Using IPUMS data to Study Work, Health, and Aging

Sarah Flood

Pre-conference workshop presented by IPUMS and NDiRA

DATA AVAILABLE



USA

U.S. Census and American Community Survey microdata from 1850 to the present.



CPS

Current Population Survey microdata including basic monthly surveys and supplements from 1962 to the present.



INTERNATIONAL

World's largest collection of census microdata with over 100 countries, contemporary and historical.



TIME USE

U.S. and international time use data for 1965 to the present.



NHGIS

Tabular U.S. Census data and GIS boundary files from 1790 to the present.



GLOBAL HEALTH

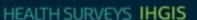
Health survey data Data on the for Africa and Asia, including harmonized data collections for DHS and PMA.



HIGHER ED

science and engineering workforce in the U.S. from 1993 to the present.





Health survey data Data tables from the National from popula Health Interview Survey and the Medical Expenditure Panel Survey.

from population and housing censuses as well as agricultural censuses from around the world with corresponding GIS files.

Pre-conference workshop presented by **IPUMS** and **NDîRA**



USA

U.S. Census and American Community Survey microdata from 1850 to the present.



CPS

Current
Population Survey
microdata
including basic
monthly surveys
and supplements
from 1962 to
the present.



INTERNATIONAL

World's largest collection of census microdata with over 100 countries, contemporary and historical.



TIME USE

U.S. and international time use data for 1965 to the present.



HEALTH SURVEYS

Health survey data from the National Health Interview Survey and the Medical Expenditure Panel Survey.

Census Other Surveys

Health Surveys

USA

CPS

NHIS

International

ATUS

MEPS

Time

Target population

Data characteristics

Topics

Time

Target population

Data characteristics

Topics

Region

State

Sub-state (> 100k)

Sub-state (< 100k)

Region

State

Sub-state (> 100k)

Sub-state (< 100k)

USA

CPS

ATUS

NHIS

MEPS

International

Region State **USA USA CPS CPS ATUS ATUS NHIS MEPS**

Sub-state Sub-state (> 100k) (< 100k)

International International

ATUS ATUS NHIS MEPS International International International

Sub-state

(> 100k)

USA

Sub-state

(< 100k)

Region State **USA USA CPS CPS**

Time

Target population

Data characteristics

Topics

USA

International

USA

International

CPS NHIS

USA

International

CPS NHIS

> ATUS MEPS

Time

Target population

Data characteristics

Topics

All Persons

Civilian, Non-institutionalized

All Persons

Civilian, Non-institutionalized

USA

CPS

International

ATUS

NHIS/MEPS

Time

Target population

Data characteristics

Topics

Microdata

Summary data

Cross-sectional

Longitudinal

Microdata

Summary data

USA

NHIS

CPS

MEPS

ATUS

International

Cross-sectional

Longitudinal

Summary data Microdata **USA NHIS MEPS CPS ATUS** International Longitudinal Cross-sectional **CPS USA NHIS MEPS MEPS CPS** ATUS

Time

Target population

Data characteristics

Topics

Int'l Health USA **ATUS CPS** SES Housing Work Income Poverty

	USA	CPS	Int'l	ATUS	Health
Fertility	•		•		
Household or Family					
Health Insurance	•		•		
Health behaviors					
Disability	•				
Migration					

Benefits of these data

- Large-scale, nationally representative
- Data on all* household members
- Expansive temporal and geographic coverage
- Rich sociodemographic data
- Varying detail on work, aging, and health

Pre-conference workshop presented by **IPUMS** and **NDiRA**

IPUMS USA

Decennial census data, 1790 to present

American Community Surveys, 2000 to present

Pre-conference workshop presented by **IPUMS** and **NDiRA**

IPUMS USA

- U.S. decennial censuses
 - Household- and county-level (1790-1840)
 - Full count (1850-1940)
 - Samples (1950-2010)
- American Community Surveys (2000-)
- Supplemental Datasets/Resources

Decennial Census Data

- Linking full-count census data across time
 - https://usa.ipums.org/usa/mlp/mlp.shtml
 - 1850-1940
 - Next up: 1950!
- Potential
 - Additional linkages with health surveys and administrative records
 - Examination of early-life environmental exposures on later-life health, demographic, and economic outcomes

Table 1. Number of linked individuals across censuses

Census years	Linked persons		
1850-1860	5,879,611		
1860-1870	8,949,304		
1870-1880	13,829,390		
1880-1900	13,782,479		
1900-1910	30,017,630		
1910-1920	37,682,985		
1920-1930	45,326,985		
1930-1940	52,490,126		
1920-1930	45,326,985		

The American Community Survey

- Replaced the long form of the Decennial Census
 - Demonstration stage: 2000 to 2004
 - Full implementation 2005, group quarters added 2006
- Rolling sample design (point in time)
- Single-year files: 1% of US population
- Multi-year files
 - 1% of population from each of 5 years
 - Best option for very small population subgroups or specific geographic locations

- 1970 present:
 - State
 - Some Metropolitan Areas
 - County Groups
 - Public Use Microdata Areas (PUMAs)
- ACS 5-year files allow for more reliable estimates of smaller geographies

Pre-conference workshop presented by **IPUMS** and **NDîRA**

Topical Coverage

Geography	Group quarters	Housing characteristics	Household economic characteristics	Household appliances	Age
Race	Marital Status	Birthplace	Nationality	Family Interrelationship	Education
Work	Income	Disability	Migration Status	Commuting	Health Insurance

Disability Variables

- Hearing difficulty: deaf or having serious difficulty hearing
- Vision difficulty: blind or having serious difficulty seeing, even when wearing glasses
- Cognitive difficulty: Because of a physical, mental, or emotional problem, having difficulty remembering, concentrating, or making decisions
- Ambulatory difficulty: Having serious difficulty walking or climbing stairs
- Self-care difficulty: Having difficulty bathing or dressing
- Independent living difficulty: Because of a physical, mental, or emotional problem, having difficulty doing errands alone such as visiting a doctor's office or shopping

Pre-conference workshop presented by **IPUMS** and **NDiRA**

IPUMS CPS

Current Population Survey data, 1962 to present

Current Population Survey

- Monthly survey of housing units conducted jointly by the U.S. Census Bureau and the Bureau of Labor Statistics
- Primary source of information about the U.S. labor force
- Fantastic resource for investigating short- and long-term change within and across population subgroups on a variety of topics
 - Fielded monthly since 1976
 - Short-term panel component
- Premier data source for understanding how the COVID-19 pandemic is affecting the well-being of the U.S. population

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Rich Topical Coverage Monthly

Employment

Unemployment

Earnings

Demographics

Socioeconomic Status

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Vast Supplement Coverage

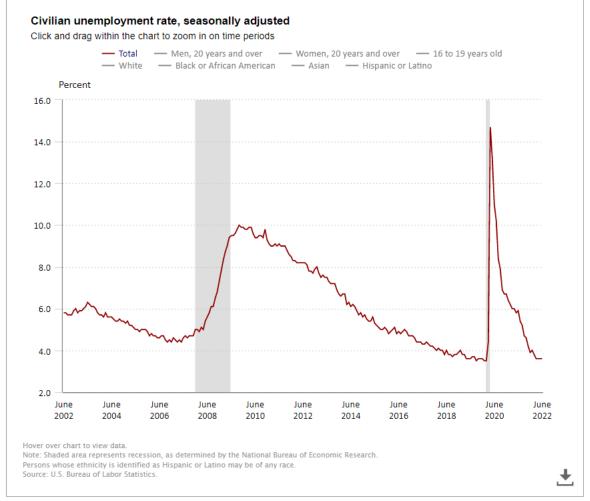
ASEC	Agricultural Worker	Banking	Child Support	Civic Engagement
Computer & Internet Use	Contingent Workers	Contingent Work	Displaced Worker	Education
Fertility	Food Security	Immunization	Job Tenure	Participation in the Arts
Tobacco Use	Veterans	Volunteer	Voting and Registration	Work Schedules

ASEC

- Used extensively by researchers and policy makers
- Contains data on income, poverty, taxes, health insurance, government subsidies, and migration
- Initial focus of IPUMS
 - Would provide information about the years between Censuses
 - Would be as comparable as possible with IPUMS USA

Basic Monthly Data





Linking Opportunities

- Large sample sizes
- Good coverage of subpopulations
- Short-run changes in employment and families; reactions to births and losses (death, divorce)
- Combine rich sources of information about different topics
 - Work and volunteering; veterans and employment

Linking Challenge: Rotation Pattern

- CPS is NOT a longitudinal survey that follows one or several cohorts as they age
- CPS does have a panel component where individuals are observed up to 8 times over 16 months, but the way they move through the survey is in a 4-8-4 rotation
 - 4: In January-April 2019
 - 8: Out May-December 2019
 - 4: In January-April 2020

Pre-conference workshop presented by **IPUMS** and **NDiRA**

IPUMS Time Use

American Time Use Survey data, 2003 to present
US time diary data, 1930 to present
International time diary data, 1965 to present

American Time Use Survey

- Cross sectional data from 2003+
- Ongoing data collection by US Census Bureau for Bureau of Labor Statistics
- Respondents are selected after completing the CPS rotation
- One day, one diary per household
 - But have information about other household members
 - Weekends are oversampled (50/50)

ATUS Time Diary

- 24 hour time diary (4am-4am)
- 6 digit activity coding scheme (400+ activities identified)
- Secondary activities
 - Childcare: all years
 - Eldercare: 2011+
 - Eating: 2006-2008, 2014-2016
 - Drinking: 2006-2008
- Locations/Modes of transportation
- Co-presence of others

What are time diary data?

- Structured narrative accounts of a 24-hour period
- What people do as well as where, with whom (and more recently how people feel about what they do and the use of smart devices)
- Suited to large sample surveys but retains properties of qualitative narratives
- Flexibly capture largely accurate activity reports

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Structure of Time Diary Data

START	STOP	ACTIVITY	LOCATION	WITH WHOM
4:00	6:00	Sleep	n/a	n/a
6:00	6:10	Personal Care	n/a	n/a
6:10	6:15	Dressing	Home	Alone
6:15	6:30	Dressing Household Adult	Home	Spouse
6:30	6:45	Preparing meal	Home	Alone
6:45	6:55	Eating meal	Home	Spouse
6:55	7:30	Cleaning up after meal	Home	Alone
7:30	10:00	Watching television	Home	Spouse
10:00	11:00	Napping	Home	n/a
11:00	11:05	Personal Care	Home	n/a
11:05	12:00	Yardwork	Home	Alone

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Topical Coverage

Physical Adult care Child care Volunteering activity Housework Television Paid work Leisure and home watching repairs Sleep and Spending time Spending time Transportation with others sleeplessness alone

ATUS Modules

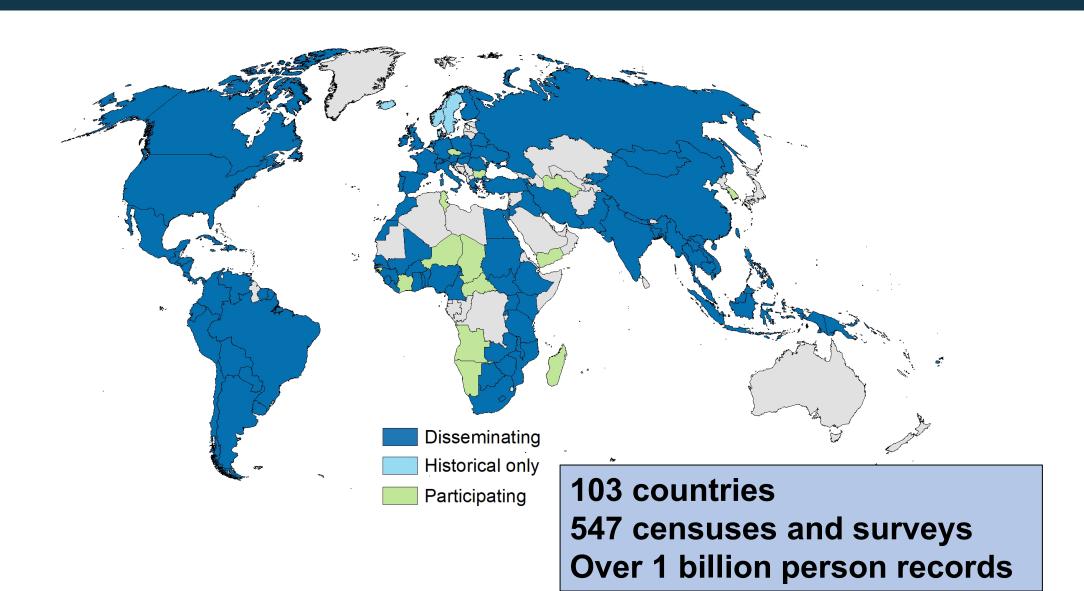
- Eating and Health
 - 2006-2008, 2014-2016
 - Shopping/meal preparation variables, poverty, food program participation
- Well-Being
 - 2010, 2012, 2013, 2021
 - Life satisfaction, momentary well-being
- Leave
 - 2011, 2017-2018
 - Access to and use of paid/unpaid leave

Pre-conference workshop presented by **IPUMS** and **NDiRA**

IPUMS International

Census and labor force data from more than 100 countries

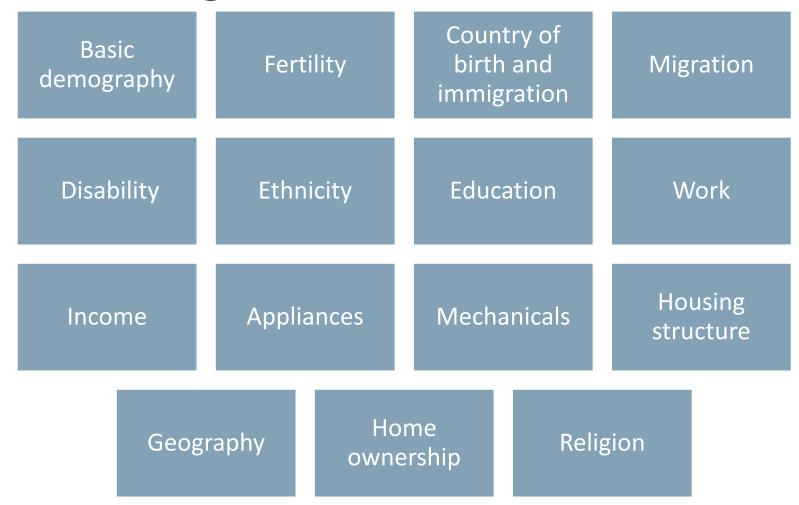
Pre-conference workshop presented by **IPUMS** and **NDiRA**



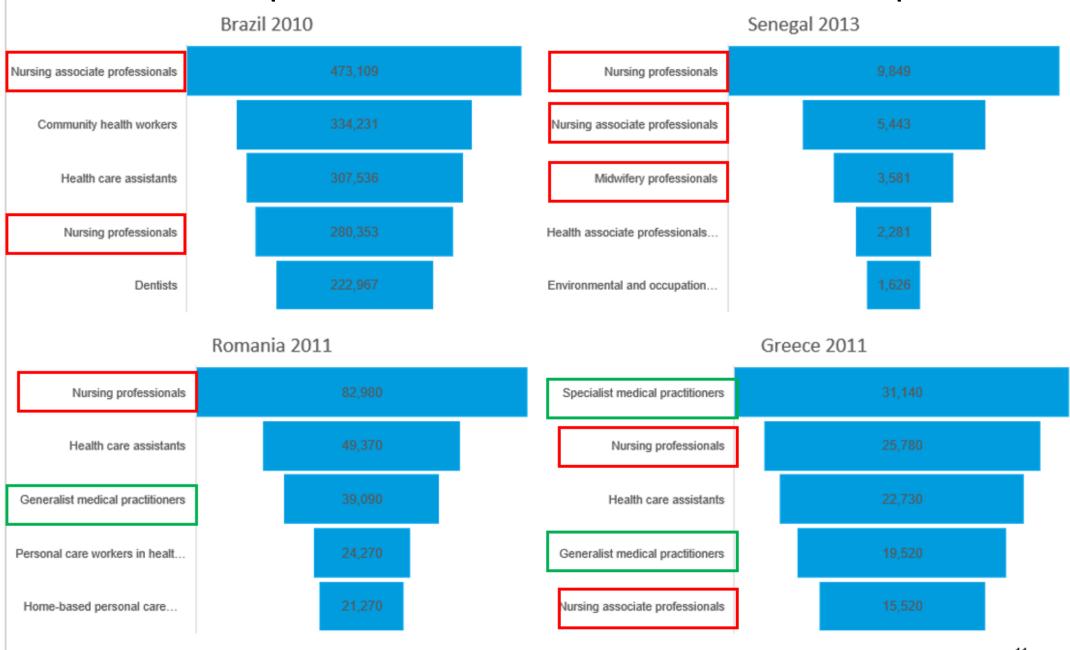
IPUMS INTERNATIONAL NUMBER OF SAMPLES PER COUNTRY 2021											
Argentina	5	Cuba	2	Iceland	5	Mali	3	Peru	2	Suriname	2
Armenia	2	Denmark	2	India NSS	6	Mauritius	3	Philippines	4	Sweden	4
Austria	5	Dominican Republic	5	Indonesia	9	Mexico	8+LFS	Poland	4	Switzerland	4
Bangladesh	3	Ecuador	6	Iran	2	Mongolia	2	Portugal	4	Tanzania	3
Belarus	2	Egypt	2	Iraq	1	Morocco	4	Puerto Rico	6	Thailand	4
Benin	4	Ethiopia	3	Ireland	12	Mozambique	2	Romania	4	Togo	3
Bolivia	4	Fiji	6	Israel	4	Myanmar	1	Russia	2	Trinidad & Tobago	5
Botswana	4	Finland	1	Italy	2 + LFS	Nepal	2	Rwanda	3	Turkey	3
Brazil	6	France	8	Jamaica	3	Netherlands	4	Saint Lucia	2	Uganda	3
Burkina Faso	3	Germany	5	Jordan	1	Nicaragua	3	Senegal	3	Ukraine	1
Cambodia	4	Ghana	3	Kenya	5	Nigeria (GIS)	1	Sierra Leone	2	United Kingdom	9
Cameroon	3	Greece	5	Kyrgyz Republic	2	Norway	5	Slovak Republic	3	United States	14
Canada	11	Guatemala	5	Lao PDR	1	Pakistan	3	Slovenia	1	Uruguay	46
Chile	6	Guinea	3	Lesotho	2	Palestine	3	South Africa	5	Venezuela	4
China	2	Haiti	3	Liberia	2	Panama	6	South Sudan	1	Vietnam	3
Colombia	5	Honduras	4	Malawi	3	Papua New Guinea	3	Spain	4 + LFS	Zambia	3
Costa Rica	5	Hungary	5	Malaysia	4	Paraguay	5	Sudan	1	Zimbabwe	1

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Topical Coverage



WHO example: health workforce occupations

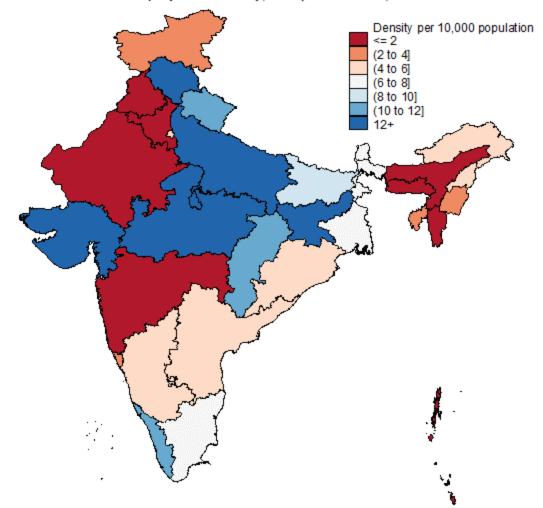


Source: Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D020.V7.0.

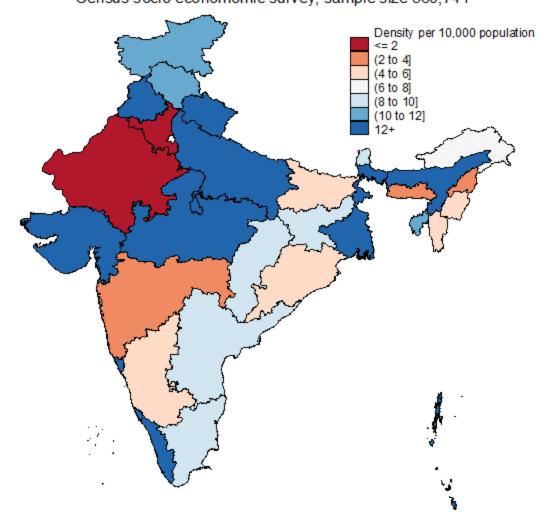
WHO example: change in health workforce (India)

Density of Nursing and Midwifery personnel* in India 2004

Employment survey, sample size 602,833



Census socio economomic survey, sample size 560,741

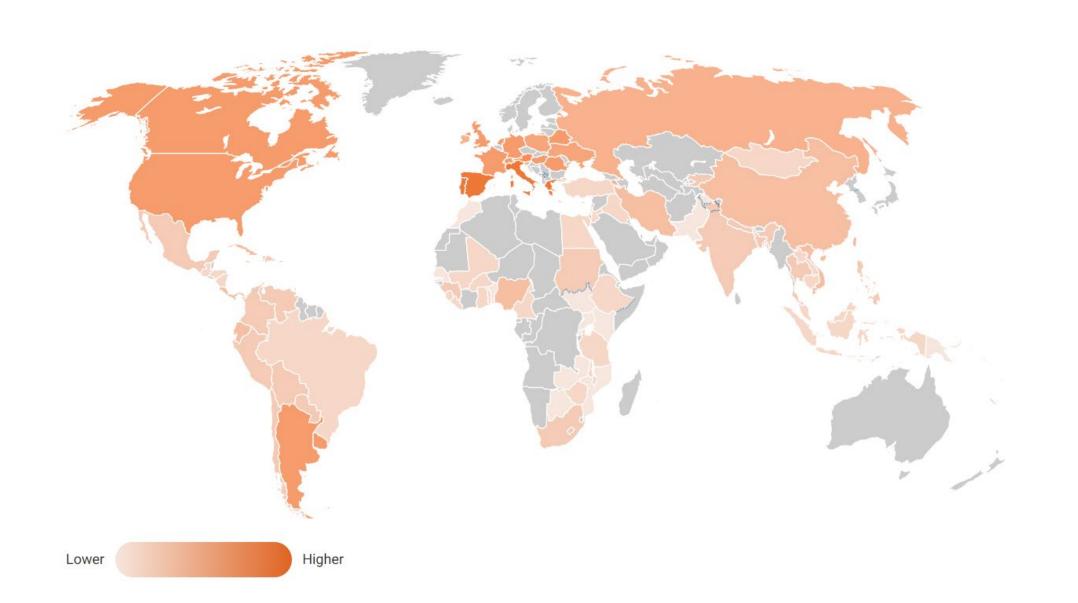


^{*} Professionals and associates

Density of Nursing and Midwifery personnel* in India 2009

^{*} Nurses, Midwives and health visitors

UNFPA example: COVID-19 vulnerability based on age



Pre-conference workshop presented by **IPUMS** and **NDiRA**



Bureau of Statistics, Bangladesh

Central Bureau of Statistics, Sudan



Pre-conference workshop presented by **IPUMS** and **NDiRA**

IPUMS Health Surveys

National Health Interview Survey data from 1960s to present

Medical Expenditure Panel Survey data from 1996 to present



IPUMS.ORG [2] | NHIS [2] | MEPS [2]













HARMONIZED DATA FROM U.S. NATIONAL HEALTH SURVEYS

IPUMS Health Surveys provide free individual-level survey data for research purposes from two leading sources of selfreported health and health care access information: the National Health Interview Survey (NHIS) and the Medical Expenditure Panel Survey (MEPS).

NHIS

NATIONAL HEALTH INTERVIEW SERIES

The National Health Interview Survey (NHIS) provides harmonized annual microdata from the 1960s to the present.

GET DATA

MEPS

MEDICAL EXPENDITURE PANEL SURVEY

The Medical Expenditure Panel Survey (MEPS) provides harmonized microdata from the longitudinal survey of U.S. health care expenditures and utilization, covering the period 1996 to the present.

GET DATA

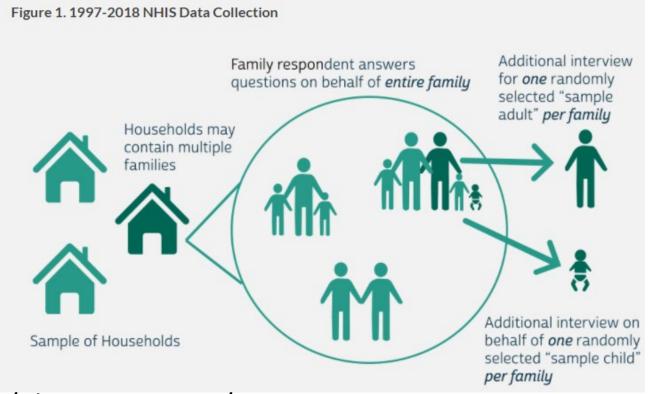
National Health Interview Survey

- Long-running health survey -- data for 1963-present
- Incredible resource for snapshots or repeated series of snapshots of health inequities
- NHIS linkage to admin sources (e.g., NDI, SSA, HUD)

Pre-conference workshop presented by **IPUMS** and **NDiRA**

NHIS is a Complex Survey

• Different sets of questions are fielded to subsets of the sample



http://blog.popdata.org/nhis-1997-2018/

Pre-conference workshop presented by **IPUMS** and **NDiRA**



Format from 1997 - 2018

Sample of households

Households can contain multiple families





Pre-conference workshop presented by **IPUMS** and **NDiRA**



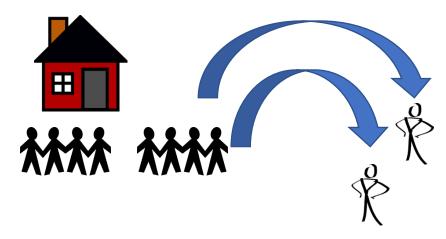
One person – family respondent – interviewed about health of all family members



- Demographics,
- Health status,
- Chronic conditions,
- Health insurance,
- Injuries

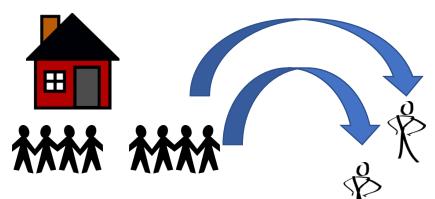
One person – family respondent – interviewed about health of all family members

Pre-conference workshop presented by **IPUMS** and **NDiRA**



One sample adult and sample child per family selected to complete an additional interview

Pre-conference workshop presented by **IPUMS** and **NDiRA**



- BMI,
- Mental Health,
- Select access to care and health behaviors measures
- Supplements
- (For adults) sexual orientation, job info

One sample adult and sample child per family selected to complete an additional interview

NHIS 2019 Redesign, In Brief

- 1997-2018 survey design changed substantially in 2019
- In part motivated by declining response rates across all US household surveys
- Involved extensive cuts to questionnaire content
- Also involved dropping full household data collection so that now data collection is on only one randomly selected adult and child per household

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Topical Coverage

General health	Mortality	AIDS	Cancer testing
Cancer family history	Conditions (e.g., asthma, diabetes, hypertension)	Health behaviors (e.g., smoking, alcohol use, nutrition, exercise)	Health education/knowledge
Limitations	Medical care/access	Health insurance	Vaccinations
Mental health	Alternative health care	Injuries	Sociodemographic characteristics

The NHIS-Linked Mortality File

- The NCHS data linkage center links NHIS survey data records to the NDI
- The NHIS-LMF linkage adds variables on the date and cause of death
- The public use NHIS-LMF are limited to adults age 18 and older at the time of the survey

Medical Expenditure Panel Survey

- Leading source of national estimates of health care expenditures in the United States
- Repeated measures on health, health care utilization and quality, collected in 5 in-person interviews over 2 years
- Data for 1996-forward
- 12-13K households and 30-35K persons interviewed per year
- Re-interviews a subsample of households participating in the previous year's NHIS
 - Can be linked to NHIS using <u>restricted</u> linking keys

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Topical Coverage

General health

Conditions (e.g., cancer, diabetes)

Medical care/access

Medical visits

Health insurance

Medical expenditures

Child mental health, preventative care

Income and taxes

Employment

Detailed information linking conditions and medical care

Sociodemographic characteristics

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Best in class awards

- Paid work: CPS
- Unpaid work: CPS/Time Use
- Health: NHIS
- Global coverage: International
- Geographic detail: USA/International
- Population Subgroups: USA/International

Geography

Time

Target population

Data characteristics

Topics

Pre-conference workshop presented by **IPUMS** and **NDiRA**

Questions?

Pre-conference workshop presented by **IPUMS** and **NDiRA**



U.S. Census and American Community Survey microdata from 1850 to the present.



CPS

Current
Population Survey
microdata
including basic
monthly surveys
and supplements
from 1962 to
the present.



INTERNATIONAL

World's largest collection of census microdata with over 100 countries, contemporary and historical.



U.S. Census and American Community Survey microdata from 1850 to the present.



CPS

Current
Population Survey
microdata
including basic
monthly surveys
and supplements
from 1962 to
the present.



INTERNATIONAL

World's largest collection of census microdata with over 100 countries, contemporary and historical.



TIME USE

U.S. and international time use data for 1965 to the present.



HEALTH SURVEYS

Health survey data from the National Health Interview Survey and the Medical Expenditure Panel Survey.

TIME USE

U.S. and international time use data for 1965 to the present.

HEALTH SURVEYS

Health survey data from the National Health Interview Survey and the Medical Expenditure Panel Survey.