

# IPUMS

## International

A Collaboration of the Minnesota Population Center  
and National Statistical Agency Partners

### Goals:

- Preserve census microdata and documentation from the 1960s to the present
- Integrate anonymized, high-precision samples and corresponding documentation
- Disseminate census microdata and metadata to researchers world-wide, without cost

### Accomplishments since 1999

- 85 National Statistical Offices endorsed the IPUMSi memorandum of understanding
- 270 sets of census microdata entrusted
- 158 samples (55 countries, 325 million persons) integrated and available to researchers
- Many datasets recovered from old tapes, primitive disc drives and even floppy discs!
- Over 3,300 registered researchers
- Many conference papers, articles, theses and books already in preparation

### How IPUMSi works

- A vast body of census microdata and documentation is being assembled at the University of Minnesota Population Center, thanks to a global collaboratory of National Statistical Offices, international organizations, universities and researchers.
- Accredited researchers obtain free access to all documents and anonymized samples.
- Microdata and documents are integrated to facilitate comparative research.
- Each user extracts a selection of microdata, then downloads and analyzes it using his/her own software and computer resources.
- Researchers provide copies of publications to IPUMSi for transmittal to National Statistical Office partners.

#### More information inside:

- List of National Statistical Agency partners

*Please make good use of the IPUMS system so National Statistical Offices will continue to entrust census microdata for the IPUMS-International database!*

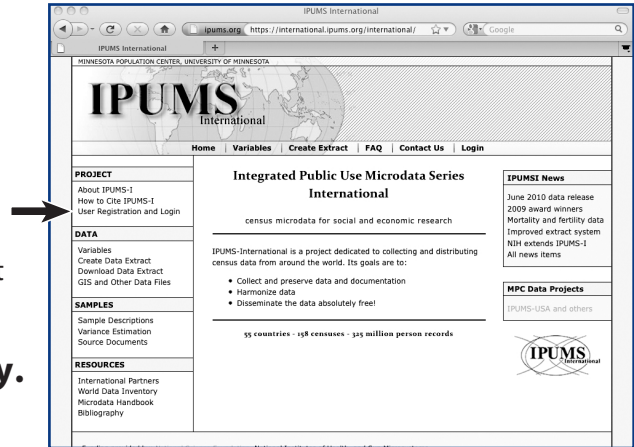
- 10 Tips for using IPUMS-International
- How to use the IPUMSi dynamic metadata system
- How to create a data extract

# 10 Tips for Using IPUMS-International

(browse <https://international.ipums.org/international> while reading the following):

## Tip 1. Register to use the microdata.

- Click "User Registration and Login" on the IPUMSi home page; then click "Apply for Access," submit email address, and complete the form. Check each box to show that you will protect the data against unauthorized use.

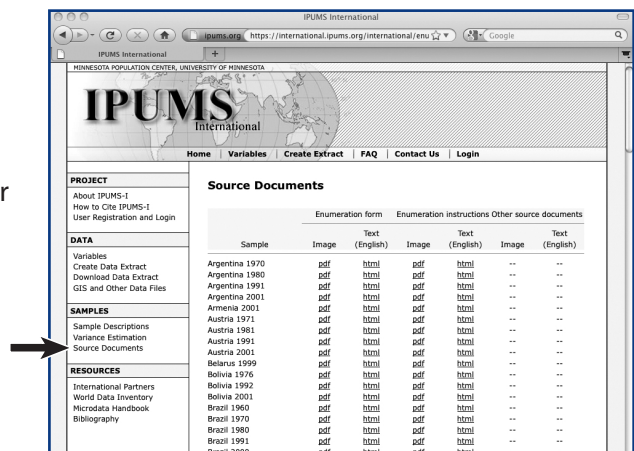


## Tip 2. Protect statistical confidentiality.

- Faithfully abide by each of the conditions of use as stated on the registration form.
  - Do not "share" IPUMSi extracts with others, unless they are duly registered users (such as co-researchers on a project.)
- Violations will lead to denial of access for the researcher and recall of all microdata. In cases of lax oversight, access may be denied to all users at the researcher's institution.
- Please report suspected violations to [ipumsi@pop.umn.edu](mailto:ipumsi@pop.umn.edu).

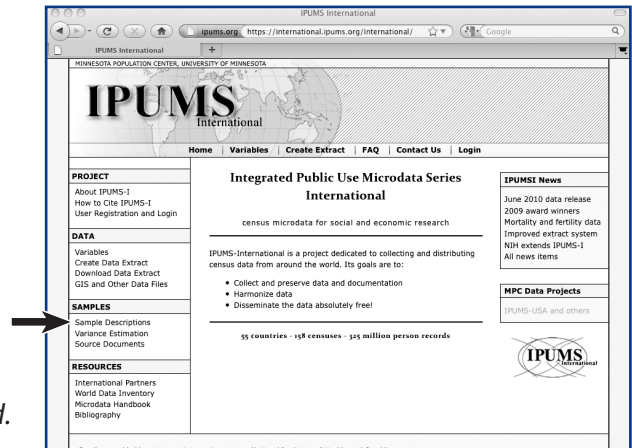
## Tip 3. Study the documentation thoroughly.

- Comprehensive documentation is provided in the official language:
  - Click "Census Questionnaires" on the IPUMSi home page then the image ("pdf") for census forms and instructions for the desired country and census year.
- IPUMSi metadata describe, for each country:
  - Censuses - title, agency, population enumerated, universe, type, unit, field work period, forms used, etc.
  - Samples - source, design, sample unit, fraction, size, weights, etc.
  - Variables - availability, universe, codes, enumeration text, unharmonized codes, description, comparability, etc.
- English texts and translations are available for all documents:
  - Click "html" on the "Census Questionnaires" page.
  - Click "enumeration text" on any variable description page to invoke the dynamic metadata system (see "Features" below).



#### Tip 4. Construct extracts judiciously.

- Request only microdata samples needed for a specific research project.
- Some samples contain 10 million or more person records. If 100,000 cases are sufficient, request that number (see “customize sample size” on the next page).
  - » *Prior extracts may be reviewed and revised.* (See “Download Data Extract” on the IPUMSi home page to see a menu of your previous extracts.)



#### Tip 5. Use weights (expansion factors) where necessary.

- » *Not all samples have the same design. To avoid incorrect inferences, use weights.*

#### Tip 6. Analyze carefully.

- Use proper statistical techniques.
  - » *Take into account bias and errors in census operations as well as sampling errors.*
- IPUMSi sample statistics may differ from official figures for many reasons (due to loss of original data, omission of certain classes of individuals such as residents of collective households, etc.).

#### Tip 7. Cite properly.

- Cite both IPUMS-International and the respective National Statistical Agency(ies) (see list on the back page). Links to all statistical agencies are available at “Citation of IPUMS”.

*Minnesota Population Center. Integrated Public Use Microdata Series — International: Version 4.0. Minneapolis: University of Minnesota, 2008.  
 Researchers should also acknowledge the statistical agencies that originally produced the data, listed here.*

#### Tip 8. Share publications

- Post citations on the IPUMi home page (click “Bibliography”).
- Email copies to [ipums@pop.umn.edu](mailto:ipums@pop.umn.edu) or forward to the respective National Statistical Office.

#### Tip 9. Click “FAQ” on the IPUMSi home page to read replies to “Frequently Asked Questions”.

#### Tip 10. Email [ipums@pop.umn.edu](mailto:ipums@pop.umn.edu) for help with questions, problems or errors.

# How to “Create an Extract” in Six Steps

Registered users obtain data by clicking “Create Data Extract” and completing these 6 steps.

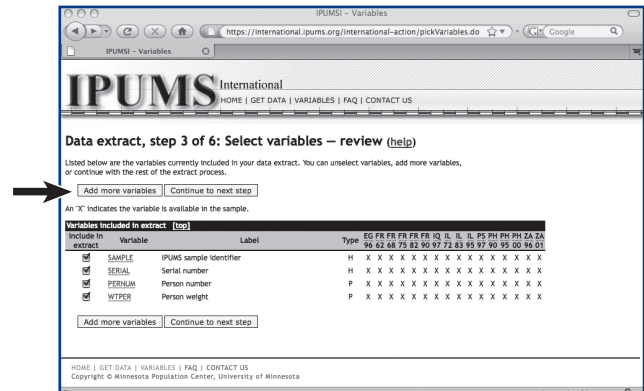
## Step 1. Login: Enter email address and password.

## Step 2. Select samples, countries and census years:

- Click as many boxes on the sample selection menu as needed.
- Click “continue”.

## Step 3. Select variables:

- Click “add more variables”.
- Roll mouse over “household”, “person”, or “A-Z” button, and click a selection.
  - » *Click the box for each variable of interest.*
- Repeat a and b until all variables of interest are selected.
- To see all variables in a single menu, roll mouse over “Options and Help” and click “view one group”.
- If unharmonized variables are of interest, click “switch to unharmonized variables”.
- A unique set of variables will be offered for each census (e.g., seemingly similar variables, “age”, “sex” are, nevertheless, designated with unique variable names).
- Once all variables of interest are selected, review all by clicking “Review variable selection”.
- Click “add more variables” (see “a” above) or “continue to next step”.



## Step 4. “Variable options”: Select cases for inclusion/omission and attach characteristics.

This step is optional.

- Click “select cases” for any variable to be used to include/exclude cases in the extract (e.g., click “sex” to be able to select only females for inclusion in the extract).
- Click “attach characteristics” of household head, mother, father and/or spouse for any selected variable.
- Click “continue to next step”.
- Attach characteristics of household head, mother, father and/or spouse; then click “continue to next step”.
- “Case selection”:
  - Choose the conditions to be included in the extract (e.g., click “female” to include only females in the extract).
  - To include in the extract all co-resident persons of the selected case, click “include all persons in the household”.
  - Click “continue to next step”.

## Step 5. Customize sample sizes.

**Tip: Ask yourself the question: “Is the projected extract file size (in megabytes) too big for what I want to do?” If not, skip this section.”**

- a. Click “continue to next step” (and skip this section), if the projected size of the dataset is OK.
- b. Otherwise, click “customize sample sizes”, if you wish to reduce the size of the samples.
  - » *To make all samples the same size, enter a number in the first row. If the requested size is larger than the maximum, the maximum is automatically entered.*
  - b-1. Enter a number in any of the boxes—either for the number of cases or % density of each sample.
  - b-2. To estimate the projected extract file size, click “recalculate”.
  - b-3. To restore the maximum projected size, click “reset selections”.
  - b-4. When finished, click “continue to next step”.

## Step 6. Submit selections to construct the extract. Before submitting the selections, review:

- a. Review the selections of samples, variables, options, attached characteristics, case selection, and sample sizes.
- b. At this stage, as in all previous stages, click any highlighted variable name to recall definitions, descriptions, etc.
  - » *If any changes are desired, click the appropriate item:*
    - “Samples”: to change the sample selections.
    - “Variables”: to change the variable selections.
    - “Variable options”: to remove/add case selections or attached characteristic variables.
    - “Attached characteristics”: to change selections of head, spouse, mother, father
    - “Sample size”: to change customized sizes.
- c. In the empty text box, type a brief description or name for this extract, as a reminder to yourself.

### The extract will be registered with IPUMS.

- d. If you need to recall an extract to add samples, variables, etc., you may do so by clicking “Download Data Extract” on the home page. A list of your extracts will appear, including data and description (after login with email and password)
  - » *Thus, as the 2010 round samples are added to IPUMS, it will be easy to update your own integrated database by revising a previous extract.*
- a. Finally, click “Submit extract” to complete the request. When the extract is ready, an email will be sent to the registered address.
- b. Click the link in the email, enter your email and password and download the data extract, codebook, and SAS, STATA, or SPSS dictionary.
- c. Unzip the data and construct the database with the corresponding dictionary.
  - » *For SPSS, enter in the command file (xxx.sps) the folder information for the exact location of the downloaded data file.*

# Features of the IPUMS-International website

## A. IPUMS-International dynamic metadata system (open access; no registration required)

For a tutorial, go to [www.hist.umn.edu/~rmccaa/IPUMSI](http://www.hist.umn.edu/~rmccaa/IPUMSI) then click "Dynamic Metadata".

While browsing the IPUMS-International website, click any box (similar to ) to select that item (sample or variable) to include in an extract. Click any checked box to de-select.

1. Click "samples" for descriptions of the samples currently available, from Argentina to Vietnam, including microdata for Cambodia, Canada, Chile, China, Colombia, Costa Rica and Cuba—to name only the "C's".
2. From the IPUMS-International home page, click "variables" for descriptions of available variables.
3. From the "variables menu" page:
  - Click "select samples" to choose samples for the desired countries and census years.
  - Click "person", "dwelling", or the alphabetic list to see a listing of variables.
  - If unharmonized variables are of interest, toggle "switch to unharmonized variables".

**Note: Unharmonized variables are unique to a single census and are similar to the original source variables. They are useful for checking integrated variables. Note, too, that variables that are exclusive to a single census are not integrated, and consequently are available only in non-harmonized form.**

4. From a list on the variables page (e.g., "Core Demographic"), click any variable name to view the metadata.
5. From this or any variable description page:
  - Click "codes," then "case-count view" for counts (number of person or households) in each sample
  - Click "enumeration text" to compare the wording and instructions to enumerators for the corresponding variable and selected samples. **Compare questions for any combination of countries and censuses in the database. Use this tool to ascertain whether integrated variables and codes are appropriate for your research needs.**

## B. Register for access to the microdata (click "User Registration and Login.")

1. With IPUMSi, the researcher downloads the microdata as "extracts".
2. IPUMSi extracts are analyzed using statistical software and computer resources familiar to the researcher. (SAS, spss, STATA command files are provided for basic setup.)
3. The advantage of extracts is that the researcher selects the desired countries, census years, sample sizes, sub-populations, and variables.
4. To access data, first click "apply for access", then complete the registration form, briefly describe the research plan, and agree to the terms and conditions (protect statistical confidentiality of the microdata, safeguard the microdata, provide a copy of publications, etc.). Read each condition of use and click each box to indicate your agreement to abide by each restriction. The registration form is quite detailed to restrict access to bona-fide researchers. Approval is usually forthcoming in 1-7 days.

## C. Create an extract (restricted to bona-fide, registered researchers).

## Why IPUMSi datasets are not distributed on DVDs or by FTP:

- The database is expanding rapidly, with the inclusion of data and documents for more countries, censuses, and variables.
- The database is already too huge to ship on a CD or DVD like a book, video or other product obtained from the web.
- The dynamic metadata system helps navigate both integrated and original source documentation and cannot easily be replicated on other media.
- New samples, variables, and enhancements are added frequently—the database is dynamic.
- Note: Where internet access is limited National Statistical Office partners are provided—upon request—a DVD containing an integrated set of samples, variables and documentation for their country to be disseminated to national researchers.

## More is coming to IPUMSi! Plans for 2010 - and beyond - include:

- Expand archiving ...
  - » *backward to encompass all published census documents in addition to the unpublished*
  - » *forward to include scanned images and other forms of completed questionnaires*
  - » *outward to provide easy access to census microdata and documentation, world-wide*
- Integrate 2010 round censuses on a top priority when the microdata become available
- Continue integration of samples for participating countries (see list of International Partners)
- Invite other National Statistical Offices and organizations to participate in IPUMS-International
- Increase access to more researchers
- Enhance documentation with modules on sampling errors, official publications, etc.
- GIS features to complement microdata samples (preliminary version available now)
- A tabulator of integrated samples (to manage extract requests)
- Expansion of the Minnesota Population Center's **Secure Data Laboratory** for access to microdata that cannot be disseminated over the internet

User support is available at [ipums@pop.umn.edu](mailto:ipums@pop.umn.edu).  
For additional questions or comments,  
please email [rmccaa@umn.edu](mailto:rmccaa@umn.edu).

<https://www.international.ipums.org/international>

## International Partners

**Argentina** - National Institute of Statistics & Censuses  
**Armenia** - National Statistical Service  
**Austria** - Statistics Austria  
**Bangladesh** - Bureau of Statistics  
**Belarus** - Ministry of Statistics & Analysis  
**Bolivia** - National Institute of Statistics  
**Botswana** - Central Statistics Office  
**Brazil** - Institute of Geography & Statistics  
**Bulgaria** - National Statistics Institute  
**Cambodia** - National Institute of Statistics  
**Canada** - Statistics Canada  
**Chile** - National Institute of Statistics  
**China** - National Bureau of Statistics  
**Colombia** - National Administrative Department of Statistics  
**Costa Rica** - National Institute of Statistics & Censuses  
**Cuba** - National Statistics Office  
**Czech Republic** - Czech Statistical Office  
**Dominican Republic** - National Statistics Office  
**Ecuador** - National Institute of Statistics & Censuses  
**Egypt** - Central Agency for Public Mobilization & Statistics  
**El Salvador** - General Directorate of Statistics & Censuses  
**Ethiopia** - Central Statistical Agency  
**Fiji Islands** - Bureau of Statistics  
**France** - National Institute of Statistics & Economic Studies  
**Germany** - Federal Statistical Office  
**Ghana** - Ghana Statistical Services  
**Greece** - National Statistical Office  
**Guatemala** - National Institute of Statistics  
**Guinea (Conakry)** - National Statistics Directorate  
**Guinea-Bissau** - National Statistics Institute  
**Haiti** - Institute of Statistics & Informatics  
**Honduras** - National Institute of Statistics  
**Hungary** - Central Statistical Office  
**India** - Ministry of Statistics & Programme Implementation  
**Indonesia** - Statistics Indonesia  
**Iraq** - Central Organization for Statistics & Information Technology  
**Ireland** - Central Statistical Office  
**Israel** - Central Bureau of Statistics  
**Italy** - National Institute of Statistics  
**Jordan** - Department of Statistics  
**Kenya** - National Bureau of Statistics  
**Kyrgyz Republic** - National Statistical Committee  
**Lesotho** - Bureau of Statistics  
**Madagascar** - National Institute of Statistics  
**Malawi** - National Statistical Office  
**Malaysia** - Department of Statistics  
**Mali** - National Directorate of Statistics & Informatics

<https://international.ipums.org/international>

Minnesota Population Center  
 50 Willey Hall  
 225 - 19th Avenue South  
 Minneapolis, MN 55455  
 612-624-5818

Major funding for IPUMS-International is provided by the  
 National Science Foundation and the National Institutes of Health.  
 The University of Minnesota is an equal opportunity educator and employer.  
 ©2008 - 2009 Regents of the University of Minnesota. All rights reserved.

**Mauritius** - Central Statistical Office  
**Mexico** - National Institute of Statistics & Geography  
**Mongolia** - National Statistical Office  
**Morocco** - High Commission of Planning  
**Mozambique** - National Institute of Statistics  
**Nepal** - Central Bureau of Statistics  
**Netherlands** - Statistics Netherlands  
**Nicaragua** - National Institute of Information for Development  
**Niger** - National Statistics Institute  
**Pakistan** - Statistics Division  
**Palestine** - Central Bureau of Statistics  
**Panama** - Census & Statistics Directorate  
**Paraguay** - General Directorate of Statistics, Surveys, & Censuses  
**Peru** - National Institute of Statistics & Informatics  
**Philippines** - National Statistics Office  
**Portugal** - National Institute of Statistics  
**Romania** - National Institute of Statistics  
**Rwanda** - National Institute of Statistics  
**Saint Lucia** - Government Statistics Department  
**Sénégal** - Direction de la Prévision et de la Statistique  
**Sierra Leone** - Statistics Sierra Leone  
**Slovenia** - Statistical Office of the Republic of Slovenia  
**South Africa** - Statistics South Africa  
**Spain** - National Institute of Statistics  
**Sudan** - Central Bureau of Statistics  
**Switzerland** - Federal Statistical Office  
**Tanzania** - National Bureau of Statistics  
**Thailand** - National Statistical Office  
**Turkey** - Turkish Statistical Institute  
**Turkmenistan** - National Institute of Statistics & Information  
**Uganda** - Bureau of Statistics  
**Ukraine** - State Committee of Statistics of Ukraine  
**United Kingdom** - Office of National Statistics  
**United States** - Bureau of the Census  
**Uruguay** - National Institute of Statistics  
**Venezuela** - National Institute of Statistics  
**Vietnam** - General Statistics Office  
**Zambia** - Central Statistical Office

### International Organizations:

- African Centre for Statistics (ACS), Addis Ababa
- Arab Institute for Training & Research in Statistics (AITRS), Amman
- Centre d'Estudis Demogràfics (CED), Barcelona
- Centre Population et Développement (CEPED), Paris
- DataFirst Archive & Center for Actuarial Research, University of Cape Town
- Demographic Center of Latin America & the Caribbean (CELADE), Santiago, Chile
- DRS Data Services Ltd., Milton Keynes, UK
- Institut National d'Études Démographiques (INED), Paris
- Inter-American Statistical Institute (IASI), Panama City & Buenos Aires
- International Household Survey Network (IHSN), World Bank, Washington, DC
- Organization for Economic Co-Operation and Development (OECD), Paris