



SECOND MEETING
ARAB TASK FORCE ON POPULATION AND
HOUSING CENSUSES
2020 ROUND
Cairo, 22-23 January 2017

IPUMS International

An Overview

By Ragui Assaad
Minnesota Population Center
University of Minnesota

IPUMS?

Integrated - consistent codes, labels, and documentation

Public Use - free, anonymized, downloadable

Microdata - individual-level

Series - pooled data over time and place



IPUMS

International

[Home](#) | [Select Data](#) | [FAQ](#) | [Help](#) | [Login](#)

PROJECT

About IPUMS-I
How to Cite IPUMS-I
User Registration and Login

DATA

Browse and Select Data
Download Your Data Extract
Geography and GIS
Supplemental Data Files
Analyze Data Online

SAMPLES

Sample Descriptions
Source Documents
Variance Estimation

RESOURCES

Help
User Forum
Bibliography
International Partners
Teaching with IPUMS

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!

82 countries - 277 censuses - 614 million person records

Source data for IPUMS-International are generously provided by participating National Statistical Offices

IPUMSI News

NEW! 19 samples added
NSF extends IPUMS project thru 2019
Geography variables revised
Online tabulator
Class accounts
[Latin America project extended](#)
Integrated DHS project funded
... All news items

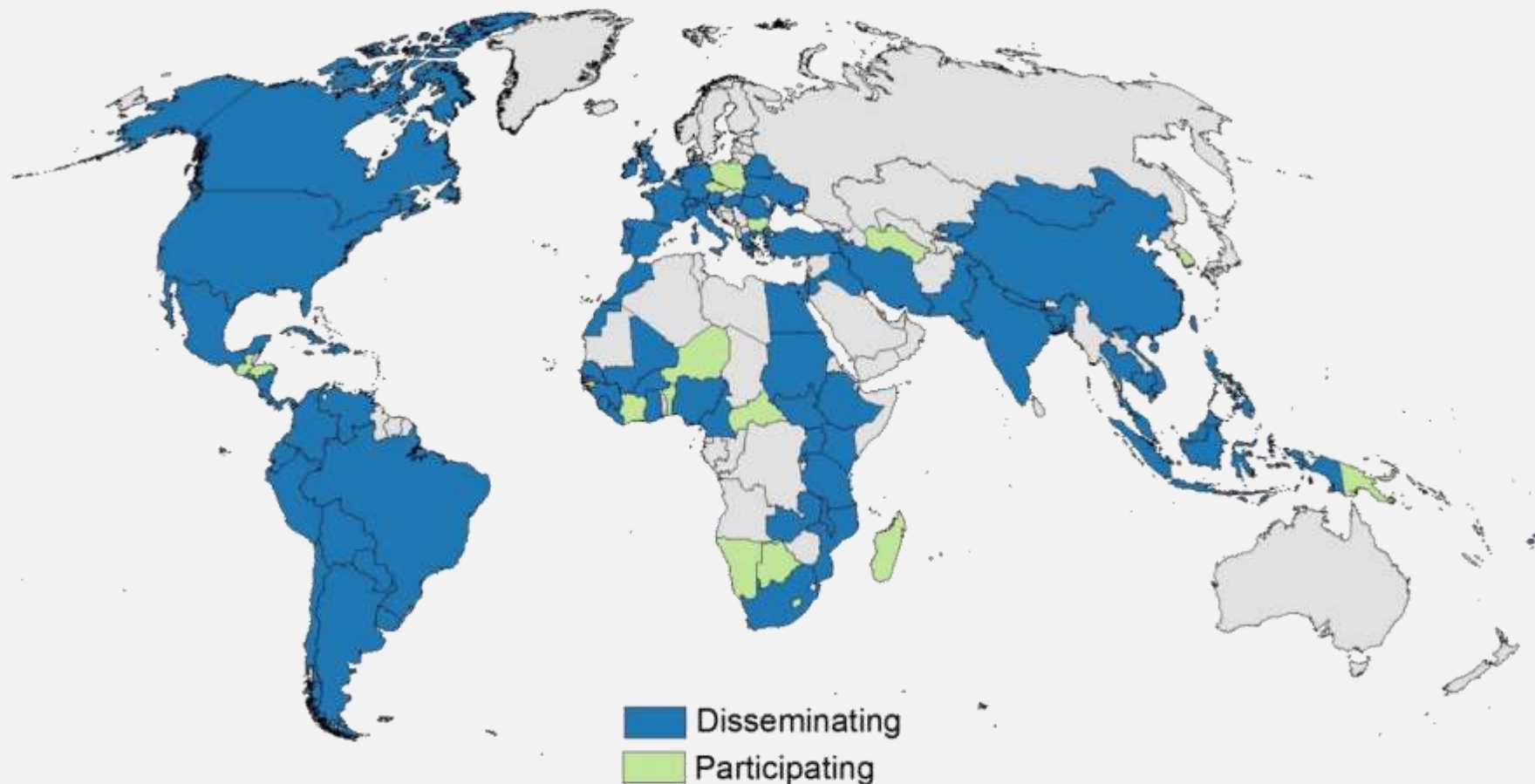
MPC Data Projects

[IPUMS-USA and others](#)



Sign up
to receive News of interest to IPUMS-International data users and producers.

Microdata Availability



82 countries – 277 censuses – 614 million person records

IPUMS Samples per Country

Argentina	5	Fiji	5	Malaysia	4	Senegal	2
Armenia	2	France	8	Mali	2	Sierra Leone	1
Austria	5	Germany	4	Mexico	7	Slovenia	1
Bangladesh	3	Ghana	3	Mongolia	2	South Africa	4
Belarus	1	Greece	4	Morocco	3	South Sudan	1
Bolivia	3	Guinea	2	Mozambique	2	Spain	4
Brazil	6	Haiti	3	Nepal	1	Sudan	1
Burkina Faso	3	Hungary	4	Netherlands	3	Switzerland	4
Cambodia	2	India	5	Nicaragua	3	Tanzania	2
Cameroon	3	Indonesia	9	Nigeria	5	Thailand	4
Canada	4	Iran	1	Pakistan	3	Turkey	3
Chile	5	Iraq	1	Palestine	2	Uganda	2
China	2	Ireland	9	Panama	6	Ukraine	1
Colombia	4	Israel	1	Paraguay	5	UK	2
Costa Rica	5	Italy	1	Peru	2	USA	7
Cuba	1	Jamaica	3	Philippines	3	Uruguay	6
Dominican Republic	5	Jordan	1	Portugal	4	Venezuela	4
Ecuador	6	Kenya	5	Puerto Rico	6	Vietnam	3
Egypt	2	Kyrgyz Republic	2	Romania	4	Zambia	3
El Salvador	2	Liberia	2	Rwanda	2		
Ethiopia	3	Malawi	3	Saint Lucia	2		

Primary Goals

1. Preserve and Archive
2. Confidentialize
3. Harmonize
4. Document
5. Disseminate



**Data collection and
preservation**

Khartoum, CBS-Sudan

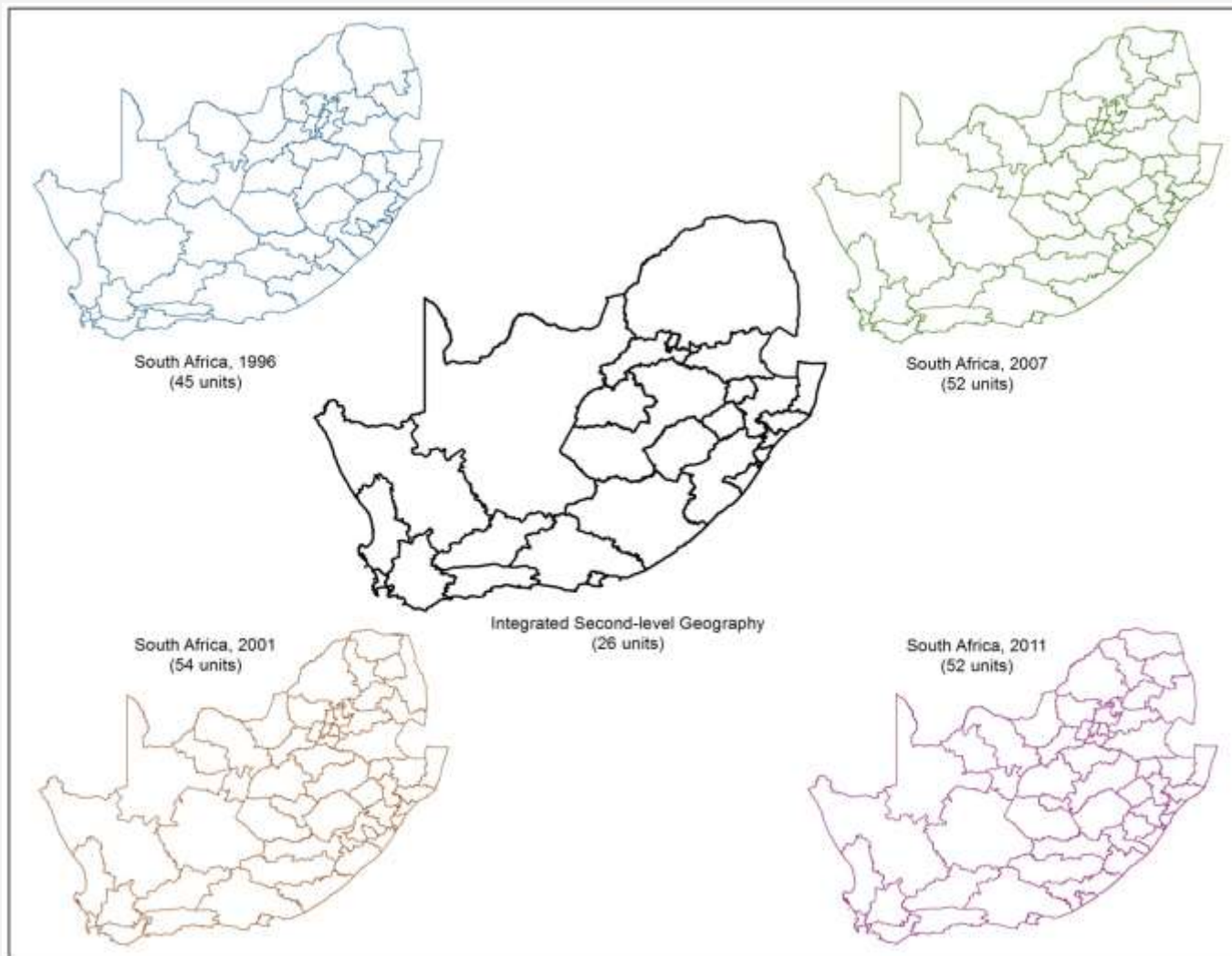
Confidentiality Measures

- Suppress low-level geographic variables.
- Recode geographic units to ensure small localities cannot be identified (usually $< 20,000$ persons).
- Swap a small percentage of cases between geographic areas.
- For recent censuses:
 - Recode cells representing very small numbers of persons in the population.
 - Top- or bottom-code continuous variables.
- Suppress categories or entire variables as requested by the NSO.

Data Enhancements

- Harmonized variables
 - Consistent codes and labels across samples
- Integrated samples
- Harmonized and year-specific geography
- Constructed family interrelationship variables
- Comprehensive documentation

Geographic harmonization




Family Interrelationship Variables

Pernum	Relate	Age	Sex	Marst	Chborn	Spouse's Location
1	head	46	male	married	n/a	2
2	spouse	44	female	married	3	1
3	aunt	77	female	widow	7	0
4	child	15	female	single	0	0
5	child	13	female	single	n/a	0
6	child	11	male	single	n/a	0

Pernum	Relate	Age	Sex	Marst	Chborn	Mother's Location	Father's Location
1	head	46	male	married	n/a	0	0
2	spouse	44	female	married	3	0	0
3	aunt	77	female	widow	7	0	0
4	child	15	female	single	0	2	1
5	child	13	female	single	n/a	2	1
6	child	11	male	single	n/a	2	1

Census Questionnaire Image (Mexico 2000)

I. Características de la vivienda

1. PAREDES	2. TECHOS	3. PISOS
<p>¿De qué material es la mayor parte de las paredes o muros de esta vivienda?</p> <p>CIRCULE UN SOLO CÓDIGO</p> <p>Material de desecho 1</p> <p>Lámina de cartón 2</p> <p>Lámina de asbesto o metálica 3</p> <p>Carrizo, bambú o palma 4</p> <p>Embarro o bajareque 5</p> <p>Madera 6</p> <p>Adobe 7</p> <p>Tabique, ladrillo, block, piedra, cantera, cemento o concreto 8</p>	<p>¿De qué material es la mayor parte del techo de esta vivienda?</p> <p>CIRCULE UN SOLO CÓDIGO</p> <p>Material de desecho 1</p> <p>Lámina de cartón 2</p> <p>Lámina de asbesto o metálica 3</p> <p>Palma, tejamanil o madera 4</p> <p>Teja 5</p> <p>Losa de concreto, tabique, ladrillo o terrado con viguería 6</p>	<p>¿De qué material es la mayor parte del piso de esta vivienda?</p> <p>CIRCULE UN SOLO CÓDIGO</p> <p>Tierra 1</p> <p>Cemento o firme 2</p> <p>Madera, mosaico u otros recubrimientos 3</p>
<p>4. COCINA</p> <p>¿Esta vivienda tiene un cuarto para cocinar?</p> <p>CIRCULE UN SOLO CÓDIGO</p> <p>Sí 1</p> <p>No 2 </p> <p>En el cuarto donde cocinan, ¿también duermen?</p> <p>CIRCULE UN SOLO CÓDIGO</p> <p>Sí 3</p> <p>No 4</p>	<p>5. NÚMERO DE CUARTOS</p> <p>¿Cuántos cuartos se usan para dormir sin contar pasillos?</p> <p>ANOTE CON NÚMERO</p> <p>Sin contar pasillos ni baños, ¿cuántos cuartos tiene en total esta vivienda? Cuente la cocina.</p> <p>ANOTE CON NÚMERO</p>	<p>6. DISPONIBILIDAD DE AGUA</p> <p>¿En esta vivienda tienen:</p> <p>LEA LAS OPCIONES HASTA OBTENER UNA RESPUESTA AFIRMATIVA Y CIRCULE UN SOLO CÓDIGO</p> <p>agua entubada dentro de la vivienda? 1</p> <p>agua entubada fuera de la vivienda, pero dentro del terreno? 2</p> <p>agua entubada de llave pública (o hidrante)? 3</p> <p>agua entubada que acarrean de otra vivienda? .. 4</p> <p>agua de pipa? 5</p> <p>agua de un pozo, río, lago, arroyo u otra? 6</p> <p>PASA</p>

Text of Census Questionnaire (Mexico 2000)

5. Number of Rooms

How many rooms are used for sleeping without counting hallways?

_____ Write the number

Without counting the hallways or bathrooms how many total rooms are in this dwelling? Count the kitchen

_____ Write the number

6. Access to water

Read all of the options until you get an affirmative answer.

Circle only one answer

- 1 Running water inside the dwelling
- 2 Running water outside the dwelling but on the land
- 3 Running water from a public faucet or hydrant
- 4 Running water that is carried from another dwelling
- 5 Tanked in by truck
- 6 Water from a well, river, lake, stream or other

Answers 3, 4, 5, 6 continue with number 8

7. Water supply

How many days of the week is water available?

Circle only one answer

- 1 Daily
- 2 Every third day
- 3 Twice a week
- 4 Once a week
- 5 Occasionally

Documentation

The screenshot displays the IPUMS International website interface. At the top left is the IPUMS International logo. A navigation bar contains links for Home, Select Data, FAQ, Help, and Login. On the right, a 'Data Cart' widget shows 'Your data extract' with '0 variables' and '2 samples', and a 'VIEW CART' button. The main content area is titled 'WATSUP' and includes sub-headers 'Water supply' and 'Group: Utilities — HOUSEHOLD'. Below this is a tabbed menu with 'Questionnaire Text' selected. The content shows 'Questionnaire Text' for 'Mexico 2000' and 'Mexico 2010'. A specific entry for 'Mexico 2000 — source variable MX00A018 — Availability of water' is highlighted, with a 'top' link. Below this is a 'Questionnaire form' section with a 'view entire document' link and options for 'text' and 'image'. The questionnaire text includes the heading '6. Access to water', instructions to read all options and circle one answer, and a list of six options: 1 Running water inside the dwelling, 2 Running water outside the dwelling but on the land, 3 Running water from a public faucet or hydrant, 4 Running water that is carried from another dwelling, 5 Tanked in by truck, and 6 Water from a well, river, lake, stream or other. A note at the bottom states 'Answers 3, 4, 5, 6 continue with number 8'.

IPUMS International

Home Select Data FAQ Help Login

Data Cart
Your data extract
0 variables
2 samples
VIEW CART

WATSUP
Water supply
Group: Utilities — HOUSEHOLD

Codes Description Comparability Universe Availability Questionnaire Text Source Variables

Questionnaire Text

[Mexico 2000](#)
[Mexico 2010](#)

Mexico 2000 — source variable MX00A018 — Availability of water [top](#)

Questionnaire form view entire document: [text](#) [image](#)

6. Access to water

Read all of the options until you get an affirmative answer.
Circle only one answer.

- 1 Running water inside the dwelling
- 2 Running water outside the dwelling but on the land
- 3 Running water from a public faucet or hydrant
- 4 Running water that is carried from another dwelling
- 5 Tanked in by truck
- 6 Water from a well, river, lake, stream or other

Answers 3, 4, 5, 6 continue with number 8

Documentation

Data Cart

Your data extract

0 variables
2 samples

[VIEW CART](#)

WATSUP

Water supply

Group: [Utilities](#) — HOUSEHOLD

[Add to cart](#)

[Change samples](#)

[Codes](#)

[Description](#)

[Comparability](#)

[Universe](#)

[Availability](#)


[Questionnaire Text](#)

[Source Variables](#)

Description

WATSUP describes the physical means by which the housing unit receives its water. The primary distinction is whether or not the household had piped (running) water.

Documentation



Home Select Data FAQ Help Login

Data Cart
Your data extract
0 variables
2 samples
VIEW CART →

WATSUP

Water supply
Group: [Utilities](#) — HOUSEHOLD

Codes Description **Comparability** Universe Availability Questionnaire Text Source Variables

Comparability — Index

[GENERAL](#) [Mexico](#)

Comparability — General

The basic distinction, at the first digit, is whether or not the household had access to piped water. This distinction is made across all samples. There are inconsistencies, however, in the treatment of piped water that is accessed only by going outside of the dwelling. In some samples this would be reported as lack of piped water, but in others this information is included as subcategories of "piped water."

A complication arises in cases where the presence of piped water is derived from a larger question about the source of water for the dwelling. In censuses that framed the question in this way, piped water implicitly means water from a public network. Water from other sources, such as a well, could have been piped into the dwelling but not register as piped water in WATSUP.

In essence, "no piped water" can mean: 1) no access to piped water -- common in many developing countries; 2) no access to piped water from a public distribution network; or 3) no water piped into the dwelling -- common among developed countries.

Piped water outside the building or lot and water explicitly from a public tap are given distinct codes, but they are not consistently distinguished.

WATSUP does not indicate the ultimate source of the water: whether it was from a public distribution system, well, stream, etc. That information is often available in the unharmonized source variables for the various samples.

Home
Ownership

Relation
to Head

Age

Marital
Status

Occupation

```
0000019000110000100000011000100711221000001992221233450999
0000019000110000200000022100210552221017001042221233100999
0000019000110000300200133000300312110000001002440011000211
0000019000110000400200133000300291110000001992222111000762
0000020000110000100000011000100441221000001992440011000532
0000020000110000200000022100210402221028001032331111000532
0000020000110000300200133000300121110000001992221233300999
0000020000110000400200133000300082110000001992112000000999
0000021000110000100300011000100441110000001992331111000123
0000021000110000200300044410500431110000001992440011000744
```

↓ ↓ ↓ ↓ ↓

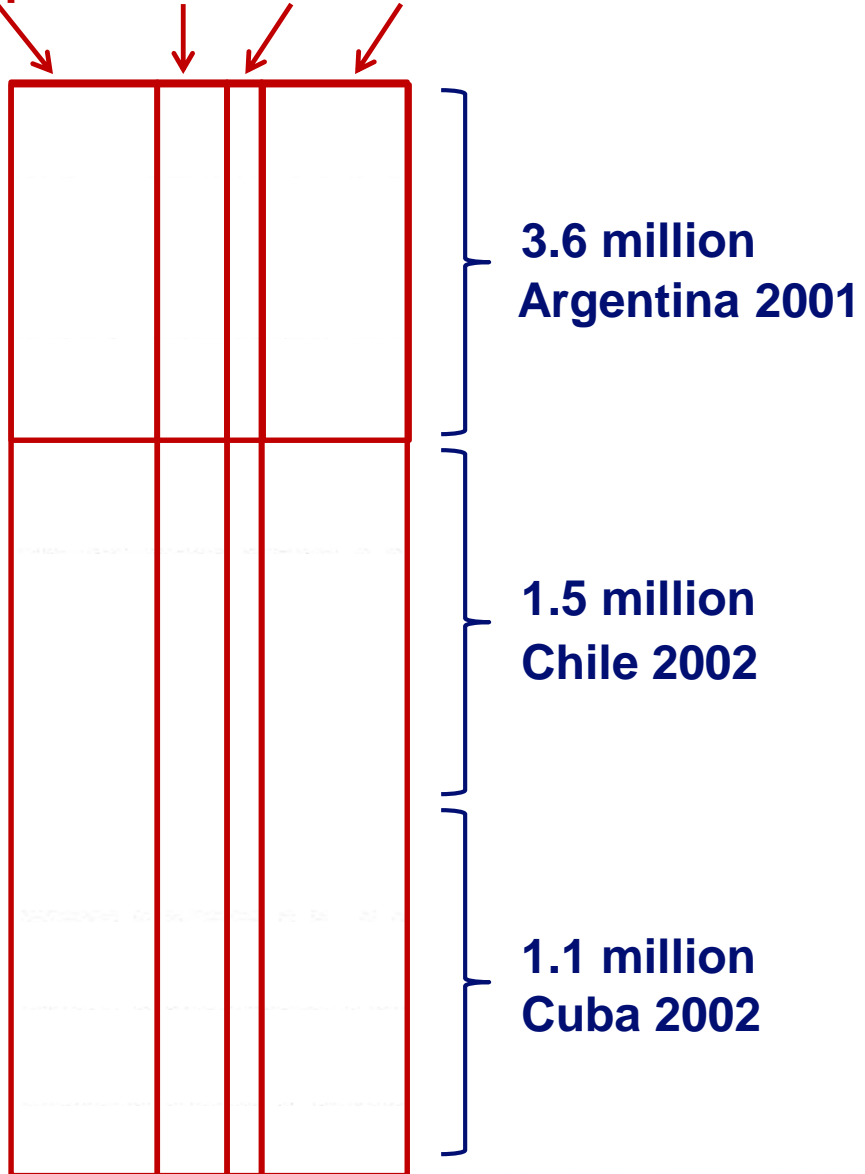
```
1001000071210999
1002100055210999
1003000031100211
1003000029100762
1001000044210532
1002100040210532
1003000012100999
1003000008100999
1001000044100123
1004410043100744
```

Customized Data Extract

Pooled Data Extracts

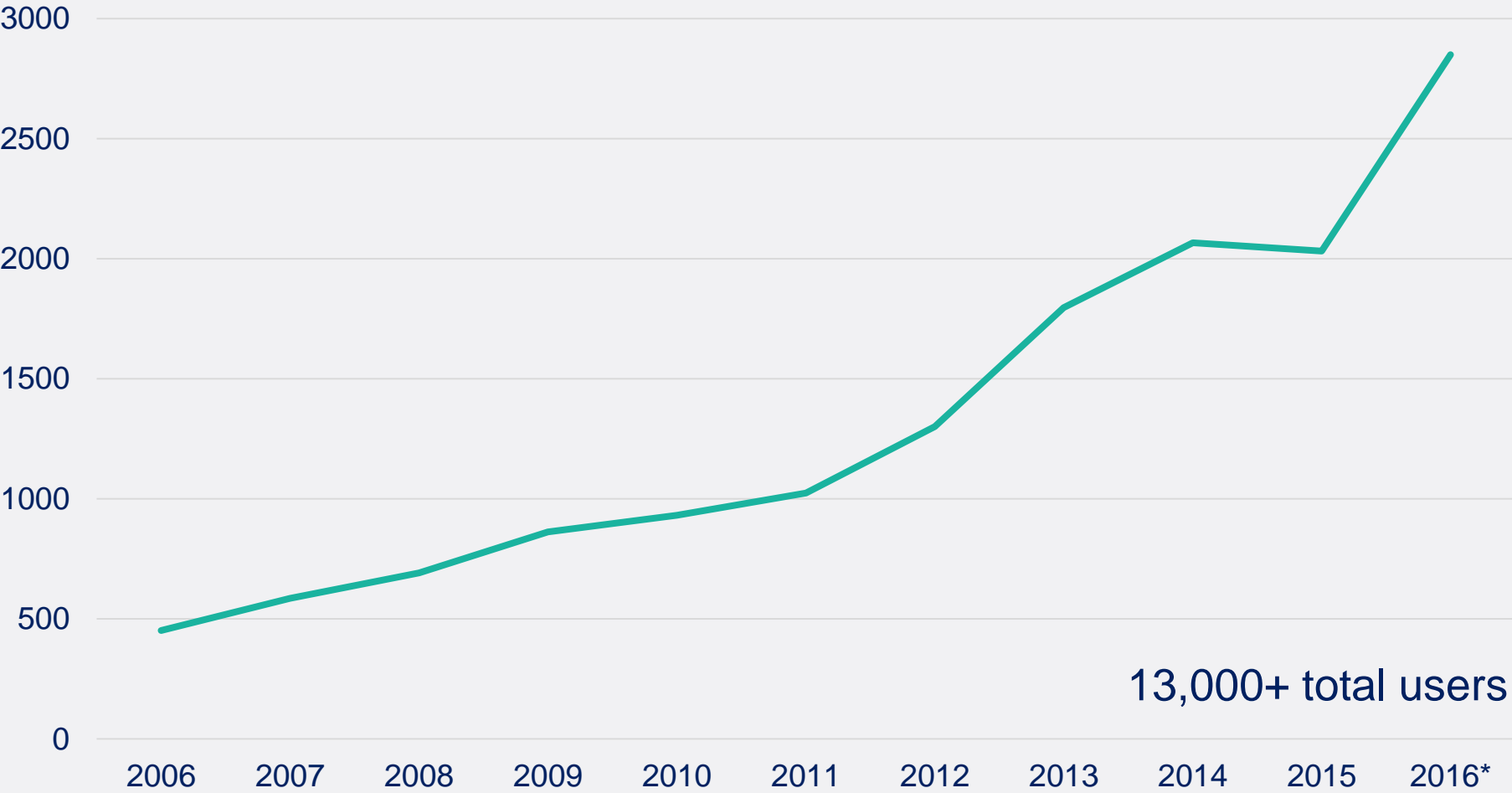
sample water sex education

1 dataset
3 censuses
4 variables
6.2 million records
Harmonized codes



User base & data use

New users by year



Users: institutions

North America

- University of Minnesota **(1)**
- University of California, Berkeley **(2)**
- Harvard University **(3)**
- University of Michigan **(4)**
- University of Chicago **(6)**
- Columbia University **(7)**
- Dartmouth College **(8)**
- University of California, Davis **(9)**
- Wellesley College **(10)**

Central and South America

- El Colegio de México
- Universidade Federal de Minas Gerais, Brazil
- Universidad de Valle, Colombia
- Pontificia Universidad Católica de Chile
- Universidad Nacional de la Plata, Argentina

Asia & Oceania

- National University of Singapore
- University of Malaya
- Hong Kong University of Science and Technology
- Xiamen University, China
- Seoul National University
- University of Queensland, Australia

Africa

- University of Witwatersrand, South Africa
- Obafemi Awolowo University, Nigeria
- University of Cape Town
- Makerere University, Uganda
- Cairo University
- University of Nairobi

Europe

- London School of Economics
- University of Oxford
- University of Cambridge
- Paris School of Economics
- European University Institute, Florence
- Universidad Carlos III de Madrid
- Universitat Autònoma de Barcelona

International Organizations

- The World Bank **(5)**
- United Nations
- Inter-American Development Bank
- World Health Organization

Users: discipline and status

Discipline		Status	
Economics	6000+	Student	50%
Demography	2000+	Faculty	15%
Sociology	1000+	Researcher	15%
Statistics	600+	Other academic	5%
Public Policy	600+	Non-academic	5%
Geography	400+		
History	150+		
Non-academic	100+		

Commonly-requested variables

Top 10

Educational attainment

Marital status

Employment status

Relationship to head

Age

Sex

Class of worker

Years of schooling

Ownership of dwelling

Urban-rural status

11 to 20

School attendance

Occupation, ISCO

Literacy

Industry, general recode

Country of birth

Nativity status

Children ever born

Occupation, unrecoded

Industry, unrecoded

Religion

21 to 30

Water supply

Children surviving

Electricity

Sewage

Migration status, 5 years

Earned income

Total income

Spouses location in HH

Number of rooms

Toilet

Extracts: objective

Anticipated result

Research article	35%
Thesis	25%
Class assignment	20%
Policy report	5%
Teaching materials	3%
Book	2%

Average extract

3 countries
6 samples
35 variables



Thank You!

ipums.org
ipums@umn.edu