

WP.45
ENGLISH ONLY

**UNITED NATIONS ECONOMIC COMMISSION
FOR EUROPE (UNECE)**

EUROPEAN COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

**STATISTICAL OFFICE OF THE EUROPEAN
UNION (EUROSTAT)**

Joint UNECE/Eurostat work session on statistical data confidentiality
(Tarragona, Spain, 26-28 October 2011)

Topic (viii): International projects, groups and forum dealing with data access, release and related methodologies

Trans-Border access to Census Microdata: The IPUMS-IECM partnership, where a single license agreement opens access to microdata for more than 60 countries to researchers world-wide free of cost

Prepared by Albert Esteve, Centre d'Estudis Demografics (Barcelona, Spain)
and the Minnesota Population Center (Minneapolis, USA)

Trans-Border access to Census Microdata: The IPUMS-IECM partnership, where a single license agreement opens access to microdata for more than 60 countries to researchers world-wide free of cost*

Note by the Centre d'Estudis Demogràfics (Barcelona, Spain)
and the Minnesota Population Center (Minneapolis, USA)

“You have to do due diligence, something to assure yourself that the people you’re giving your data to can be trusted.”

--<http://www.nytimes.com/2011/09/09/us/09breach.html?hp>

I. SUMMARY

1. The IPUMS-International project has negotiated uniform agreements (Appendix A) with 98 national statistical offices (21 in Europe) to disseminate integrated census microdata to accredited researchers world-wide at no cost through a single licensing agreement (Appendix B). As of June 2011, the integration of microdata and metadata for the censuses of 62 countries (185 samples, 397 million person records) is available to almost 5,000 registered researchers from 94 countries. Thanks to sustained funding by the National Science Foundation, the National Institutes of Health (USA), and Research Infrastructures Action (EU) as well as exceedingly generous cooperation from National Statistical Offices worldwide, the database is expanding at the rate of 5-10 additional countries per year (see Table 1). The project (www.ipums.org/international) is led by the University of Minnesota Population Center (MPC) in partnership—for the censuses of Europe—with the Centre d'Estudis Demogràfics (CED), Autonomous University of Barcelona (www.iecm-project.org).

2. Usage statistics of 4,519 researchers making 24,699 extracts totalling 85,505 samples and 891,267 variables are analyzed to rank the samples and variables in greatest demand. The most striking finding is that 40 variables account for 60% of those requested and among these are sixteen IPUMS constructed variables, four of which are imputed from household information: spouse's location in household, mother's location, father's location, and the rules for inferring locations. High precision, richly detailed household samples are essential for high quality research based on census microdata.

II. Trans-Border access and IPUMS-International: a massive, widely used, global resource for restricted access census microdata

3. 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),

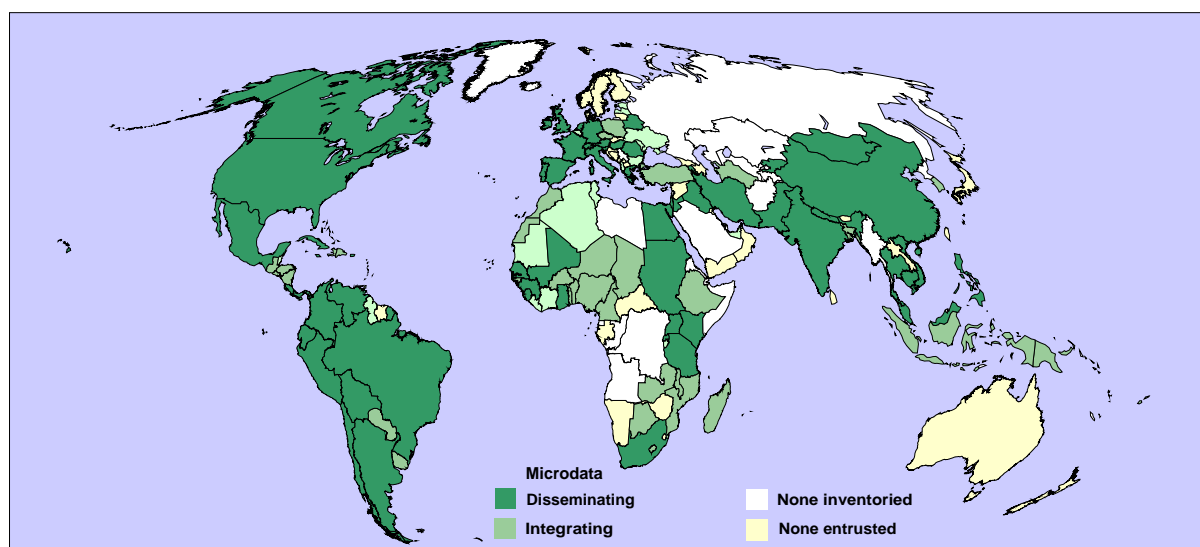
* Research for this paper was funded in part by the National Institutes of Health of the United States, grant HD047283 European and Asian census microdata harmonization project (IPUMS-EurAsia) and Harmonizing Integrated European Census Microdata (HIECM), funded by the European Union, Research Infrastructures Action, FP6-026033.

reside 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 trans-border access is essential if they are to conduct research using census microdata of their country of birth, whether comparative or not.

4. All census microdata in the IPUMS-International system is accessible to researchers, regardless of country of birth or of residence. IPUMS strives to maintain a balance between confidentiality protections, utility of microdata, and freedom of research. To protect the privacy and confidentiality of the microdata, we rely not only on technical measures, but also on legal and administrative controls. Researchers wishing to access the microdata must complete a lengthy application form and agree to specific conditions of use designed to protect the statistical confidentiality and good use of the microdata. Applicants are carefully vetted to prevent access to researchers who are unqualified or who lack a research need. Before an individual account is activated, we do “due diligence” to establish the research bona fides of every applicant.

5. IPUMS-International (www.ipums.org/international) archives, integrates, and disseminates high precision, richly detailed microdata from national population and housing censuses. This massive data infrastructure—totalling more than 397 million anonymized, integrated person records representing more than 100 million households—encompasses 62 countries and 185 censuses.

Table 1. Status of census microdata for countries participating in the IPUMS-International collaboratory, July 2011:
3 shades of green--Integrated (dark), Integrating (medium), and Negotiating (light).



6. Twenty-one European countries participate in the IPUMS/IECM initiatives (number of sets of microdata contributed in parentheses):

16 are integrated: Armenia (1), Austria (4), Belarus (1), France (7), Germany (4—includes GDR, FRG and 4 microcensuses to be integrated), Greece (4), Hungary (4), Ireland (8), Italy (1), the Netherlands (3), Portugal (3), Romania (3), Slovenia (1), Spain (3), Switzerland (4),

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

and the *United Kingdom (2—to be expanded to 6).

5 are being integrated: Bulgaria (0), the Czech Republic (2), Poland (1 to be expanded to 4), Turkey (3 to be launched in 2012), and Ukraine (0).

7. The 2012 IPUMS launch is scheduled to incorporate samples for Turkey (1985, 1990, 2000)—as well as five other non-European nations (El Salvador, Indonesia, Morocco, Mexico, and Uruguay). Additional launches are planned for successive years, integrating 2010 round census samples as expeditiously as they become available.

8. 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.

9. With IPUMS no two extracts are alike. Each extract is custom-tailored. The researcher, by means of a series of selections, places an order, indicating:

- 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 along with all other 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 a system file in SPSS, SAS or STATA. Copies of original source metadata are available from the web-site, as well as integrated metadata in interactive form.

10. At the UN-ECE Expert Group Meeting on Statistical Data Confidentiality, November 2005, we explained the IPUMS-International data dissemination procedure as follows²:

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 has weathered the test of time, and indeed as usage soars, the rapid acceleration of internet transmission speeds has validated the IPUMS approach.

² Robert McCaa and Albert Esteve, "[*IPUMS-Europe: Confidentiality measures for licensing and disseminating restricted access census microdata extracts to academic users*](#)," *Monographs of official statistics: Work session on statistical data confidentiality*. Luxembourg: Office for Official Publications of the European Communities, p. 40.

III. IPUMS-International Usage Statistics (through April 18, 2011)

11. 85,505 sample (24,699 extracts) have been downloaded from the IPUMS-International site, averaging over 1,500 samples per country for the 55 countries represented in the database (Table 2). Nonetheless, usage by country varies greatly. The smallest number of extracts, 153, was registered for the 2001 census of Nepal. The greatest number, 9,338 was registered for the six censuses of Mexico. The top ranked sample was Mexico 2000, with 2,945 extracts.

12. Mexico, Brazil and Colombia predominate in usage not only because their samples offer lots of variables and a long series covering a half century of dramatic demographic transformations, but also due to the fact that many Latin American emigrants reside in the United States (or Spain) and thus it is possible to analyze these populations in a single integrated database, whether the researcher resides in the country of birth or in the two most important countries of emigration—the USA and Spain. In addition, all the Latin American samples, as well as those for the United States and Spain, are high precision, household samples with richly detailed, extensive information on migration, economic, social and demographic variables of both individuals and households.

Rank	Country	Sample %*	Variables (n)*	Years of census samples	Extracts
1	Mexico	10	120	1960p, 70, 90, 95, 2000, 05	9,338
2	Brazil	5	106	1960, 70, 80, 91, 2000	6,889
3	United States	5	92	1960, 70, 80, 90, 2000, 05	6,171
4	Colombia	10	120	1964p, 72, 85, 93, 2005	4,629
5	France	33	94	1962, 68, 75, 82, 90, 99, 06	3,861
9	Canada	2.5	59	1971p, 81p, 91p, 2001p	2,539
13	Greece	10	89	1971, 81, 91, 2001	2,031
14	Spain	5	99	1981, 91, 2001	2,013
20	Austria	10	75	1971, 81, 91, 2001	1,608
21	Hungary	5	74	1970, 80, 90, 2001	1,517
23	Portugal	5	96	1981, 91, 2001	1,400
24	Romania	10	97	1976, 92, 2002	1,375
30	United Kingdom	3	47	1991, 2001p	960
31	The Netherlands	1	33	1960p, 71p, 2001p	810
36	Switzerland	5	79	1970, 80, 90, 2000	485
38	Belarus	10	84	1999	437
41	Italy	5	81	2001	370
50	Slovenia	10	80	2002	242
Total samples extracted for 55 countries (159 samples)					85,505
*2000 round census; refers to all integrated variables, including IPUMS constructed variables. "p" = person sample; all other samples are of households					

13. 4,519 researchers have qualified for access to the IPUMS-International database, representing 94 countries. The mean number of extracts per user is 5.5 and the mean number of samples per extract is 3.5. In total more than 85,000 samples have been extracted, totalling 891,267 variables. The fact that the average extract consists of a mere ten variables shows that IPUMS-International users are researchers, not hoarders. Their extract requests are parsimonious, limited to specifically what is needed to address well-defined research questions.

14. In June 2010, 736 integrated variables were available to researchers at that date. Table 3 lists the 32 most commonly extracted. The top 8 encompass 4 demographic variables (marital status,

relationship to head, age and sex), 2 economic (employment status and class of worker), and one each social (educational attainment) and technical (person weight).

<u>Rank</u>	<u>Label</u>	<u>Extracts</u>	<u>Mnemonic</u>	<u>Comment</u>
1	Educational attainment	19,307	EDATTAN	
2	Age (single years to 85+)	19,009	AGE	Grouped age n=3,838
3	Employment status	18,490	EMPSTAT	
4	Marital status	18,214	MARST	
5	Person weight	17,511	WTPER	Technical variable
6	Relationship to head	15,783	RELATE	
7	Sex	14,595	SEX	
8	Class of work	12,583	CLASSWK	
9	Ownership of dwelling	8,050	OWNRSHP	
10	Occupation ISCO recode	8,004	OCCISCO	
11	School attendance	7,919	SCHOOL	
12	Years of schooling	7,576	YRSCHL	
13	Literate	7,290	LIT	
14	Urban/rural	7,098	URBAN	
15	Industry-general code	7,044	INDGEN	
16	Household weight	6,656	WTHH	Technical variable
17	Children ever born	6,363	CHBORN	
18	Nativity (native/foreign born)	6,332	NATIVTY	
19	Occupation	6,246	OCC	
20	Country of birth	6,153	BPLCTRY	
21	Religion	6,075	RELIG	
22	Industry	5,670	IND	
23	Location of spouse in household	5,007	SPLOC	Constructed (household)
24	Rule for locating spouse	4,171	SPRULE	Constructed (household)
25	Location of mother in household	4,153	MOMLOC	Constructed (household)
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	Constructed (household)
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	

15. One of the more surprising rankings is the presence of four IPUMS constructed variables among the top 30 requested: location in household of spouse, mother, and father and the rule used in locating the spouse—the “LOC” variables to experienced IPUMS users. Researchers exploit these variables to study the joint characteristics of spouses and characteristics of parents relative to their children. The variables are constructed by inference 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.

16. The IPUMS-I “Top 40” institutions in terms of data usage includes many of the world’s premier universities and research organizations (see Appendix C), 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 activity.) 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 14 countries with microdata in the IPUMS-I database but as yet no national researchers use them. The 14 are: Armenia, Belarus, Ghana, 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 these is no obligation to do so.

17. Future initiative: trans-border on-line tabulator. The Center for Demographic Studies, in collaboration with the Minnesota Population Center, is designing new tools to foster trans-border dissemination of European census microdata. As part of this effort, an online tabulator of integrated variables is currently under development. This will add a fast and convenient tool to explore the contents of the database before making an extract. Researchers will be able to do exploratory analysis or obtain some basic figures without the need to download microdata. The tabulator is built upon REDATAM, a trusted and secure UN software to provide access to census microdata (www.eclac.org/redatam/). Researchers will have access to a single sample or to multiple samples, as needed. Access will be restricted to registered researchers. The microdata will be identical to the integrated samples disseminated by IPUMS-International and will be subjected to the same statistical disclosure controls.

IV. Conclusion

17. When we began a decade ago, we dreamed of integrating samples for 21 countries in ten years. Thanks to the generous cooperation of National Statistical Offices and undreamed of technological innovations, that number has tripled, and work continues. 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 our success. Participating statistical agencies are invited to entrust metadata and microdata for the 2010 census round at their earliest convenience. Those that are not yet participating in the IPUMS/IECM initiatives are invited to consider doing so. Researchers who have yet to access the IPUMS/IECM microdata are invited to peruse the metadata and use the microdata should their research needs require.

Appendix A. Example of Uniform Memorandum of Understanding between the University of Minnesota and National Statistical Offices (Italy, 2006).

Letter of Understanding	
Integrated Public Use Microdata Series International and L'ISTITUTO NAZIONALE DI STATISTICA (ISTAT)	
<p><u>Purpose.</u> 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.</p> <ol style="list-style-type: none"> <u>Ownership.</u> 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. <u>Use.</u> 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. <u>Authorization.</u> 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. <u>Restriction.</u> 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. <u>Confidentiality.</u> 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. <u>Security.</u> Users will implement security measures to prevent unauthorized access to microdata acquired from Integrated Public Use Microdata Series International or its partners. <u>Publication.</u> 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 	<p>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.</p> <ol style="list-style-type: none"> <u>Violations.</u> 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. <u>Sharing.</u> 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. <u>Jurisdiction.</u> 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. <u>Order of Precedence.</u> 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. <p>Date: <u>2/21/06</u></p> <p>Signed: <u>[Signature]</u> Regents of the University of Minnesota By: Kevin J. McKoskey, Sponsored Projects Administration</p> <p>Date: <u>23.01.2006</u></p> <p>Signed: <u>[Signature]</u> <small>Rev. Jan. 27, 2005</small></p>

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

<p style="font-size: small;">IPUMS International Page 1 of 1</p> <p>Application to Use Restricted Microdata</p> <p>IPUMS-International microdata are available free of charge, but their use imposes responsibilities upon the user. To access the data, a prospective user must submit an electronic authorization form (this form) identifying the user by name, electronic address, and institutional affiliation.</p> <p>The investigator must state the purpose of the proposed project and agree to abide by the regulations specified below. If multiple investigators are involved in a project, all must register separately.</p> <p>Once a user is approved, a message will be sent by email granting access to the system. The notification licenses the user to acquire microdata from Integrated Public Use Microdata Series International or other authorized distributors. No titles or other rights are conveyed to the user.</p> <p>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.</p> <p>Information provided on this form will be kept confidential. All information on this form is required for registration unless otherwise indicated by an asterisk.</p> <p style="text-align: center;">PERSONAL INFORMATION</p> <p style="text-align: center;">*****</p> <p style="text-align: center;">INSTITUTIONAL AFFILIATION</p> <p>IPUMS-International staff must confirm the identity of prospective users. To speed the processing of your application, please provide as much of the following</p> <p>Name of institution or employer <input style="width: 100%;" type="text"/></p> <p>Your email address at institution (*) <input style="width: 100%;" type="text"/></p> <p>Web link showing your affiliation with institution (*) <input style="width: 100%;" type="text"/></p> <p>Email address of employer, supervisor, or instructor (*) <input style="width: 100%;" type="text"/></p> <p>Phone number of institution (*) <input style="width: 100%;" type="text"/></p> <p>Does your institution have an Institutional Review Board (IRB), or Office for Human Subject Protections, Professional Conduct or similar committee?</p> <p style="margin-left: 20px;"> <input type="radio"/> No <input type="radio"/> Yes; Name of board or office <input style="width: 100%;" type="text"/> </p> <p>RESEARCH PROJECT</p> <p>Please provide at least 75 words <i>in English</i> describing your research project or educational use for the data. This description will be used to evaluate your application.</p> <p style="text-align: center;">*****</p> <p>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. (*)</p> <input style="width: 100%; height: 20px;" type="text"/>	<p style="text-align: center;">USAGE LICENSE</p> <p>Please check all of the following boxes to indicate that you have read about the limitations of the IPUMS-International data and you agree to abide by the conditions of use. The purpose of this license is to specify the terms and conditions 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.</p> <p>Data must not be redistributed without authorization.</p> <p>All data extracted from the IPUMS-International database are intended solely for the use of the licensee. Under IPUMS-International agreements with collaborating agencies, redistribution of the data to third parties is prohibited. Each member of a research team using the data must apply for access and be licensed individually.</p> <p>The microdata are intended only for scholarly research and educational purposes.</p> <p><input type="checkbox"/> These microdata are provided for the exclusive purposes of teaching and scholarly research, and may not be used for any other purposes without explicit written approval from the relevant official statistical authority.</p> <p>Commercial use and redistribution of the microdata is strictly prohibited.</p> <p><input type="checkbox"/> Users are prohibited from using microdata acquired from the Integrated Public Use Microdata Series International or other authorized distributors in the pursuit of any commercial or income-generating venture either privately, or otherwise.</p> <p>Use of the microdata must follow strict rules of confidentiality.</p> <p><input type="checkbox"/> Users will maintain the confidentiality of persons and households. Any attempt to ascertain the identity of persons or households from the microdata is prohibited. Alleging that a person or household has been identified in these data is also prohibited. Statistical results that might reveal the identity of persons or entities may not be reported or published in any form.</p> <p>The microdata must always be safely secured.</p> <p><input type="checkbox"/> Users will implement security measures to prevent unauthorized access to microdata acquired from Integrated Public Use Microdata Series International, its partners or authorized distributors. Upon the completion of this research, data may be retained only if they can be safely secured. If security cannot be guaranteed, the microdata must be destroyed.</p> <p>Scholarly publications are permitted, and must be cited appropriately.</p> <p><input type="checkbox"/> 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.</p> <p>Any violation of this license agreement will result in disciplinary action, including possible loss of employment.</p> <p><input type="checkbox"/> 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.</p> <p><input type="checkbox"/> User agrees to notify ipums@pop.umn.edu regarding errors in the data.</p>
---	--

Appendix C. IPUMS-I Top 40 University or 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	International 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	Organization 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