. L'extrait est remis au chercheur dans un seul fichier commun-quel que soit le nombre de pays ou de recensements. Alors que les sélections, les métadonnées sont aisément disponibles en "naviguant" dans la documentation dans un ordre quelconque. Pour passer une commande le chercheur choisit :

- pays (ou pays)
- année(s) de recensement
- les variables (âge, sexe, niveau d'éducation, etc.)
- sous-populations (par exemple, les femmes chefs de ménages âgés de moins de vingt cinq ans). Notez qu'il y a une option pour sélectionner un individu ou toutes les personnes co-résidentes dans le ménage)
- et la densité de l'échantillon (soit en tant que un pour cent ou le nombre d'enregistrements).

Le moteur d'extraction de IPUMS remplit la demande en générant un ensemble de données ne contenant que les microdonnées demandées et l'ensemble correspondant de métadonnées compatibles DDI ainsi qu'un dictionnaire approprié à la construction d'un fichier SPSS, SAS ou STATA. Des copies des métadonnées de source originale sont disponibles à partir du site Web, ainsi que les métadonnées intégrées sous forme interactive. Celles-ci aussi peuvent être téléchargées librement.

Le fait que IPUMS-International distribue électroniquement sous forme d'extraits de microdonnées sur mesure, adaptés au(x) pays, année(s) de recensement(s), sous-population(s), et aux variables, selon les besoins individuels du chercheur, offre des incitations supplémentaires pour garder jalousement les extraits. Du fait que les ensembles complets de données ne sont pas distribués sur CD ou autre support, la tendance à partager des données avec des personnes non autorisées est considérablement réduite, si elle n'est pas complètement éliminée. Le moteur d'extraction IPUMS ajoute encore plus de valeur à chaque extrait au moyen de quatre outils uniques :

**a. Sélectionner des cas (voir «sous-populations" ci-dessus).** Filtre les échantillons pour sélectionner précisément les cas de recherche d'intérêt. Par exemple, si le chercheur souhaite obtenir rien que les femmes économiquement actives, âgées de 15-19 ans, nées dans un pays différent du pays de résidence actuel, le moteur d'extraction IPUMS va générer un ensemble de

---

<sup>1</sup> Statistiques communiquées à l'auteur par le Secrétariat de l'Union Internationale pour l'Etude Scientifique de la Population, 14 Septembre 2011

données précisément pour ces caractéristiques. Par ailleurs le moteur d'extraction inclura dans l'extrait tous les individus co-résidents dans le même foyer avec la personne sélectionnée.

**b. Attacher les caractéristiques.** Le moteur d'extraction joint les caractéristiques des mères, des pères, des conjoints et des chefs de ménage aux individus co-résidents. Cette caractéristique facilite l'analyse des enfants selon les caractéristiques de leurs mères, des pères et / ou chefs de ménage. Cette fonctionnalité est utile pour analyser la fertilité des enfants propres, l'homogamie maritale, et une foule d'autres sujets où les caractéristiques communes de deux ou plusieurs membres du ménage sont nécessaires. Par exemple, pour le récent séminaire international sur «Les mariages transfrontaliers» (Séoul, République de Corée), il était facile de générer un ensemble de données sur le site IPUMS-I qui représente 12 millions d'étrangers individuellement mariées ou en union avec les co-résidents natifs conjoints nés pour 51 pays (Esteve, Garcia et McCaa, 2011; voir aussi Parker, 2011).

**c. Personnaliser la taille de l'échantillon.** Le système d'extraction offre un outil permettant de personnaliser la taille de chaque échantillon en termes de nombre absolu de personnes ou de ménages ou de pourcentages. Si le chercheur désire un échantillon de seulement 50 000 ménages, il suffit d'entrer "50" dans le tableau correspondant et le moteur d'extraction va construire un échantillon systématique. Les pondérations appropriées (facteurs d'expansion) sont automatiquement calculées à la volée et incluses dans l'extrait.

**d. Regroupées dans un seul fichier.** Peu importe le nombre de pays ou de recensements, toutes les microdonnées relatives à chaque extrait sont regroupées dans un seul fichier. Comparer les pays ou les années de recensement est facile parce que chacun est traité comme une variable au lieu de fichiers séparés. .

**11. Usage.** Les microdonnées sont libérées de l'enceinte des bureaux de statistique afin de permettre à certains des plus brillants statisticiens, démographes, économistes, chercheurs en sciences sociales et décideurs politiques sur la planète d'extraire des valeurs inestimées à partir des données sans coût (Meyer, McCaa, et Lam, 2011). Car de nombreux chercheurs nationaux n'ont jamais été en mesure d'utiliser les microdonnées de leur pays jusqu'à ce que IPUMS-International ait ouvert la porte. D'autres chercheurs sont confrontés à des ennuis, et font parfois face à des frais exorbitants.

A la date de Novembre 2011, 166 chercheurs représentant 24 pays africains étaient inscrits à l'accès aux microdonnées sur les sites IPUMS-International/AICMD. L'Afrique du Sud s'est classée numéro 1 avec 56 chercheurs, suivie par le Kenya (32), l'Egypte (15), l'Ouganda (14), le Ghana (8), le Nigeria (7), la Tanzanie (6), l'Algérie et le Cameroun (4 chacun), le Sénégal (3), le Burkina Faso, le Malawi et la Sierra Leone (2 chacun) et le Botswana, la République du Congo, la Côte d'Ivoire, le Maroc, Maurice, le Mozambique, la Namibie, la Réunion, la Zambie et le Zimbabwe (1 chacun).

Lorsque la plus récente, analyse complète de l'utilisation a été réalisée en avril 2011, 85.505 échantillons (24.699 extraits) avaient été téléchargés depuis le site IPUMS-International, faisant en moyenne plus de 1500 échantillons par pays pour les 55 pays représentés dans la base de données en ce moment. L'Afrique du Sud, avec 3 échantillons intégrés, a déplacé le Kenya de la première place avec 1 940 extraits, suivie par l'Ouganda (1299 extraits), le Rwanda (1015), la Guinée, le Ghana, la Tanzanie, l'Egypte, le Mali et le Sénégal.

Le tableau 2 énumère les 32 variables les plus fréquemment extraites. Les huit premiers englobent quatre variables démographiques (statut matrimonial, lien avec le chef de ménage, âge et sexe), deux économiques (statut d'emploi et catégorie de travailleur), et une sociale (niveau d'éducation) et technique (poids de l'individu).

Un constat le plus frappant est que 40 des variables comptent pour 60% de celles qui sont demandées et, parmi celles-ci, seize sont des variables construites par IPUMS, dont quatre, les variables "LOC", sont uniques à IPUMS-I. Les variables LOC sont imputées à partir des informations des ménages : situation du conjoint dans le ménage, localisation de la mère, localisation du père, et les règles pour inférer les localisations. Les chercheurs exploitent ces variables pour étudier les caractéristiques communes des conjoints et les caractéristiques des parents par rapport à leurs enfants. Les variables sont construites par l'imputation de la variable relation avec le chef de ménage, avec les variables âge, sexe, état matrimonial, l'ordre des individus listés dans le ménage, et quelques autres variables. L'usage intensif des variables "LOC" indique leur grande importance pour l'analyse des individus par rapport aux caractéristiques de leurs conjoints, les mères et les pères. Les variables LOC constituent une très importante valeur ajoutée à chaque échantillon de ménages (Sobek et Kennedy, 2009).

Les Institutions "top 40" en termes d'utilisation de données de IPUMS-I comprennent de nombreuses grandes universités du monde et des organismes de recherche (voir Annexe F), dispersés à travers quatorze pays. Dans 46 pays, nous trouvons un total de 501 institutions avec des chercheurs faisant dix ou plusieurs extraits. (En outre, aux États-Unis, il ya 295 institutions à ce niveau d'utilisation). Un nombre surprenant d'extraits sont faites par des chercheurs de pays ne disposant pas de microdonnées dans le système IPUMS-I. Le top 10 de ceux-ci sont : Singapour (494 extraits), Belgique (250), Australie (229), Japon (170), Fédération de Russie (58), République de Corée (45), République tchèque (42), Suède (41), Hong Kong SAR (40), et Nouvelle-Zélande (40). L'autre face de la médaille, ce sont 13 pays avec dans la base de microdonnées IPUMS-I, qu'aucun chercheur national n'utilise encore. Les 13pays sont les suivants : Arménie, Bélarus, Guinée, Iraq, Jordanie, Kirghizistan, Mali, Mongolie, Népal, Pérou, Rwanda, Sainte-Lucie et Slovénie. Bien sûr, les chercheurs de ces pays-au lieu d'accéder électroniquement aux microdonnées à partir du site de IPUMS-I peuvent acquérir des copies des échantillons intégrés sur des CD fournis par le Centre de Population du Minnesota aux Offices nationaux de statistique correspondants. Nous conseillons aux partenaires des ONS d'inscrire tous ces utilisateurs et de les admonester pour qu'ils respectent les conditions d'utilisation de IPUMS-I, mais il n'ya aucune obligation de le faire parce que l'ONS est responsable de toute diffusion.

**12. Échantillons du cycle de recensement de 2010.** Les Microdonnées relatives au cycle de recensements l'année 2010 sont en cours de processus de traitement «accéléré» par IPUMS-international sans aucun délai. Les microdonnées reçues en date du 1<sup>er</sup> Septembre sont traitées pour un lancement le 1er Juin de l'année suivante. À ce jour, la plupart des partenaires de IPUMS-International sont désireux de confier les microdonnées dès que le traitement des données est terminée. Les échantillons de recensement de 2010 de l'Indonésie et du Mexique seront lancés en 2012. En 2011, des échantillons du recensement de 2006 de l'Égypte, de 2008 du Malawi, du Soudan et du Sud-Soudan, et de 2004 de la Sierra Leone sont intégrés. Il est à espérer qu'ils seront bientôt suivis par les échantillons provenant de recensements effectués au cours des cinq dernières années dans d'autres pays africains.

#### **IV. Éthique.**

**13. La transparence statistique.** Confier les microdonnées du recensement démontre au monde entier que l'agence statistique officielle « n'a rien à cacher » et est prête à soumettre ses plus précieuses microdonnées à un examen externe.

**14. La liberté académique.** Les microdonnées diffusées par IPUMS-International ne sont sous aucune contrainte préalable à la publication ou la diffusion des résultats.

**15. Réduire la fraude** et la tentation d'exagérer les résultats de recherche. Du fait qu'aucun chercheur n'a accès exclusif aux microdonnées, les résultats peuvent être facilement reproduits pour tester les conclusions d'autres chercheurs.

**16. Les résultats des recherches.** Les chercheurs postent des liens vers leurs conclusions dans la bibliographie IPUMS-International avec un accès ouvert: <http://bibliography.ipums.org/>. La bibliographie de IPUMS-International recense plus de 500 entrées.

### **Réflexions.**

Pour que les produits de recensement répondent aux besoins des utilisateurs, le projet IPUMS requiert une gamme formidable et une quantité de métadonnées et de microdonnées. Néanmoins, le retour sur investissement est considérable pour les bureaux de statistique qui sont soulagés d'un grand nombre de tâches les plus lourdes et des responsabilités dans la diffusion de microdonnées pour les chercheurs. Par ailleurs, en s'appuyant sur des procédures standard utilisées par la majorité des bureaux de statistique du monde, il y a la sécurité dans les chiffres. Les bureaux de statistique isolés qui diffusent des microdonnées sur une base ad hoc encourent des risques importants et des responsabilités ainsi que d'importantes ressources humaines et des coûts de matériels, pour un rendement relativement faible par rapport au nombre d'utilisateurs. Le projet IPUMS offre d'importantes économies d'échelle avec les plus hautes normes de sécurité et diffuse des métadonnées intégrées et des microdonnées qui facilitent grandement de la recherche scientifique de qualité.

Quand nous avons commencé il y a une décennie, nous avons rêvé d'intégrer des échantillons relatifs à 21 pays en dix ans. Nous avons triplé cet objectif (y compris le Sud-Soudan en tant que 63<sup>ème</sup> pays), grâce à la généreuse collaboration des offices nationaux de statistiques et aux innovations technologiques insoupçonnées. Le nombre d'utilisateurs et la quantité d'utilisation ont également dépassé de loin nos attentes. Pour la deuxième décennie, nous rêvons de doubler le nombre d'utilisateurs et de doubler à nouveau le nombre d'échantillons. Des échantillons de haute précision pour le cycle des recensements de 2010 seront cruciaux pour que le succès continue. Les Bureaux de statistique participants sont invités à confier des métadonnées et des microdonnées pour le cycle de recensement de 2010 sans tarder. Quant aux agences qui ne participent pas encore aux initiatives IPUMS, elles sont invitées à envisager de le faire. Les chercheurs qui ont déjà accès aux microdonnées IPUMS sont invités à parcourir les métadonnées et à utiliser les microdonnées selon leurs besoins de recherche.

### **Références.**

Alexander, J.T.; Davern, M.; and Stevenson, B. 2010. "Inaccurate Age and Sex Data in the [United States] Census PUMS Files: Evidence and Implications," *Public Opinion Quarterly*, 10 (Aug 10), pp. 1-10. doi: 10.1093/poq/nfq033

- Conference of European Statisticians. 2007. "Annex 1.23 Case study: Access to anonymized census microdata samples via the IPUMS-International and the Integrated European Census Microdata websites," *Managing Statistical Confidentiality and Microdata Access: Principles and Guidelines on Good Practice*. Geneva: United Nations Economic Commission for Europe. See online edition: <http://www.unece.org/stats/publications/> pp. 98-104.
- Editors. 2010. "Official statistics and microdata – access and confidentiality". *Statistical Journal of the IAOS* 26:55-56.
- Esteve, A., J. Garcia and R. McCaa. 2011. "Comparative perspectives on Marriage and International Migration, 1970-2000: findings from IPUMS-International census microdata samples," *Seminar on Global Perspectives on Marriage and International Migration*, Seoul, South Korea: IUSSP Scientific Panel, Oct. 20-21.
- Kennedy, S., R. McCaa, M. Sobek, and L. Cleveland. 2011. "The quality of constructed family and household relationships in African Census Samples," *Sixth African Population Conference*, Ouagadougou, Burkina Faso, 5-9 December.
- McCaa, R. and A. Esteve. 2005. "[IPUMS-Europe: Confidentiality measures for licensing and disseminating restricted access census microdata extracts to academic users](#)," *Joint UNECE/Eurostat Work Session on Statistical Confidentiality*, Geneva, Nov. 9-11.
- McCaa, R. and A. Esteve. 2009. "[Entrusting census microdata and metadata for timely integration and dissemination via the IPUMS-EurAsia and IECM initiatives, 2010-2014](#)," *Census Outputs to Meet User Needs*. Geneva: United Nations Economic Commission for Europe, Oct. 28-30.
- McCaa, R., S. Ruggles and M. Sobek. 2010. "[IPUMS-International statistical disclosure controls: 159 census microdata samples in dissemination, 100+ in preparation](#)," in J. Domingo-Ferrer and E. Magkos (Eds.): *Privacy in Statistical Data 2010*, LNCS 6344. Springer, Heidelberg, pp.74-84.
- McCaa, R. and W. Thomas. 2009. "[IPUMS-International: lessons from 10 years of archiving and disseminating census microdata](#)," *International Statistical Institute IPM100*. Durban, South Africa.
- Minnesota Population Center. 2011. *Integrated Public Use Microdata Series – International: Version 6.0*. Minneapolis: University of Minnesota: <https://www.ipums.org/international>.
- Meier, A., R. McCaa and D. Lam. 2011. "[Creating statistically literate global citizens: The use of IPUMS-International integrated census microdata in teaching](#)". *Statistical Journal of the IAOS* 27(3):145-156.
- [Parker, J.] 2011. "Herr and Madame, Señor and Mrs. Research at last begins to cast some light on the extent, causes and consequences of cross-border marriages," *Economist*, Nov. 12. <http://www.economist.com/node/21538103>
- Sobek, M and S. Kennedy. 2009 The development of family interrelationship variables for international census data, Minnesota Population Center. [https://international.ipums.org/international/resources/misc\\_docs/pointer\\_working\\_paper\\_2009.pdf](https://international.ipums.org/international/resources/misc_docs/pointer_working_paper_2009.pdf) .
- Trewin, D. 2007. "A Review of IPUMS-International." Unpub. [http://www.hist.umn.edu/~rmccaa/IPUMSI/trewin\\_ipums\\_report.pdf](http://www.hist.umn.edu/~rmccaa/IPUMSI/trewin_ipums_report.pdf)
- Woolfrey, L. 2009. "African Microdata Access Survey 2009," Unpub.

**Tableau 1. Echantillons IPUMS/AICMD**

|   | Pays,<br>recensement | (%) | Ménages           | Individus         | Pondéré | De jure<br>/de fait | Date du<br>recensement | Plus petit<br>niveau<br>géographique | Ménages<br>collectifs | Remarques  |
|---|----------------------|-----|-------------------|-------------------|---------|---------------------|------------------------|--------------------------------------|-----------------------|--|
| 1   | Egypt 1996           | 10  | 1,270,787         | 5,902,243         | oui     | de fait             | 18/19-11-96            | district                             | non                   |  |
| 2   | Egypt 2006           | 10  | 1,740,414         | 7,282,434         | non     | de fait             | 20/21-11-06            | district                             | oui                   |  |
| 3   | Ghana 2000           | 10  | 397,097           | 1,894,133         | non     | de fait             | 26-03-00               | district                             | oui                   |  |
| 4   | Guinea 1983          | 10  | 110,777           | 457,837           | non     | les deux            | 1/2/1983               | préfecture                           | oui                   |  |
| 5   | Guinea 1996          | 10  | 108,793           | 729,071           | non     | les deux            | 1/12/1996              | préfecture                           | oui                   |  |
| 6   | Kenya 1989           | 5   | 224,861           | 1,074,098         | non     | de fait             | 25-10-89               | district                             | oui*                  |  |
| 7   | Kenya 1999           | 5   | 317,106           | 1,407,547         | non     | de fait             | 25-08-99               | district                             | oui*                  |  |
| 8   | Malawi 1987          | 10  | 186,270           | 798,669           | non     | de fait             | 1/21-09-87             | district                             | oui                   |  |
| 9   | Malawi 1998          | 10  | 229,005           | 991,393           | non     | de fait             | 1/21-06-98             | district                             | oui                   |  |
| 10  | Malawi 2008          | 10  | 299,864           | 1,341,977         | non     | les deux            | 1/21-06-08             | district                             | oui                   |  |
| 11  | Mali 1987            | 10  | 136,515           | 785,384           | non     | les deux            | 14-04-87               | district                             | oui                   |  |
| 12  | Mali 1998            | 10  | 161,880           | 991,330           | non     | les deux            | 14-04-98               | district                             | oui                   |  |
| 13  | Rwanda 1991          | 10  | 153,041           | 742,918           | non     | les deux            | 1/8/1991               | province                             | oui                   |  |
| 14  | Rwanda 2002          | 10  | 191,719           | 843,392           | non     | les deux            | 15/16-08-02            | province                             | oui                   |  |
| 15  | Senegal 1988         | 10  | 79,904            | 700,199           | non     | les deux            | 20-05-88               | département                          | non                   |  |
| 16  | Senegal 2002         | 10  | 107,999           | 994,562           | non     | les deux            | n.a.                   | département                          | non                   |  |
| 17  | Sierra Leone 2004    | 10  | 82,518            | 494,298           | non     | de fait             | 4/12/2004              | chiefdom                             | oui                   |  |
| 18  | South Africa 1996    | 10  | 993,801           | 3,621,164         | oui     | de fait             | 10/10/1996             | district                             | oui                   | 1,3% de l'échantillon n'est pas dans les ménages |
| 19  | South Africa 2001    | 10  | 991,543           | 3,725,655         | oui     | de fait             | 10/10/2001             | Municipalité                         | oui                   |  |
| 20  | South Africa 2007    | 2   | 345,170           | 1,047,657         | oui     | les deux            | 7/2/2007               | Municipalité                         | oui                   |  |
| 21  | Sudan 2008           | 15  | 1,015,408         | 5,609,295         | oui     | de fait             | 22-04-08               | Comté                                | non                   | Y compris le Soudan et le Sud Soudan             |
| 22  | Tanzania 1988        | 10  | 472,616           | 2,310,424         | oui     | de fait             | 27-08-88               | District                             | non                   |  |
| 23  | Tanzania 2002        | 10  | 841,768           | 3,732,735         | oui     | de fait             | 22-08-02               | District                             | oui                   |  |
| 24  | Uganda 1991          | 10  | 339,166           | 1,548,460         | oui     | de fait             | 11/1/1991              | Comté                                | oui                   |  |
| 25  | Uganda 2002          | 10  | 529,271           | 2,497,449         | non     | les deux            | 12/9/2002              | Comté                                | oui                   |  |
| <b>TOTAL</b>  |                      |     | <b>11,327,293</b> | <b>51,524,324</b> |         |                     |                        |                                      |                       |  |
| Source: <a href="http://ecastats.uneca.org/aicmd/en-us/samples.aspx">http://ecastats.uneca.org/aicmd/en-us/samples.aspx</a> |                      |     |                   |                   |         |                     |                        |                                      |                       |  |

**Table 2. 32 most popular variables in IPUMS-International (85,505 Sample Extracts)**

| <u>Rank</u> | <u>Label</u>                    | <u>Extracts</u> | <u>Mnemonic</u> | <u>Comment</u>      |
|-------------|---------------------------------|-----------------|-----------------|---------------------|
| 1           | Educational attainment          | 19,307          | EDATTAN         |                     |
| 2           | Age (single years to 85+)       | 19,009          | AGE             | Grouped age n=3,838 |
| 3           | Employment status               | 18,490          | EMPSTAT         |                     |
| 4           | Marital status                  | 18,214          | MARST           |                     |
| 5           | Person weight                   | 17,511          | WTPER           | Technical variable  |
| 6           | Relationship to head            | 15,783          | RELATE          |                     |
| 7           | Sex                             | 14,595          | SEX             |                     |
| 8           | Class of work                   | 12,583          | CLASSWK         |                     |
| 9           | Ownership of dwelling           | 8,050           | OWNRSHP         |                     |
| 10          | Occupation ISCO recode          | 8,004           | OCCISCO         |                     |
| 11          | School attendance               | 7,919           | SCHOOL          |                     |
| 12          | Years of schooling              | 7,576           | YRSCHL          |                     |
| 13          | Literate                        | 7,290           | LIT             |                     |
| 14          | Urban/rural                     | 7,098           | URBAN           |                     |
| 15          | Industry-general code           | 7,044           | INDGEN          |                     |
| 16          | Household weight                | 6,656           | WTHH            | Technical variable  |
| 17          | Children ever born              | 6,363           | CHBORN          |                     |
| 18          | Nativity (native/foreign born)  | 6,332           | NATIVTY         |                     |
| 19          | Occupation                      | 6,246           | OCC             |                     |
| 20          | Country of birth                | 6,153           | BPLCTRY         |                     |
| 21          | Religion                        | 6,075           | RELIG           |                     |
| 22          | Industry                        | 5,670           | IND             |                     |
| 23          | Location of spouse in household | 5,007           | SPLOC           | IPUMS unique        |
| 24          | Rule for locating spouse        | 4,171           | SPRULE          | IPUMS unique        |
| 25          | Location of mother in household | 4,153           | MOMLOC          | IPUMS unique        |
| 26          | Number of children surviving    | 4,074           | CHSURV          |                     |
| 27          | Place of residence 5 years ago  | 4,064           | MGRATE5         |                     |
| 28          | Location of father in household | 3,983           | POPLOC          | IPUMS unique        |
| 29          | Total household income          | 3,965           | INCTOT          | Household variable  |
| 30          | Earned income                   | 3,655           | INCEARN         |                     |
| 31          | Number of rooms                 | 3,465           | ROOMS           |                     |
| 32          | Consensual union                | 3,443           | CONSENS         |                     |

## Appendix A.

### Lettre d'accord: IPUMS-International (Maroc, 2009).

**Lettre d'accord entre  
Integrated Public Use Microdata Series International et  
le Haut-commissariat au Plan du Royaume du Maroc**

**Objet:** L'objet de cette lettre est de spécifier les termes et conditions régissant la diffusion par **Integrated Public Use Microdata Series International**, Université du Minnesota, des macro- et micro-données fournies par le **Haut-commissariat au Plan du Royaume du Maroc**.

1. **Propriété.** Le **Haut-commissariat au Plan du Royaume du Maroc** est le propriétaire et le détenteur des droits de propriété intellectuelle (incluant les droits de copie) des macro- et micro-données de son pays acquises auprès de lui par l'Université du Minnesota pour être distribuées par **Integrated Public Use Microdata Series International**.
2. **Usage.** Ces données sont destinées à l'usage exclusif de l'enseignement ainsi que de la recherche et édition scientifiques. Elles ne sauraient être utilisées à aucune autre fin, sauf accord écrit explicite du **Haut-commissariat au Plan** donné au préalable.
3. **Autorisation.** Pour consulter ou obtenir copie des micro-données intégrées du Maroc auprès d'**Integrated Public Use Microdata Series International**, l'utilisateur potentiel doit tout d'abord déposer un formulaire d'autorisation électronique l'identifiant comme chercheur principal, indiquant ses nom, adresse électronique et l'institution dont il relève. Obligation lui est faite d'exposer les fins poursuivies par son projet de recherche et de se conformer au règlement ci-inclus. Ce projet une fois approuvé, un mot de passe lui permettra d'obtenir des données auprès des serveurs et autres moyens électroniques de diffusion d'**Integrated Public Use Microdata Series International**, du **Haut-commissariat au Plan** ou d'autres distributeurs autorisés. L'approbation du projet entraîne pour l'utilisateur une licence d'acquérir les macro- ou micro-données intégrées du Maroc auprès d'**Integrated Public Use Microdata Series International** ou autres distributeurs autorisés. Aucun autre titre ou droit n'est concédé à l'utilisateur.
4. **Restrictions.** Il est interdit aux utilisateurs de données obtenues d'**Integrated Public Use Microdata Series International** ou autres distributeurs autorisés d'en faire quelque usage commercial que ce soit ou d'en tirer profit, à titre privé ou sous une autre forme.
5. **Confidentialité.** Les utilisateurs traiteront de façon rigoureusement confidentielle tout ce qui touche aux personnes et ménages. Toute tentative d'identification d'une personne, d'une famille, d'un ménage, d'un lieu de résidence, d'une organisation ou d'une entreprise à partir de ces micro-données est strictement proscrite. Il est également interdit de prétendre que telle ou telle personne ou entité a été identifiée à partir de ces données.
6. **Sécurité.** Les utilisateurs prendront les mesures de sécurité appropriées pour prévenir tout accès non autorisé aux micro-données recueillies auprès d'**Integrated Public Use Microdata Series International** ou de ses partenaires.

7. **Publication.** Est autorisée la publication des données et de l'analyse qui en est faite, après recherche à partir des macro- et micro-données concernant le Maroc, dans le cadre de communications scientifiques, d'articles de journaux de recherche ou autres parutions du même type. Il est exigé des auteurs qu'ils citent le **Haut-commissariat au Plan** et l'**Integrated Public Use Microdata Series International** comme sources des données concernant le Maroc. Ils doivent également mentionner que les résultats qu'ils ont obtenus et les opinions qu'ils émettent n'engagent que l'auteur/utilisateur.
8. **Violations.** Toute violation de la licence d'utilisation (§ 3) peut se conclure par un blâme professionnel, par la perte de l'emploi et/ou par des poursuites au civil. L'Université du Minnesota, des organisations scientifiques nationales et internationales et le **Haut-commissariat au Plan** veilleront à la bonne exécution des dispositions de cet accord.
9. **Partage des tâches.** L'**Integrated Public Use Microdata Series International** fournira au **Haut-commissariat au Plan** des copies électroniques de sa documentation et des données correspondant aux micro-données qu'il aura intégrées. Il transmettra également les rapports que lui auront adressés les utilisateurs des données.
10. **Jurisdiction.** Les désaccords qui puissent surgir seront traités par biais de la conciliation, la transaction et une attitude amicale. Si l'accord s'avérait impossible par ces moyens, un Tribunal d'Accord serait établi pour juger la situation d'accord avec la loi. Ce Tribunal serait composé par un arbitre sélectionné par le Tribunal International d'Arbitrage (TIA) . L'accord devra s'accorder aux principes généralement acceptés de la Loi Internationale.
11. **Ordre de priorité.** En cas de conflit entre un terme ou une condition de cette Lettre d'Accord et un terme ou une condition d'un contrat auquel cette Lettre est rattachée, le terme ou condition de cette Lettre d'Accord devra prévaloir.

Date: 04 mai 2009

Signature :

**Le Haut-commissariat au Plan du Royaume du Maroc**

Représenté par Monsieur Ahmed LAHLIMI ALAMI, Haut-commissaire au Plan

Date: 04 mai 2009

Signature :

**Integrated Public Use Microdata Series International-Université du Minnesota**

Représenté par Monsieur Robert McCaa, Professeur d'Histoire et Coordinateur des projets internationaux

**Appendix B. Snippets of Application Form to Use Restricted Microdata disseminated by IPUMS-International. See: <https://international.ipums.org/international-action/register/0>**

|   |  |
|---|--|
| <p>IPUMS International <span style="float: right;">Page 1 of 1</span></p>   | <p style="text-align: center;"><b>USAGE LICENSE</b></p>  |
| <p><b>Application to Use Restricted Microdata</b></p> <p>IPUMS-International microdata are available free of charge, but their use imposes responsibilities upon the user. To access the data, a prospective user must submit an electronic authorization form (this form) identifying the user by name, electronic address, and institutional affiliation.</p> <p>The investigator must state the purpose of the proposed project and agree to abide by the regulations specified below. If multiple investigators are involved in a project, all must register separately.</p> <p>Once a user is approved, a message will be sent by email granting access to the system. The notification licenses the user to acquire microdata from Integrated Public Use Microdata Series International or other authorized distributors. No titles or other rights are conveyed to the user.</p> <p><b>Legal notice:</b> Submission of this application constitutes a legally binding agreement between the applicant, the applicant's institution, the University of Minnesota, and the relevant official statistical authorities. Submitting false, misleading or fraudulent information constitutes a violation of this agreement. Misusing the data by violating any of the conditions detailed below also constitutes a violation of this agreement and may lead to professional censure, loss of employment, or civil prosecution under relevant national and international laws, and to sanctions against your institution, at the discretion of the University of Minnesota and the official statistical authorities.</p> <p>Information provided on this form will be kept confidential.<br/>All information on this form is required for registration unless otherwise indicated by an asterisk.</p> | <p>Please check all of the following boxes to indicate that you have read about the limitations of the IPUMS-International data and you agree to abide by the conditions of use. The purpose of this license is to specify the terms and condition under which integrated microdata samples distributed by Integrated Public Use Microdata Series International of the University of Minnesota may be used. <b>Note: The license is valid for one year and may be renewed.</b></p> <p><b>Data must not be redistributed without authorization.</b></p> <p>All data extracted from the IPUMS-International database are intended solely for the use of the licensee. Under IPUMS-International agreements with collaborating agencies, redistribution of the data to third parties is prohibited. Each member of a research team using the data must apply for access and be licensed individually.</p> <p><b>The microdata are intended only for scholarly research and educational purposes.</b></p> <ul style="list-style-type: none"><li><input type="checkbox"/> These microdata are provided for the exclusive purposes of teaching and scholarly research, and may not be used for any other purposes without explicit written approval from the relevant official statistical authority.</li></ul> <p><b>Commercial use and redistribution of the microdata is strictly prohibited.</b></p> <ul style="list-style-type: none"><li><input type="checkbox"/> Users are prohibited from using microdata acquired from the Integrated Public Use Microdata Series International or other authorized distributors in the pursuit of any commercial or income-generating venture either privately, or otherwise.</li></ul> <p><b>Use of the microdata must follow strict rules of confidentiality.</b></p> <p>Users will maintain the confidentiality of persons and households. Any attempt to ascertain the identity of persons or households from the microdata is prohibited. Alleging that a person or household has been identified in these data is also prohibited. Statistical results that might reveal the identity of persons or entities may not be reported or published in any form.</p> <p><b>The microdata must always be safely secured.</b></p> <ul style="list-style-type: none"><li><input type="checkbox"/> Users will implement security measures to prevent unauthorized access to microdata acquired from Integrated Public Use Microdata Series International, its partners or authorized distributors. Upon the completion of this research, data may be retained only if they can be safely secured. If security cannot be guaranteed, the microdata must be destroyed.</li></ul> |
| <p style="text-align: center;"><b>PERSONAL INFORMATION</b></p> <p style="text-align: center;">*****</p> <p style="text-align: center;"><b>INSTITUTIONAL AFFILIATION</b></p> <p>IPUMS-International staff must confirm the identity of prospective users. To speed the processing of your application, please provide as much of the following</p>   |  |

Name of institution or employer

Your email address at institution (\*)

Web link showing your affiliation with institution (\*)

Email address of employer, supervisor, or instructor (\*)

Phone number of institution (\*)

Does your institution have an Institutional Review Board (IRB), or Office for Human Subject Protections, Professional Conduct or similar committee?

No

Yes; Name of board or office

#### RESEARCH PROJECT

Please provide at least 75 words *in English* describing your research project or educational use for the data. This description will be used to evaluate your application.

.....

If your research is funded by someone other than your employer, indicate the name of the granting institution, title of grant, and other pertinent information. (\*)

#### Scholarly publications are permitted, and must be cited appropriately.

- The publishing of research results based on IPUMS-International microdata is permitted in communications such as scholarly papers, journals and the like. The authors of these communications are required to cite Integrated Public Use Microdata Series-International

and the relevant official statistical authority as the source of the microdata, and to indicate that the results and views expressed are those of the author. Users are requested to provide the IPUMS-International staff with a full citation for any publications resulting from their work with these data.

#### Any violation of this license agreement will result in disciplinary action, including possible loss of employment.

- Violation of this agreement will lead to revocation of this license, recall of all microdata acquired, a motion of censure to the relevant professional organization(s) and civil prosecution under national or international statutes, at the discretion of the Regents of the University of Minnesota and the official statistical agencies. Sanctions likewise may be taken against the institution with which the violator is affiliated.
- User agrees to notify [ipums@pop.umn.edu](mailto:ipums@pop.umn.edu) regarding errors in the data.

**Appendix C. Educational attainment harmonized codes: 23 African census samples integrated in IPUMS-International**  
**“X” indicates that the code is present in the respective sample**

| Code | Label                                  | Country | EG | GH | GN | GN | KE | KE | MW | MW | MW | ML | ML | RW | SN | SN | SL | ZA | ZA | ZA | SD | TZ | TZ | UG | UG |  |
|------|--|---------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|--|
|      |  | Census  | 06 | 00 | 83 | 96 | 89 | 99 | 87 | 98 | 08 | 87 | 98 | 02 | 88 | 02 | 04 | 96 | 01 | 07 | 08 | 88 | 02 | 91 | 02 |  |
| 0    | NIU (not in universe)                  |         | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |  |
| 100  | LESS THAN PRIMARY COMPLETED            |         | X  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  |  |
| 110  | Non schooling                          |         | .  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |  |
| 120  | Some primary                           |         | .  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |  |
| 130  | Primary (4 years)                      |         | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  |  |
|      | PRIMARY COMPLETED, LESS THAN SECONDARY |         |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |  |
|      | Primary completed                      |         |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |  |
| 211  | Primary (5 years)                      |         | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  |  |
| 212  | Primary (6 years)                      |         | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |  |
|      | Lower secondary completed              |         |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |  |
| 221  | General and unspecified track          |         | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |  |
| 222  | Technical track                        |         | .  | .  | .  | .  | .  | .  | .  | .  | .  | X  | X  | .  | .  | X  | .  | .  | .  | .  | .  | .  | .  | X  | X  |  |
|      | SECONDARY COMPLETED                    |         |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |  |
|      | General or unspecified track           |         |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |  |
| 311  | General track completed                |         | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | .  | X  | X  | X  | X  | X  | X  | X  | X  |  |
| 312  | Some college/university                |         | .  | X  | X  | X  | X  | X  | X  | X  | .  | X  | X  | X  | X  | .  | .  | .  | .  | .  | .  | X  | X  | X  | X  |  |
| 320  | Technical track                        |         | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | X  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  |  |
| 321  | Secondary technical degree             |         | .  | X  | X  | X  | .  | .  | .  | .  | .  | X  | X  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  |  |
| 322  | Post-secondary technical               |         | X  | X  | .  | X  | .  | .  | .  | .  | X  | X  | .  | .  | .  | X  | .  | .  | .  | X  | .  | .  | .  | X  |    |  |
| 400  | UNIVERSITY COMPLETED                   |         | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |  |
| 999  | UNKNOWN/MISSING                        |         | X  | .  | X  | X  | X  | .  | X  | .  | X  | X  | X  | X  | .  | X  | X  | .  | X  | X  | X  | X  | X  | X  | .  |  |

## Appendix D. 42 (of 60) Integrated Household Variables: Availability for 13 African Countries (25 Censuses)

| Variable Name | Label                                   | Country<br>Census year | Country |    |    |    |    |    |    |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|---------------|---|------------------------|---------|----|----|----|----|----|----|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|               |   |                        | EG      | EG | GH | GN | GN | KE | KE | MWM | MW | ML | ML | RW | RW | SN | SN | SL | ZA | ZA | ZA | SD | TZ | TZ | UG | UG |    |
|               |   | Sum                    | 96      | 06 | 00 | 83 | 96 | 89 | 99 | 87  | 98 | 08 | 87 | 98 | 91 | 02 | 88 | 02 | 04 | 96 | 01 | 07 | 08 | 88 | 02 | 91 | 02 |
| 1 RECTYPE     | Record type                             | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 2 CNTRY       | Country [preselected]                   | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 3 YEAR        | Year [preselected]                      | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 4 SAMPLE      | IPUMS sample identifier [preselected]   | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 5 SERIAL      | Household serial number [preselected]   | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 6 PERSONS     | Number of person records in the house   | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 7 WTHH        | Household weight                        | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 8 SUBSAMP     | Subsample number                        | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 9 GQ          | Group quarters status                   | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 10 REGIONW    | Continent and region of country         | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 11 composite  | Place of Residence: Level 1             | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 12 HHTYPE     | Household classification                | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 13 NFAMS      | Number of families in household         | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 14 NCOUPLS    | Number of married couples in household  | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 15 NMOTHR     | Number of mothers in household          | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 16 NFATHRS    | Number of fathers in household          | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 17 HEADLOC    | Head's location in household            | 25                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 18 UNREL      | Number of unrelated persons             | 24                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 19 OWNRSHP    | Ownership of dwelling                   | 24                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 20 ELECTRC    | Electricity                             | 24                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 21 TOILET     | Toilet                                  | 24                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 22 WATSUP     | Water supply                            | 23                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 23 URBAN      | Urban-rural status                      | 21                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 24 FUELCK     | Cooking fuel                            | 21                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 25 ROOMS      | Number of rooms                         | 21                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 26 composite  | Place of Residence: Level 2             | 19                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 27 FLOOR      | Floor material                          | 16                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 28 WALL       | Wall or building material               | 16                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 29 ROOF       | Roof material                           | 16                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 30 RADIO      | Radio in household                      | 14                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 31 ANYMORT    | Any deaths in household last year       | 13                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 32 PHONE      | Telephone availability                  | 12                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 33 TV         | Television set                          | 11                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 34 MORTNUM    | Number of deaths in household last year | 10                     | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 35 REFRIG     | Refrigerator                            | 9                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 36 CELL       | Cellular phone availability             | 8                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 37 AUTOS      | Automobiles available                   | 8                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 38 COMPUTR    | Computer                                | 8                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 39 SEWAGE     | Sewage                                  | 7                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 40 TRASH      | Trash disposal                          | 7                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 41 BATH       | Bathing facilities                      | 6                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 42 composite  | Place of Residence: Level 3             | 5                      | 1       | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |

## Appendix E. 88 (of 108) Integrated Person Variables: Availability for 13 African Countries (25 Censuses)

| Variable     | Country   | EG  | EG | GH | GN | GN | KE | KE | MWM | MW | ML | ML | RW | RW | SN | SN | SL | ZA | ZA | ZA | SD | TZ | TZ | UG | UG |    |    |
|--------------|---|-----|----|----|----|----|----|----|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|              |   |     |    |    |    |    |    |    |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 96 | 06 |
| Name         | Label   | Sum | 49 | 44 | 52 | 42 | 61 | 60 | 56  | 51 | 50 | 64 | 49 | 54 | 48 | 65 | 52 | 57 | 72 | 55 | 62 | 69 | 66 | 56 | 64 | 56 | 78 |
| 1 PERNUM     | Person number [preselected]                     | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 2 WTPER      | Person weight [preselected]                     | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 3 MOMLOC     | Mother's location in household                  | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 4 POPLOC     | Father's location in household                  | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 5 SPLOC      | Spouse's location in household                  | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 6 PARRULE    | Rule for linking parent                         | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 7 SPRULE     | Rule for linking spouse                         | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 8 STEPMOM    | Probable stepmother                             | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 9 STEPPOP    | Probable stepfather                             | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 10 POLYMAL   | Man with more than one wife linked              | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 11 POLY2ND   | Woman is second or higher order wife            | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 12 FAMUNIT   | Family unit membership                          | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 13 FAMSIZ    | Number of own family members in household       | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 14 NCHILD    | Number of own children in household             | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 15 NCHLT5    | Number of own children under age 5 in household | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 16 ELDCH     | Age of eldest own child in household            | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 17 YNGCH     | Age of youngest own child in household          | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 18 RELATE    | Relationship to household head                  | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 19 AGE       | Age   | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 20 AGE2      | Age, grouped into intervals                     | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 21 SEX       | Sex   | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 22 MARST     | Marital status                                  | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 23 EMPSTAT   | Employment status                               | 25  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 24 NATIVTY   | Nativity status                                 | 24  | .  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 25 composite | Place of birth, geographic level 1              | 23  | 1  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 26 EDATTAN   | Educational attainment, international re        | 23  | .  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 27 OCCISCO   | Occupation, ISCO general, 1-digit               | 23  | 1  | 1  | 1  | 1  | 1  | 1  | .   | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 28 SCHOOL    | School attendance                               | 22  | .  | 1  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 29 OCC       | Occupation, unrecoded                           | 22  | .  | 1  | 1  | 1  | 1  | 1  | .   | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 30 CLASSWK   | Class of worker                                 | 22  | 1  | 1  | 1  | 1  | 1  | .  | .   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 31 CHBORN    | Children ever born                              | 21  | .  | .  | 1  | .  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 32 CITIZEN   | Citizenship                                     | 21  | 1  | 1  | 1  | 1  | 1  | .  | .   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 33 LIT       | Literacy  | 21  | 1  | 1  | 1  | 1  | 1  | 1  | .   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 34 CHSURV    | Children surviving                              | 20  | .  | .  | .  | .  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 35 YRSCHL    | Years of schooling                              | 20  | .  | .  | 1  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 36 BPLCTRY   | Country of birth                                | 19  | .  | .  | .  | 1  | 1  | 1  | 1   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 37 composite | Place of previous residence 1/5/10 year         | 19  | 1  | 1  | 1  | .  | 1  | 1  | 1   | 1  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 38 MORTMOT   | Mortality status of mother                      | 18  | .  | .  | .  | .  | 1  | 1  | .   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 39 INDGEN    | Industry, general recode                        | 18  | 1  | 1  | 1  | 1  | .  | .  | .   | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 40 IND       | Industry, unrecoded                             | 18  | 1  | 1  | 1  | 1  | .  | .  | .   | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 41 CHBORNF   | Number of female children ever born             | 17  | .  | .  | 1  | .  | 1  | 1  | 1   | 1  | 1  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 42 CHBORNM   | Number of male children ever born               | 17  | .  | .  | 1  | .  | 1  | 1  | 1   | 1  | 1  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 43 MORTFAT   | Mortality status of father                      | 17  | .  | .  | .  | .  | 1  | 1  | .   | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |
| 44 NATION    | Country of citizenship                          | 17  | 1  | 1  | .  | 1  | 1  | .  | .   | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | .  | 1  | 1  | 1  | 1  | .  |    |
| 45 DISABLE   | Disability status                               | 17  | 1  | 1  | .  | .  | 1  | .  | .   | .  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |    |

**Appendix E (continued). 88 (of 108) Integrated Person Variables: Availability for 13 African Countries (25 Censuses)**

| Variable     | Country                                   | Census year |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|--------------|---|-------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|              |   | 96          | 06 | 00 | 83 | 96 | 89 | 99 | 87 | 98 | 08 | 87 | 98 | 91 | 02 | 88 | 02 | 04 | 96 | 01 | 07 | 08 | 88 | 02 | 91 | 02 |    |
| Name         | Label                                     | Sum         | 49 | 44 | 52 | 42 | 61 | 60 | 56 | 51 | 50 | 64 | 49 | 54 | 48 | 65 | 52 | 57 | 72 | 55 | 62 | 69 | 66 | 56 | 64 | 56 | 78 |
| 46 POLYGAM   | Polygamous union                          | 16          | 1  | 1  | .  | 1  | 1  | 1  | 1  | .  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | .  | .  | .  | .  | 1  |
| 47 RELIG     | Religion                                  | 16          | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | 1  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | 1  | 1  |
| 48 DISBLND   | Blind or vision-impaired                  | 16          | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 49 CHSURVF   | Number of female children surviving       | 15          | .  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | .  | .  | 1  | 1  | .  | 1  | 1  | .  | .  | .  | 1  | 1  | 1  | .  | 1  |
| 50 CHSURVM   | Number of male children surviving         | 15          | .  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | .  | .  | 1  | 1  | .  | 1  | 1  | .  | .  | .  | 1  | 1  | 1  | .  | 1  |
| 51 DISDEAF   | Deaf or hearing-impaired                  | 15          | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | 1  | 1  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 52 DISMUTE   | Mute                                      | 14          | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | 1  | 1  | .  | 1  | .  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 53 BRTHLYR   | Number of births last year                | 13          | .  | .  | 1  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | .  | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | 1  | 1  | 1  | .  |
| 54 MGYRS1    | Years residing in current locality        | 12          | 1  | 1  | .  | .  | 1  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 55 RESIDENT  | Residence status: de facto, de jure [pre: | 11          | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  |
| 56 DISEMP    | Employment disability                     | 11          | 1  | 1  | 1  | 1  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | 1  | 1  | 1  | 1  | .  | 1  |
| 57 CHDEAD    | Number of children dead                   | 10          | .  | .  | .  | .  | .  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 58 DEADFEM   | Number of female children dead            | 10          | .  | .  | .  | .  | .  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 59 DEADMAL   | Number of male children dead              | 10          | .  | .  | .  | .  | .  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 60 BIRTHYR   | Year of birth                             | 9           | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | .  | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  |
| 61 composite | Language, mother tongue                   | 9           | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  |
| 62 ISCO88A   | Occupation, ISCO-1988, 3-digit            | 9           | .  | 1  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | 1  | .  | 1  | 1  | 1  | .  | .  | .  | 1  | 1  |
| 63 DISLOWR   | Disability affecting lower extremities    | 9           | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | .  | .  | .  | 1  | .  | 1  | 1  | 1  | .  | 1  | .  | .  | .  | 1  | 1  |
| 64 DISMNTL   | Mental disability                         | 9           | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| 65 CONSENS   | Consensual union                          | 8           | .  | .  | 1  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | 1  | 1  | 1  | .  | .  | .  | 1  | .  |
| 66 HOMECHD   | Number of own children in household       | 8           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 67 HOMEFEM   | Number of own female children in hou      | 8           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 68 HOMEMAL   | Number of own male children in house      | 8           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 69 AWAYCHD   | Number of own children living elsewh      | 8           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 70 AWAYFEM   | Number of own female children living      | 8           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 71 AWAYMAL   | Number of own male children living el     | 8           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | 1  | 1  | 1  | 1  |
| 72 LSTBSEX   | Sex of last birth                         | 7           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | 1  | .  | 1  | 1  | 1  | 1  |
| 73 BTHSURV   | Children surviving from births last yea   | 7           | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | 1  | .  | .  | 1  | 1  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | .  |
| 74 composite | Ethnicity                                 | 7           | .  | .  | 1  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | 1  | 1  |
| 75 MGRATEP   | Migration status, previous residence      | 7           | 1  | 1  | .  | .  | 1  | .  | .  | .  | .  | 1  | .  | 1  | .  | .  | .  | .  | 1  | .  | 1  | .  | .  | .  | .  | .  | .  |
| 76 DISUPPR   | Disability affecting upper extremities    | 7           | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | 1  | 1  | 1  | .  | .  | 1  | .  | .  | .  | .  | 1  |
| 77 BIRTHMO   | Month of birth                            | 6           | 1  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  |
| 78 LSTBMTH   | Month of last birth                       | 6           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | 1  |
| 79 LSTBYR    | Year of last birth                        | 6           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | 1  |
| 80 SPKENG    | Speaks English                            | 6           | .  | .  | 1  | .  | 1  | .  | .  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | .  |
| 81 MGCTRY1   | Country of previous residence             | 6           | .  | .  | .  | .  | 1  | .  | .  | .  | .  | 1  | .  | 1  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | 1  | 1  |
| 82 DISPSYC   | Psychological disability                  | 6           | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | 1  | .  | 1  | 1  | .  | .  | .  | .  | 1  |
| 83 MGRATE1   | Migration status, 1 year                  | 5           | .  | .  | .  | .  | .  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | 1  | .  | .  | .  |
| 84 DISORIG   | Origin of disability                      | 5           | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  |
| 85 LSTSTAT   | Mortality status of last birth            | 4           | .  | .  | .  | .  | .  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  |
| 86 RACE      | Race or color                             | 4           | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  |
| 87 composite | Second language spoken                    | 4           | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  |
| 88 EMPSECT   | Sector of employment                      | 4           | 1  | 1  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | 1  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  | .  |

**Appendix F. IPUMS-I Top 40 University/Research Institutions by Number of Extracts**

| Rank | Institution                                  | N   | Rank | Institution   | N   |
|------|--|-----|------|---|-----|
| 1    | University of Michigan                       | 742 | 21   | University of North Carolina – Chapel Hill            | 203 |
| 2    | Columbia University                          | 701 | 22   | Universite Montesquieu-Bordeaux IV, France            | 196 |
| 3    | Universitat de Barcelona, Spain              | 615 | 23   | University of California - San Diego                  | 189 |
| 4    | Harvard University                           | 589 | 24   | University of Utah                                    | 189 |
| 5    | Inter - American Development Bank            | 499 | 25   | World Health Organization, Switzerland                | 183 |
| 6    | Arizona State University                     | 495 | 26   | University of Virginia                                | 182 |
| 7    | National University of Singapore, Singapore  | 467 | 27   | Michigan State University                             | 178 |
| 8    | World Bank                                   | 408 | 28   | Intl. Institute for Applied Systems Analysis, Austria | 165 |
| 9    | University of California - Berkeley          | 362 | 29   | University of Sussex, U.K.                            | 158 |
| 10   | Universidade Federal de Minas Gerais, Brazil | 314 | 30   | London School of Economics, U.K.                      | 157 |
| 11   | University of Chicago                        | 285 | 31   | Dartmouth College                                     | 155 |
| 12   | Universidad del Valle, Colombia              | 270 | 32   | University of Guelph, Canada                          | 148 |
| 13   | Institute for Health Metrics & Evaluation    | 260 | 33   | Institut de Recherche pour le Developpement, France   | 148 |
| 14   | Princeton University                         | 237 | 34   | Banco de la Republica, Colombia                       | 145 |
| 15   | University of Wisconsin - Madison            | 234 | 35   | Yale University                                       | 143 |
| 16   | Brown University                             | 229 | 36   | University of Tübingen, Germany                       | 143 |
| 17   | University of Vienna, Austria                | 229 | 37   | Org. of Economic Cooperation & Development, Fr.       | 140 |
| 18   | University of Pittsburgh                     | 227 | 38   | Catholic University Leuven, Brussels                  | 139 |
| 19   | University of Delaware                       | 213 | 39   | Brigham Young University                              | 138 |
| 20   | El Colegio de México, México                 | 214 | 40   | University of Queensland, Australia                   | 136 |

Source: IPUMS-International User Statistics Database, April 18, 2011 (excludes IPUMS's home, the University of Minnesota)