**Title: IPUMS and AICMD Add Significant Value to African Census Microdata.**

**Authors:**
**Robert McCaa (rmccaa@umn.edu), Professor Emeritus of Population History; Ambassador, IPUMS-International, Minnesota Population Center**
**Patricia Kelly-Hall (pkelly@umn.edu), Coordinator, IPUMS-International, Minnesota Population Center**

**Web-sites:**
**www.ipums.org/international**
**http://ecastats.uneca.org/aicmd/Home/tabid/40/language/en-US/Default.aspx**

**"Dissemination [means] opening up the value inherent in our data"**
Seminar on Emerging Trends in Data Communication and Statistics, New York Feb. 19, 2010
Walter Radermacher (President, Eurostat) and Pieter Everaers (Director, Eurostat)

## Introduction.

African researchers and policy makers need access to census microdata to analyze challenging social, demographic, and economic transformations underway throughout the continent (McCaa, Esteve, Ruggles and Sobek 2006). IPUMS-International (www.ipums.org/international) offers statistical agencies worldwide a one stop solution to disseminating anonymized microdata samples. The project asumes the responsibilities, risks, and costs for recovering, archiving, anonymizing, integrating, and disseminating microdata worldwide (McCaa and Thomas 2009). Researchers and policy makers access extracts of the data free of cost—regardless of country of birth, residence, or nationality—thanks to sustained funding by the National Science Foundation and the National Institutes of Health (USA).

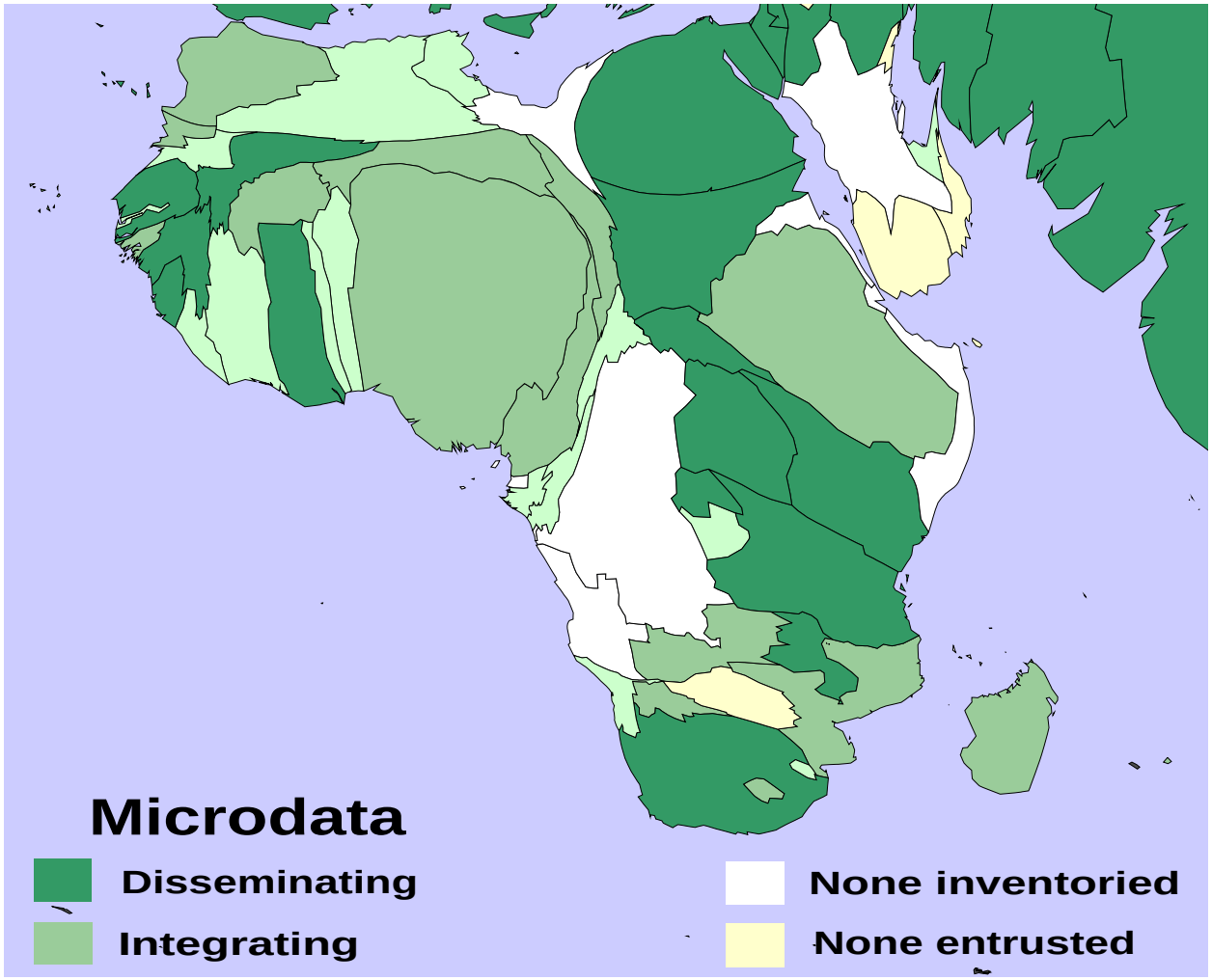
Begun in 1999, IPUMS-International enjoys the satisfaction of more than 5,000 researchers in some 100 countries and of over 90 national statistics office-partners. As of June 2011, the global project has completed the anonymization and integration of population microdata for 62 countries worldwide, totaling 185 samples and 397 million person records. Thanks to the exceedingly generous cooperation of NSOs, the IPUMS-International database is expanding at the rate of 5-10 additional countries per year. By 2015, it is likely that coverage will extend to 85 or more countries, encompassing over four-fifths of the world's population.

The project is led by the Minnesota Population Center (MPC). In Africa, the African Center for Statistics' AICMD initiative partners with the MPC in a pan-African collaboratory of National Statistical Offices, universities, and major research centers (http://ecastats.uneca.org/aicmd/Home/tabid/40/language/en-US/Default.aspx). Twenty-nine African National Statistics Offices, encompassing 85% of the continent's population, entrust census microdata and national surveys to IPUMS under a uniform memorandum of understanding (Appendix A, Figure 1 and Table 1). Access is governed by means of a stringent licensing agreement (Appendix B).

**Avoid marginalization; entrust 2010 round census microdata without delay.** The Journal of Official Statistics, in a special issue on microdata access and confidentiality, cautioned that "*there is a real risk that NSOs may become marginalized if they are unable to meet the requirements of researchers in this increasingly important area of [micro]data provision*"

(Editors, 2010:55). African statistical agencies not currently cooperating with the IPUMS/ AICMD initiatives are invited to contact the authors of this paper. Participating statistical agencies are encouraged to entrust microdata for the 2010 round of censuses without undue delay to facilitate timely access to anonymized, integrated samples. Doing so will increase stakeholders and, in the words of the President of Eurostat (see epigraph), open "up the value inherent in" census microdata.

**Figure 1. Cartogram weighted by population size of African countries participating in the IPUMS-International collaboratory, November 2011: 3 shades of green--integrated (dark), integrating (medium), and negotiating (light).**




The fifty African countries with one million or more inhabitants may be divided into four groups, according to participation in the IPUMS/AICMD initiative:

**a. 14 with microdata in dissemination** (number of microdatasets in parentheses):: Egypt (2), Ghana (2), Guinea (2), Kenya (2), Malawi (3), Mali (2), Rwanda (2), Senegal (3), Sierra Leone (1), South Africa (3), South Sudan (1), Sudan (4), Tanzania (2), and Uganda (2).

**b. 15 with microdata being integrated**: Botswana (3), Burkina Faso (3), Cameroun (3), Cape Verde (2), Chad (2), Ethiopia (3), Lesotho (2), Madagascar (1), Mauritius (2), Morocco (3), Mozambique (2), Niger (2), Nigeria (2 General Household Surveys), Togo (2) and Zambia (2).

**c. 15 are *not* participating,** despite a persistent campaign dating back to 2005: Algeria, Benin, Burundi, Central African Republic, Comoros, Republic of Congo, Cote d'Ivoire, Gabon, The Gambia, Liberia, Mauritania, Namibia, Swaziland, Tunisia, and Zimbabwe.

**d. 6 do not, at present, have census microdata:** Angola, Congo Democratic Republic, Djibouti, Eritrea, Libya, and Somalia.

The 2012 IPUMS launch is scheduled to integrate samples for Morocco (1982, 1994 and 2004) as well as six nations outside Africa (El Salvador, Indonesia, Mexico, Nicaragua, Turkey, and Uruguay). More African countries will be included in future launches, as expeditiously as the 2010 round census microdata become available. Unfortunately, as of late 2011, seven countries still have not entrusted microdata for enumerations conducted 2, 3, 4, and even 5 years ago: Cameroun (2005), Chad (2009), Ethiopia (2007), Kenya (2009), Lesotho (2006), Mozambique (2007), and Nigeria (2006).

**IPUMS-I add value to census microdata in 4 broad categories.** IPUMS-International seeks to open the value inherent in census microdata for academic researchers, policy makers and public officials. IPUMS-I is *not* simply a conduit for passing census samples from NSOs along to researchers. Instead, typically, two or more years of labor are invested by a sizeable group of experienced MPC personnel to prepare anonymized, integrated microdata and metadata for dissemination. In the twenty-first century, handing along a copy of the source microdata and a codebook is *not* sufficient for high quality research. This paper discusses 16 value-addeds offered by IPUMS-International. These may be categorized into four groups: statistical confidentiality, integration, dissemination and ethics.

## I. Statistical Confidentiality and Security.

**1. Microdata Security and Statistical Confidentiality.** Neither the MPC nor its employees ever disseminate original source microdata. These data are exceedingly sensitive and for that reason only seasoned, specially trained, full time researchers are allowed to work with the data until they are anonymized. MPC employees are subject to civil fine (up to US$250,000), criminal prosecution and loss of employment for violating University of Minnesota security procedures. The University legal authority assumes responsibility for protecting the total security and confidentiality of data entrusted to the MPC.

IPUMS-I is one of only a handful of academic organization disseminating census microdata that has been audited by external review. The reviewer, Mr. Dennis Trewin, the respected Australian statistician and chair of the UNECE Task Force Managing Statistical Confidentiality and Microdata Access, cited the computing environment of the MPC as "best practice," "the standard of the best statistical offices", etc. Mr. Trewin conducted a week-long, on-site inspection of the MPC and concluded his report as follows (Trewin 2007):

> Without question IPUMS-International meets the four Core Principles outlined in CES [Conference of European Statisticians] (2007). It is cited in CES (2007) as a Case Study of good practice. This review confirms its status as good practice for Data Repositories. Indeed it is likely to provide the best practice for a Data Repository for international statistical data [emphasis added].

At the UN-ECE Expert Group Meeting on Statistical Data Confidentiality, November 2005, we explained the IPUMS-International data dissemination security procedures as follows (McCaa and Esteve 2005):

> When the extract is ready (usually in a matter of minutes), the researcher is notified by email that the data should be retrieved within 72 hours. A link is provided to a password-protected site for downloading the specific extract. The

> data are encrypted during transmission using 128-bit SSL (Secure Sockets Layer) encryption standard, matching the level used by the banking and other industries where security and confidentiality are essential. The researcher may then securely download the file, decompress it and proceed with the analysis using the supplied integrated metadata consisting of variable names and labels.

This method of dissemination continues to weather the test of time, and indeed as usage soars, the rapid acceleration of internet transmission speeds has validated IPUMS-I security protocols.

**2. Statistical disclosure control protections.** The microdata are subjected to strong, uniform legal, administrative, and technical statistical disclosure controls providing greater protections for the group as a whole than for any single statistical office that chooses to "go it alone" (McCaa, Ruggles and Sobek 2010). The most important technical statistical disclosure control is the suppression of records by the use of sub-sampling. All the values in the records outside the sample are suppressed. Second comes the suppression of names and low level geographical detail.

Each statistical authority balances the confidentiality/utility trade-off by instructing the IPUMS-I project as to the minimum threshold for identifiable geographical units. For many countries, the threshold is commonly set at 20,000 inhabitants. Others place it as high as 100,000 (United States) or in the most extreme case (Netherlands) all administrative geography is suppressed. We are gratified that in some cases our statistical agency partners have reconsidered earlier decisions, offering higher precision samples (Mexico 1990 increased from one to ten percent) and greater detail. In the case of Colombia, the geographical threshold, initially set at 100,000, was reduced to 20,000 after Colombian geographers vigorously complained. The Colombian statistical agency not only reduced the threshold, but also harmonized the geo-statistical identifiers so that all the census microdata samples for Colombia could be disseminated with a single set of geographical codes.

Additional statistical disclosure protections are provided by randomly ordering the records and swapping the geographical identifiers of an undisclosed number of households. This means that no one can state with certainty that an individual or household has been identified.

In consultation with the national statistical office, some variables may be top-coded, others may be subjected to global recoding, deletion of digits for hierarchical variables (occupation, industry, geography), or the suppression of a variable entirely. Decisions are made in consultation with the corresponding national statistical authority. Weight variables and expansion factors are usually not an issue because most of the samples are implicitly stratified so that all records carry an identical weight.

**3. Restricted access.** Access to the IPUMS-International microdata is restricted—despite the "P" in IPUMS. Would-be users must submit a detailed electronic application both to establish research bona-fides and to explain need for access. An essential part of the process is to agree to ten stringent restrictions on condition of use—prohibiting redistribution, restricting to scholarly use, prohibiting commercial user, protecting confidentiality, assuring security, enforcing strict rules of confidentiality, permitting scholarly publication, citing properly, threatening disciplinary action for violations, and the reporting of errors. In other words, the IPUMS-I is a "trusted user" access system.

Agreeing to the conditions of use binds both the researcher and the researcher's institution. The Legal Counsel of the University of Minnesota is poised to strike at the first indication of misuse. Both the individual researcher and the researcher's institution are

responsible for maintaining security and enforcing the license agreement. Violations are likely to lead to sanctions against both the individual and the offending researcher's institution.

IPUMS-I resolves the conundrum of managing the broadest possible access to sensitive microdata while protecting statistical confidentiality. Many statistical agencies have long wanted to make census microdata available to researchers, but lack the substantial material a nd human resources required to implement and manage secure systems. IPUMS-I is the only academic organization disseminating census microdata that is cited as good practice by the Conference of European Statisticians Task Force on Managing Statistical Confidentiality and Microdata Access (2007).

Although access to the IPUMS-International microdata is free of cost, usage is restricted to bona-fide researchers who agree to abide by stringent conditions of use (Appendix B). IPUMS disseminates extracts, custom-tailored to the precise research needs of each user. The average IPUMS extract consists of a mere 10 variables.

This contrasts with the practices of most statistical offices where census microdata are disseminated as complete sets, consisting of a data dictionary and an entire sample containing all variables and all person records. Typically, under the traditional approach, when requests are fulfilled, each researcher receives exactly the same set of data and documentation. Given the massive size of the IPUMS-International database, disseminating the full set of variables and unvarying size of samples is impractical

Despite the the stringent conditions of use and restrictions more than five thousand researchers—representing over 100 countries and 900 institutions—are approved for access to the IPUMS-I database. More than one-third of IPUMS-I trusted users request access to microdata for a single country. A large fraction of these are resident abroad and seek access to data for their own country of identity.

## II. Integration

**4. Comprehensive source metadata.** Researchers must have ready access to the original census documentation in the official language. At a minimum, census questionnaires, enumerator instructions or training manuals, and codebooks are required. Additional metadata regarding the organization, preparation work, and actual census taking is also valuable to the IPUMS-International project and is catalogued and archived with all other documents received. Original hardcopy or PDF documents are preferred for published metadata materials. Our goal is to provide an archived collection of high-quality PDF files for all forms of metadata pertaining to census microdata. Source documents are made available to researchers in English and the official language of the original.

**5. Integrated, DDI compatible metadata**—cross national and over time. To facilitate the research process microdata are integrated for all censuses and countries, including detailed descriptions of each census, sample and variable.

**a. Censuses and samples**. IPUMS metadata offer detailed descriptions of each census in the database, listing the title, year, universe, de jure/de facto, enumeration unit, official census day, forms, field work period and type, respondent and estimates of undercount, if any. Images of census enumeration forms and instructions manuals are available in the official language and the text in English translation. Each sample is described with regard to source, sample design, sampling unit, sample fraction, number of person records, sample weights, dwelling or housing units, vacant dwellings, households, group quarters and special populations.

**b. Variable descriptions, source texts, and codes**. IPUMS metadata define each integrated variable and describe basic characteristics: availability by census, universe of the variable or question, codes, source (enumeration) text, and non-harmonized variables used for integration. Access to this information is through clickable hypertext on the IPUMS website. A general comparability discussion is provided for every variable, with country or census specific discussions focusing on departures from standard practice. The purpose of these discussions is to highlight important contrasts. Clicking "Enumeration text" leads to source questions and corresponding instructions in English for each selected census. Additional clicks yield views of the original documentation in image form so that researchers may study lay-out and actual wording in the official language.

Researchers navigate the integrated metadata to quickly examine the actual wording of census questions and instructions to enumerators for any combination of countries and census years. the MPC integration team applies XML tags to the census documents, associating the variables in the microdata with the concepts in the text. The tagged material is then imported into a database. Once this step is completed, metadata may be retrieved dynamically for any combination of countries and census years, variable-by-variable. Initially this tool was developed to speed the work of the integration team. Once its utility became apparent, we harnessed the dynamic metadata system to the web-site, to permit open access to the metadata. Metadata are custom-tailored to each researcher's request and available as a single, searchable file in DDI or text format.

**6. Integrated microdata.** The principal benefit of IPUMS to researchers and NSOs alike is integration of multiple microdata samples for each country—typically beginning with the earliest census for which microdata exist or are recoverable. For decades, many NSOs have provided a sample for the most recent census, but few re-examine earlier censuses to harmonize successive datasets or to draft new documentation to facilitate comparative analysis of two or more censuses. Modern statistical offices construct a census sample and a data dictionary for national researchers. Five or ten years later, with the ensuing census, the process is repeated with little guidance on enhancing the comparability of successive census datasets.

We must reiterate that the IPUMS project does *not* disseminate census files entrusted by national statistical offices. Instead high-precision census samples are anonymized (McCaa et. al. 2006) and integrated, variable-by-variable, using a composite coding system (Esteve and Sobek, 2003). Samples are integrated both chronologically and cross-nationally. Integrated metadata are constructed by means of meticulous study of comprehensive original source documentation and after extensive analysis of the microdata. Thousands of hours are devoted to analyze, discuss, debate, draft, test and re-test until the microdata integration is validated for dissemination to researchers. The process is repeated with each annual launch of additional census samples into the IPUMS database.

The basic goal of the IPUMS-I harmonization effort is to simplify the use of the data while losing no meaningful information. This is challenging because to make the data simple for comparative analysis across time and space, it is necessary to develop comparable coding schemes. Microdata are integrated so that identical concepts (variables, categories) have identical codes. To avoid the loss of important information for those samples that have even more detail, IPUMS-I uses a composite coding strategy to retain all original detail, and at the same time provide comparable codes across samples. With composite codes, researchers may easily compare across time and space, yet nuances in changes of meaning are readily discernible. The first digit, which we call the "general code," provides information that is available across all

samples (the lowest common denominator data). The next one or two digits provides additional information available in a substantial subset of the samples. Trailing digits provide detail that is only rarely available. Where information is not available for a particular sample, a zero place-holder is assigned to that digit.

As an example of the IPUMS method of integrating variables, consider the concept "educational attainment," the single most widely used variable in the IPUMS-International database. Most census microdata with information on this measure indicate whether the respondent completed primary, secondary or higher schooling or no schooling at all. Thus the first digit of the IPUMS-International composite code consists of four categories (1-4), plus a missing data code (9) and not-in-universe code (0, for children too young to attend or others to whom the question was not addressed). Some census samples contain further information indicating, for example, those who attended, primary, secondary or even tertiary schooling, but did not complete the course of study. The second digit captures this information.

Successful international integration must document such distinctions so that researchers may readily be informed of these and thousands of other details. Appendix C illustrates the detailed and general coding schemes for the educational attainment variable for African countries (represented by its two-digit ISO 3166 code) and censuses (represented by a two-digit year code with century omitted).

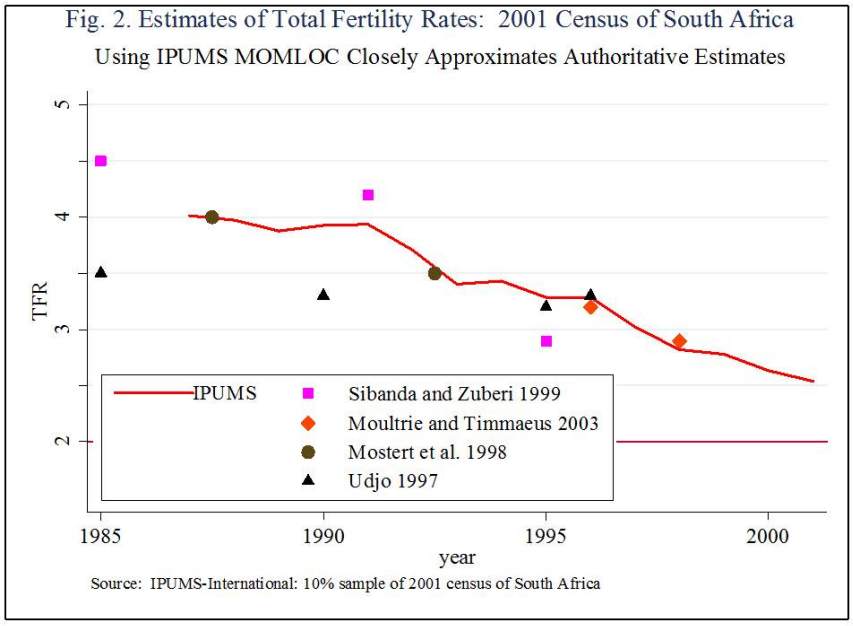
**7. IPUMS-I Value-added variables.** Appendix D and E lists integrated household and person variables for 14 African censuses. There are three types of IPUMS-I value-added variables:

**a. Technical variables**: Record type, Country, Year, IPUMS sample identifier, Household serial number, Number of person records in household, Household weight, Subsample number, Group quarters status, Continent, region of country, Residence at first administrative level, expansion factors (sample weights—for households and persons), etc.

**b. Summary household and family variables**: Household classification, Number of families in household, Number of married couples in household, Number of mothers in household, Number of fathers in household, Head's location in household, Number of unrelated persons, Family unit membership, Number of own family members in household, Number of own children in household, Number of own children under age 5 in household, Age of eldest own child in household, Age of youngest own child in household, etc.

**c. Pointer ("LOC") variables** to identify co-resident spouses, children and their parents: Mother's, Father's and Spouse's location in household, Rule for linking parent(s) and spouse(s), Probable stepmother, Probable stepfather, Man with 2+ wives linked, Second+ order wife, etc. (Sobek and Kennedy, 2009).

MOMLOC (mother's pointer variable) is particularly powerful for estimating fertility using the reverse survival method. To restrict MOMLOC to biological mothers, when STEPMOM is greater than zero then MOMLOC should be reset to zero. Figure 2 demonstrates that using MOMLOC to match mothers to children for computing own-child estimates of the total fertility rate produces a series of Total Fertility Rates that is striking similar to other authoritative estimates—despite the fact that South Africa households are unusually complex (Kennedy, McCaa, Sobek, and Cleveland, 2011).




**8. Integrated boundary files.** IPUMS-International provides boundary files (shapefiles) to facilitate national and international data mapping. The researcher downloads the boundary files and manipulates them with software as desired. Two levels of GIS boundary files are available: a world file with country level boundaries, and country level files with top administrative level boundaries (states, provinces, etc.). Users can create maps with IPUMS-International data using a statistical software program and ArcMap (a GIS mapping software). For more information about linking IPUMS-International data to world and select country maps, see: https://international.ipums.org/international/gis.shtml .

## III. Dissemination

**9. Trans-border access**: Trans-Border access to microdata is essential in today's global world, where researchers are highly mobile. Consider, for example, the field of demography, where one-fifth of the membership of the global professional society, the International Union for the Scientific Study of Population (IUSSP), resides outside their country of birth. For the 506 members of the IUSSP resident in the USA, thirty percent were not born there. Of Chinese born demographers, almost one-third reside outside China. For German and Dutch born IUSSP members the fraction stands at two-fifths.[1] For many professional demographers—and many social science researchers in general--trans-border access is essential if they are to conduct research using census microdata of their country of birth, whether comparative or not. The IPUMS-International portal provide a uniform experience for accessing samples for all countries (for the AICMS all African countries) and census years, regardless of the researcher's country of birth, residence or nationality.

**10. Custom-tailored extracts pooled into a single file**. With IPUMS-I no two extracts are alike. Each extract is custom-tailored by the researcher by means of a series of point-and-

[1] Statistics communicated to the author by the Secretariat of the International Union for the Scientific Study of Population, September 14, 2011.

click selection screens. The extract is delivered to the researcher in a single, pooled file—regardless of the number of countries or censuses. While making selections, metadata are readily available to "surf" the documentation in any order desired. To place an order the researcher selects:

- country (or countries)
- census year(s)
- variables (age, sex, educational attainment, etc.)
- sub-populations (e.g., female heads of households aged less than twenty five years old. Note that there is an option for selecting the individual or all co-resident persons in the selected household)
- and sample density (either as a percent or number of cases).

The IPUMS extract engine fulfils the request by generating a dataset containing only the requested microdata and the corresponding set of DDI compatible metadata as well as a codebook suitable for constructing an SPSS, SAS or STATA file. Copies of original source metadata are available from the web-site, as well as integrated metadata in interactive form. These also may be downloaded freely.

The fact that IPUMS-International distributes microdata electronically as custom extracts, tailored as to country(ies), census year(s), subpopulation(s), and variables, according to the individual needs of the researcher, provides additional incentives for jealously guarding extracts. Since complete datasets are not distributed on CD or other medium, the inclination to share data with unauthorized individuals is greatly reduced, if not completely eliminated.

The IPUMS extract engine adds even more value to each extract by means of four unique tools:

a. **Select cases** (see "sub-populations" above)**.** Filter the samples to select precisely the cases of research interest. For example, if the researcher wishes to research only economically active females, aged 15-19, born in a country different from the current country of residence, the IPUMS extract engine will generate a dataset precisely to these characteristics. Moreover the extract engine will include in the extract all individuals co-resident in the same household with the selected individual.

b. **Attach characteristics.** The extract engine attaches characteristics of mothers, fathers, spouses and household heads to co-resident individuals. This feature facilitates the analysis of children by characteristics of their mothers, fathers, and/or household heads. The feature is useful for analyzing own-child fertility, marital homogamy, and a host of other topics where the joint-characteristics of two or more members of the household are required. For example, for the recent international seminar on "Cross Border Marriages" (Seoul, Republic of Korea), it was easy to generate a dataset from the IPUMS-I website representing 12 million foreign born individuals married or in union with co-resident native born spouses for 51 countries (Esteve, Garcia and McCaa, 2011; see also Parker, 2011).

c. **Customize sample size.** The extract system offers a tool to customize the size of each sample in terms of the absolute number of persons or households or in percentages. If the researcher desires a sample of only 50,000 households, simply enter "50" in the corresponding table and the extract engine will construct a systematic sample. The appropriate weights (expansion factors) are automatically computed on-the-fly and included in the extract.

d. **Pooled into a single file.** Regardless of the number of countries or censuses, all the microdata for each extract are pooled into a single file. Comparing countries or census years is easy because each is treated as a variable instead of separate files. .

**11. Usage.** Microdata are liberated from the confines of statistical offices to permit some of the brightest statisticians, demographers, economists, social scientists and policy makers on the planet to extract un-dreamed of value from the data without cost (Meyer, McCaa, and Lam, 2011). For, national researchers many were never able to use the microdata of their country until IPUMS-International opened the door. Other researchers faced troublesome, sometimes exorbitant fees.

As of November 2011, 166 African researchers representing 24 countries were registered to access microdata from the IPUMS-International/AICMD websites. South African ranked number 1 with 56 researchers, followed by Kenya (32), Egypt (15), Uganda (14), Ghana (8), Nigeria (7), Tanzania (6), Algeria and Cameroun (4 each), Senegal (3), Burkina Faso, Malawi and Sierra Leone (2 each) and Botswana, Congo Republic, Cote d'Ivoire, Mauritius, Morocco, Mozambique, Namibia, Reunion, Zambia and Zimbabwe (1 each).

When the most recent, full-scale analysis of usage was conducted in April 2011, 85,505 samples (24,699 extracts) had been downloaded from the IPUMS-International site, averaging over 1,500 samples per country for the 55 countries represented in the database at that time. South Africa, with 3 samples integrated, displaced Kenya from the top spot with 1,940 extracts trailed by Uganda (1,299 extracts), Rwanda (1,015), Guinea, Ghana, Tanzania, Egypt, Mali, and Senegal.

Table 2 lists the 32 most commonly extracted variables. The top eight encompass four demographic variables (marital status, relationship to head, age and sex), two economic (employment status and class of worker), and one each social (educational attainment) and technical (person weight).

A most striking finding is that 40 variables account for 60% of those requested and among these are sixteen IPUMS constructed variables, four of which, the "LOC" variables, are unique to IPUMS-I. The LOC variables are imputed from household information: spouse's location in household, mother's location, father's location, and the rules for inferring locations. Researchers exploit these variables to study the joint characteristics of spouses and characteristics of parents relative to their children. The variables are constructed by imputation from the relationship to head variable, age, sex, marital status, order of individuals listed in the household, and a few other variables. The heavy usage of the "LOC" variables indicates their great importance for analyzing individuals in relation to characteristics of their spouses, mothers and fathers. The LOC variables constitute an exceedingly important value-added to each household sample (Sobek and Kennedy, 2009).

The IPUMS-I "top 40" institutions in terms of data usage includes many of the world's premier universities and research organizations (see Appendix F), scattered across fourteen countries. In 46 countries, we find a total of 501 institutions with researchers making ten or more extracts. (In addition, in the United States, there are 295 institutions at this level of usage.) A surprising number of extracts are made by researchers from countries with no microdata in the IPUMS-I system. The top 10 of these are: Singapore (494 extracts), Belgium (250), Australia (229), Japan (170), Russian Federation (58), Republic of Korea (45), Czech Republic (42), Sweden (41), Hong Kong SAR (40), and New Zealand (40). On the opposite side of the coin are 13 countries with microdata in the IPUMS-I database but as yet no national researchers use them. The 13 are: Armenia, Belarus, Guinea, Iraq, Jordan, Kyrgyzstan, Mali, Mongolia, Nepal, Peru, Rwanda, Saint Lucia, and Slovenia. Of course, researchers from these countries—instead of accessing microdata electronically from the IPUMS-I website—may acquire copies of the integrated samples on CDs supplied by the Minnesota Population Center to the corresponding

National Statistical Office. We advise NSO partners to register any such users and admonish them to respect the IPUMS-I conditions of use, but there is no obligation to do so because the NSO is responsible for any such dissemination.

**12. 2010 round census samples**. Microdata for 2010 round of censuses are "fast-tracked" processing by IPUMS-International without due delay. Microdata received by September 1 are processed for launch June 1st of the following year. To date, most the IPUMS-International partners are eager to entrust microdata as soon as data processing is completed. Samples from the 2010 censuses of Indonesia and Mexico will be launched in 2012. In 2011, samples of the 2006 census of Egypt, 2008 of Malawi, Sudan and South Sudan, and the 2004 of Sierra Leone. Hopefully, they will be followed soon by the samples from censuses taken in the past five years in other African countries.

## IV. Ethics.

**13. Statistical transparency.** Entrusting census microdata demonstrates to the world that the Official Statistical agency "has nothing to hide" and is willing to subject its most treasured microdata to external scrutiny.

**14. Academic freedom**. Microdata disseminated by IPUMS-International are under no prior restraints on publication or dissemination of results.

**15. Reduce fraud** and the temptation to exaggerate research findings. Since no researcher has exclusive access to the microdata, results may readily be replicated to test the findings of other researchers.

**16. Research results.** Researchers post links to their findings in the IPUMS-International open-access bibliography: http://bibliography.ipums.org/ . The bibliography for IPUMS-International lists over 500 entries.

## Reflections.

As census outputs to meet user needs, the IPUMS project requests a formidable range and amount of metadata and microdata. Nonetheless the return on the investment is substantial Statistical offices are relieved of many of the most burdensome tasks and responsibilities in disseminating microdata to researchers. Moreover, by relying on standard procedures used by a majority of the world's statistical offices, there is safety in numbers. The isolated statistical office that disseminates microdata on an ad hoc basis incurs substantial risks and responsibilities as well as significant human resource and material costs, for a relatively small return with respect to number of users. The IPUMS project offers substantial economies of scale with the highest standards of security and disseminates integrated metadata and microdata that greatly facilitates sound scientific research.

When we began a decade ago, we dreamed of integrating samples for 21 countries in ten years. We tripled that goal (including South Sudan as no. 63)—thanks to the generous cooperation of National Statistical Offices and undreamed of technological innovations. The number of users and the amount of use also far exceeded our expectations. For the second decade, we dream of doubling the number of users and doubling again the number of samples. High precision samples for the 2010 round of censuses will be crucial to continued success. Participating statistical agencies are invited to entrust metadata and microdata for the 2010 census round without delay. Those agencies not yet participating in the IPUMS initiatives are

invited to consider doing so. Researchers who have yet to access the IPUMS microdata are invited to peruse the metadata and use the microdata should their research needs require.

**References.**


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 .

Trewin, D. 2007. "A Review of IPUMS-International." Unpub. http://www.hist.umn.edu/~rmccaa/IPUMSI/trewin_ipums_report.pdf

Woolfrey, L. 2009. "African Microdata Access Survey 2009," Unpub.

**Table 1. IPUMS/AICMD samples**

| | Country, census | (%) | Households | Persons | Weighted | De jure /de facto | Census date | Smallest geography | Collective dwellings | Notes |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Egypt 1996 | 10 | 1,270,787 | 5,902,243 | yes | de facto | 18/19-11-96 | district | no | |
| 2 | Egypt 2006 | 10 | 1,740,414 | 7,282,434 | no | de facto | 20/21-11-06 | district | yes | |
| 3 | Ghana 2000 | 10 | 397,097 | 1,894,133 | no | de facto | 26-03-00 | district | yes | |
| 4 | Guinea 1983 | 10 | 110,777 | 457,837 | no | both | 1/2/1983 | prefecture | yes | |
| 5 | Guinea 1996 | 10 | 108,793 | 729,071 | no | both | 1/12/1996 | prefecture | yes | |
| 6 | Kenya 1989 | 5 | 224,861 | 1,074,098 | no | de facto | 25-10-89 | district | yes* | |
| 7 | Kenya 1999 | 5 | 317,106 | 1,407,547 | no | de facto | 25-08-99 | district | yes* | |
| 8 | Malawi 1987 | 10 | 186,270 | 798,669 | no | de facto | 1/21-09-87 | district | yes | |
| 9 | Malawi 1998 | 10 | 229,005 | 991,393 | no | de facto | 1/21-06-98 | district | yes | |
| 10 | Malawi 2008 | 10 | 299,864 | 1,341,977 | no | both | 1/21-06-08 | district | yes | |
| 11 | Mali 1987 | 10 | 136,515 | 785,384 | no | both | 14-04-87 | district | yes | |
| 12 | Mali 1998 | 10 | 161,880 | 991,330 | no | both | 14-04-98 | district | yes | |
| 13 | Rwanda 1991 | 10 | 153,041 | 742,918 | no | both | 1/8/1991 | province | yes | |
| 14 | Rwanda 2002 | 10 | 191,719 | 843,392 | no | both | 15/16-08-02 | province | yes | |
| 15 | Senegal 1988 | 10 | 79,904 | 700,199 | no | both | 20-05-88 | department | no | |
| 16 | Senegal 2002 | 10 | 107,999 | 994,562 | no | both | n.a. | department | no | |
| 17 | Sierra Leone 2004 | 10 | 82,518 | 494,298 | no | de facto | 4/12/2004 | chiefdom | yes | |
| 18 | South Africa 1996 | 10 | 993,801 | 3,621,164 | yes | de facto | 10/10/1996 | district | yes | 1.3% of sample not in households |
| 19 | South Africa 2001 | 10 | 991,543 | 3,725,655 | yes | de facto | 10/10/2001 | municipality | yes | |
| 20 | South Africa 2007 | 2 | 345,170 | 1,047,657 | yes | both | 7/2/2007 | municipality | yes | |
| 21 | Sudan 2008 | 15 | 1,015,408 | 5,609,295 | yes | de facto | 22-04-08 | county | no | Includes Sudan and South Sudan |
| 22 | Tanzania 1988 | 10 | 472,616 | 2,310,424 | yes | de facto | 27-08-88 | district | no | |
| 23 | Tanzania 2002 | 10 | 841,768 | 3,732,735 | yes | de facto | 22-08-02 | district | yes | |
| 24 | Uganda 1991 | 10 | 339,166 | 1,548,460 | yes | de facto | 11/1/1991 | county | yes | |
| 25 | Uganda 2002 | 10 | 529,271 | 2,497,449 | no | both | 12/9/2002 | county | yes | |
| | **TOTAL** | | **11,327,293** | **51,524,324** | | | | | | |

Source: http://ecastats.uneca.org/aicmd/en-us/samples.aspx

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

| Rank | Label | Extracts | Mnemonic | Comment |
|---|---|---|---|---|
| 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. IPUMS Memoranudm of Understanding with CBS Kenya, 2002.**

**MEMORANDUM OF UNDERSTANDING (MOU) BETWEEN INTEGRATED PUBLIC USE MICRODATA SERIES INTERNATIONAL (IPUMS), MINNESOTA UNIVERSITY AND CENTRAL BUREAU OF STATISTICS (CBS)**

Purpose: The purpose of this letter of understanding is to specify the terms and conditions under which integrated metadata and microdata provided by **Central Bureau of Statistics of Kenya** shall be disseminated by **Integrated Public Use Microdata Series International** of the University of Minnesota.

1. Ownership: The **Central Bureau of Statistics** of Kenya is the owner and licensee of the intellectual property rights (including copyright) in metadata and microdata supplied to the University of Minnesota to be distributed by **Integrated Public Use Microdata Series International**.

2. Use: These data are provided for the exclusive purposes of teaching, academic research and publishing, and may not be used for any other purposes without the explicit written approval, in advance, of the **Central Bureau of Statistics**.

3. Authorization: To access or obtain copies of integrated microdata of Kenya from **Integrated Public Use Microdata Series International**, a prospective user must first submit an electronic authorization form identifying the user, that is, principal investigator by name, electronic address, and institution. The principal investigator must state the purpose of the proposed project and agree to abide by the regulations contained herein. Once a project is approved, a password will be Issued and data may be acquired from servers or other electronic dissemination media maintained by **Integrated Public Use Microdata Series International**, the **Central Bureau of Statistics**, or other authorized distributors. Once approved, the user is licensed to acquire integrated metadata and microdata of Kenya from **Integrated Public Use Microdata Series International** or other authorized distributors. No titles or other rights are conveyed to the user.

4. Restriction: Users are prohibited from using Kenyan data 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.

5. Confidentiality: Users will maintain the absolute confidentiality of persons and households as provided for under the Statistics Act Cap.112 of the Laws of Kenya. Any attempt to ascertain the identity of persons or households from the microdata is strictly prohibited. Alleging that a person or household has been identified in these data is also prohibited.

6. Security: Users will implement security measures to prevent unauthorized access to Kenyan microdata acquired from **Integrated Public Use Microdata Series International** or its partners.

7. Publication: The publishing of data and analysis resulting from research using metadata or microdata of Kenya is permitted in communications such as scholarly papers, journals and the like. The authors of these communications are required to cite the **Central Bureau of Statistics** and **Integrated Public Use Microdata Series International** as the sources of Kenyan data, and to indicate that the results and views expressed are those of the author/user.

8. Sharing: **Integrated Public Use Microdata Series International** will provide electronic copies to the **Central Bureau of Statistics** of documentation and data related to its integrated microdata as well as timely reports of authorized users.

9. Violations: Violation of this agreement may lead to professional censure and/or civil prosecution.

*Signed:* _____

Edward F. Wink, Assoc. V.P.
University of Minnesota
Sponsored Projects Admin.
200 Oak Street SE, Suite 450
Minneapolis, MN 55455-2070

*Signed:* _____

**Director of Statistics**
**P.O. Box 30266**
**NAIROBI, Kenya**

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



## Application to Use Restricted Microdata

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.

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.

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.

**Legal notice**: 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.

Information provided on this form will be kept confidential.
All information on this form is required for registration unless otherwise indicated by an asterisk.

**PERSONAL INFORMATION**

.....

**INSTITUTIONAL AFFILIATION**

IPUMS-International staff must confirm the identity of prospective users. To speed the processing of your application, please provide as much of the following

**USAGE LICENSE**

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. **Note: The license is valid for one year and may be renewed.**

**Data must not be redistributed without authorization.**

☐ 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.

**The microdata are intended only for scholarly research and educational purposes.**

☐ 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.

**Commercial use and redistribution of the microdata is strictly prohibited.**

☐ 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.

**Use of the microdata must follow strict rules of confidentiality.**

☐ 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.

**The microdata must always be safely secured.**

☐ 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.

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 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 | No 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 | 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 | | Country | EG | EG | GH | GN | GN | KE | KE | MW | MW | MW | ML | ML | RW | RW | SN | SN | SL | ZA | ZA | ZA | SD | TZ | TZ | UG | UG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 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 | 37 | 41 | 32 | 27 | 28 | 29 | 30 | 28 | 28 | 38 | 30 | 31 | 29 | 38 | 31 | 42 | 40 | 34 | 37 | 38 | 39 | 26 | 31 | 33 | 40 |
| 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 househc | 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 | NMOTHRS | 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 | . |
| 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 |
| 27 | FLOOR | Floor material | 16 | . | . | 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 |
| 29 | ROOF | Roof material | 16 | . | . | 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 |
| 31 | ANYMORT | Any deaths in household last year | 13 | . | . | . | . | 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 |
| 33 | TV | Television set | 11 | 1 | 1 | . | . | . | . | . | . | . | 1 | . | . | . | 1 | 1 | 1 | 1 | . | 1 | 1 | 1 | . | . | . | 1 |
| 34 | MORTNUM | Number of deaths in household last ye: | 10 | . | . | . | . | . | . | . | . | 1 | 1 | 1 | 1 | . | 1 | . | 1 | 1 | . | . | 1 | 1 | 1 | . | . | . |
| 35 | REFRIG | Refrigerator | 9 | 1 | 1 | . | . | . | . | . | . | . | 1 | . | . | . | . | 1 | 1 | 1 | . | 1 | 1 | 1 | . | . | . | . |
| 36 | CELL | Cellular phone availability | 8 | . | 1 | . | . | . | . | . | . | . | . | . | . | . | 1 | . | . | 1 | 1 | 1 | 1 | 1 | . | . | . | 1 |
| 37 | AUTOS | Automobiles available | 8 | 1 | 1 | . | . | . | . | . | 1 | . | 1 | . | . | . | 1 | . | 1 | 1 | . | . | . | 1 | . | . | . | . |
| 38 | COMPUTR | Computer | 8 | 1 | 1 | . | . | . | . | . | . | . | . | . | . | . | 1 | . | 1 | 1 | . | 1 | 1 | 1 | . | . | . | . |
| 39 | SEWAGE | Sewage | 7 | 1 | 1 | . | . | . | 1 | 1 | . | . | . | . | . | . | . | 1 | 1 | . | . | 1 | . | . | . | . | . | . |
| 40 | TRASH | Trash disposal | 7 | . | . | . | . | 1 | . | . | . | . | . | . | . | . | 1 | . | 1 | 1 | . | 1 | 1 | . | . | . | . | 1 |
| 41 | BATH | Bathing facilities | 6 | 1 | 1 | 1 | . | . | . | . | . | . | . | . | . | . | . | . | . | 1 | . | . | . | . | . | . | 1 | 1 |
| 42 | composite | Place of Residence: Level 3 | 5 | . | . | . | . | . | | | . | . | . | 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 | MW | MW | MW | ML | ML | RW | RW | SN | SN | SL | ZA | ZA | ZA | SD | TZ | TZ | UG | UG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 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 |
| 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 1 |
| 13 | FAMSIZE | Number of own family members in hou | 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 | 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 | 1 |
| 15 | NCHLT5 | Number of own children under age 5 ir | 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 | 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 | 1 |
| 17 | YNGCH | Age of youngest own child in househol | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 | 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 |
| 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 | 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 |
| 37 | composite | Place of previous residence 1/5/10 yea | 19 | 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 |
| 39 | INDGEN | Industry, general recode | 18 | 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 |
| 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 |
| 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 |
| 43 | MORTFAT | Mortality status of father | 17 | . | . | . | . | . | 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 |

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

| | Variable | | Country | EG | EG | GH | GN | GN | KE | KE | MW | MW | MW | ML | ML | RW | RW | SN | SN | SL | ZA | ZA | ZA | SD | TZ | TZ | UG | UG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 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 |
| 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 |
| 52 | DISMUTE | Mute | 14 | 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 | . | . |
| 54 | MGYRS1 | Years residing in current locality | 12 | 1 | 1 | . | . | 1 | . | 1 | . | . | 1 | 1 | 1 | 1 | 1 | . | . | . | . | . | . | 1 | . | . | 1 | 1 |
| 55 | RESIDENT | Residence status: de facto, de jure [pres | 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 | . |
| 57 | CHDEAD | Number of children dead | 10 | . | . | . | . | . | 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 |
| 59 | DEADMAL | Number of male children dead | 10 | . | . | . | . | . | 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 |
| 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 |
| 64 | DISMNTL | Mental disability | 9 | 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 |
| 67 | HOMEFEM | Number of own female children in hou | 8 | . | . | . | . | . | 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 |
| 69 | AWAYCHD | Number of own children living elsewha | 8 | . | . | . | . | . | 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 |
| 71 | AWAYMAL | Number of own male children living el | 8 | . | . | . | . | . | 1 | 1 | . | . | . | . | . | . | . | . | . | 1 | . | . | 1 | 1 | 1 | 1 | . | 1 |
| 72 | LSTBSEX | Sex of last birth | 7 | . | . | . | . | . | 1 | 1 | . | . | . | . | . | . | . | . | . | 1 | . | . | 1 | . | 1 | 1 | . | 1 |
| 73 | BTHSURV | Children surviving from births last yea | 7 | . | . | . | . | . | . | . | 1 | 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 |
| 77 | BIRTHMO | Month of birth | 6 | 1 | . | . | . | 1 | . | . | . | . | 1 | . | . | 1 | 1 | . | . | . | . | . | . | . | . | . | . | 1 |
| 78 | LSTBMTH | Month of last birth | 6 | . | . | . | . | . | 1 | 1 | . | . | . | . | . | . | . | . | . | . | . | 1 | 1 | . | . | . | 1 | 1 |
| 79 | LSTBYR | Year of last birth | 6 | . | . | . | . | . | 1 | 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 | 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)