IPUMS Global Health

Free IPUMS Survey and Census Data for International Health Research

Prepared for a 2024 CUGH pre-conference satellite session
March 7, 2023
Presented by Anna Bolgrien, PhD & Mehr Munir, MA

Workshop Agenda

- Introduction to IPUMS
  - IPUMS DHS
  - IPUMS MICS
  - IPUMS PMA
  - IPUMS International (census data)
- Additional IPUMS Global Health features
- Web demonstration
- Consultations & hands-on exercises

With Support From

NIH (NIA | NICHD | NIGMS | NIEHS)

[R01 HD099182], [R01 HD046697], [R01 AG062601], [R01 AG057679], [R01 HD043392], [R01 HD057929], [R01 AG041831], [R01 HD053654], [R01 HD069471], [R01 GM123642], [R01 HD067258], [R01 HD047283], [R01 HD083829], [R01ES034788]

NSF

[2020002], [1852842], [1825768]

Gates Foundation

Please properly cite IPUMS MICS if using the documentation or downloading our syntax!

INTRODUCTION TO IPUMS

What is IPUMS?

- US & international census & survey data
- To improve data accessibility
- Integrated documentation and standardized data to streamline analysis
- Available free of charge

Microdata

Summary data

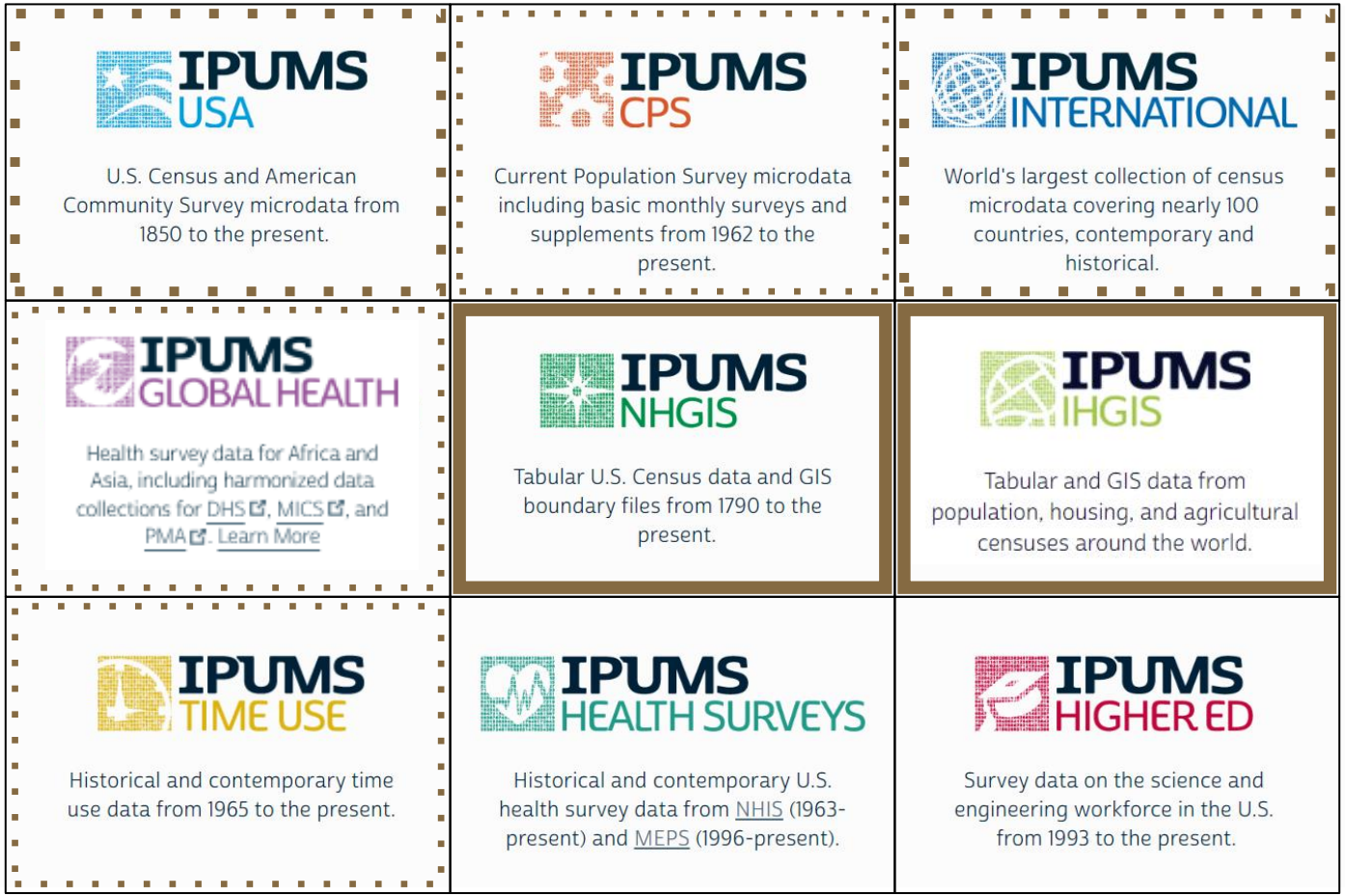
U.S. data

International data

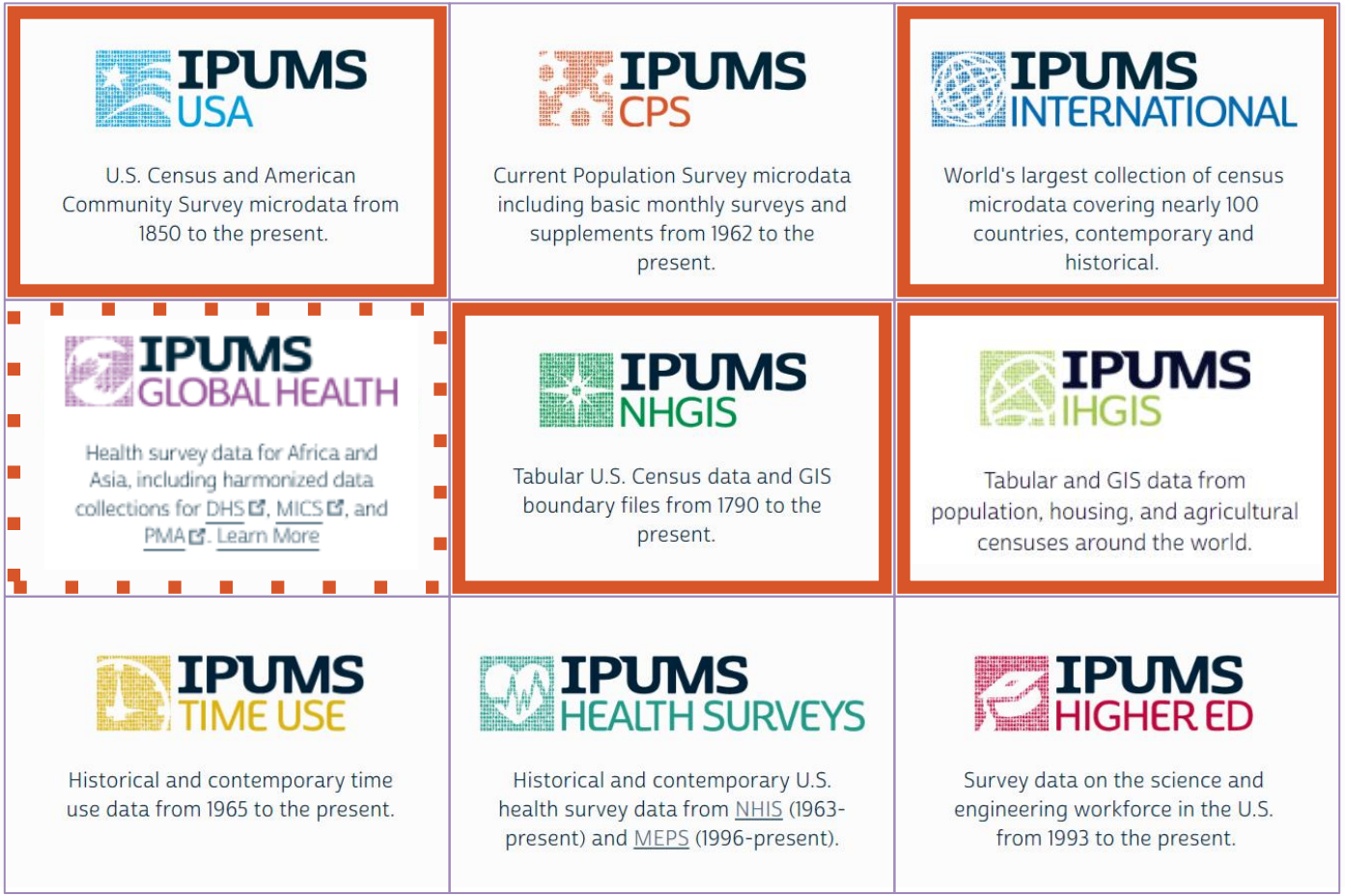
Census data

Survey data

Sub-state geographic detail



GIS data


IPUMS USA
U.S. Census and American Community Survey microdata from 1850 to the present.
IPUMS CPS
Current Population Survey microdata including basic monthly surveys and supplements from 1962 to the present.
IPUMS INTERNATIONAL
World's largest collection of census microdata covering nearly 100 countries, contemporary and historical.
IPUMS GLOBAL HEALTH
Health survey data for Africa and Asia, including harmonized data collections for DHS, MICS, and PMA. Learn More
IPUMS NHGIS
Tabular U.S. Census data and GIS boundary files from 1790 to the present.
IPUMS IHGIS
Tabular and GIS data from population, housing, and agricultural censuses around the world.
IPUMS TIME USE
Historical and contemporary time use data from 1965 to the present.
IPUMS HEALTH SURVEYS
Historical and contemporary U.S. health survey data from NHIS (1963-present) and MEPS (1996-present).
IPUMS HIGHER ED
Survey data on the science and engineering workforce in the U.S. from 1993 to the present.

Why use IPUMS?

- Integration & harmonization
- Data & documentation available via web-based dissemination system
- Create and download custom extracts
- Online data analysis tools
- Data enhancements
- User support

IPUMS Integration Process

Microdata Harmonization

- A single, consistent data series from datasets collected in different times and places
- Codes group broadly comparable categories while retaining sample-specific detail
- Denote potential comparability issues

Basics of Integration

- Consistent codes and variable names
- Thorough and accessible documentation at the variable level
  - Comparability of concepts and universes across time and/or place

Original

Burkina Faso 2018	Kenya 2018	India 2018
0 = Never	0 = Never	0 = Never
1 = Primary	1 = Primary	1 = Primary
2 = Secondary 1 cycle	4 = University	2 = Secondary
3 = Secondary 2 cycle	5 = College	3 = Higher
4 = Tertiary	6 = Post-primary/vocational	4 = Postgrad
	7 = Secondary A level	

Integrated

Original

		Burkina Faso 2018	Kenya 2018	India 2018
		0 = Never	0 = Never	0 = Never
		1 = Primary	1 = Primary	1 = Primary
		2 = Secondary 1 cycle	4 = University	2 = Secondary
		3 = Secondary 2 cycle	5 = College	3 = Higher
		4 = Tertiary	6 = Post-primary/vocational	4 = Postgrad
			7 = Secondary A level	

Integrated

Original

		Burkina Faso 2018	Kenya 2018	India 2018
	Never attended	0 = Never	0 = Never	0 = Never
	Primary	1 = Primary	1 = Primary	1 = Primary
	Post-primary/vocational		6 = Post-primary/vocational	
	Secondary			2 = Secondary
	Cycle 1	2 = Secondary 1 cycle		
	Cycle 2	3 = Secondary 2 cycle		
	A level		7 = Secondary A level	
	Tertiary or higher	4 = Tertiary		3 = Higher
	College		5 = College	
	University		4 = University	
	Postgrad			4 = Postgrad

Integrated		Original		
		Burkina Faso 2018	**Kenya 2018**	**India 2018**
100	Never attended	0 = Never	0 = Never	0 = Never
200	Primary	1 = Primary	1 = Primary	1 = Primary
300	Post-primary/vocational		6 = Post-primary/vocational	
400	Secondary			2 = Secondary
420	Cycle 1	2 = Secondary 1 cycle		
430	Cycle 2	3 = Secondary 2 cycle		
450	A level		7 = Secondary A level	
600	Tertiary or higher	4 = Tertiary		3 = Higher
610	College		5 = College	
620	University		4 = University	
650	Postgrad			4 = Postgrad

GLOBAL HEALTH

Data Collections

IPUMS DHS

DEMOGRAPHIC AND HEALTH SURVEYS

Integrated Demographic and Health Surveys from The DHS Program, currently covering Africa and South Asian surveys from the 1980s to the present.

GET DATA

IPUMS PMA

PERFORMANCE MONITORING FOR ACTION

Integrated PMA surveys on fertility, contraception, hygiene, and health, administered frequently to monitor trends in select high-fertility countries.

GET DATA

IPUMS MICS

MULTIPLE INDICATOR CLUSTER SURVEYS

Integrated MICS surveys covering child health and well-being.

GET DATA

mics.ipums.org

Harmonization for
900+ variables

89 countries 211 samples 1,089 datasets

MICS

Multiple Indicator Cluster Survey

Developed by UNICEF in the 1990s

To assist countries in filling data gaps on children's and women's well-being for tracking progress toward

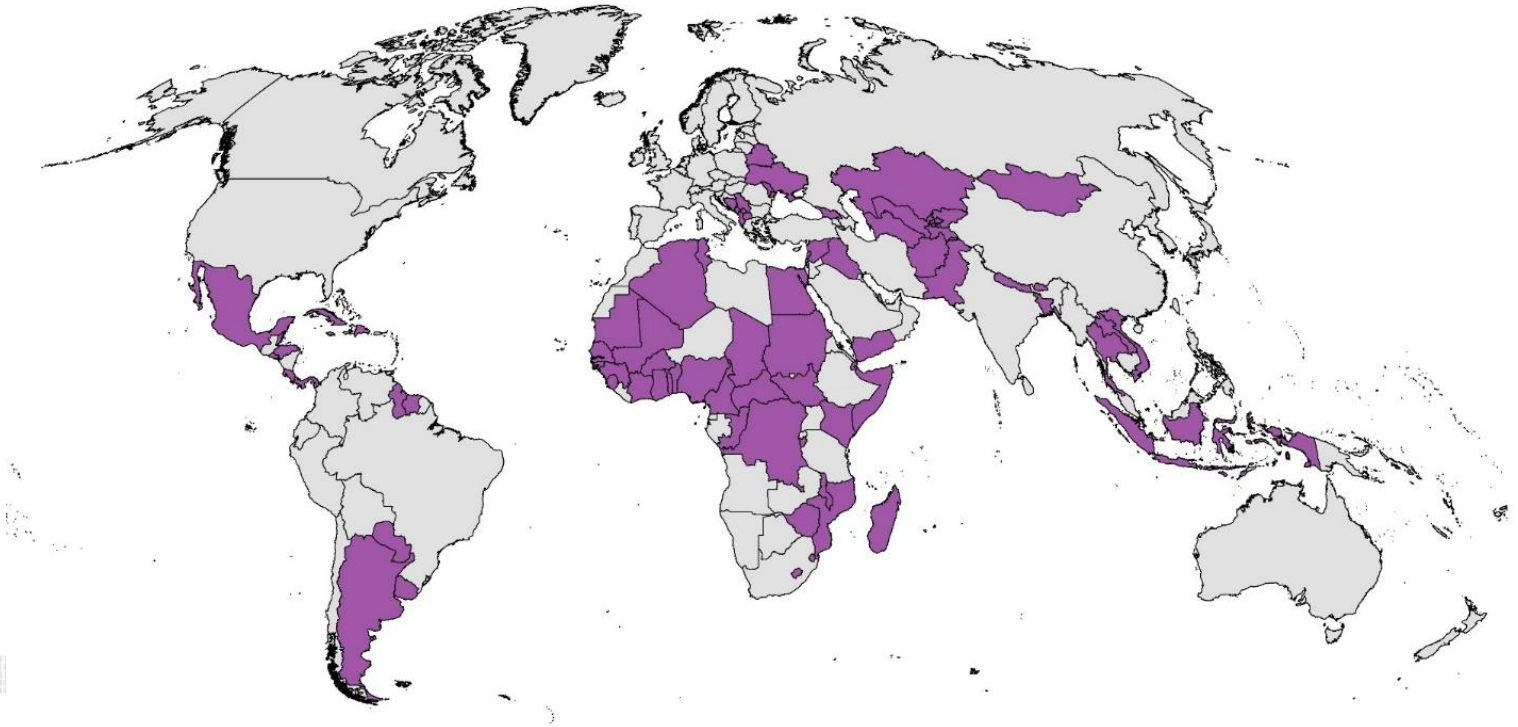
- World Summit for Children Goals
- MDGs
- SDGs

Images: Copyright

Advantages of using IPUMS MICS

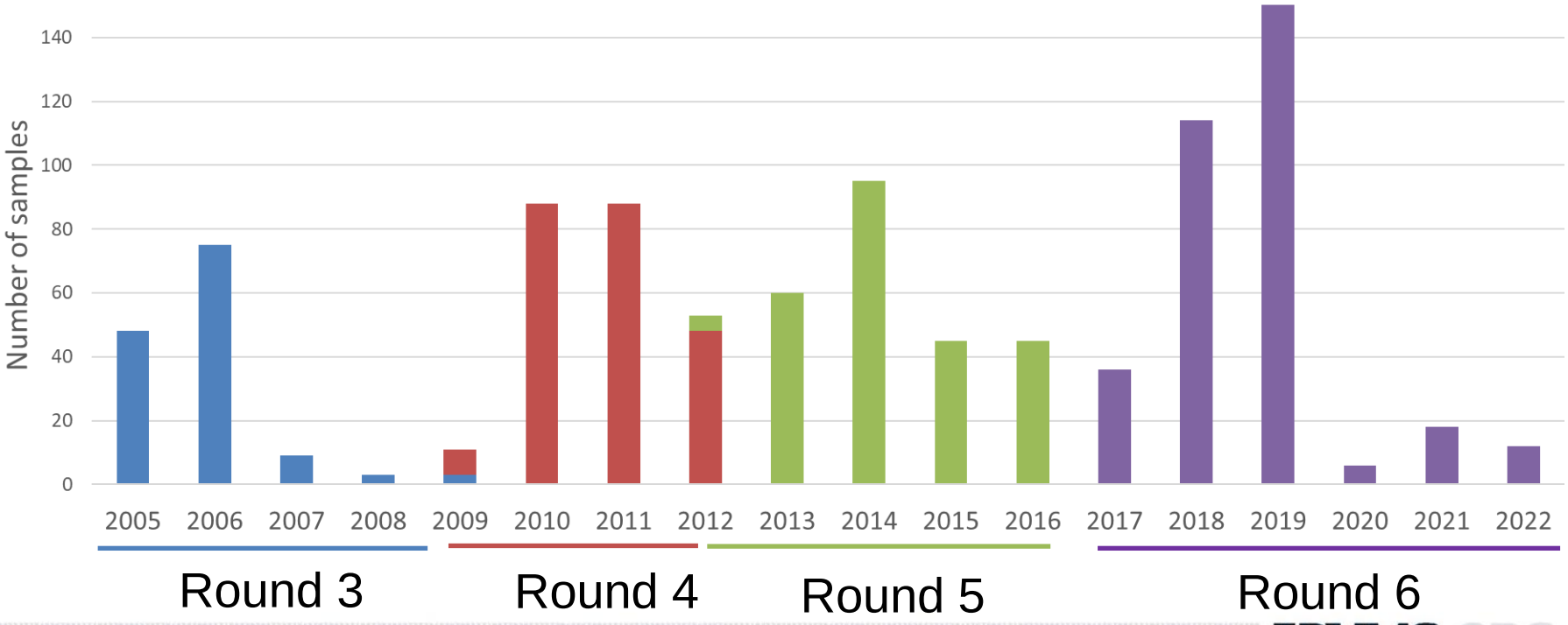
- Comparability across time and place
  - Approximately 60% of common variables currently harmonized
- Harmonized geography
- Easily browse topics and variables
- Documentation on each variable's universe, question wording (in English), comparability issues
- Stata data files

89 IPUMS MICS Countries



.ORG

IPUMS MICS Samples Over Time



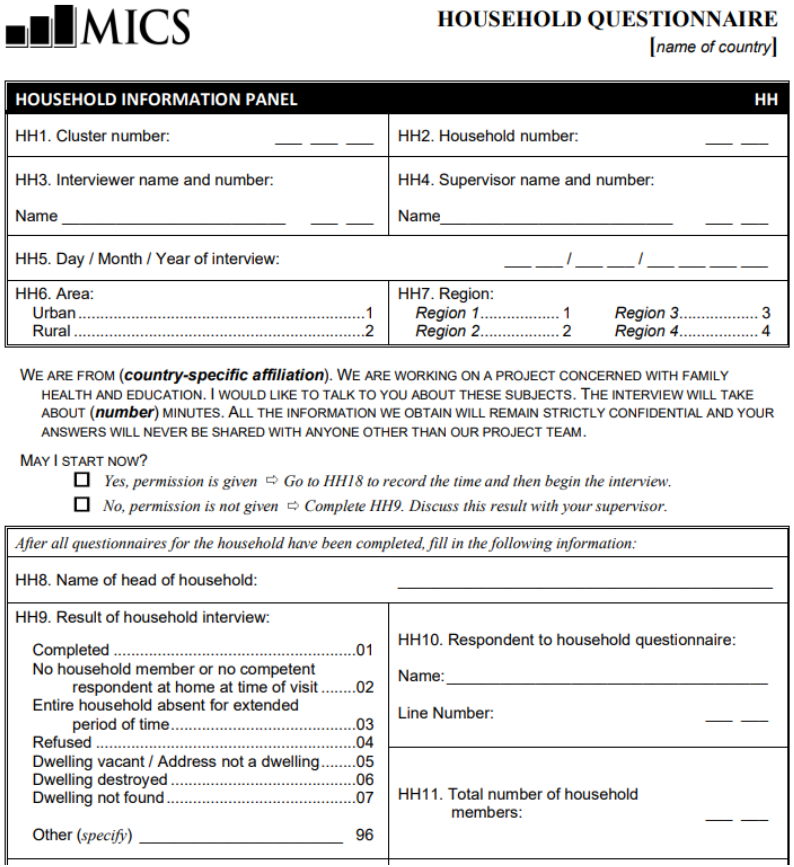
Sample Design

Nationally representative sample of households

Typically, around 10,000 – 30,000 HHs

Multi-stage cluster sampling of HHs

Similar to Demographic and Health Surveys (DHS)



MICS

HOUSEHOLD QUESTIONNAIRE

[*name of country*]

HOUSEHOLD INFORMATION PANEL	**HH**
HH1. Cluster number: ___ ___ ___	HH2. Household number: ___ ___
HH3. Interviewer name and number: Name _____ ___ ___	HH4. Supervisor name and number: Name_____ ___ ___
HH5. Day / Month / Year of interview:	___ ___/___ ___/___ ___ ___ ___
HH6. Area: Urban ..... 1 Rural 2	HH7. Region: *Region 1* ..... 1 *Region 3* ..... 3 *Region 2* ..... 2 *Region 4* 4

WE ARE FROM (***country-specific affiliation***). WE ARE WORKING ON A PROJECT CONCERNED WITH FAMILY HEALTH AND EDUCATION. I WOULD LIKE TO TALK TO YOU ABOUT THESE SUBJECTS. THE INTERVIEW WILL TAKE ABOUT (***number***) MINUTES. ALL THE INFORMATION WE OBTAIN WILL REMAIN STRICTLY CONFIDENTIAL AND YOUR ANSWERS WILL NEVER BE SHARED WITH ANYONE OTHER THAN OUR PROJECT TEAM.

MAY I START NOW?

☐ *Yes, permission is given* ⇨ *Go to HH18 to record the time and then begin the interview.*

☐ *No, permission is not given* ⇨ *Complete HH9. Discuss this result with your supervisor.*

After all questionnaires for the household have been completed, fill in the following information:	
HH8. Name of head of household:	_____
HH9. Result of household interview: Completed ..... 01 No household member or no competent respondent at home at time of visit ..... 02 Entire household absent for extended period of time ..... 03 Refused ..... 04 Dwelling vacant / Address not a dwelling ..... 05 Dwelling destroyed ..... 06 Dwelling not found ..... 07 Other (*specify*) _____ 96	HH10. Respondent to household questionnaire: Name: _____ Line Number: ___ ___ HH11. Total number of household members: ___ ___

Household Questionnaire

HH Member Roster
HH Characteristics
WASH
Water Quality

Household Questionnaire

Individual Women (Age 15-49)

Demography
Fertility
Birth History
Contraception & Unmet Need
Maternal and Newborn Health

Household Questionnaire

Media & Technology
Sexual Behavior
HIV
Domestic Violence
Life Satisfaction

Individual Women (Age 15-49)

Individual Men (Age 15-49)
(all or sub-sample of HHs)

Household Questionnaire

Demography
Illness
Child Functioning
ECE & Development
Anthropometry
Child discipline

Individual Women (Age 15-49)

Individual Men (Age 15-49)
(all or sub-sample of HHs)

Children Under-5 (Age 0-4)

Household Questionnaire

Demography
Literacy and Numeracy
Parental Involvement
Child Discipline
Child Labor

Individual Women (Age 15-49)

Individual Men (Age 15-49)
(all or sub-sample of HHs)

Children Under-5 (Age 0-4)

Children Age 5-17
(one per HH)

Household Questionnaire

Household Members

- **Individual Women (Age 15-49)**
- **Individual Men (Age 15-49)** (all or sub-sample of HHs)
- **Children Under-5 (Age 0-4)**
- **Children Age 5-17** (one per HH)

Research with IPUMS MICS

- Comparative research across countries (e.g., tracking SDGs)
- Studying trends over time
- Pooling data to study small populations
  - Orphans
  - Children with particular disabilities

Newly released at IPUMS MICS

New variables

Child labor
Alcohol and tobacco use
Victimization
Household energy use
HH assets

New samples

Afghanistan 2022-2023
Fiji 2021 – NEW COUNTRY
Vietnam 2020-2021
Nigeria 2021
Uzbekistan 2021-2022
Thailand 2022
Pakistan (Balochistan) 2019-2020
Pakistan (Khyber Pakhtunkhwa) 2019

More new variables and samples will be released in Summer 2024!

IPUMS DHS

dhs.ipums.org

What is the DHS?

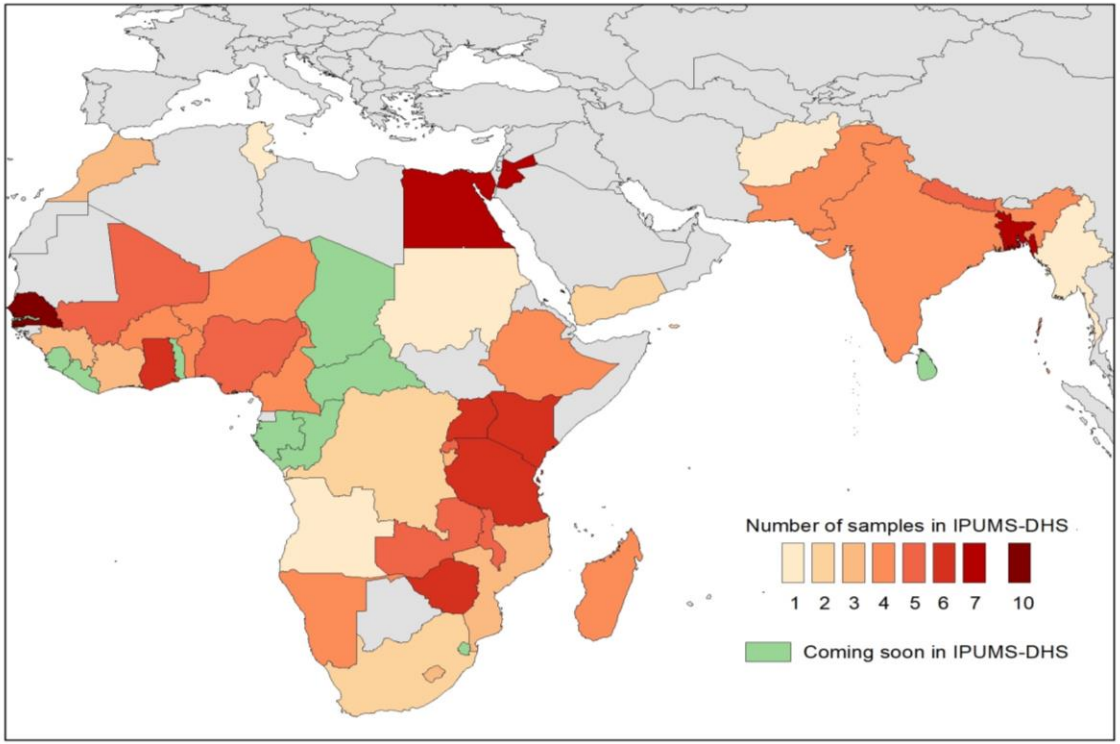
- Demographic and Health Surveys (DHS)
- 1980's to the present
- "Collect and disseminate accurate and representative data on population, health, HIV, and nutrition"
- Over 400 surveys in more than 90 countries
- Leading source of information on health in Low and Middle Income countries (LMICs)

What is IPUMS DHS?

- Website with over 20,000 harmonized variables
- Microdata drawn from DHS public use files, recoded
- Consistent variable names and codes for “non-standard” variables
- Additional features (contextual variables, reproductive calendar

Samples and Countries in IPUMS DHS

- 180 samples
- 45 countries
- Most sub-Saharan African countries
- North Africa, Middle East
- South Asia



Contextual Variables in IPUMS DHS

Physical and Environmental

Ecoregion

Soil

Normalized Difference Vegetation Index (NDVI)

Precipitation

Temperature (minimum and maximum)

Economic and Social

Livelihood index

Population density

Malaria

Agricultural

Cropland

Pastureland

Crops harvested (17 major crops)

Crop production (17 major crops)

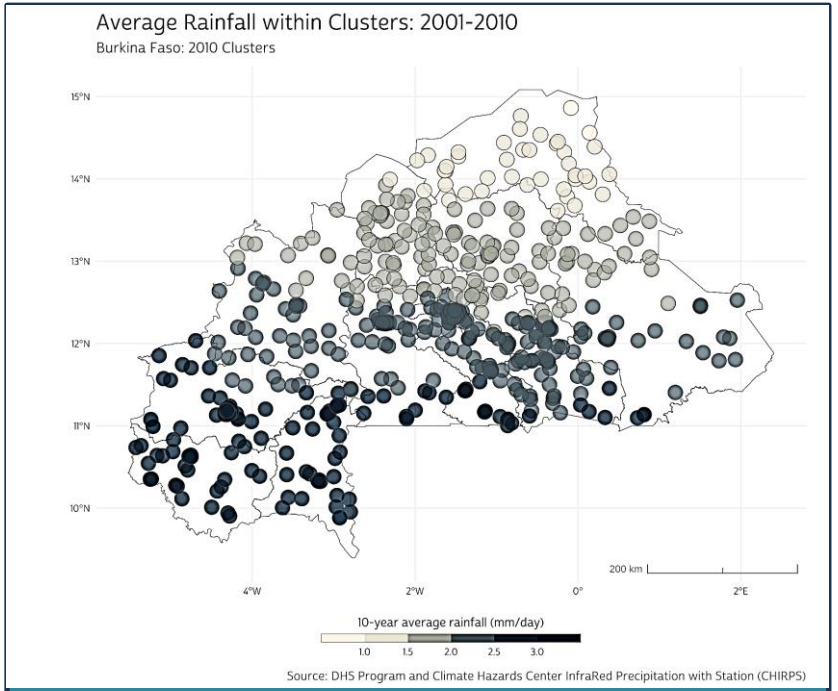


Image from Climate change and health research hub at IPUMS DHS

Constructing Contextual Variables in IPUMS DHS

- Use GPS sample cluster points (displaced w/in 5-10 kilometers)
- Draw 5 or 10 km buffer zone around sample points with ARC GIS
- Bring in non-DHS information with geographic location info
- Associate that non-DHS information with all survey members in a sample cluster

More information: "Contextual data in IPUMS DHS: physical and social environment variables linked to the Demographic and Health Surveys," Population and Environment (May 2020).

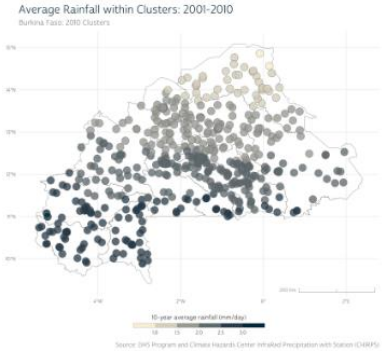
Climate Change and Health Research Hub

https://tech.popdata.org/dhs-research-hub/

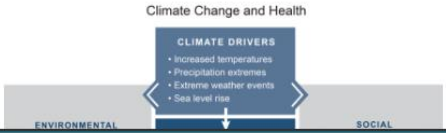
ATTACHING CHIRPS PRECIPITATION DATA TO DHS SURVEYS

CHIRPS | PRECIPITATION | CLIMATE | AGRICULTURE | TERRA | SF | GGSPATIAL | R

Remotely sensed precipitation data can provide important context for understanding the health outcomes reported in the DHS



DEVELOPING ROBUST CONCEPTUAL MODELS IN CLIMATE CHANGE AND HEALTH RESEARCH



research-hub/posts/2024-02-04-dhs-chirps/

CATEGORIES

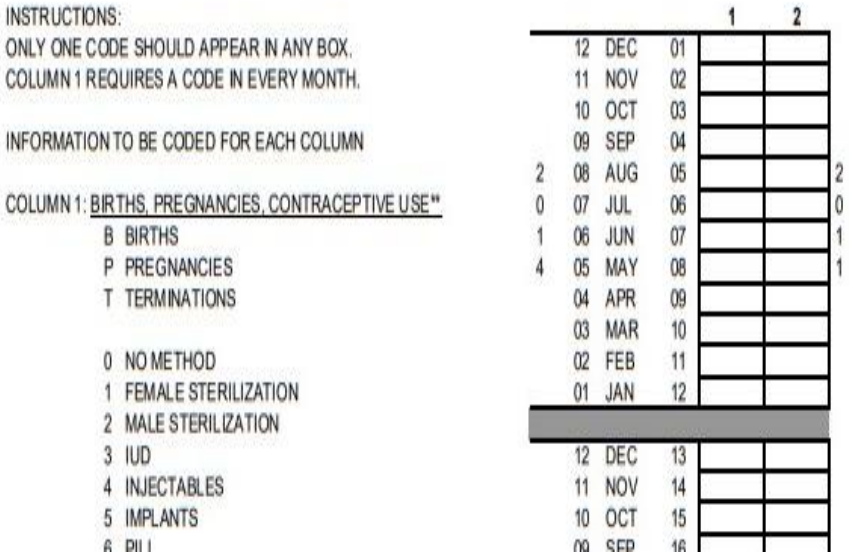
All (4)
Agriculture (1)
CHIRPS (1)
Climate (2)
Importing data (1)
Methodology (1)
Precipitation (1)
R (3)
Research concepts (1)
Tips (2)
ggspatial (1)
sf (1)
terra (1)

Reproductive Calendar Data in IPUMS DHS

Women of childbearing age were asked about the timing/dates of

- Pregnancies
- Births
- Pregnancy terminations
- Use of specific contraceptive methods

in period preceding the survey interview (5 years)



INSTRUCTIONS:
ONLY ONE CODE SHOULD APPEAR IN ANY BOX.
COLUMN 1 REQUIRES A CODE IN EVERY MONTH.

INFORMATION TO BE CODED FOR EACH COLUMN

COLUMN 1: BIRTHS, PREGNANCIES, CONTRACEPTIVE USE**

B BIRTHS
P PREGNANCIES
T TERMINATIONS

0 NO METHOD
1 FEMALE STERILIZATION
2 MALE STERILIZATION
3 IUD
4 INJECTABLES
5 IMPLANTS
6 PILL

				1	2	
	12	DEC	01			
	11	NOV	02			
	10	OCT	03			
	09	SEP	04			
2	08	AUG	05			2
0	07	JUL	06			0
1	06	JUN	07			1
4	05	MAY	08			1
	04	APR	09			
	03	MAR	10			
	02	FEB	11			
	01	JAN	12			
	12	DEC	13			
	11	NOV	14			
	10	OCT	15			
	09	SEP	16			

Reproductive Calendar Data in IPUMS DHS

Allows study of topics such as

- Contraceptive failure rates
- Length of use, reason for stopping use of FP methods
- Prevalence of births delivered before 9 months of pregnancy
- Relationship between IPV and miscarriage
- Prevalence and timing of abortion

For more information, see **"How to Use Simplified Reproductive Calendar Data from the Demographic and Health Survey,"** Studies in Family Planning, March 2023; Elizabeth Heger Boyle, Nir Rotem, and Miriam L. King,

PMA

PERFORMANCE MONITORING FOR ACTION

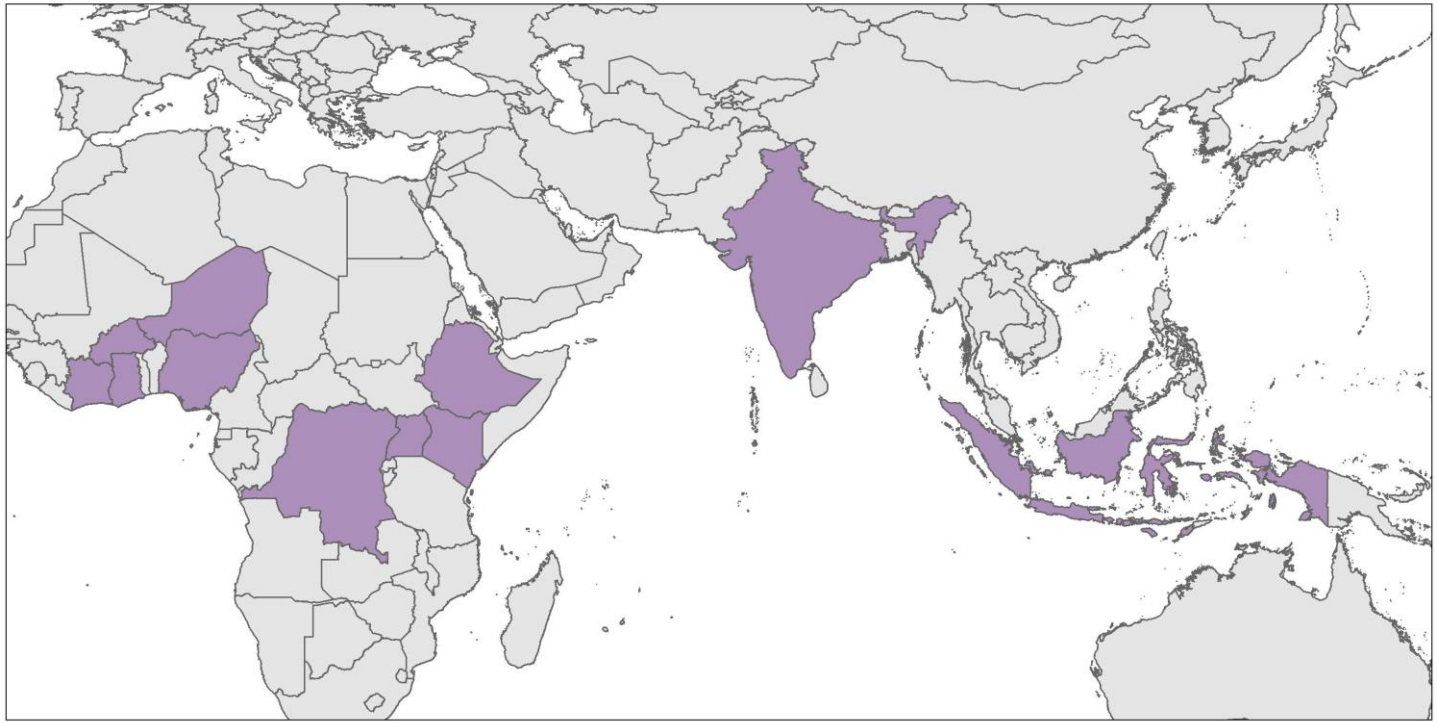
IPUMS PMA Overview

Performance Monitoring for Action (PMA)

- High frequency, recent surveys on family planning, sexual and reproductive health since 2013
- Data from 9 countries in Africa and Asia
- Originally designed to monitor progress towards FP2020 goals
- Funded by the Bill & Melinda Gates Foundation

200+ SAMPLES · 9000+ VARIABLES · 2 MILLION RECORDS

PMA Countries



CHOOSE THE TOPIC FOR DATA BROWSING

PERSON	FAMILY PLANNING DESCRIPTION	NUTRITION DESCRIPTION
	CLIENT EXIT INTERVIEW DESCRIPTION	COVID-19 DESCRIPTION
SERVICE DELIVERY POINT	FAMILY PLANNING DESCRIPTION	NUTRITION DESCRIPTION
INFANT	MATERNAL AND NEWBORN HEALTH DESCRIPTION	

MS.ORG

Household and Female

- Cross-sectional (2013-2018)
- Longitudinal and cross-section (2019-onward)
- Household
  - wealth, water, sanitation
- Female
  - Fertility, family planning, reproductive health

Topic Module

- Migration
- Gender-based Violence
- Economic empowerment
- Menstrual hygiene
- Health insurance
- Abortion
- Contraceptive Side Effects
- Contraceptive calendar

Service Delivery Point

- Same enumeration areas as surveyed households
- Family planning service provision
- Stock-outs
- Other health services
  - Antenatal, abortion, post abortion, post natal

COVID-19

- Phone survey in summer 2020
  - Burkina Faso, DRC, Kenya, Nigeria
  - Can be linked with longitudinal panel
- Information sources and trust level
- Health care access impacts
- Social distancing measures and attitudes

Maternal and Newborn Health

- Ethiopia cohort studies (2016- 2017 and 2019-2021)
- Antenatal care
- Pregnancy and delivery complications
- Infant health issues and treatment
- Infant vaccination

Client Exit Interviews

- Short interviews of family planning clients at health facilities
- Can be linked to facility data
- Method preference
- Quality of care
- Follow up interviews

Nutrition

- Burkina Faso and Kenya
- Children under 5 diet and biomarkers (MUAC)
- Household food security
- Antenatal nutrition
- Breastfeeding assistance

https://tech.popdata.org/pma-data-hub/

IPUMS PMA DATA ANALYSIS HUB Search... Posts Courses ▾ R Tips ▾ Community ▾ About

DATA ANALYSIS HUB

March 22, 2023
Matt Gunther

ABORTION INCIDENCE WITH THIRD PARTY REPORTING

DATA ANALYSIS | ABORTION | MEASUREMENT

Recent PMA surveys from Côte d'Ivoire and Nigeria include questions about the abortion experiences of both respondents and other women of reproductive age with whom they share personal information.

Feb. 15, 2023
Matt Gunther
Devon Kristiansen

COMPARING MEASURES OF ABORTION INCIDENCE WITH A SHINY DASHBOARD

DATA ANALYSIS | ABORTION | SHINY | DPLYR

New abortion data from PMA includes women who reported doing something to remove a pregnancy, or to regulate a late period because of a suspected pregnancy.

CATEGORIES

- Articles (52)
- Abortion (3)
- ACLED (1)
- across (2)
- Agriculture (1)
- Animation (1)
- Armed Conflict (1)
- Bar Charts (2)
- bootstraps (1)
- CHIRPS (1)
- Client Exit Interviews (6)
- Climate (2)
- Cluster Sampling (2)
- Codebook (1)
- Contraceptive Calendar (3)
- COVID-19 (5)
- cowplot (1)
- cur_data (1)
- Data Analysis (18)

IPUMS International

HARMONIZED INTERNATIONAL CENSUS DATA FOR SOCIAL SCIENCE AND HEALTH RESEARCH

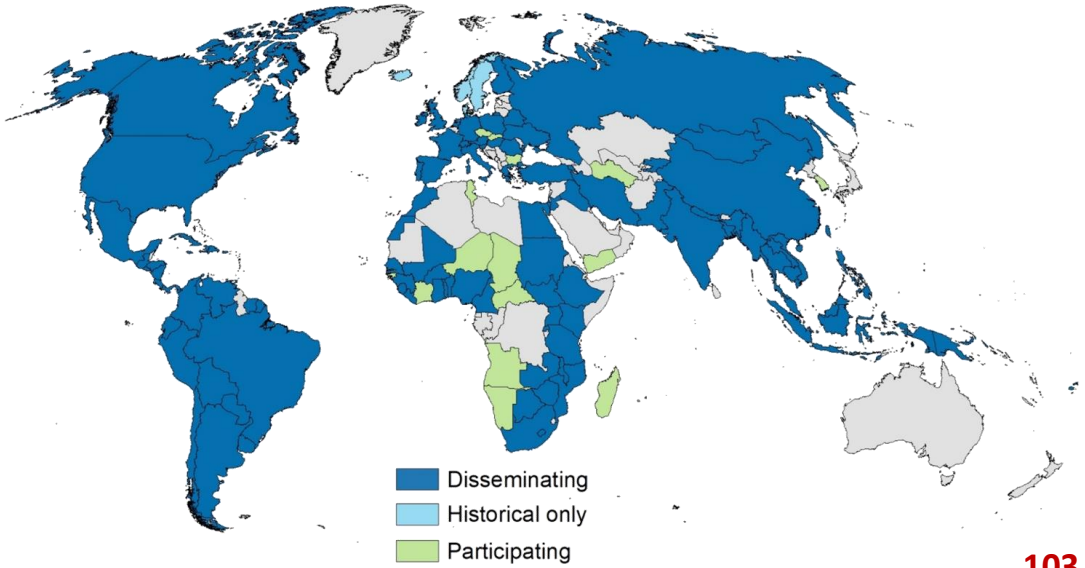
103 COUNTRIES — 547 CENSUSES & SURVEYS — OVER 1 BILLION PERSON RECORDS

Features of IPUMS International

- Census (and survey) microdata samples
- Anonymized
- Harmonized with consistent coding
- **Spatially harmonized geography and GIS files**
- Data extract system with customizable extracts
- **Online data analysis**
- **FREE** access to researchers!

IPUMS microdata availability

IPUMS International Data provided by the National Statistical Institutes, partners in the census integration and dissemination project



Coming soon!
Cambodia (2019)
Ivory Coast (1988, 1998)
Laos (1995, 2015)
Mexico (2020)
Peru (2017)
Switzerland (2011)
Vietnam (2019)
Puerto Rico (2015, 2020)
United States (2015, 2020)
England & Wales (1961, 1971)
Denmark (1880, 1885)

1,190,012,299 person records

103 countries; 547 censuses and surveys

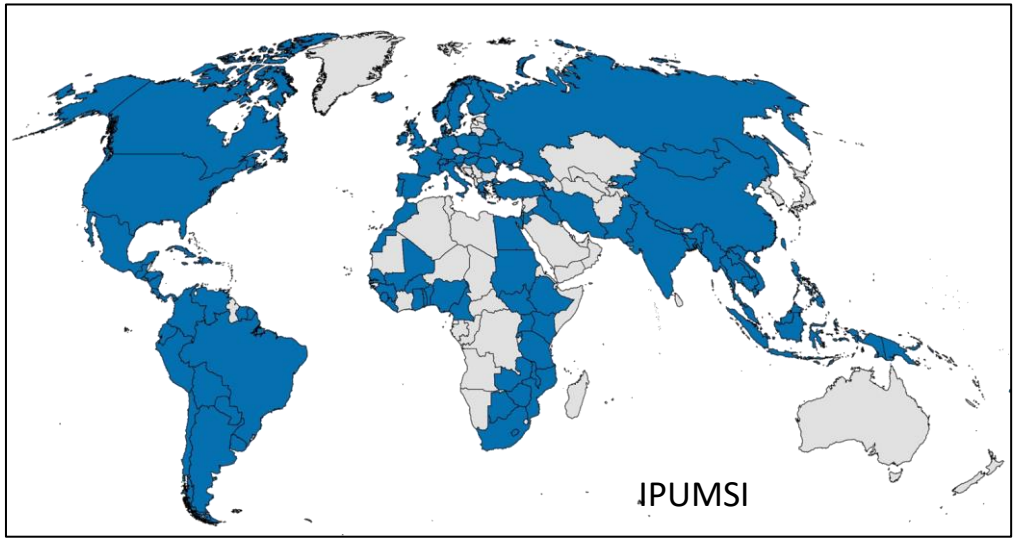
Commonly requested variables

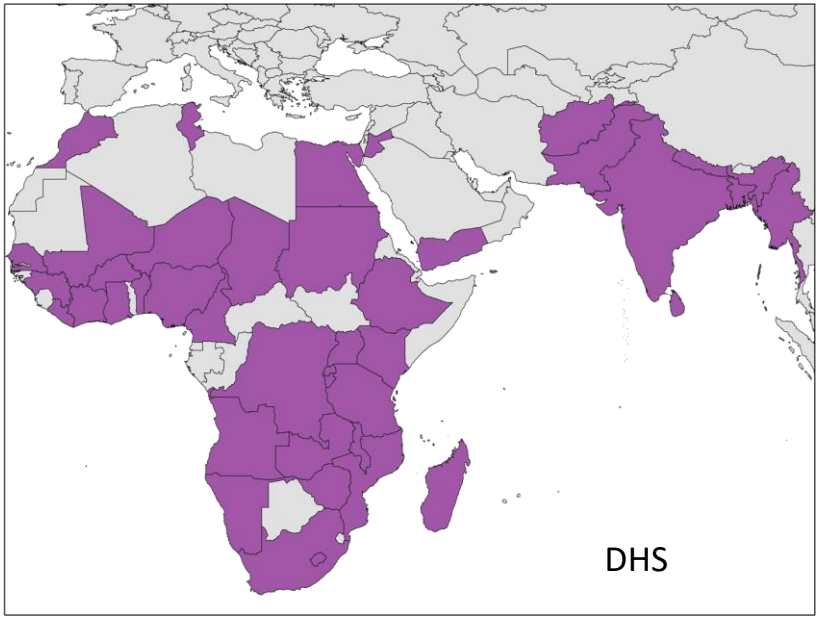
Top 10	11 to 20	21 to 30
Employment status	Industry	Spouse's location in HH
Education Attainment	Urban-rural status	Country of birth
Marital Status	Ownership of dwelling	Father's location in HH
Age	Years of schooling	Family size
Sex	Children ever born	Children surviving
Relationship to head	Religion	Number own children <5 in HH
Class of worker	Household size	Group quarters status
School attendance	Nativity status	Age of eldest child
Occupation	Number own children in HH	Age of youngest child
Literacy	Mother's location in HH	Total Income

1,900+ Harmonized variables

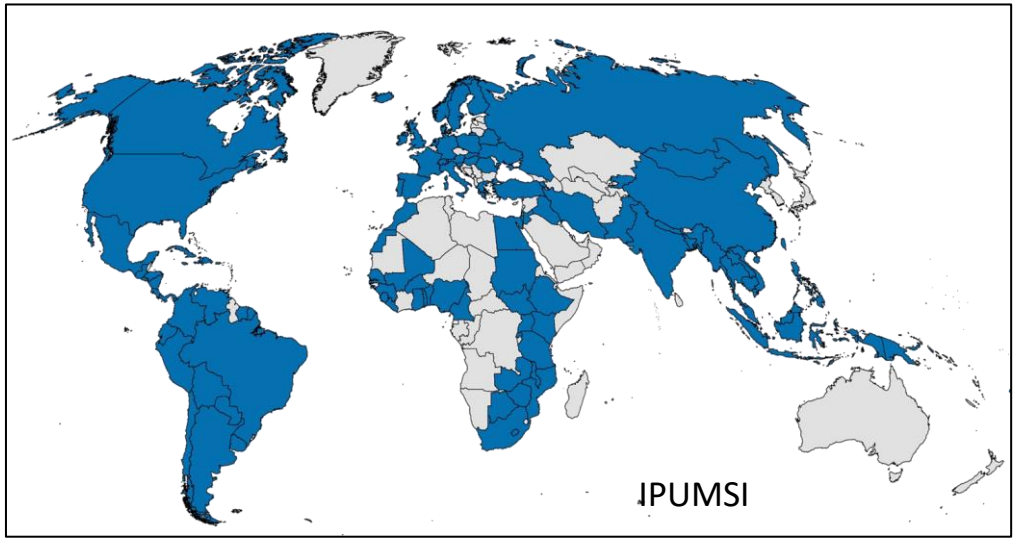
98,000+ Source variables

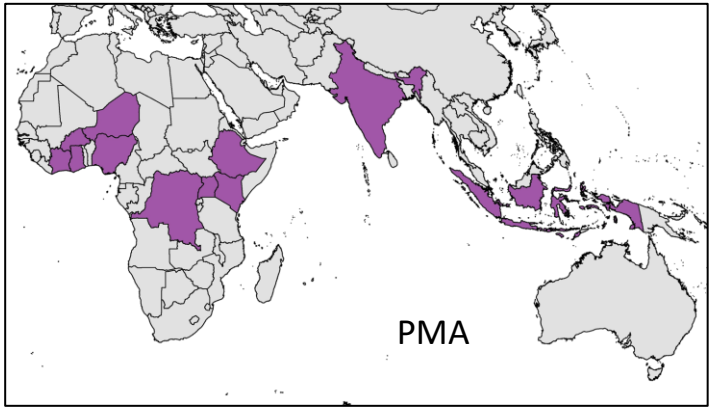
IPUMSI and other global health projects

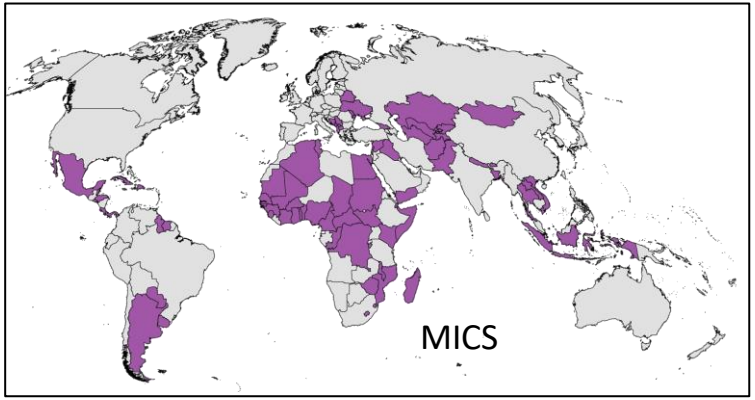




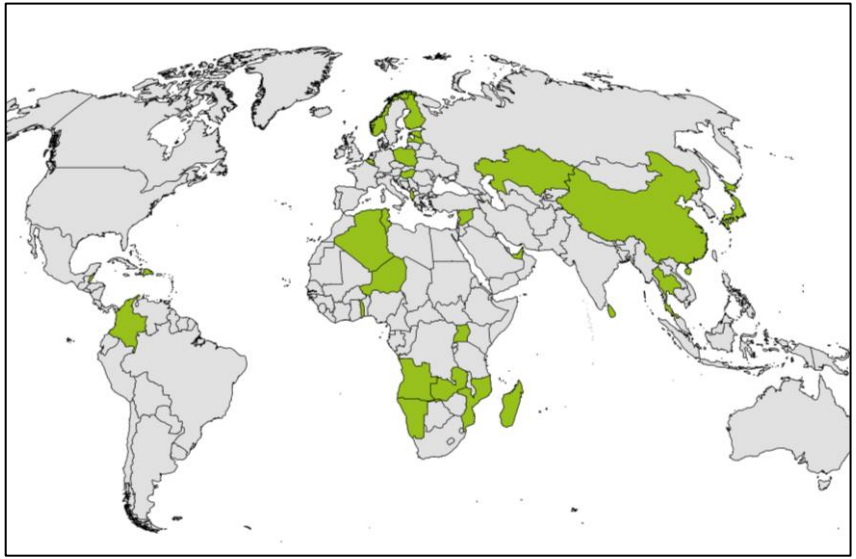
IPUMSI and other global health projects



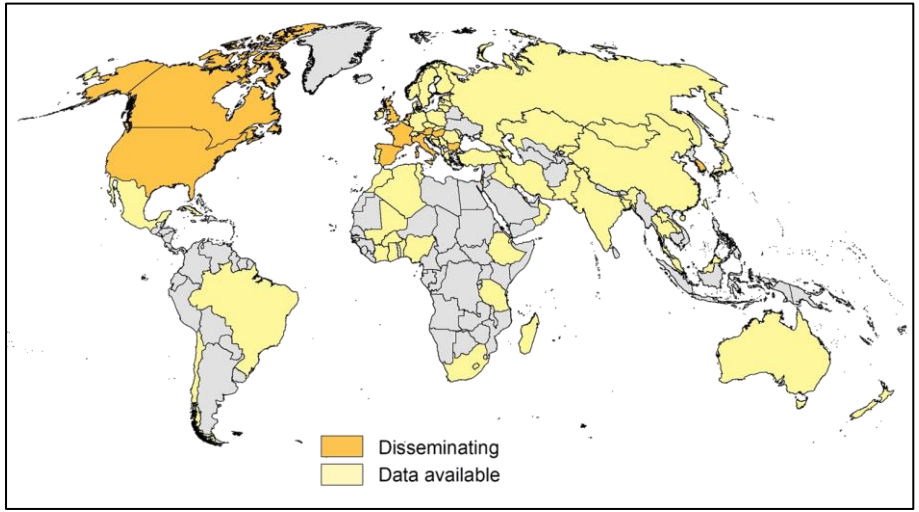




Other IPUMS projects - international

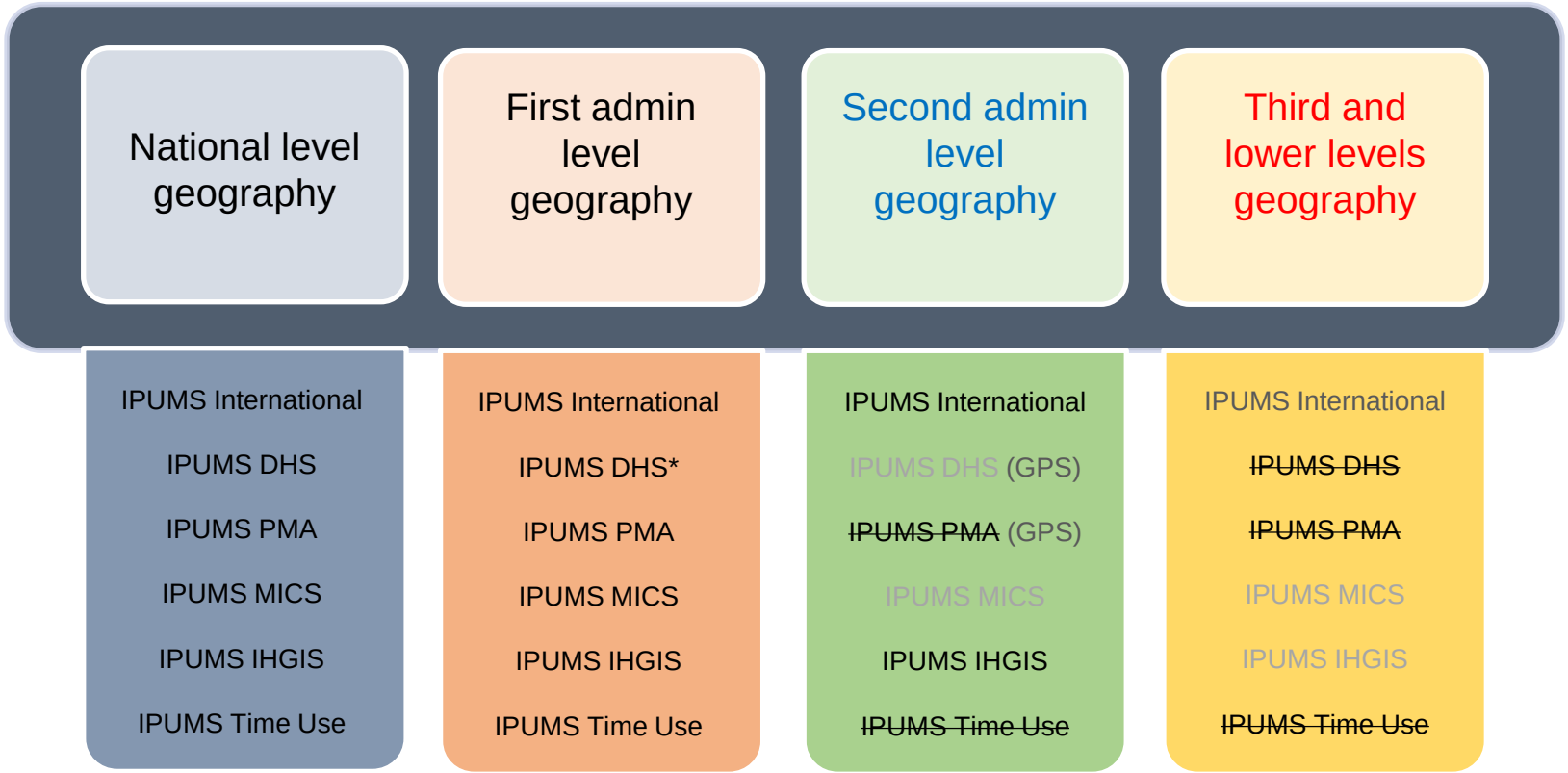


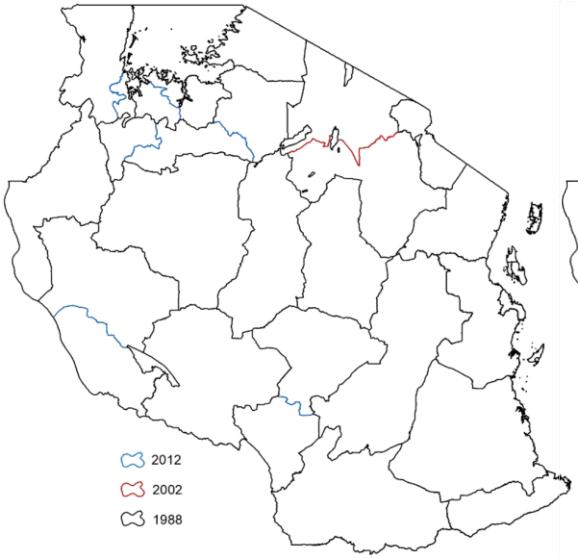
IPUMS IHGIS

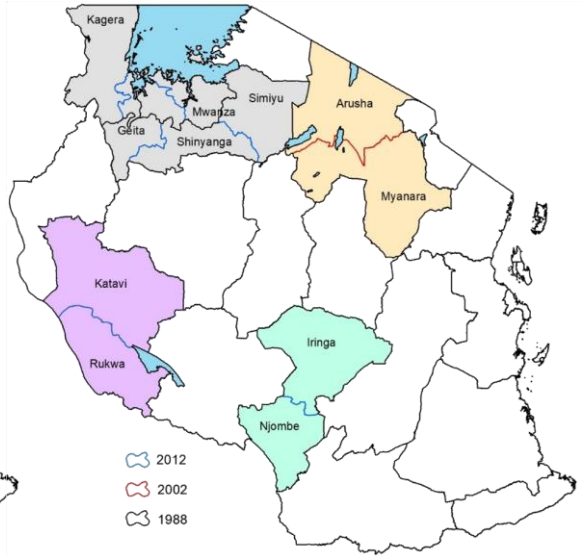


IPUMS TIME USE

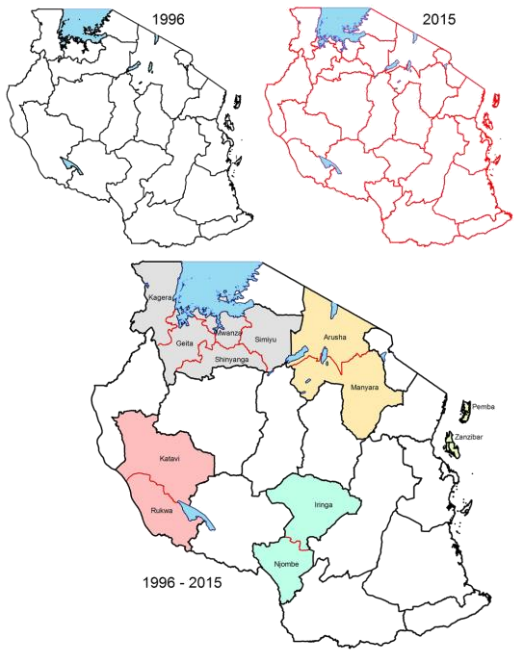
Crosswalks and data interoperability





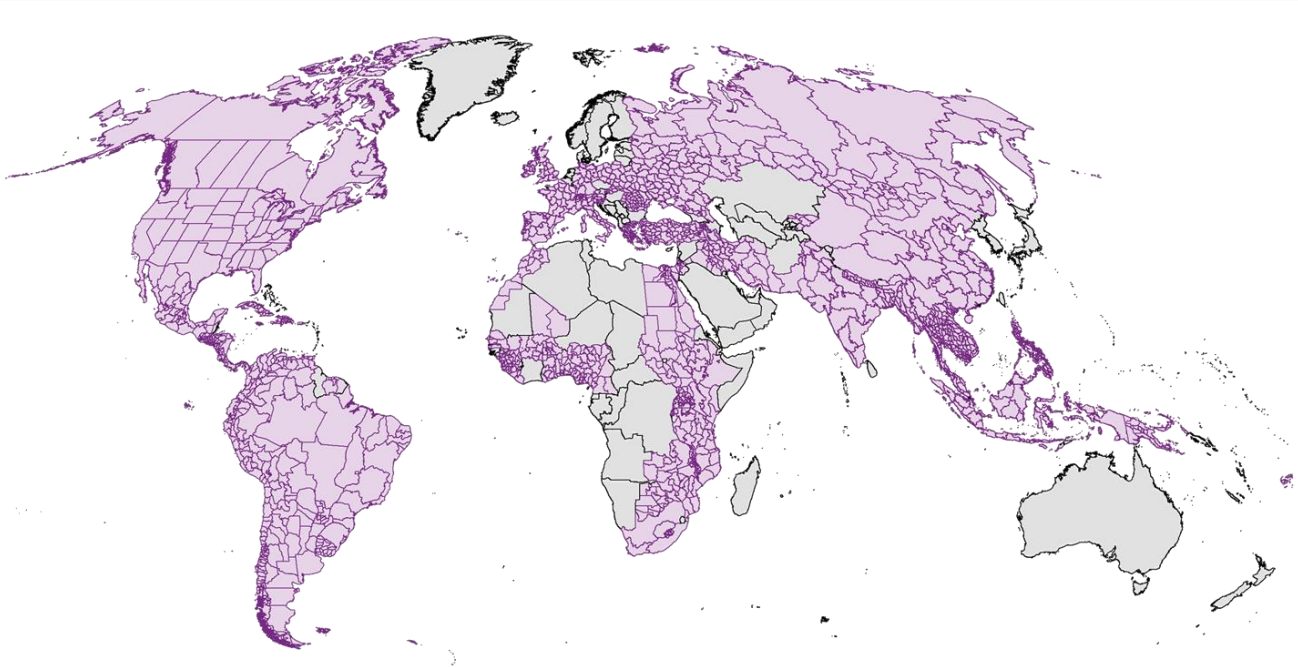


Tanzania, 1988 – 2012 – IPUMS International

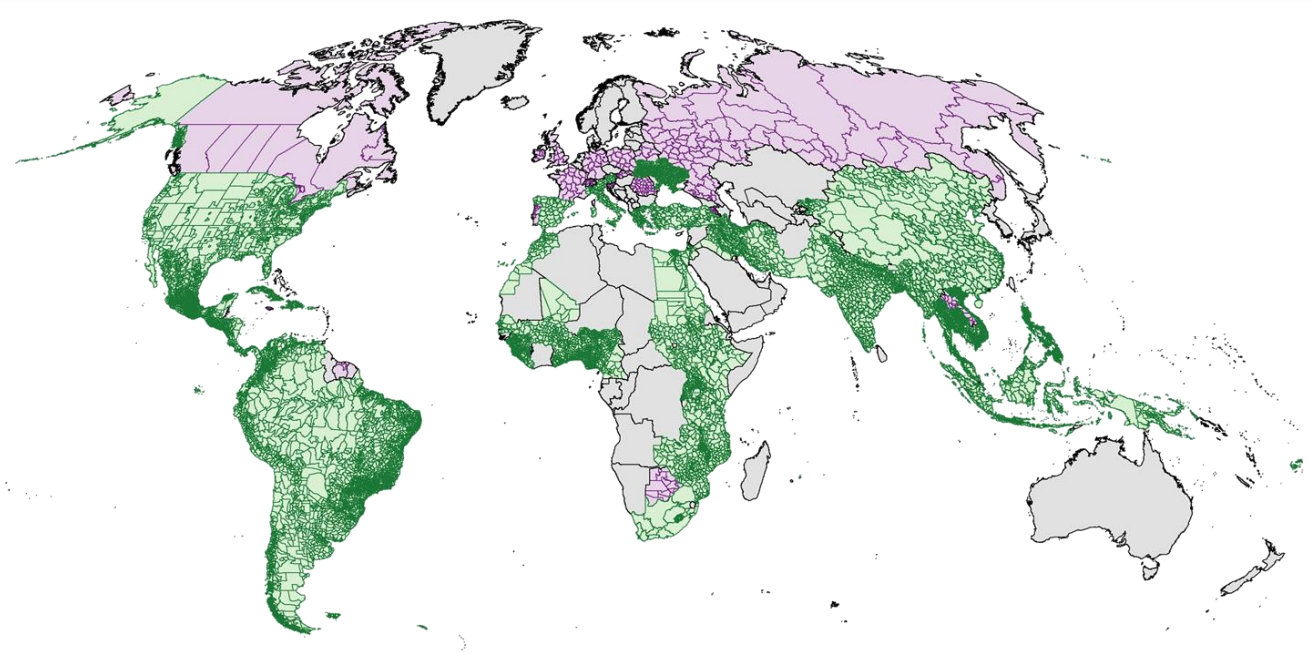


Tanzania, 1996 – 2015 – IPUMS DHS

GIS Files availability – Level 1



GIS Files availability – Level 2



IPUMS GLOBAL HEALTH

User base for “IPUMS Global Health”

- Over 7500 registered users from 115 countries
- Broad range of disciplines
- 31 percent of users are students
- Dozens of students have used IPUMS Global Health for their theses

Similarities

Question wording*

Definitions of indicators

Universe

Sample design

Differences

Different countries

Starting year

- DHS: 1980s
- MICS: 2000s
- PMA: 2013

Freq of data collection

- DHS & MICS: 5-10 years
- PMA: Annual

Other IPUMS differences

- Registration

  MICS -> UNICEF

  DHS -> The DHS Program

  IPUMSI and PMA

  USA and other US projects

- Creating data extracts and downloading data

  IPUMS MICS = Stata syntax + original data from UNICEF

  All other projects: choice of format and already harmonized data

  **Will demonstrate MICS today*

	IPUMS DHS	IPUMS MICS	IPUMS PMA
Household Members	x	x	x
Women	x	x	x
Birth History	x	x	
FGMC Daughters	x	Coming soon	
Children 0-4	x	x	x
Household	Coming soon	x	
Men	x	x	
Calendar	x		x
Children 5-17		x	
Service Delivery			x
Infant Panel			x
Longitudinal			x

Resources

- IPUMS Website: ipums.org
- IPUMS Blog: https://blog.popdata.org/
- IPUMS User Forum: https://forum.ipums.org/
- IPUMS Youtube: https://www.youtube.com/user/mpcipums

- Demonstration of walkthrough of IPUMS MICS
  - https://www.youtube.com/watch?reload=9&v=tm3rJKpxQt0&feature=youtu.be