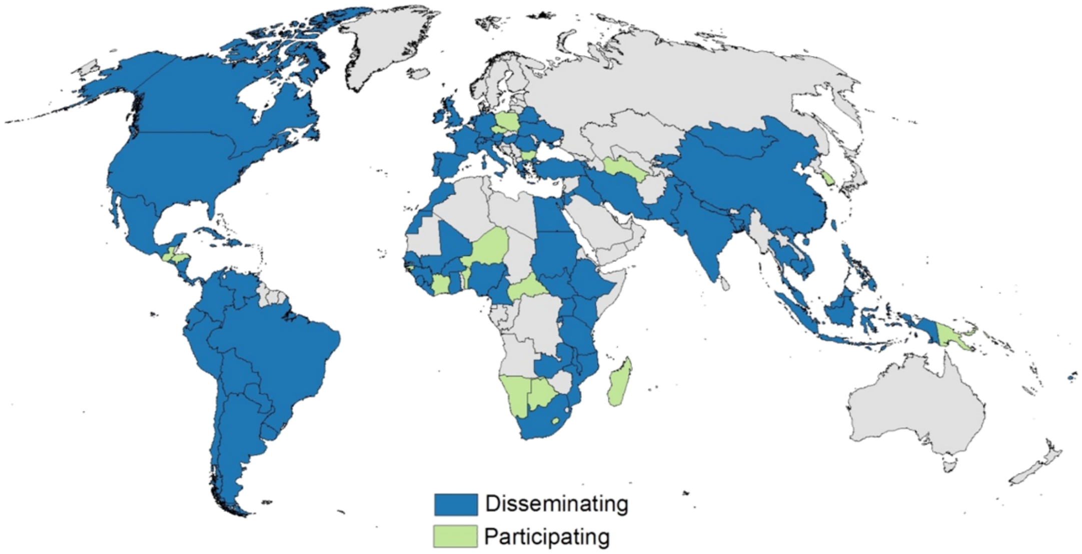
IPUMS International Workshop:
Promoting Understanding of Statistics about Society with free online data from international censuses

Patricia Kelly Hall and Lara Cleveland
Minnesota Population Center
University of Minnesota
clevelan@umn.edu

The Integrated Public Use Microdata Series-International (IPUMS International) is a data infrastructure project which disseminates high-precision census microdata samples to researchers world-wide free of cost. The samples in IPUMS are drawn from the very data source used to create official statistics by each country for policy and planning purposes. In partnership with most of the world's national statistical agencies, as well as data archives, research centers, and international organizations, IPUMS International has assembled the most comprehensive collection of census microdata in the world (Figure 1). The microdata records describe more than one-half billion persons nested within families and households, spanning five continents and more than 80 percent of the world's population (Appendix A). IPUMS-International lowers barriers to cross-national and cross-temporal research and teaching by converting international census microdata to a uniform format, providing comprehensive documentation, and making custom downloadable data files available through a user-friendly Web-based access system.

For each person, data include detailed information about geographic location, demographic characteristics, and economic activities. Individuals are nested within families and households, thereby preserving information about inter-relationships within residential groups. The data cover a broad range of population characteristics, including education and literacy, fertility history, child mortality, migration and place of former residence, marital status and consensual unions, disabilities, characteristics of the building (floor, roof, etc.), and a host of other characteristics.

Figure 1. IPUMS-International disseminates population data from 82 countries and 277 censuses



In: J. Engel (Ed.), *Promoting understanding of statistics about society.*
Proceedings of the Roundtable Conference of the International Association of Statistics Education (IASE),
July 2016, Berlin, Germany. ©2016 ISI/IASE iase-web.org/Conference_Proceedings.php

IPUMS-International converts census microdata for multiple countries into a consistent format, supplying comprehensive documentation and making the data available through a web-based data dissemination system. Along with supplying unique access to these nationally representative datasets, the principal advantage of IPUMS-International is its replacement of sample-specific variable codes with new integrated codes that are consistent across time and space. This "variable integration" ensures that identical concepts have identical codes, simplifying comparative analysis of multiple samples. More than 700 integrated variables are included in the IPUMS-International database, and the website displays at a glance which variables are included in each sample. Original or "source" variables are also available.

Guidelines from international organizations encourage consistency in census question wording and coding. However, each country's statistical office ultimately decides the subjects covered, the question wording, who was asked a question (i.e., the question universe), and the response categories included in their national census. Inevitably, then, other issues of comparability not covered by IPUMS-International's variable harmonization arise for researchers doing comparative analysis of census data. Sample descriptions and variable-specific documentation on the IPUMS-International website are designed to highlight possible comparability problems, so users can make informed judgments or adjustments and avoid inadvertent errors.

Workshop activities

Participants at the IPUMS-International workshop were trained in navigating the interactive metadata system, building customized datasets using the web dissemination system, analyzing census data online with the data tabulator, and other IPUMS tools designed specifically for classroom use. The User Guide (Appendix B) provides additional instructions and helpful tips for accessing data through the IPUMS International website.

Interactive metadata

IPUMS-International provides harmonized English-language documentation on each sample. This documentation covers enumeration procedures and instructions; definitions of households, dwellings, group quarters, and other enumeration units; guidance on the variability in sample design; and scanned images of original-language versions of the questionnaires. IPUMS also provides descriptions of the sources for each variable, including question wording and instructions (in the original and translated into English), universe definitions, frequency distributions, and variable codes. Comparability discussions describe any deviations of particular censuses from the standard variable definition and address differences over time and across countries. Participants at the workshop explored sample documentation and navigated the variable metadata system, which allows users to filter the information displayed to only those elements relevant to a given research project, as defined by the user. Participants were provided with metadata-related exercises designed to encourage responsible research and informed dataset creation. Exercises highlighted the value of variable documentation and the importance of considering issues of comparability in cross-national and cross-temporal research.

Customized data downloads

The IPUMS data access system allows users to merge datasets, select variables, define population subsets, and draw subsamples tailored to their specific needs. Workshop participants learned to build customized datasets containing the samples, variables and cases of their choosing. They were shown how to select cases based on individual and household characteristics and draw representative subsamples.

Workshop participants also learned how to best take advantage of their personal IPUMS portals. Each registered user of IPUMS-International has a private, password-protected extract history page. This page contains the statistical package syntax files and data for download of recent extract requests, as well as the extract syntax file and description (if user provided) for each data order ever requested by that user. With one click on the "resubmit" button, a researcher can regenerate the same extract. The "Revise" button opens the syntax file so the researcher can modify the data request. This is particularly useful in the classroom where a complex classroom exercise or exam can be re-used for a new class by modifying the data request with a different country or year.

Online tabulator

Workshop participants were trained in the IPUMS-International online data tabulator.

The IPUMS International website features a robust online tabulation system available to registered IPUMS International data users. Researchers can quickly analyze data files for individual samples, pooled samples from multiple census years within a country, or all pooled samples from a world region. The data tabulator supports descriptive and inferential analyses, including frequencies and cross-tabulations, comparison of means, and regression analysis. The interface allows users to recode existing variables, construct new variables, or exclude specified values.

Classroom features

The IPUMS-International workshop highlighted IPUMS-International features of special interest to statistics educators, including classroom accounts and the value of harmonized data for teaching. The IPUMS international user register system also includes a classroom feature. Course instructors can apply to register a class for a specified duration of time. Upon approval, the instructor receives a code for their students. Students get facilitated registration and are automatically assigned to membership in the class. Instructors can push common data extracts to students enrolled in the class. Students also have full IPUMS user capabilities for the duration of the quarter or semester and can also make their own customized data extracts.

Educators also find the uniform variables and coding schemes in IPUMS useful in teaching. Uniform coding means that a statistical algorithm developed to answer a question with one sample (country and year), can be readily applied to other samples. This feature is useful for facilitating student exploration of new contexts. It can also come in handy for developing exam materials, since a problem set used for an exam one semester can be re-run to output a different set of answers (using a different sample) for the same exam in a different semester.

With a few tools to aid statistical educators in understanding the data, IPUMS can be used to effectively engage students in accessing real-world data to understand social problems. Students will be interested in how to critically evaluate news reports involving statistics and, importantly, in how to make sure their own analyses are rigorous enough to qualify as accurate. Students and instructors may have to be prepared to consult references and conduct investigations outside the mathematical world of traditional statistical instruction in order to fully explore the questions raised by examining statistics about the social world, even those as simple as the age distribution comparisons shown in this paper. That alone makes teaching with real world data both exciting and daunting. Partnerships between statistically trained subject matter scholars and statistics educators are fertile ground for engaging students and training well-rounded social statisticians. For more information, visit https://international.ipums.org.

.

Appendix A. IPUMS Integrated Microdata Samples
https://international.ipums.org 86 countries 675 million person records (+ = 2016 launch)

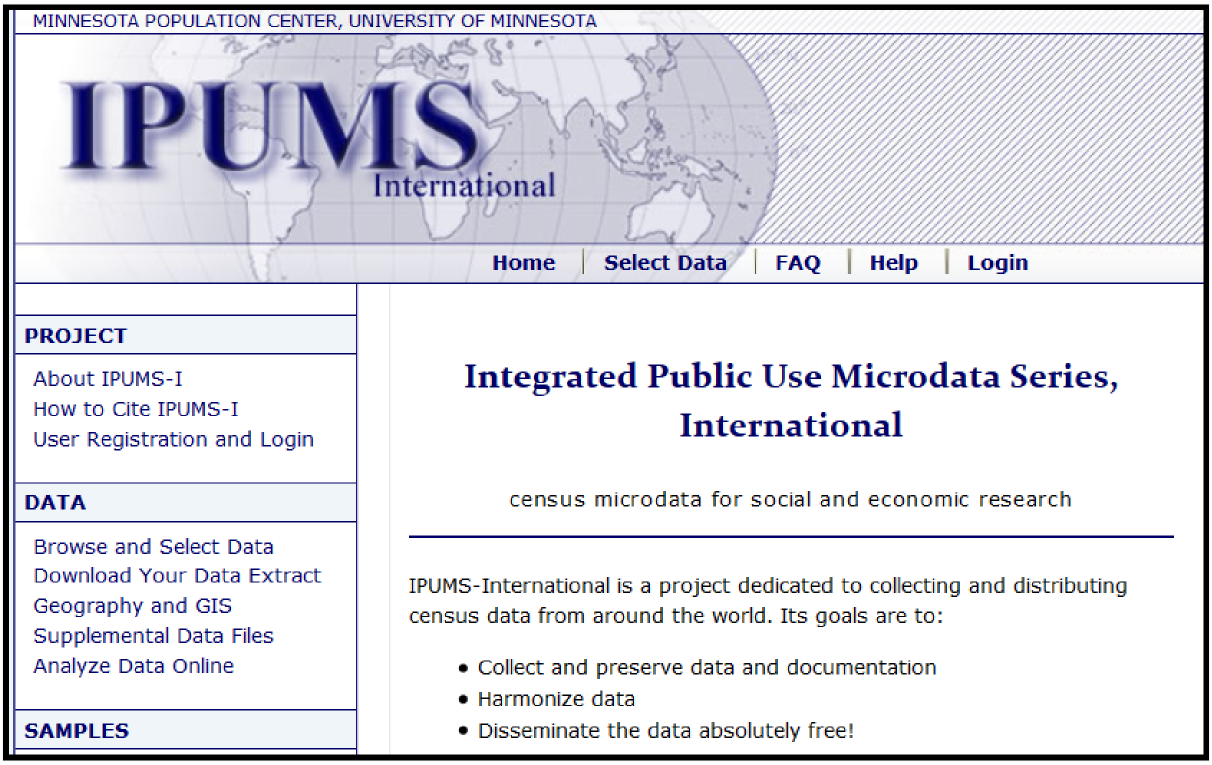
Census	%	Persons	Census	%	Persons	Census	%	Persons	Census	%	Persons
1970 Argentina	2	466,892	**1960 Chile**	1	88,184	**1962 France**	5	2,320,901	**2005 Indonesia-cont.**	0.5	1,090,892
1980	10	2,667,714	1970	10	890,481	1968	5	2,487,778	2010	10	23,603,049
1991	10	4,286,447	1982	10	1,133,062	1975	5	2,629,456	**2006 Iran**	2	1,299,825
2001	10	3,626,103	1992	10	1,335,055	1982	5	2,631,713	*2011+*	*2*	*1,482,000*
2010	10	3,966,245	2002	10	1,513,914	1990	4.2	2,360,854	**1997 Iraq**	10	1,944,278
2001 Armenia	10	326,560	**1982 China**	1	10,039,191	1999	5	2,934,758	**1971 Ireland**	10	296,878
2011	10	301,831	1990	1	11,835,947	2006 (RR)	33	19,973,287	1979	10	337,686
1971 Austria	10	749,894	*2000+*	*1*	*11,804,000*	2011	33	20,541,337	1981	10	344,291
1981	10	756,556	**1964 Colombia**	2	349,652	**1970 Germany**	5	3,094,845	1986	10	355,020
1991	10	780,512	1973	10	1,988,831	1971 DR	25	4,089,856	1991	10	353,149
2001	10	803,471	1985	10	2,643,125	1981 DR	25	4,278,563	1996	10	365,323
2011	10	839,501	1993	10	3,213,657	1987	5	3,160,224	2002	10	410,688
1991 Bangladesh	10	10,580,904	2005	10	4,006,168	**1984 Ghana**	15	1,309,352	2006	10	440,314
2001	10	12,442,115	**1963 Costa Rica**	6	82,345	2000	10	1,894,133	2011	10	474,535
2011	5	7,205,720	1973	10	186,762	2010	10	2,466,289	**1972 Israel**	10	315,608
1999 Belarus	10	990,706	1984	10	241,220	**1971 Greece**	10	845,483	1983	10	403,474
2009+	*10*	*941,000*	2000	10	381,500	1981	10	923,108	1995	10	556,365
1976 Bolivia	10	461,699	2011	10	430,082	1991	10	951,875	**2001 Italy**	5	2,990,739
1992	10	642,368	**2002 Cuba**	10	1,118,767	2001	10	1,028,884	**1982 Jamaica**	10	223,667
2001	10	827,692	**1960 Dominican Rep**	6.6	201,556	*2011+*	*10*	*1,057,000*	1991	10	232,625
1981+ Botswana	*10*	*97,000*	1970	6.8	272,090	**1983 Guinea**	10	457,837	2001	10	205,179
1991+	*10*	*133,000*	1981	8.5	475,829	1996	10	729,071	**2004 Jordan**	10	510,646
2001+	*10*	*169,000*	2002	10	857,606	**1971 Haiti**	10	434,869	**1969 Kenya**	3.3	659,310
2011+	*10*	*202,000*	2010	10	943,784	1982	2.5	128,770	1979	5	1,033,769
1960 Brazil	5	3,001,439	**1962 Ecuador**	3	136,443	2003	10	838,045	1989	5	1,074,098
1970	5	4,953,759	1974	10	648,678	**1970 Hungary**	5	515,119	1999	5	1,407,547
1980	5	5,870,467	1982	10	806,834	1980	5	536,007	2009	10	3,841,935
1991	5.8	8,522,740	1990	10	966,234	1990	5	518,240	**1999 Kyrgyzstan**	10	476,886
2000	6	10,136,022	2001	10	1,213,725	2001	5	510,502	2009	10	564,986
2010	5	9,693,058	2010	10	1,448,233	*2011+*	*5*	*497,000*	**1974 Liberia**	10	150,256
1985 Burkina Faso	10	884,797	**1986+ Egypt**	*15*	*6,799,000*	**1983 India - NSSO**	0.1	623,494	2008	10	348,057
1996	10	1,081,046	1996	10	5,902,243	1987	0.1	667,848	**1987 Malawi**	10	798,669
2006	10	1,417,824	2006	10	7,282,434	1993	0.1	564,740	1988	10	991,393
1998 Cambodia	10	1,141,254	**1992 El Salvador**	10	510,760	1999	0.1	596,688	2008	10	1,341,977
2008	10	1,340,121	2007	10	574,364	2004	0.1	602,833	**1970 Malaysia**	2	175,997
1976 Cameroon	10	736,514	**1984 Ethiopia**	10	3,404,306	*2010+*	*0.1*	*460,000*	1980	2	182,601
1987	10	897,211	1994	10	5,044,598	**1971 Indonesia**	0.5	634,642	1991	2	347,892
2005	10	1,772,359	2007	10	7,434,086	1976	0.2	281,170	2000	2	435,300
1971 Canada	1	214,019	**1966 Fiji Islands**	10	47,579	1980	5	7,234,577	**1987 Mali**	10	785,384
1981	2	486,875	1976	10	57,214	1985	0.3	605,858	1998	10	991,330
1991	3	809,654	1986	10	72,158	1990	0.5	912,544	2009	10	1,451,856
2001	2.5	801,055	1996	10	77,382	1995	0.3	718,837			
2011+	*3*	*926,000*	2007	10	84,323	2000	10	20,112,539	**(continued)**		

Table 1. IPUMS Integrated Microdata Samples https://international.ipums.org (continued)
86 countries 675 million person records (+ = 2016 launch)

Census	%	Persons	Census	%	Persons	Census	%	Persons	Census	%	Persons
1960 Mexico	1.5	502,800	**1993 Peru**	10	2,206,424	**2008 Sudan**	17	5,066,530	**1963 Uruguay**	10	256,171
1970	1	483,405	2007	10	2,745,895	**1970 Switzerland**	5	312,538	1975	10	279,994
1990	10	8,118,242	**1990 Philippines**	10	6,013,913	1980	5	317,803	1985	10	295,915
1995	0.4	332,061	1995	10	6,864,758	1990	5	342,797	1996	10	315,920
2000	11	10,099,182	2000	10	7,417,810	2000	5	364,086	2006	6	256,866
2005	10	10,284,550	***1978+ Poland***	*10*	*3,577,000*	**1988 Tanzania**	10	2,310,424	2011	10	328,425
2010	10	11,938,402	*1988+ coming*	*10*	*3,894,000*	2002	10	3,732,735	**1971 Venezuela**	10	1,158,527
2015+	*10*	*11,292,000*	*2002+ September*	*10*	*3,824,000*	*2011+*	*10*	*4,497,000*	1981	10	1,441,266
1989 Mongolia	10	190,631	*2011+ 2016*	*5*	*2,000,000*	**1970 Thailand**	2	772,169	1990	10	1,803,953
2000	10	243,725	**1981 Portugal**	5	492,289	1980	1	388,141	2001	10	2,306,489
1982 Morocco	5	1,012,873	1991	5	491,755	1990	1	485,100	**1989 Vietnam**	5	2,626,985
1994	5	1,294,026	2001	5	517,026	2000	1	604,519	1999	3	2,368,167
2004	5	1,482,720	2011	5	528,870	***1970+ Trinidad & T***	*10*	*69,000*	2009	15	14,177,590
1997 Mozambique	10	1,551,517	**1970 Puerto Rico**	1	27,212	*1980+*	*10*	*105,000*	**1990 Zambia**	10	787,461
2007	10	2,047,048	1980	5	160,219	*1990+*	*10*	*113,000*	2000	10	996,117
2001 Nepal	11	2,583,245	1990	5	177,655	*2000+*	*10*	*112,000*	2010	10	1,321,973
1960 Netherlands	1.2	143,251	2000	5	189,828	*2011+*	*10*	*117,000*	**Candidates for 2017 and beyond:**		
1971	1.2	159,203	2005 (PRCS)	1	35,416	**1985 Turkey**	5	2,554,364	**2010 round censuses**		
2001	1.2	189,725	2010	1	36,032	1990	5	2,864,207	**More countries/places**		
1971 Nicaragua	10	189,469	**1977 Romania**	10	1,937,021	2000	5	3,444,456	Angola		
1995	10	435,728	1992	10	2,238,578	**1991 Uganda**	10	1,548,460	Australia		
2005	10	515,485	2002	10	2,137,967	2002	10	2,497,449	Belgium		
2006 Nigeria - GHS	0.1	83,700	*2011+*	*10*	*1,990,000*	**2001 Ukraine**	10	4,889,288	Benin		
2007	0.1	85,183	**1991 Rwanda**	10	742,918	**1991 UK**	1	541,894	Bulgaria		
2008	0.1	107,425	2002	10	843,392	2001	3	1,843,525	Cote d' Ivoire		
2009	0.1	77,896	*2012+*	*10*	*1,038,369*	**1960 USA**	1	1,799,888	Finland		
2010	0.1	72,191	**1980 Saint Lucia**	10	11,451	1970	1	2,029,666	Guatemala		
1973 Pakistan	2	1,453,332	1991	10	13,382	1980	5	11,343,120	Guinea Bissau		
1981	10	8,433,058	**1988 Senegal**	10	700,199	1990	5	12,501,046	Honduras		
1998	10	13,102,024	2002	10	994,562	2000	5	14,081,466	Japan		
1997 Palestine	10	259,191	**2004 Sierra Leone**	10	494,298	2005 (ACS)	1	2,878,380	Korea, Republic of		
2007	10	227,067	**2002 Slovenia**	10	179,632	2010	1	3,061,692	Madagascar		
1960 Panama	5	53,553	**1996 South Africa**	10	3,621,164	**1963 Uruguay**	10	256,171	Mauritius		
1970	10	150,473	2001	10	3,725,655	1975	10	279,994	Myanmar		
1980	10	195,577	2007	2	1,047,657	1985	10	295,915	Namibia		
1990	10	232,737	2011	10	4,418,594	1996	10	315,920	Niger		
2000	10	284,081	**2008 South Sudan**	7	542,765	2006	6	256,866	Nigeria PES		
2010	10	341,118	**1981 Spain**	5	2,084,221	2011	10	328,425	Papua New Guinea		
1962 Paraguay	5	90,236	1991	5	1,931,458	**1971 Venezuela**	10	1,158,527	Russia		
1972	10	233,669	2001	5	2,039,274	1981	10	1,441,266	Tunisia		
1982	10	301,582	2011	10	4,107,465	1990	10	1,803,953	Turkmenistan		
1992	10	415,401				2001	10	2,306,489	Yemen		
2002	10	516,083							Zimbabwe, etc.		

Appendix B

User's Guide



Integrated Public Use Microdata Series, International

census microdata for social and economic research

IPUMS-International is a project dedicated to collecting and distributing census data from around the world. Its goals are to:

- Collect and preserve data and documentation
- Harmonize data
- Disseminate the data absolutely free!

Register for access

Original census forms, enumerator instructions, harmonized variable codes and descriptions, and other critical metadata are available to everyone in an interactive web system. Access to the microdata is restricted to registered users. Follow the directions below to become a registered user.

- Click **User Registration and Login** on the IPUMS-I home page.
- To preview the application form, click **View application form.**
- When you have collected the needed information, click **Apply for Access**.
- Submit your email address and password, which takes you to the application form.
- Briefly describe your research plan. Show that you have a non-commercial research project that requires access to international census microdata, as required by our partner legal agreements.
- Click each tick box and agree to all terms and conditions for access to the data, such as protecting confidentiality, not sharing the data, and citing the data properly.
- Submit the application for review.

Notification of approval (or denial) is emailed within a few days, after the application is reviewed.

Study the online integrated metadata

- For information on sampling and censuses, click **Sample Descriptions**.
- **Source Documents** provides original forms and instructions and their English translations.
- If only some countries or years interest you, in the **Select Data** browsing window, use the **Select Samples** option to limit the information displayed.

www.international.ipums.org

How to make a data extract

You can make as many extracts as you need. In each extract, include only variables required for one project, to simplify your analysis.

- After login, click **Select Data** (on top menu bar), then click **Select Samples.**
- Click a box for each sample (country and year) you need, then click **Submit sample selections.**
- To see the variables in each sample, browse in the **Select Data** window and choose **Household** or **Person** variable display. From the drop-down menu, choose a variable group (e.g., **Demographic**).
- Click the variable name (e.g., EDATTAIN) for more information, such as codes, question wording, comparability, and universes. To add a variable to your extract, check the yellow circle to the left of the variable name or click the **add to cart** box in the variable's description.
- Repeat until all variables of interest are selected. Review selections by clicking **View Cart**.

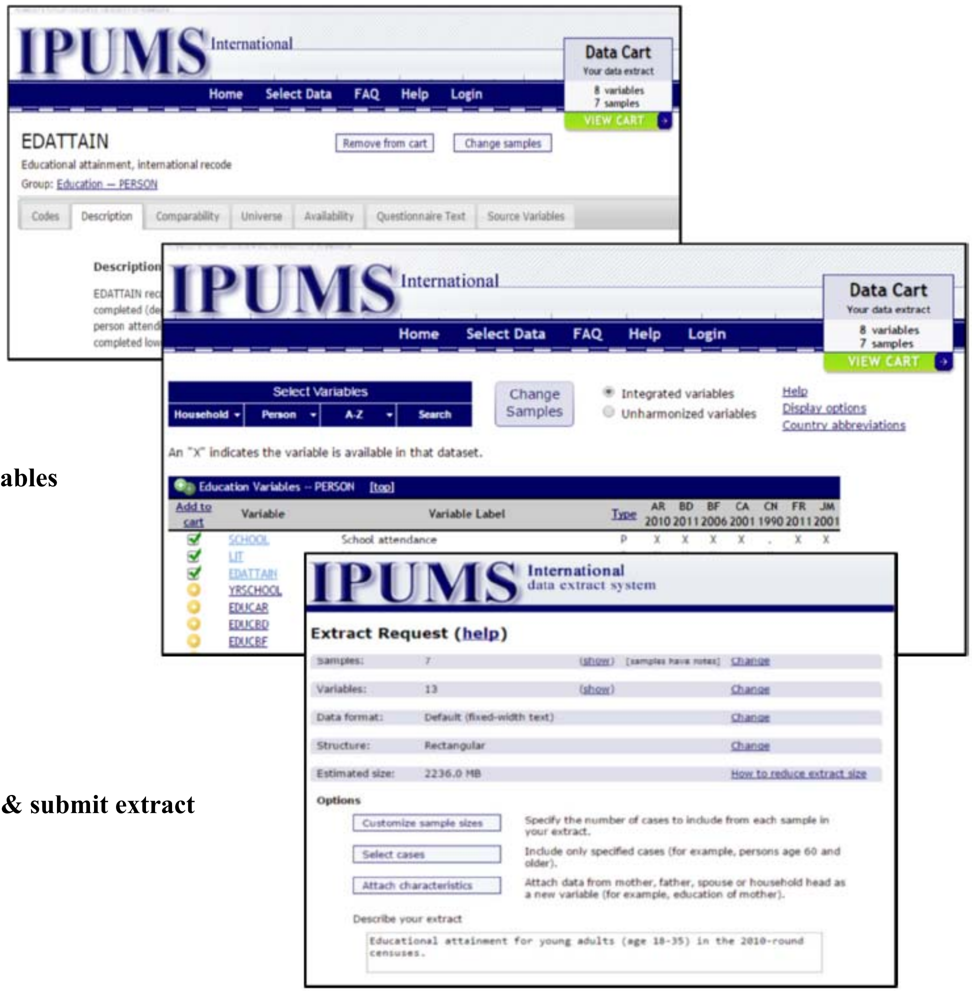


Figure 1.
Study the metadata

Figure 2.
Select samples & variables

Figure 3. Customize & submit extract

EDATTAIN

Educational attainment, international recode

Group: Education -- PERSON

Description

EDATTAIN rec
completed (de
person attend
completed low

● Integrated variables
○ Unharmonized variables

An "X" indicates the variable is available in that dataset.

Education Variables -- PERSON [top]

Add to cart	Variable	Variable Label	Type	AR 2010	BD 2011	BF 2006	CA 2001	CN 1990	FR 2011	JM 2001
✓	SCHOOL	School attendance	P	X	X	X	X	.	X	X
✓	LIT									
✓	EDATTAIN									
○	YRSCHOOL									
○	EDUCAR									
○	EDUCBD									
○	EDUCBF									

Extract Request (help)

Samples:	7	(show) [samples have notes]	Change
Variables:	13	(show)	Change
Data format:	Default (fixed-width text)		Change
Structure:	Rectangular		Change
Estimated size:	2236.0 MB		How to reduce extract size

Options

Customize sample sizes — Specify the number of cases to include from each sample in your extract.

Select cases — Include only specified cases (for example, persons age 60 and older).

Attach characteristics — Attach data from mother, father, spouse or household head as a new variable (for example, education of mother).

Describe your extract

Educational attainment for young adults (age 18-35) in the 2010-round censuses.

Review & submit your extract request:

- On the **Extract Request** menu, click **Submit extract**. You may customize sample size, select cases, or attach characteristics (see Tips & Tools on the back).
- To obtain an extract in SPSS, SAS, Stata or cvs format, click **Change** on the **Data Format** line, click the desired option, and click **Submit.**
- Describe your extract so you remember its contents later (in case you want to resubmit the job).
- When everything is as you like it, click **Submit extract.**

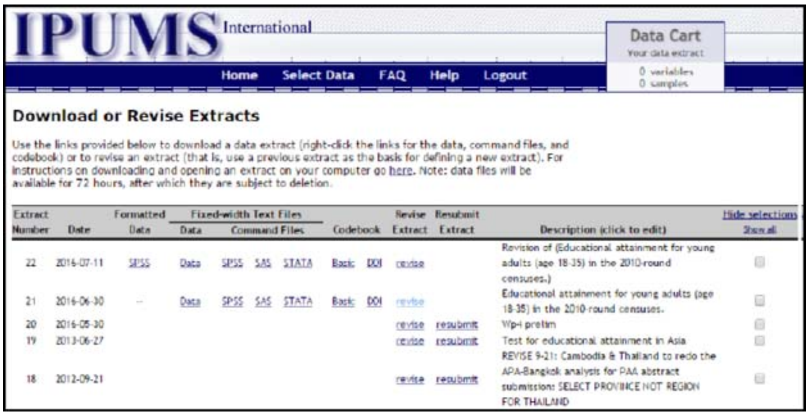
How to download & use your data

Retrieve your microdata extract for analysis:

You will receive an email at your registered-user address when your extract is ready. Click **Download data extract** or the link in the email notice. Enter your email and password. You will be taken to your private data carrel, which records all your extract requests (see Figure 4).

- Download the data extract, codebook and, if desired, a SAS, Stata, or SPSS syntax file.
- Unzip the data (which requires WINZIP, 7z, RAR, or similar software).
- Open the SPSS, SAS, Stata, or ASCII (.dat) file. If ASCII, construct the database with the syntax file provided.
- Edit the top of your command file to tell the program where you have stored the data.

Figure 4. Personal Data Carrel



Download or Revise Extracts

Use the links provided below to download a data extract (right-click the links for the data, command files, and codebook) or to revise an extract (that is, use a previous extract as the basis for defining a new extract). For instructions on downloading and opening an extract on your computer go here. Note: data files will be available for 72 hours, after which they are subject to deletion.

Extract		Formatted	Fixed-width Text Files					Revise	Resubmit		Hide selections
Number	Date	Data	Data	Command Files			Codebook	Extract	Extract	Description (click to edit)	Show all
22	2016-07-11	SPSS	Data	SPSS	SAS	STATA	Basic DDI	revise		Revision of (Educational attainment for young adults (age 18-35) in the 2010-round censuses.)	□
21	2016-06-30	--	Data	SPSS	SAS	STATA	Basic DDI	revise		Educational attainment for young adults (age 18-35) in the 2010-round censuses.	□
20	2016-05-30							revise	resubmit	Wp-i prelim	□
19	2013-06-27							revise	resubmit	Test for educational attainment in Asia	□
18	2012-09-21							revise	resubmit	REVISE 9-21: Cambodia & Thailand to redo the APA-Bangkok analysis for PAA abstract submission: SELECT PROVINCE NOT REGION FOR THAILAND	□

To resubmit or revise an extract:

From the main menu, click on **Download Your Data Extract,** which brings you to your virtual data carrel (Figure 4). To replicate your original extract, click on **resubmit**. To change your original extract request, click on **revise**, make your new variable/sample selections and click **Submit extract**.

Follow good data analysis practices:

- Use the weights (expansion factors) that are automatically included in every extract. Not all samples share the same design. To make correct inferences, use weights.
- Use proper statistical techniques. Take into account response biases and errors in census operations as well as sampling errors.
- Remember that IPUMS-I sample statistics may differ from official figures for many reasons (e.g., loss of portions of original data, omission of collective households).
- Honor all conditions of use. Protect statistical confidentiality, cite properly, report publications. (https://bibliography.ipums.org/user_submissions/new).

For more information:

- Click **help** on pages throughout the website and read the **FAQ**.
- Email ipums@umn.edu with questions, problems, or to report suspected violations of conditions of use.
- IPUMS-International project management:

Lara Cleveland	Project Manager	clevelan@umn.edu
Rodrigo Lovaton	Research Scientist	lovat003@umn.edu
Kristen Jeffers	Senior Data Analyst	kjeffers@umn.edu
Matt Sobek	Data Science Services Director	sobek@umn.edu

Tips & Tools

Unharmonized variables: These are variables for which codes ***have not*** been harmonized across countries or years. Therefore, variables with similar content, such as "sex," may have different codes and labels from one sample to the next.

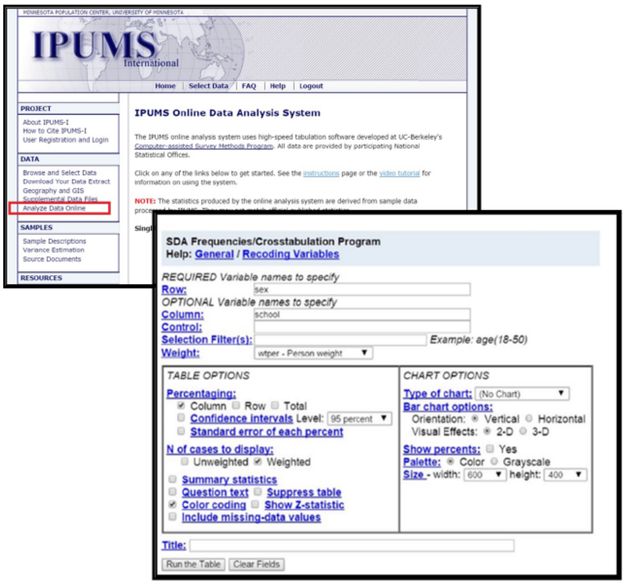
To reduce the size of an extract: If your extract size is too large, you have several options.

- Customize sample size. Change the percent or number of cases for each sample.
- Eliminate variables or samples**:** **revise** your extract, click on **Samples** and/or **Variables**, and make your selections. To deselect a sample or variable, simply click on the checked box.
- Restrict cases: Click on **Select cases**, choose a variable and select one or more values for that variable. For example, eliminate persons under 15 and over 68 if you are studying paid labor.

To analyze data online: If you want to make a table without downloading a data extract, you can use the online data tabulator (Figure 5.) This tool provides minimal information about the variables so be sure to use the web metadata to determine exactly what you want before running a table.

In addition to making tables, this tool also can generate additional statistical output. Be sure to click on the **weighted** box under **N of cases to display** since all data in IPUMS-International are sample data.

Figure 5. Online Data Tabulator



IPUMS Online Data Analysis System

The IPUMS online analysis system uses high-speed tabulation software developed at UC-Berkeley's Computer-assisted Survey Methods Program. All data are provided by participating National Statistical Offices.

Click on any of the links below to get started. See the instructions page or the video tutorial for information on using the system.

NOTE: The statistics produced by the online analysis system are derived from sample data proce

Singl

SDA Frequencies/Crosstabulation Program
Help: General / Recoding Variables

REQUIRED Variable names to specify
Row: sex
OPTIONAL Variable names to specify
Column: school
Control:
Selection Filter(s): Example: age(18-50)
Weight: wtper - Person weight ▼

TABLE OPTIONS
Percentaging:
☑ Column □ Row □ Total
□ Confidence intervals Level: 95 percent ▼
□ Standard error of each percent
N of cases to display:
□ Unweighted ☑ Weighted
□ Summary statistics
□ Question text □ Suppress table
☑ Color coding □ Show Z-statistic
□ Include missing-data values

CHART OPTIONS
Type of chart: (No Chart) ▼
Bar chart options:
Orientation: ● Vertical ○ Horizontal
Visual Effects: ● 2-D ○ 3-D
Show percents: □ Yes
Palette: ● Color ○ Grayscale
Size - width: 600 ▼ height: 400 ▼

Title:

Run the Table | Clear Fields

To attach characteristics: For each variable in your extract, you can create new variables by household **Head**, **Father**, **Mother**, or **Spouse**. The extract system then attaches the value for that person (e.g., spouse's age, mother's years of schooling) to each individual record in the household. Note: Attached characteristics are assigned only when the relevant reference person (spouse, parent) is present within the household. Otherwise, the new variable will have missing values.

More features: Access these additional tools and datasets from the left menu of the home page.

- **Variance Estimation.** Documentation on the IPUMS-I stratified sample design.
- **Geography and GIS.** Downloadable boundary files for each country, as well as additional information on geography variables.
- **Supplemental Data Files.** Available for a subset of samples, contain downloadable data on mortality, migration and/or fertility events for households for a period preceding a census.
- **Bibliography.** Citations of publications by researchers (Project: "IPUMS-International").

IPUMS-International is an initiative of the MPC-Minnesota Population Center, with funding from the University of Minnesota, the NSF-National Science Foundation and the NIH-National Institutes of Health.

Source data for IPUMS-International are generously provided by participating National Statistical Offices.
When your project is complete, cite both national statistical office(s), which provided the data, as well as IPUMS-International, which harmonized and disseminated the data. Proper citation language is provided on the home page at **How to Cite IPUMS-I**.

July 2016