**Creating Statistically Literate Global Citizens: The Use of IPUMS-International Integrated Census Microdata in Teaching**

Forthcoming: Statistical Journal of the IAOS http://www.iospress.nl/loadtop/load.php?isbn=18747655

Journal of the International Association for Official Statistics

Authors:

Ann Meier, Associate Professor of Sociology, Department of Sociology and the Minnesota Population Center, University of Minnesota. meierann@umn.edu .

Corresponding author:

Robert McCaa, Professor of Population History, Department of History and Minnesota Population Center, and Principal Investigator of IPUMS-International projects, University of Minnesota.

email: rmccaa@umn.edu .

Address: Minnesota Population Center

50 Willey Hall

225 19th Ave S.

Minneapolis, MN 55455 USA

Tel. 1+952.334.7867 (cell)

FAX: 1+612.826.8375

David Lam, Professor of Economics and Director of the Population Studies Center, University of Michigan. davidl@umich.edu .

**Abstract.**

Census microdata are ideal for developing statistical literacy of university students. Access, particularly to internationally comparable microdata, has been a significant obstacle. The IPUMS-International project offers a uniform solution to providing access for policy analysts, researchers, and students to integrated microdata and metadata, while protecting statistical confidentiality. Eighty-five official statistical agencies have endorsed IPUMS-I dissemination principles and entrusted microdata for 249 censuses to the project. From June 2010, 159 integrated samples, representing 55 countries and totaling over 325 million person records, are available at no cost to researchers and their students. The database is being expanded with the addition of samples for 5-10 countries per year as well as samples for the 2010 round of censuses. This paper illustrates two approaches to using IPUMS-I census microdata in the university curriculum to promote statistical literacy among undergraduates.

**Keywords.**

Census microdata, statistical literacy, microdata access, integrated samples, university students

## 1. Census microdata and the IPUMS-International Project

Census microdata are the individual responses to census questionnaires recorded in computerized form as numeric or alphabetic codes. The data include demographic characteristics such as age, sex, marital status, relationship to head of household, migration, education, and occupation among other individual-level variables. They also include information on household characteristics such as urban/rural residence, home ownership, amenities like electricity and water supply, and number of rooms in the dwelling. In recent decades a growing number of statistical agencies have prepared census microdata files for analysis by staff and, in some cases, by academic researchers and policy makers. With ever-expanding access to computers, analysis of large census microdata files is now feasible for ordinary researchers, and even, as this article demonstrates, their students. The idea of gaining access to census microdata from around the world and across time is exciting for researchers, but it is also daunting.

The first challenge is accessing microdata from the country of interest. In the past decade, a policy revolution has taken place among official statistical agencies. Most statistical authorities now recognize census microdata as statistical products to be disseminated along with conventional publications. A good example is the dissemination policy of the Central Statistical Agency of the Federal Democratic Republic of Ethiopia, which in 2005 began to distribute a wide variety of microdata products on CD and from its website.

A second challenge in using census microdata is comparability. In order to study trends across time or to compare countries on a particular dimension, files from different years and different countries must first be made compatible. Although comparability is improving, historically there has been very little coordination between National Statistical Agencies to ensure comparable measures

are included. Moreover, even the same statistical agency sometimes asks questions differently in different census years.

The IPUMS-International project is a global initiative to anonymize, integrate, and disseminate samples of census microdata to researchers, policy makers, teachers, and their students. Thanks to major funding by the National Science Foundation and the National Institutes of Health (USA), IPUMS-I has become the largest repository of census microdata in the world with the official statistical authorities of more than 85 countries, encompassing over half of the world's population, entrusting a total of 249 censuses to the Minnesota Population Center. As of June 2010, integration of samples is complete for 55 countries, totaling 159 samples and more than 325 million person records (Table 1) [1]. Some 3,000 users, representing 76 countries, are registered and approved for access to the IPUMS-I database at www.ipums.org/international. The microdata are made available to researchers at no cost. The only conditions are that researchers must protect statistical confidentiality, use the data for non-commercial purposes, protect the data against unauthorized use, and provide copies of publications to the corresponding national statistical offices.

Table 1 near here.

IPUMS-International initiatives are undertaken only in countries where authorization is provided by means of a memorandum of understanding with the official statistical agency. The memorandum is entirely general in nature, yet it provides a legal framework for the project to proceed (see Appendix). Official statistical agencies that have not yet affiliated with the project are invited to do so by contacting the second author of this paper.

Official statistical agencies cooperate in the project for diverse reasons. First, IPUMS-I resolves the conundrum of managing the broadest possible access to sensitive microdata while protecting statistical confidentiality. Many agencies have long wanted to make census microdata

available to researchers, but lack the substantial material and 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 [2]. Likewise, IPUMS-I is the only academic organization audited by external review. Mr. Dennis Trewin, the chair of the UNECE Task Force, conducted a week-long, on-site inspection at the Minnesota Population Center and concluded his report as follows [3]:

> 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].

Second, few material or human resources are required to participate in IPUMS-I. Indeed, a standard fee is paid to compensate for the marginal costs of assembling the census microdata and documentation entrusted to the project. Upon receipt of microdata, documentation and official invoice, the National Science Foundation of the United States authorizes the Minnesota Population Center to pay US$1,000 per census for non-exclusive rights to dissemination. For datasets with more than one million person records, the fee is increased to US$5,000.

Third, IPUMS-I is responsible for constructing and managing the integrated database: validating the microdata, constructing the sample, anonymizing the data, translating, where necessary, the metadata to English, integrating both microdata and metadata into a single system for all countries, approving access to the database, certifying researcher compliance with maintaining data privacy, and responding to user queries.

Finally, as more and more national statistical agencies join the IPUMS International project, the fruits of evidence-based policy decisions and training students to be statistically literate are becoming

widely known [4]. This may motivate those agencies not yet participating to cooperate with the initiative.

IPUMS-I has two goals: first, to preserve census microdata and, second, to make anonymized, integrated sample extracts available to researchers, policy analysts and students free of charge. This article focuses on the latter goal, and offers two illustrative examples of how the IPUMS-I effort facilitates statistical literacy and general global awareness among the next generation.

To make census microdata useful for research, they must be thoroughly documented and integrated. While the idea of integrating census data is not new, progress toward true integration has been slow. One example of positive movement towards integration is the massive achievement of the United Nations Statistical Division in the international harmonization of census concepts from the enumeration form to the publication of final tables. While still incomplete, the effort enjoys widespread support and cooperation from statistical agencies around the globe. Beginning in 1991, the IPUMS-USA project has worked to harmonize census data for the United States for the period since 1850 [5], and IPUMS-International has capitalized on this experience [6].

The IPUMS-International project adopts uniform coding schemes, nomenclatures, and classifications, based where possible on the United Nations Statistical Division's Principles and Recommendations for Population and Housing Censuses and other international standards such as UNESCO's International Standard Classification of Education, and similar classifications for occupations (International Labor Office), other economic activities (United Nations Statistical Division), and population and housing indicators (United Nations Economic Commission for Europe).

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 create comparable codes across

samples. This means that the harmonized measures must be the lowest common denominator, or the simplest measure, from all the samples that include a measure of a particular characteristic. For example, regarding information on educational attainment, most samples 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 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. Thus, the second digit captures this information. To avoid the loss of important information for those samples that have even more detail, IPUMS-I uses a composite coding strategy that retains all original detail, and at the same time provides comparable codes across samples. Thus, the first one or two digits, which we call the "general code," provide 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. Table 2 illustrates the detailed and general coding schemes for the educational attainment variable for five countries (represented by its two-digit ISO 3166 code) and 11 censuses (represented by a two-digit year code with century omitted).

Table 2 near here:

**IPUMS-I General and Detailed Coding Schemes, Example: Educational Attainment**

Integration means more than constructing composite coding schemes. Original source documentation is readily available in the official language and English translation so that researchers may easily study nuances of definitions and concepts from one census to another. For integrated

variables, the IPUMS website also offers abundant metadata to describe how codes are constructed and to compare definitions and concepts from country-to-country and census-to-census. The integrated variable metadata consists of three types of discussions: description, general comparability, and country-specific comparability. A dynamic navigation tool permits researchers to focus on specific countries and censuses to surf the documentation integrated into the web-site. For educational attainment (EDATTAN) the general text begins as follows (https://international.ipums.org/international-action/variables/ 173897):

> EDATTAN records the person's educational attainment in terms of the level of schooling completed (degree or other milestone). The emphasis on level completed is critical: a person attending the final year of secondary education receives the code for having completed lower secondary only -- and in some samples only primary.
>
> EDATTAN does not necessarily reflect any particular country's definition of the various levels of schooling in terms of terminology or the number of years of schooling. EDATTAN is an attempt to merge -- into a single, roughly comparable variable -- samples that provide degrees, ones that provide actual years of schooling, and those that have some of both. In addition to EDATTAN, a country-specific education classification is provided which loses no information and reflects the particular educational system of that country (for example EDUCBR [Brazil], EDUCCL [Chile], and EDUCUS [USA]).

The harmonization efforts allow researchers to use detailed measures on a single country or somewhat less detailed but comparable measures across countries, within a country across census years, or across countries and across census years. Thus, IPUMS-I integrated microdata have become a very flexible and user-friendly data source. Academic scholars and policy makers regularly use the IPUMS-I data to investigate issues such educational attainment and labor market success of return migrants in Africa [7], the economic effects of malaria eradication in Uganda [8], etc.

In addition to informing discussions about development strategies and policy decisions, IPUMS-I is now used as a teaching tool as well. Using IPUMS-I microdata in teaching serves several important pedagogical goals. First, it promotes statistical literacy among students who may learn statistical theory in the classroom, but rarely have the opportunity to apply their knowledge to real-world data. Second, it allows students to gain insight and knowledge about countries world-wide and across time. Because the IPUMS-I project contains microdata on so many countries and in so many census years, the possibilities for interesting and useful projects are endless. Below we offer two examples of how we have used IPUMS-I data in our own undergraduate classrooms to further the statistical literacy of students, and to make them more knowledgeable global citizens. By these examples, we hope that readers will see how easy and beneficial use of IPUMS-I data is to the educational enterprise.

## 2. Example 1: World Population Problems (University of Minnesota)

The first author used the IPUMS-I data in her World Population Problems course in the spring of 2009. The course is an upper level undergraduate course, enrolling mostly juniors and seniors. Almost all of the 55 students enrolled were American, but a few were first-generation immigrants and at least ten students had parents who had immigrated to the U.S. Several foreign students also enrolled in the course. Because of the topic and the diverse student body, the students and instructor had a keen interest in expanding their world views. The course objectives listed on the syllabus were as follows:

1. provide key demographic facts about the world and its major regions;
2. explain how social, economic, and cultural factors interrelate to produce demographic outcomes across the globe and in comparative perspective;

3. understand the basic social science theories that depict the relationship between population, society, and the environment; and
4. discuss the implications of population issues for the international community, the United States, and our own lives.

***2.a. Population Profile Assignments***

In addition to the use of a textbook, the instructor delivered a series of lectures on topics ranging from global population trends to fertility control to population and economic development. The highlight of the students' learning, however, was a series of assignments in which students accessed and analyzed census data from the IPUMS-I project along with other sources of data.

On the first day of class, each student was assigned a new identity: they were assigned a gender, age, and country of residence, along with a year in which they were living. For example, a female student was assigned to a new identity of a 22-year-old male from Rwanda in 2002. The instructor created mock "passports" from their new country of residence which detailed each student's new identity. Each assigned country and year corresponded with an available IPUMS-I sample (e.g. Rwanda census 2002), and three or four students were assigned the same country but different ages, genders, and/or census years. Students assumed their new identities when completing a series of assignments throughout the semester, culminating in the creation of a poster for a group poster session in the Minnesota Population Center at the end of the semester. Using the IPUMS-I data and other sources in conjunction with new student identities was meant to help students understand how population issues are experienced differently based on geography (where you live), history (when you live), and social location (your age and gender). Below we detail each assignment, discuss access to the data for students, and illustrate some of the insights gained through this learning tool.

In the first several weeks of the semester, students learned about the demographic forces of health and mortality. Students accessed tables from the United Nations World Population Prospects to document the crude death rate, infant mortality rate, and life expectancy at birth for their assigned country and year. To assess how different living conditions influence health and mortality, students used the IPUMS-I data to investigate infrastructure features such as the source of water supply and access to a flush toilet for someone of their assigned identity. Lectures discussed how such features are related to disease acquisition and spread. Students used SPSS software to generate simple descriptive statistics using the IPUMS-I data. With their statistical output, students created a health and mortality profile for someone of their assigned identity. To do this they combined statistics, graphics that displayed statistics across time (trends) or regions (comparatively), and a text evaluation of the health and mortality situation.

In the middle of the semester, the course turned to discussion of fertility and family situations around the globe. Again, students accessed data from IPUMS-I project for their country and year to understand marital status, number of children, and families per household to understand these family features for someone of their assigned identity. They combined this information with statistics on the crude birth rate, total fertility rate and population sex ratio to graph and discuss trends in and comparisons of the fertility and family profiles across the globe.

In their final assignment, students who had been assigned identities of the same country and year combined what they had learned in a group poster depicting the population profile of their assigned country. Figure 1 displays the population profile poster for Rwanda constructed by a group of students.

Figure 1 near here: **Student Poster using integrated census microdata and other sources: Rwanda**.

The posters were displayed in a public session at the Minnesota Population Center. Faculty, graduate students, and staff of the Minnesota Population Center were invited to view the posters and ask questions. Students were graded on the content and presentation of their posters as well as how well they answered questions about the population of their country. In Figure 2 students showcase their posters and field questions from those who attended the session.

Figure 2 near here: **Students showcase posters and field questions**

***2.b. Learning Outcomes***

The series of population profile assignments described above resulted in four key learning outcomes. First, students became experts on the population situation in a country other than their own. Being assigned a new identity with a mock "passport" personalized the assignments, and therefore made students invested in doing a good job. In addition, having to present what they have learned in a formal poster format and then answer questions from a live audience of faculty, staff and graduate students provided the opportunity for students to gain competency in explaining what they learned. Second, students learned how history and geography shape population issues. By comparing health, mortality, fertility and family statistics across time and/or between countries or regions, students gained a good understanding of how historical events and geographic location can influence population.

Third, students learned how to analyze secondary data by using the IPUMS-International data and SPSS software. This is a practical skill that they can carry forward into their other coursework and future professional lives. Many students expressed surprise and pride in their ability to easily generate graphs and tables based on their analysis of data. Conducting their own analysis brought dry statistics to life. We would not be surprised if several of these students pursued statistical analysis as a career path. At the very least, students gained a baseline level of statistical literacy that they did

not possess before. Finally, the availability of the IPUMS-I data helped make the world seem like a smaller place for students; it helped them feel more connected to those from other parts of the world. It is our hope that this makes them more knowledgeable global citizens.

## 3. Example 2: Economics of Population (University of Michigan)

The third author and his colleague Martha Bailey have incorporated IPUMS-I data into their course on Economics of Population at the University of Michigan. This is a course designed primarily for upper-level undergraduate majors in economics. The course meets in a lecture format on Mondays and in a computer lab on Wednesdays, with the size of the class determined by the capacity of the computer lab (typically about 35 students). The class is designed to satisfy the requirement that students take an advanced writing course in their third or fourth year, ideally within their major field of concentration. The class has been taught five times using IPUMS-I data. The data are directly linked with numerous writing assignments, including a major term paper. The course objectives are as follows:

1. use economic theory to guide empirical analysis of issues such as fertility, marriage, investments in children, and household bargaining;
2. analyze interactions between demographic change and economic change, including the effects of age structure on government programs such as education and old-age support;
3. use STATA to conduct statistical analysis of IPUMS-I census data to test economic theories of demographic behavior

### *3.a. Computer Labs with IPUMS-I*

The course assumes that students have no prior experience with the statistical package STATA, although a few students typically have used the package in prior economics or statistics courses. Students work with STATA in the early labs to learn both the basics of using STATA as well

as the basics of how census data are collected and distributed for use in research. Weekly problem sets are used to reinforce material covered in the lab. For example, labs during the first few weeks of class involve making simple tables of the distribution of children ever born for women of different ages, generating new variables such as the natural logarithm of household income, and writing simple programs in STATA to carry out these functions.

IPUMS-I is an excellent resource for teaching students a statistical package like STATA. Because most variables have been integrated across years and countries using the same variable names and variable structure, it is easy, for example, to use the 2000 United States census during the instructional lab sessions, and then have the students do a problem set that does similar analyses using the 1999 Kenya census. Once the students have been taught to append the data sets it is easy for them to do a problem set that compares two countries or two points in time for the same country.

The course assumes that students have previously taken intermediate microeconomics and introductory courses in statistics. Given these prerequisites, the course moves relatively quickly into multivariate regression analysis. One lab, for example, has students estimate regressions using children ever born as the dependent variable. The students use various combinations of the education of the mother, education of the father, and household income as the independent variables, discussing how the results change as different variables are included. This gives students experience at interpreting regression coefficients and thinking about the impact of including additional variables in a multivariate regression. Although the students usually have relatively good knowledge of statistics and econometrics, the experience from teaching the course is that it takes students several weeks and many different examples before they are comfortable with interpreting regression coefficients in real-world examples.

One advantage of the IPUMS-I data is that the large sample sizes make it possible to include many different variables in a regression, along with interactions and higher-order terms, while still producing meaningful estimates. The large sample sizes also make it possible to produce separate estimates for population subgroups such as narrowly defined age groups, regions, and racial or ethnic groups. A very attractive feature of the IPUMS-I data extraction system is that it is easy to extract samples with a target file size. For computer lab purposes we typically use extract files that are less than 50MB in size in order to keep computer processing fast during the labs. Students are encouraged to work with larger extracts for their projects, however, and it is easy for them to generate customized files from the IPUMS-I web site that meet their particular needs. For example, one student wrote an excellent paper analyzing schooling outcomes for the children of Mexican immigrants in the state of California. He was able to build an extract using only California residents in several recent U.S. censuses, giving him the statistical power necessary for his analyses while keeping the file size manageable.

***3.b. Learning Outcomes***

The course combines lectures, computer labs, problem sets and writing assignments in a systematic way, culminating in a 20-page term paper built around statistical analysis of IPUMS-I data. Students are encouraged to analyze changes in one country over time, taking advantage of the many IPUMS-I countries that have two or more censuses, or to do a comparative analysis of two or more countries. They may analyze any variables that they find interesting in the census data. Students write a short proposal in about the fourth week of the course in which they take data from one country and present two or three tables and figures analyzing some set of variables that they think might be interesting for a term paper. They get feedback on these proposals in writing and in person during office hours. The first draft is due about halfway through the semester. They get extensive feedback

on this draft, which is then revised before final submission at the end of the semester. Considerable time is spent teaching students how to make tables and figures that effectively present their statistical results. They are told to write their papers as data-intensive policy briefs, focusing on how to communicate statistical analysis in a clear and effective way.

Students have shown a great deal of creativity in using IPUMS-I data. Examples of term papers include the following: Poverty and education in Kenya; Living arrangements of the elderly in South Africa; Women's education and fertility decline in Brazil; Changes in child mortality over time and across socioeconomic groups in Ecuador; Rwanda: A study of the impact of conflict on children's literacy; Male-female earnings gaps across time and across countries; The impact of family size on educational attainment in Costa Rica.

Student feedback from the course is very positive, with well-above average scores on the systematic course evaluation forms. In open-ended comments on these forms, many students comment that the course allows them to apply the statistics they have learned in previous courses to real-world economic problems. In the process they get a good understanding of statistics along with an appreciation of the power of statistics to inform public policy. By the end of the course the best students are doing applied econometric analysis at a level comparable to economics doctoral students. Other students do not quite reach that level, but they learn how data can be used for statistical analysis, including an appreciation for the challenges involved in drawing causal inferences.

The skills acquired in the course often have immediate payoff when the students apply for jobs or graduate school. A number of the top students from the class have taken prestigious research assistant positions in Washington agencies such as the World Bank, the Federal Reserve Board, and the Urban Institute, where their statistical computing skills are immediately useful. These jobs are standard stepping stones to the top economics doctoral programs in the country. Other students

have secured jobs in government and private industry where their ability to carry out statistical analysis of large data sets is highly valued. All students, whether or not they use the skills directly in their next position, develop a better understanding of how statistics can be applied using large-scale data sets to answer important questions. Many of them learn about new parts of the world and develop an appreciation for the enormous differences in living standards within and across countries.

**4. Conclusions**

Each year the IPUMS-I database expands as high precision census samples for five to ten additional countries and data for the 2010 round of censuses are entrusted to the project. Teachers interested in using integrated census microdata are invited to register with IPUMS-International to experiment with the possibilities that the database offers for exciting student interest in statistics, population studies, econometrics and quantitative analysis in general.

In this article, we have described the IPUMS-I database and its benefits for researchers and National Statistical Agencies and illustrated the use of IPUMS-I data in two undergraduate University courses. For researchers, the IPUMS-I database offers the benefit of free access to international census data that is comparable across countries and years. For National Statistical Agencies, the IPUMS-I project offers several benefits as well. It provides a mechanism for easily disseminating their data to researchers while protecting the statistical confidentiality of the data. The project also takes the responsibility of managing access to the data, queries about the data, and data privacy compliance. Finally, there are low barriers to participation in the project. In fact, the project pays National Statistical Agencies a standard fee to compensate for the marginal costs of assembling census microdata and documentation.

There are numerous ways that these data could be used in the classroom. We assert that the use of these data in the classroom helps students gain competency in using microdata and

performing statistical analysis. We also suggest that the use of international census data helps students understand some of the opportunities and challenges of their fellow global citizens who live in different parts of the world. Ultimately, we hope that this makes them more knowledgeable global citizens than they were before working with these data.

## 5. References


[1] Minnesota Population Center. *Integrated Public Use Microdata Series – International: Version 6.0.* Minneapolis: University of Minnesota, 2010: https://www.ipums.org/international.

[2] Conference of European Statisticians. *Managing Statistical Confidentiality and Microdata Access: Principles and Guidelines of Good Practice*. New York and Geneva, 2007.

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

[4] R. McCaa, A. Esteve, S. Ruggles, and M. Sobek, Using Integrated Census Microdata for Evidence-based Policy Making: the IPUMS-International Global Initiative, in: *The African Statistical Journal* 2 (2006), 83-100.

[5] S. Ruggles, M. Sobek et. Al., *Integrated Public Use Microdata Series: Version 2.0*. Minneapolis: Historical Census Projects, University of Minnesota, 1997.

[6] A. Esteve and M. Sobek. “Challenges and Methods of International Census Harmonization. *Historical Methods*. 36(2003), 66-79.

[7] K.J. Thomas, Return Migration in Africa and the Relationship between Educational Attainment and Labor Market Success: Evidence from Uganda, in: *International Migration Review* 42(2008):652-674.

[8] C. Chase, T. Anekwe, J. Barofsky, and F. Farshad, "Economic effects of malaria eradication: evidence from an eradication experiment in Kigezi, Uganda". Working Paper, 2008

## 6. Appendix. IPUMS-International Memorandum of Understanding with Official Statistical Agency to disseminate census microdata, Example: ISTAT (Italy)

**Letter of Understanding**

**Integrated Public Use Microdata Series International**
and **L'ISTITUTO NAZIONALE DI STATISTICA (ISTAT)**

Purpose. The purpose of this letter is to specify the terms and conditions under which metadata and microdata produced by **L'ISTITUTO NAZIONALE DI STATISTICA** shall be distributed by **Integrated Public Use Microdata Series International** of the University of Minnesota.

1. Ownership. ISTAT is the owner and licensee of the intellectual property rights (including copyright) in the metadata and microdata of Italy acquired by the University of Minnesota to be distributed by **Integrated Public Use Microdata Series International**.
2. Use. These data are for the exclusive purposes of teaching, scientific research and publishing, and may not be used for any other purposes without the explicit written approval, in advance, of ISTAT.
3. Authorization. To access or obtain copies of integrated microdata of Italy from **Integrated Public Use Microdata Series International**, a prospective user must first submit an electronic authorization form identifying the user (i.e., 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**, ISTAT, or other authorized distributors. Once approved, the user is licensed to acquire integrated metadata and microdata of Italy 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 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. Any attempt to ascertain the identity of a person, family, household, dwelling, organization, business or other entity from the microdata is strictly prohibited. Alleging that a person or any other entity has been identified in these data is also prohibited.
6. Security. Users will implement security measures to prevent unauthorized access to 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 Italy is permitted in communications such as scholarly papers, journals and the like. The authors of these communications are required to cite **ISTAT**

**and Integrated Public Use Microdata Series International** as the sources of the data of Italy, and to indicate that the results and views expressed are those of the author/user.

8. Violations. Violation of the user license may lead to professional censure, loss of employment, and/or civil prosecution. The University of Minnesota, national and international scientific organizations, and ISTAT will assist in the enforcement of provisions of this accord.
9. Sharing. **Integrated Public Use Microdata Series International** will provide electronic copies to ISTAT of documentation and data related to its integrated microdata as well as timely reports of authorized users.
10. Jurisdiction. Disagreements which may arise shall be settled by means of conciliation, transaction and friendly composition. Should a settlement by these means prove impossible, a Tribunal of Settlement shall be convened which will rule upon the matter under law. This Tribunal shall be composed of an arbitrator, which shall be selected by the ICC International Court of Arbitration. This agreement shall be governed by, and construed in accordance with, generally accepted principles of International Law.
11. Order of Precedence. In the event of a conflict between a term or condition of this Letter of Understanding and a term or condition of any Contract, to which this Letter of Understanding is attached, the term or condition in this Letter of Understanding shall prevail.

Date: 2/21/06

Signed: _____

**Regents of the University of Minnesota**
By: Kevin J. McKoskey, Sponsored Projects Administration

Date: 23.01.2006

Signed: _____
Rev. Jan. 27, 2005

**Table 1. IPUMS-International Integrated Samples Available to Researchers by Country and Decade**

55 countries, 159 samples, 325,573,167 person records (July 1, 2010)

Person records (000)

| | 2000s | 1990s | 1980s | 1970s | 1960s |
|---|---|---|---|---|---|
| | | **Americas** | | | |
| Argentina | 3,626 | 4,286 | 2,668 | 467 | . |
| Bolivia | 828 | 642 | . | 462 | . |
| Brazil | 10,136 | 8,523 | 5,870 | 4,954 | 3,001 |
| Canada | 801 | 810 | 487 | 214 | . |
| Chile | 1,514 | 1,335 | 1,133 | 890 | 88 |
| Colombia | 4,118 | 3,214 | 2,643 | 1,989 | 350 |
| Costa Rica | 382 | . | 241 | 187 | 82 |
| Cuba | 1,118 | . | . | . | . |
| Ecuador | 1,214 | 966 | 807 | 649 | 136 |
| Mexico | 10,099 | 8,118 | . | 483 | 503 |
| xxx5 | 10,285 | 332 | . | . | . |
| Panama | 284 | 233 | 196 | 150 | 54 |
| Peru | 2,745 | 2,206 | . | . | . |
| Puerto Rico | 190 | 178 | 160 | 27 | . |
| Saint Lucia | . | 13 | 11 | . | . |
| United States | 14,081 | 12,501 | 11,343 | 2,030 | 1,800 |
| xxx5 | 2,878 | . | . | . | . |
| Venezuela | 2,306 | 1,804 | 1,441 | 1,159 | . |
| | | **Europe** | | | |
| Austria | 803 | 781 | 757 | 750 | . |
| Belarus | . | 991 | . | . | . |
| France | 2,935 | 2,361 | 2,632 | 2,629 | 2,321 |
| xxx5 | . | . | . | . | 2,488 |
| Greece | 1,029 | 952 | 923 | 845 | . |
| Hungary | 511 | 518 | 536 | 515 | . |
| Italy | 2,991 | . | . | . | . |
| Netherlands | 190 | . | . | 159 | 143 |
| Portugal | 517 | 492 | 492 | . | . |
| Romania | 2,138 | 2,239 | . | 1,937 | . |
| Slovenia | 180 | . | . | . | . |
| Spain | 2,039 | 1,931 | 2,084 | . | . |
| Switzerland | 364 | 343 | 318 | 313 | . |
| United Kingdom | 1,844 | 542 | . | . | . |
| | | **Asia** | | | |
| Armenia | 327 | . | . | . | . |
| Cambodia | . | 1,141 | . | . | . |
| China | . | 11,836 | 10,039 | . | . |

cont.

| | 2000s | 1990s | 1980s | 1970s | 1960s |
|---|---|---|---|---|---|
| India (NSSO) | 603 | 565 | 623 | . | . |
| xxx5 (NSSO) | . | 597 | 668 | . | . |
| Iraq | . | 1,944 | . | . | . |
| Israel | . | 556 | 403 | 316 | . |
| Jordan | 511 | . | . | . | . |
| Kyrgyz Republic | . | 477 | . | . | . |
| Malaysia | 435 | 348 | 183 | 176 | . |
| Mongolia | 244 | . | 191 | . | . |
| Nepal | 2,583 | . | . | . | . |
| Pakistan | . | 13,102 | 8,433 | 1,453 | . |
| Palestine | . | 259 | . | . | . |
| Philippines | 7,418 | 6,014 | . | . | . |
| xxx5 | . | 6,865 | . | . | . |
| Thailand | 605 | 485 | 388 | 772 | . |
| Vietnam | . | 2,368 | 2,627 | . | . |
| | | **Africa** | | | |
| Egypt | . | 5,902 | . | . | . |
| Ghana | 1,894 | . | . | . | . |
| Guinea (Conakry) | . | 729 | 458 | . | . |
| Kenya | . | 1,408 | 1,074 | . | . |
| Mali | . | 991 | 785 | . | . |
| Rwanda | 843 | 743 | . | . | . |
| Senegal | 995 | . | 705 | . | . |
| South Africa | 3,726 | 3,621 | . | . | . |
| xxx5 | 1,048 | . | . | . | . |
| Tanzania | 3,733 | . | 2,310 | . | . |
| Uganda | 2,497 | 1,548 | . | . | . |

Source: https://international.ipums.org/international/samples.shtml
Note: "xxx5" refers to second sample in decade indicated

**Table 2. IPUMS-I Coding Schemes, Example: Educational Attainment**

| Code | country (ISO 3166 2 digit code): / year of sample (2 right-most digits): | AR 70 | AR 80 | AR 91 | AR 01 | AM 01 | BY 99 | BO 76 | BO 92 | BO 01 | BR 60 | BR 70 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **General Educational Attainment Code: 1 digit** | | | | | | | | | | | |
| 0 | NIU (not in universe) | X | X | X | X | X | X | X | X | X | X | X |
| 1 | Less than primary completed | X | X | X | X | X | X | X | X | X | X | X |
| 2 | Primary completed | X | X | X | X | X | X | X | X | X | X | X |
| 3 | Secondary completed | X | X | X | X | X | X | X | X | X | X | X |
| 4 | University completed | X | X | X | X | X | X | X | X | X | X | X |
| 9 | Unknown | X | . | X | . | . | X | X | X | X | X | X |
| | **Detailed Educational Attainment Code: 3 digit** | | | | | | | | | | | |
| 0 | NIU (not in universe) | X | X | X | X | X | X | X | X | X | X | X |
| 100 | LESS THAN PRIMARY COMPLETED | . | . | . | . | X | X | . | . | . | . | . |
| 110 | No schooling | X | X | X | X | . | . | X | X | X | X | X |
| 120 | Some primary | X | X | X | X | . | . | X | X | X | X | X |
| 130 | Primary (4 years) | . | . | . | . | X | . | . | . | . | X | X |
| | PRIMARY COMPLETED, LESS THAN SECONDARY | | | | | | | | | | | |
| | Primary completed | | | | | | | | | | | |
| 211 | Primary (5 years) | . | . | . | . | . | X | X | X | X | . | . |
| 212 | Primary (6 years) | 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 |
| 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 |
| 312 | Some college/university | X | X | X | X | X | . | X | X | X | X | X |
| 320 | Technical track | . | . | . | . | . | X | . | . | . | . | . |
| 321 | Secondary technical degree | X | X | . | . | X | . | X | X | X | . | . |
| 322 | Post-secondary technical | . | X | X | X | X | . | . | . | X | . | . |
| 400 | UNIVERSITY COMPLETED | X | X | X | X | X | X | X | X | X | X | X |
| 999 | UNKNOWN/MISSING | X | . | X | . | . | X | X | X | X | X | X |

X' indicates the category is available for that sample.

Source: https://international.ipums.org/international-action/codes.do?mnemonic=EDATTAN

**Fig. 1. Student Poster from integrated census microdata and other sources: Rwanda**

# Rwanda

Mai Nguyen, Stephanie Vogt, Michele Hurry, Naqwai Davis
SOC 3511 World Population Problems, University of Minnesota

## OBJICTIVE

Our objective is to depict the population changes that have occurred in Rwanda within the last thirty years throughthe use of fertility and mortality rates. We focus on showing how the Civil War and the genocide had devastating affects on the population of Rwanda and how the country is recovering from the tragedies.

## BACKGROUND

Rwanda is a little smaller than the state of Massachusetts,which has a population of about 6.5 million people. Rwanda, in comparison, is home to approximately 10.1 million people. It is the most densely populated country in central Africa..

Rwanda is primarily a rural society. The population is made up of three ethnic groups The Hutus, The Tutsi and The Twa. The economy is fueled primarily by the agriculture such as tea and coffee exports and service sectors.

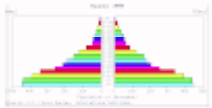
## METHODOLOGY

In order to achieve our results and conclusions we used the following methods:

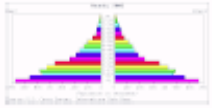
- UN data
- SPSS
- PUMS
- Graphs

## POPULATICH PYRAMIDS

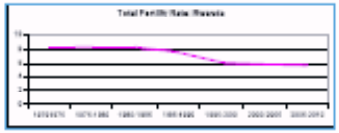
GENOCIDE



POST GENOCIDE



## DATA

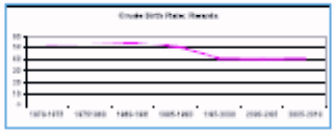


### Fertility

The fertility rate is relatively high,until 1990 which was the beginning of a civil war, leading to the genocide of roughly 800,000 people. There is a crop in fertility after 1990.

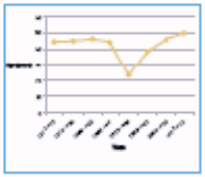
### Birth Rate

The crude birth rate rose steadily until just before the civil war began. There was a drastic drop around 1985 and leveled off in the year 1998



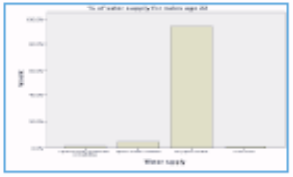
Currently conditions in Rwanda are steadily improving every year. Life expectancy is expected to reach 50 in 2010, which is over better than the life expectancy before the Genocide. Things that may have caused this rise are the improved Government and healthconditions for the people.

- The Civil War and 1994 Genocide in Rwanda had devastating effect on life expectancy for both sexes. Prior to the 1994 Genocide, the average life expectancy for both sexes was in the mid forties. That number drastically plummeted, years leading up to the genocide and also during it. At the height of the genocide, the life expectancy was in the low 20s. there are three factors that may have contributed to that drop in rate during the genocide.
- The first factor was the mass killings, it was reported that over 500,000 people died during the genocide.
- The second factor that may have affected this rate were the many rape cases that unfortunately caused many women to contract incurable sexual infections leading to their death.
- The last factor that may have contributed to this low rate was the fact that many women were widowed as well as many children were orphaned. This may haveled to poore health conditions resulting in a lower life expectancy.

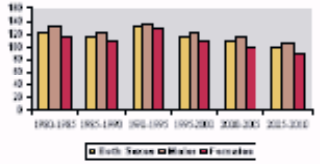


### Living Standards

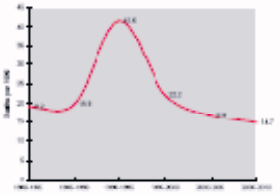
Males age 22, living in Rwanda have a higher chance of having piped water and a flushable toilet. Among the general population .5% have a flushable toilet. At age 22, 1.8% have a flushable toilet. 3% of the general population have piped water, 5.5% of males age 22 have piped water.



The infant mortality rate in Rwanda have been decreasing in the last thirty years, with the exception of the genocide during the 1990s. The decrease in infant mortality means that the population of Rwanda will continue to grow rapidly and the population will become younger.



The crude death rate in Rwanda had been decreasing prior to the genocide. The large peak in the crude death rate shows how drastic of an affect the genocide had on the mortality rates in Rwanda. The post-genocide crude death rate has decreased and is predicted to do so in the years to come.



## CONCLUSION

Overall, the Civil War and the genocide in Rwandahad a tremendous impact on the population and living standards of the country. These events increased mortality rates and decreased fertility rates thus causing a decrease in the population in the 1990s.

Today,Rwanda has overcome the affects of the genocide and is now experiencing steady fertility rates, low mortality rates, and a higher life expectancy. Factors such as improving health care and advancing technology in thecountry will contribute to the growth of the country in the years to come.

## SOURCES


- UN World Population Prospects. World Population Prospects: The 2008 Revision Population Database. "Rwanda." http://esa.un.org/unpp/index.asp?panel=2
- WHO | World Health Organization. "Rwanda." http://www.who.int/en/
- National Geographic. "Rwanda." http://travel.nationalgeographic.com/places/
- IPUMS data using SPSS. "Rwanda 2002."
- Background information "Rwanda" http://www.un.gov/rp/ei/bgn/2861.htm
- Population pyramids. "Rwanda."http://www.nationmaster.com/country/rw-rwanda/Age_distribution

**Fig. 2. Students showcase posters and field questions**

Photo credit: Wade Stebbings


Iraq: A Profile of Health and Mortality in the Country from 1960-2000
SOC 3511 World Population Problems, University of Minnesota