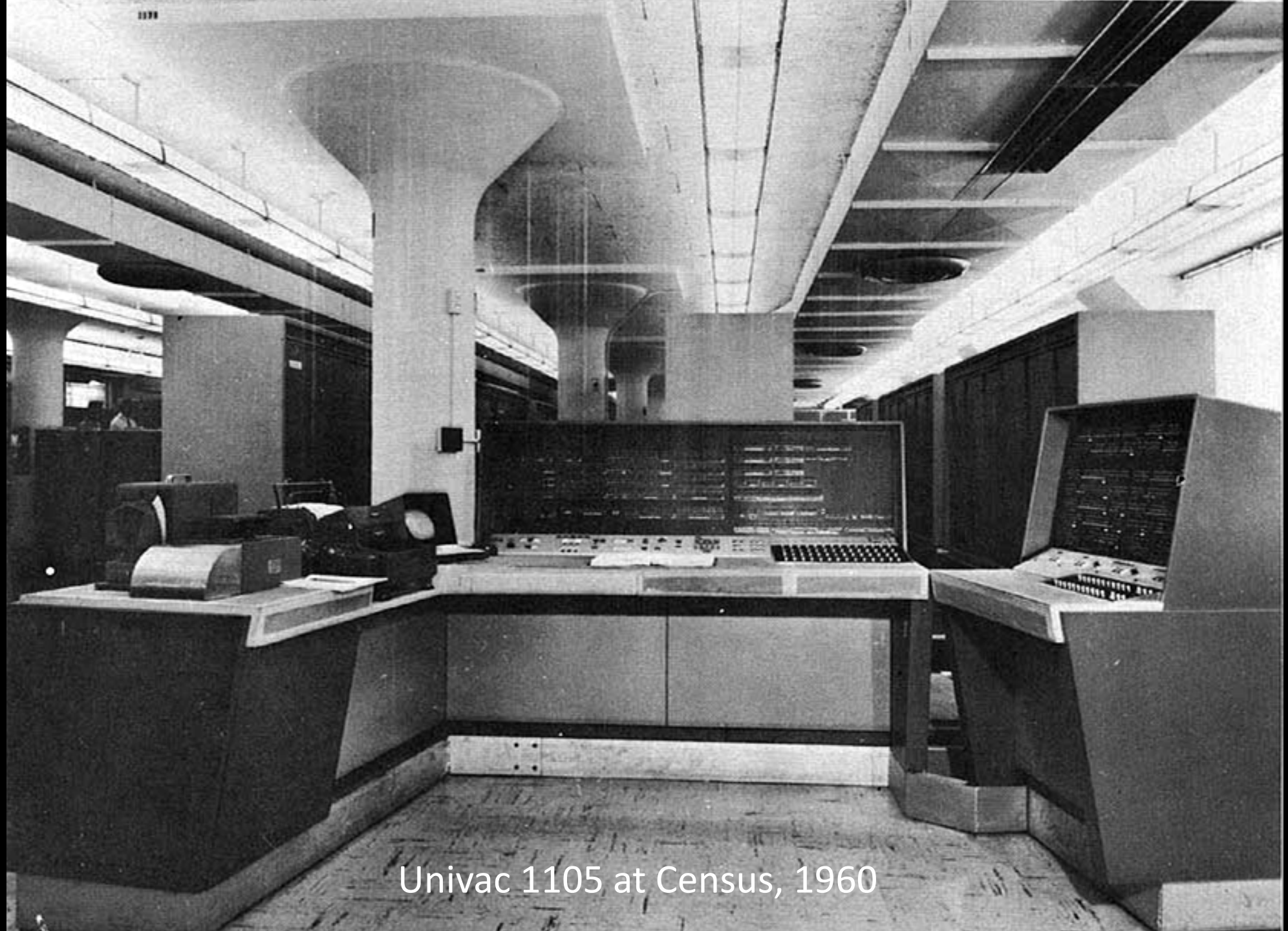
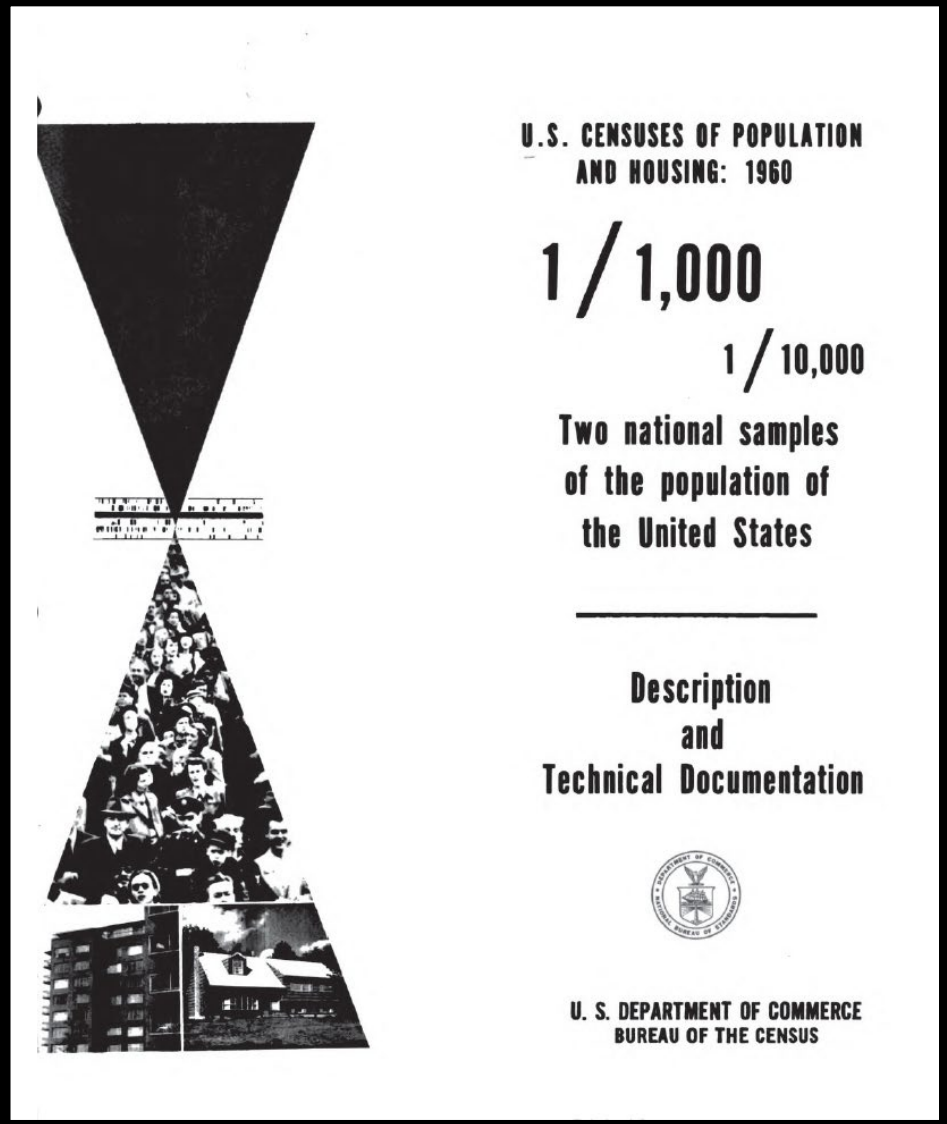
History of IPUMS

July 2022



Univac 1105 at Census, 1960

The Invention of Microdata: The 1960 Census Samples



Cover, 1960 Census Microdata Codebook

Distributed on 13 Univac Tapes
(or 180,000 punch cards)

School enrollment, 1950 Census

Table 42.—SCHOOL AND KINDERGARTEN ENROLLMENT, BY AGE, FOR THE UNITED STATES, URB

[Based on 20-percent sample. For totals of age groups from complete count, see table 38. Percent not shown where less than 0.1 or where b

Age	United States			Urban			Rural nonfarm		
	Population	Enrolled		Population	Enrolled		Population	Enrolled	
		Number	Percent		Number	Percent		Number	Percent
SCHOOL ENROLLMENT									
Total, 5 to 29 years old	58,708,860	28,984,985	49.4	35,928,820	17,178,935	47.8	12,853,930	6,296,570	49.0
5 and 6 years	5,490,200	2,160,160	39.3	3,156,225	1,297,670	41.1	1,301,890	480,065	36.9
7 to 13 years	16,801,950	16,077,270	95.7	9,363,045	9,002,225	96.1	3,921,845	3,744,725	95.5
14 and 15 years	4,267,680	3,963,575	92.9	2,332,850	2,211,225	94.8	963,525	887,380	92.1
16 and 17 years	4,175,195	3,104,265	74.4	2,334,415	1,839,365	78.8	924,350	649,040	70.2
18 and 19 years	4,344,325	1,400,700	32.2	2,686,440	980,435	36.5	932,590	238,415	25.6
20 to 24 years	11,440,095	1,480,745	12.9	7,680,440	1,215,615	15.8	2,352,865	188,195	8.0
25 to 29 years	12,189,415	798,270	6.5	8,375,405	632,400	7.6	2,456,865	108,750	4.4
KINDERGARTEN ENROLLMENT									
Total, 5 and 6 years old	5,490,200	898,970	16.4	3,156,225	742,835	23.5	1,301,890	111,695	8.6

The Power of Microdata

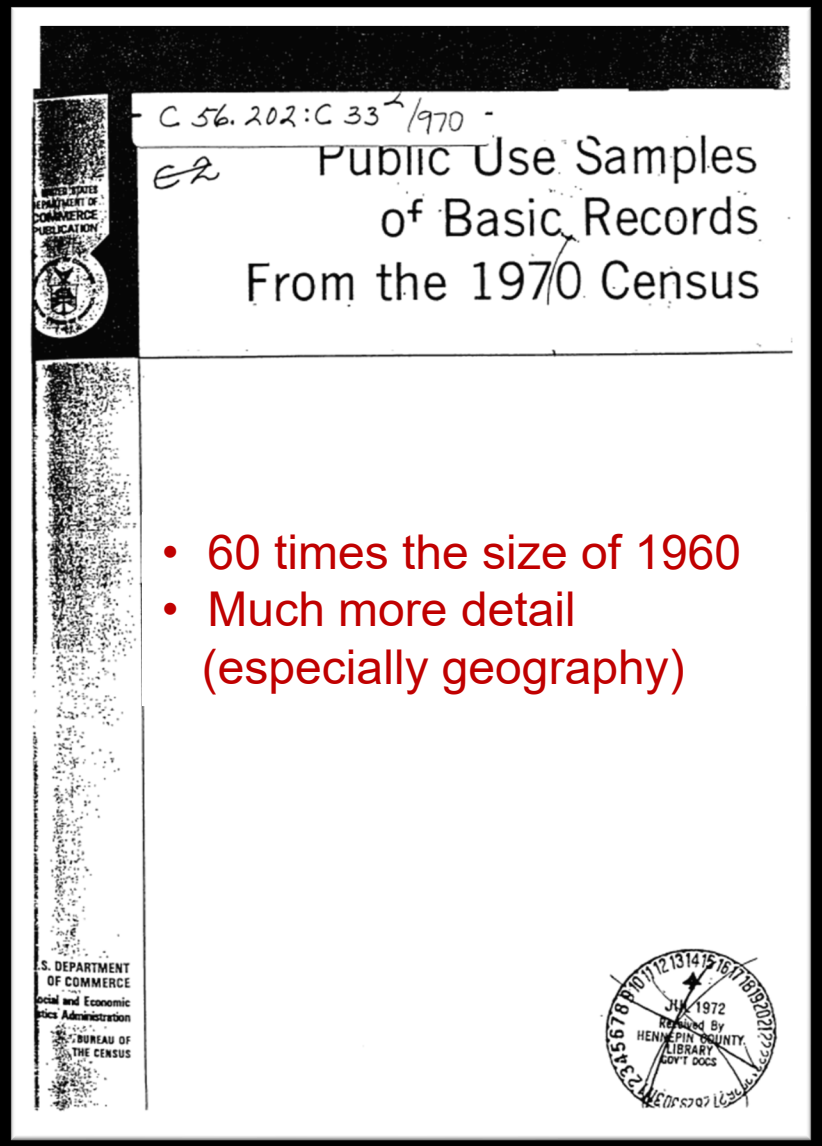
- **Customized measures:** Variables based on combined characteristics of family and household members, capitalizing on the hierarchical structure of the data
- **Multivariate analysis:** Analyze many individual, household, and community characteristics simultaneously
- **Interoperability:** Harmonize data across time and space

Age classification for school enrollment in published U.S. Census

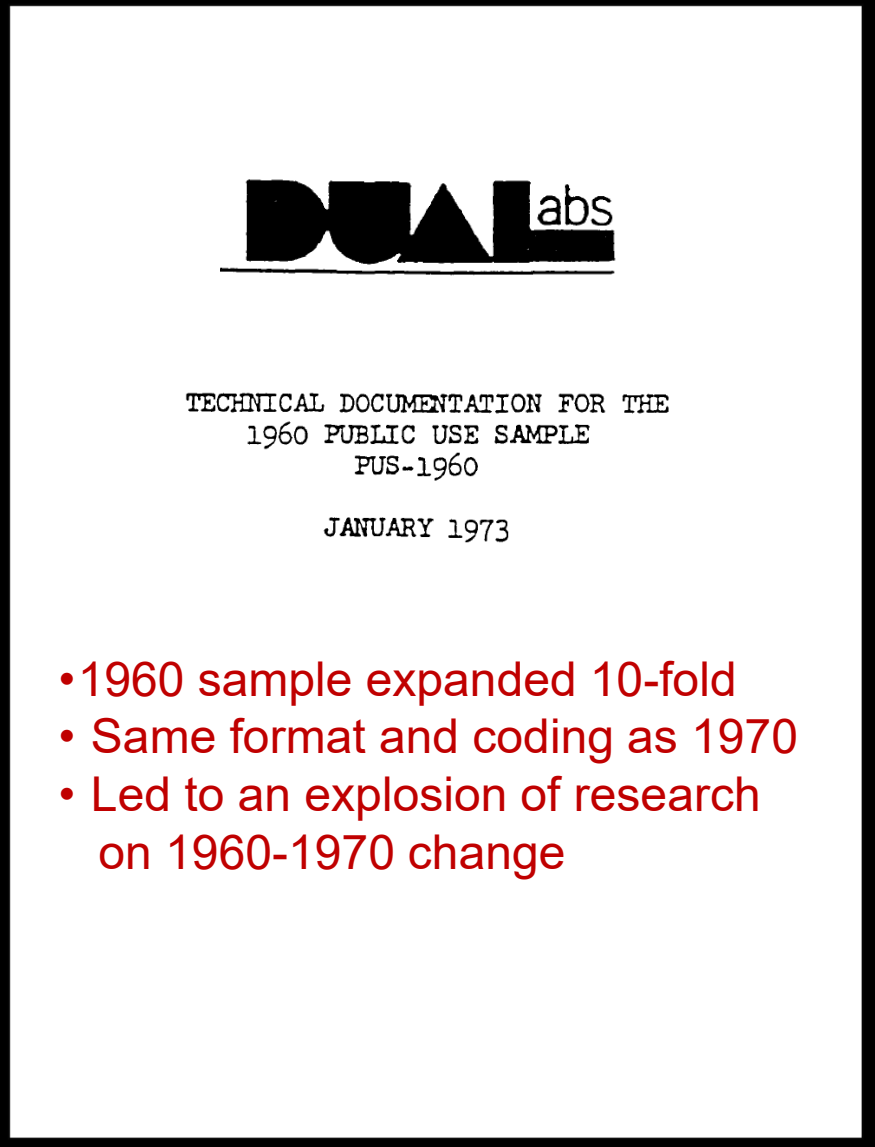
1980	1990	2000
5-6	5-6	5-9
7-13	7-9	10-14
14-15	10-14	15-17
16-17	15-17	

Relationship	Age	Sex	School	Occupation
Head/Householder	34	Male	No	Operative and kindred worke...
Spouse	34	Female	No	Telegraph operators
Child	11	Male	Yes	N/A (blank)
Child	7	Female	Yes	N/A (blank)
Parent-in-Law	76	Female	No	N/A (blank)
Head/Householder	48	Male	No	Carpenters
Spouse	47	Female	No	Bookkeepers
Child	13	Male	Yes	N/A (blank)
Child	11	Female	Yes	N/A (blank)
Head/Householder	54	Female	No	Stenographers, typists, and ...
Parent	82	Male	No	N/A (blank)
Head/Householder	64	Male	No	Laborers (nec)
Spouse	67	Female	No	N/A (blank)
Child	39	Male	No	N/A (blank)
Head/Householder	46	Male	No	Hucksters and peddlers
Spouse	42	Female	No	Barbers, beauticians, and m...
Child	18	Male	No	Laborers (nec)
Child	14	Male	Yes	N/A (blank)

The 1970 Public Use Samples



- 60 times the size of 1960
- Much more detail (especially geography)



- 1960 sample expanded 10-fold
- Same format and coding as 1970
- Led to an explosion of research on 1960-1970 change

Extending the Series Backwards

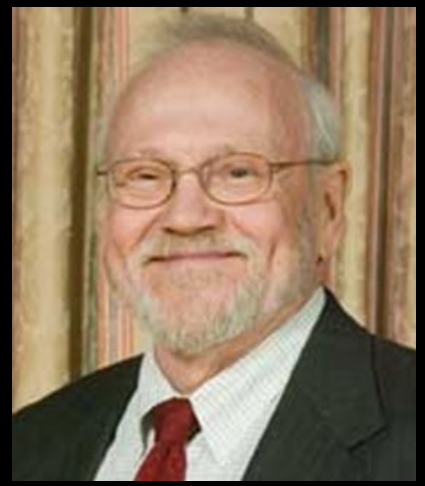
Preston

1900 Public Use Sample

- Microfilm became public in 1972 under 72-year rule
- Sample created 1976-1980 at University of Washington
- 100,000 cases (1-in-760 sample)

1910 Public Use Sample

- Microfilm released 1982
- Created 1983-1989 at the University of Pennsylvania
- 366,000 cases (1-in-250)



Winsborough

1940-1950 Public Use Microdata Samples

- Created 1978-1984 at Census Bureau and University of Wisconsin, with NSF funding
- 3.3 million cases (1-in-100 samples)
- Budget of $7.4M (32.5M in 2021 $)

Page No. 13

Supervisor's Dist. No. 2

Enumeration Dist. No. 96

Note B.—All persons will be included in the Enumeration who were living on the 1st day of June, 1880. No others will. Children BORN SINCE June 1, 1880, will be OMITTED. Members of Families who have DIED SINCE June 1, 1880, will be [illegible]

Received July 22, 1880.

Note C.—Questions Nos. 13, 14, 22 and 23 are not to be asked in respect to persons under 10 years of age.

119

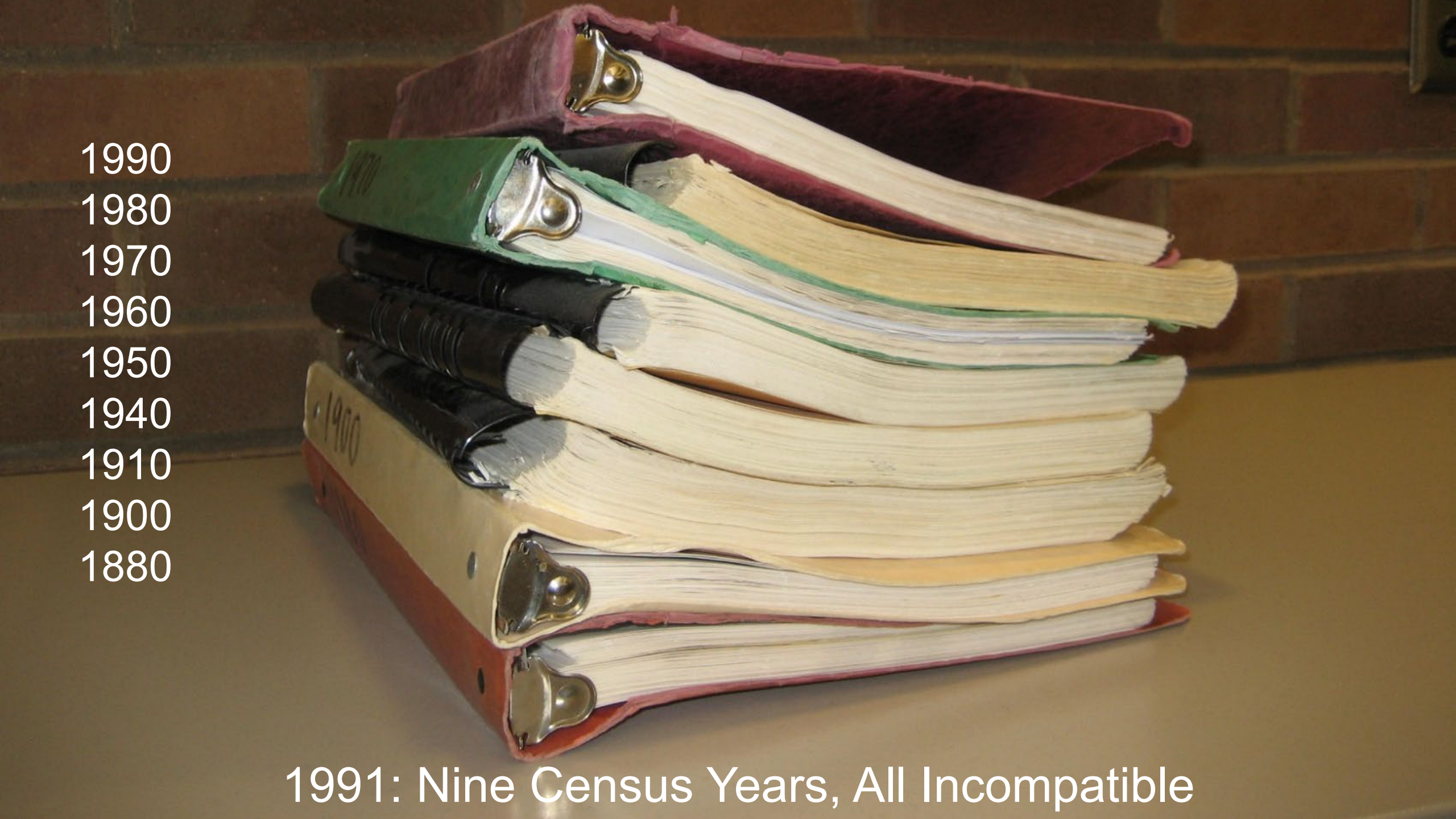
SCHEDULE 1.—Inhabitants in 1st District Princeton township Borough of Princeton, in the County of Mercer, State of New Jersey

enumerated by me on the ____ day of June, 1880.

Chas O Hedrick

Enumerator.

In Cities: Name of Street	In Cities: House Number	1 Dwelling houses numbered in order of visitation	2 Families numbered in order of visitation	3 The Name of each Person whose place of abode, on 1st day of June, 1880, was in this family.	Personal Description: 4 Color—White, W.; Black, B.; Mulatto, Mu.; Chinese, C.; Indian, I.	Personal Description: 5 Sex—Male, M.; Female, F.	Personal Description: 6 Age at last birthday prior to June 1, 1880. If under 1 year, give months in fractions, thus: 2/12.	7 If born within the Census year, give the month.	8 Relationship of each person to the head of this family—whether wife, son, daughter, servant, boarder, or other.	Civil Condition: 9 Single, /.	Civil Condition: 10 Married, /.	Civil Condition: 11 Widowed, /. Divorced, D.	12 Married during Census year, /.	Occupation: 13 Profession, Occupation or Trade of each person, male or female.	Occupation: 14 Number of months this person has been unemployed during the Census year.	Health: 15 Is the person [on the day of the Enumerator's visit] sick or temporarily disabled, so as to be unable to attend to ordinary business or duties? If so, what is the sickness or disability?	Health: 16 Blind, /.	Health: 17 Deaf and Dumb, /.	Health: 18 Idiotic, /.	Health: 19 Insane, /.	Health: 20 Maimed, Crippled, Bedridden, or otherwise disabled, /.	Education: 21 Attended school within the Census year, /.	Education: 22 Cannot read, /.	Education: 23 Cannot write, /.	Nativity: 24 Place of Birth of this person, naming State or Territory of United States, or the Country, if of foreign birth.	Nativity: 25 Place of Birth of the Father of this person, naming the State or Territory of United States, or the Country, if of foreign birth.	Nativity: 26 Place of Birth of the Mother of this person, naming the State or Territory of United States, or the Country, if of foreign birth.
		165	188	Schenck Daniel	Wm	M	40				1			Laborer											New Jersey	New Jersey	New Jersey
				Phoebe	Wm	F	32		wife		1			Keeping House									1	1	New Jersey	N.J.	N.J.
		166	189	Johnson Clara	B	F	5-1					1		Keeping House									1	1	New Jersey	N.J.	N.J.
				Jane	B	F	29		Daughter	1				Servant											New Jersey	Kentucky	N.J.
				Mary E	B	F	21		Daughter	1				At Home											New Jersey	Penn^a	N.J.
				Georgiana	Wm	F	14		Daughter	1				at Home									1		New Jersey	Penn^a	N.J.
				Viola	Wm	F	2		G. Daughter																New Jersey	Penn^a	N.J.
				Mary E	Wm	F	10/12		G. Daughter																New Jersey	Penn^a	N.J.
				[illegible]																							



1991: Nine Census Years, All Incompatible

Relationship Variable (part):

1900 Public Use Sample

72 categories

0 P03 REL RELATIONSHIP TO HEAD COLS 9-11

100	HEAD OF HOUSEHOLD	21336	21.243
108	PARTNER / COHEAD	173	.172
120	WIFE OF HEAD	16665	16.592
128	WIFE OF PARTNER/COHEAD	1	.001
129	SECOND OR THIRD WIFE OF HEAD	3	.003
130	CHILD OF HEAD	46174	45.973
131	STEP-CHILD OF HEAD	755	.752
132	ADOPTED CHILD OF HEAD	103	.103
133	SON/DAUGHTER-IN-LAW	466	.464
136	FOSTER CHILD / FOUNDLING	23	.023
140	HUSBAND / NOT HEAD	17	.017
200	RELATIVE - UNSPECIFIED	23	.023
210	PARENT OF HEAD	920	.916
211	STEP-PARENT OF HEAD	24	.024
213	PARENT-IN-LAW OF HEAD	568	.566
220	BROTHER/SISTER OF HEAD	1325	1.319
221	STEP/HALF BROTHER/SISTER	12	.012
223	BROTHER/SISTER-IN-LAW	688	.685
230	NIECE/NEPHEW	822	.818
232	ADOPTED NIECE/NEPHEW	1	.001
233	NIECE/NEPHEW-IN-LAW	4	.004
237	GRAND NIECE/NEPHEW	15	.015
240	COUSIN	108	.108
243	COUSIN-IN-LAW	1	.001
249	SECOND COUSIN	5	.005
250	AUNT/UNCLE OF HEAD	99	.099
253	AUNT/UNCLE-IN-LAW	2	.002
260	GRANDPARENT OF HEAD	27	.027
261	STEP-GRANDPARENT	1	.001
263	GRAND-PARENT-IN-LAW	2	.002
270	GRANDCHILD OF HEAD	1541	1.534
271	STEP-GRANDCHILD	33	.033

Relationship Variable:

1940 Public Use Sample

23 categories

NAME: RELATE

ITEM DESCRIPTION: Detailed relationship to household head
RECORD TYPE: P
BEGIN: 11
SIZE/SCALE: 2
UNIVERSE: All persons in households
SEE GLOSSARY TERMS: Relationship
SOURCE: Columns 8 and A of the population schedule

CODE	DESCRIPTION
1	Head
2	Wife
3	Son, daughter
4	Stepson, stepdaughter
5	Son-in-law, daughter-in-law
6	Grandson, granddaughter
7	Father, mother, stepfather, stepmother
8	Father-in-law, mother-in-law
9	Grandfather, grandmother, including in-laws
10	Brother, sister, stepbrother, stepsister, half brother, half sister
11	Brother-in-law, sister-in-law
12	Uncle, aunt, including in-laws
13	Nephew, niece, including in-laws
14	Cousin, including in-laws
15	Relative of head, n.e.c.
16	Roomer, boarder, lodger or his/her relative
17	Domestic employee
18	Nondomestic employee
19	Relative of employee
20	Partner or friend and his/her relatives
21	Ward, foster child
22	Nonrelative of head, n.e.c.
99	Inap. (group quarters)

Relationship Variables:

1980 Public Use Sample

20 unique categories

RELAT1		2	2
			Relationship
	00		Householder
			Family member other than householder:
	01		Spouse
	02		Child
	03		Brother or sister
	04		Parent
	05		Other relative (See RELAT2)
			Person not related to householder:
	06		Roomer or boarder
	07		Partner or roommate
	08		Paid employee
	09		Other nonrelative
			In group quarters:
	10		Inmate
	11		Noninmate
RELAT2		1	4
			Detailed Relationship
	0		N/A (person not listed as "other relative" of householder)
	1		Son-in-law or daughter-in-law
	2		Grandchild
	3		Father-in-law or mother-in-law
	4		Brother-in-law or sister-in-law
	5		Nephew or niece
	6		Grandparent
	7		Uncle or aunt
	8		Cousin
	9		Other person related by blood or

1991 IPUMS proposal

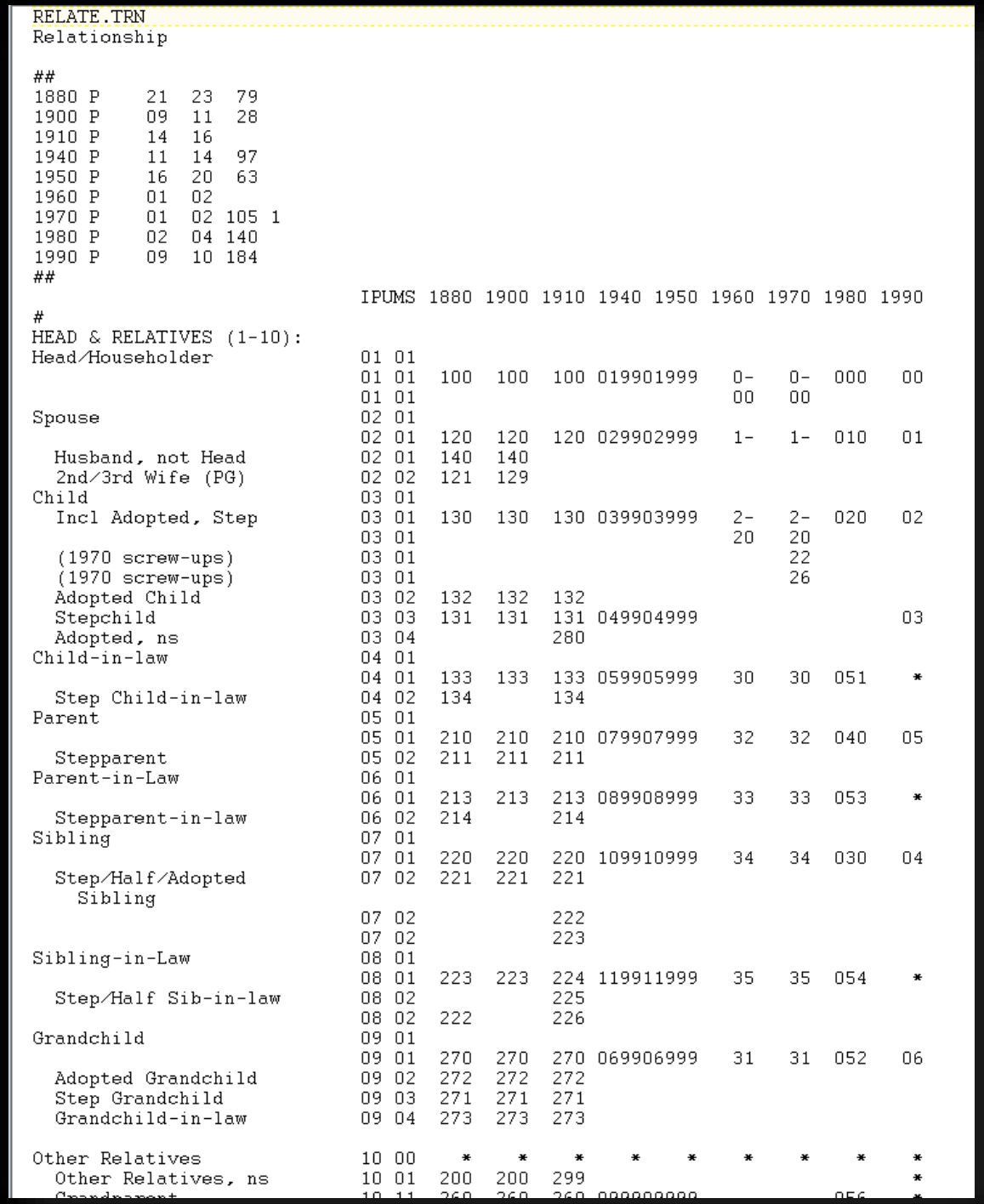
An integrated database for
1880, 1900, 1910, 1940, 1950, 1960, 1970, 1980, 1990

- ✓Harmonized codes
- ✓Consistent record layout
- ✓Integrated documentation
- ✓No loss of information.

Two Key IPUMS innovations

1991: First structured metadata system for data integration

1995: First interactive data access system, enabling pooling multiple datasets, variable selection, and subsetting



```
RELATE.TRN
Relationship

##
1880 P     21   23   79
1900 P     09   11   28
1910 P     14   16
1940 P     11   14   97
1950 P     16   20   63
1960 P     01   02
1970 P     01   02  105 1
1980 P     02   04  140
1990 P     09   10  184
##
                              IPUMS 1880 1900 1910 1940 1950 1960 1970 1980 1990
#
HEAD & RELATIVES (1-10):
Head/Householder              01 01
                              01 01  100  100  100 019901999   0-   0-  000   00
                              01 01                            00   00
Spouse                        02 01
                              02 01  120  120  120 029902999   1-   1-  010   01

  Husband, not Head           02 01  140  140
  2nd/3rd Wife (PG)           02 02  121  129
Child                         03 01
  Incl Adopted, Step          03 01  130  130  130 039903999   2-   2-  020   02
                              03 01                            20   20
  (1970 screw-ups)            03 01                                 22
  (1970 screw-ups)            03 01                                 26
  Adopted Child               03 02  132  132  132
  Stepchild                   03 03  131  131  131 049904999                  03
  Adopted, ns                 03 04       280
Child-in-law                  04 01
                              04 01  133  133  133 059905999   30   30  051    *

  Step Child-in-law           04 02  134       134
Parent                        05 01
                              05 01  210  210  210 079907999   32   32  040   05

  Stepparent                  05 02  211  211  211
Parent-in-Law                 06 01
                              06 01  213  213  213 089908999   33   33  053    *

  Stepparent-in-law           06 02  214       214
Sibling                       07 01
                              07 01  220  220  220 109910999   34   34  030   04

  Step/Half/Adopted           07 02  221  221  221
    Sibling
                              07 02             222
                              07 02             223
Sibling-in-Law                08 01
                              08 01  223  223  224 119911999   35   35  054    *

  Step/Half Sib-in-law        08 02             225
                              08 02  222        226
Grandchild                    09 01
                              09 01  270  270  270 069906999   31   31  052   06

  Adopted Grandchild          09 02  272  272  272
  Step Grandchild             09 03  271  271  271
  Grandchild-in-law           09 04  273  273  273

Other Relatives               10 00    *    *    *    *    *    *    *    *    *
  Other Relatives, ns         10 01  200  200  299                            *
  Grandparent                 10 11  260  260  260 000000000                  056    *
```

1991: First structured metadata system for data integration

1991: First structured metadata system for data integration

```
RELATE.TRN
Relationship

##
1880 P     21   23   79
1900 P     09   11   28
1910 P     14   16
1940 P     11   14   97
1950 P     16   20   63
1960 P     01   02
1970 P     01   02  105 1
1980 P     02   04  140
1990 P     09   10  184
##

#
HEAD & RELATIVES (1-10):
Head/Householder


Spouse

  Husband, not Head
  2nd/3rd Wife (PG)
Child
  Incl Adopted, Step

  (1970 screw-ups)
  (1970 screw-ups)
  Adopted Child
  Stepchild
  Adopted, ns
Child-in-law

  Step Child-in-law
Parent
```

```
IPUMS 1880 1900 1910 1940 1950 1960 1970 1980 1990

01 01
01 01  100  100  100 019901999   0-   0-  000   00
01 01                            00   00
02 01
02 01  120  120  120 029902999   1-   1-  010   01
02 01  140  140
02 02  121  129
03 01
03 01  130  130  130 039903999   2-   2-  020   02
03 01                            20   20
03 01                                 22
03 01                                 26
03 02  132  132  132
03 03  131  131  131 049904999                  03
03 04            280
04 01
04 01  133  133  133 059905999   30   30  051    *
04 02  134       134
05 01
```

Input data locations

Data quality flag locations

Original codes

Standardized composite code

Standardized labels

Excerpt from budget justification
1991 IPUMS proposal to NSF

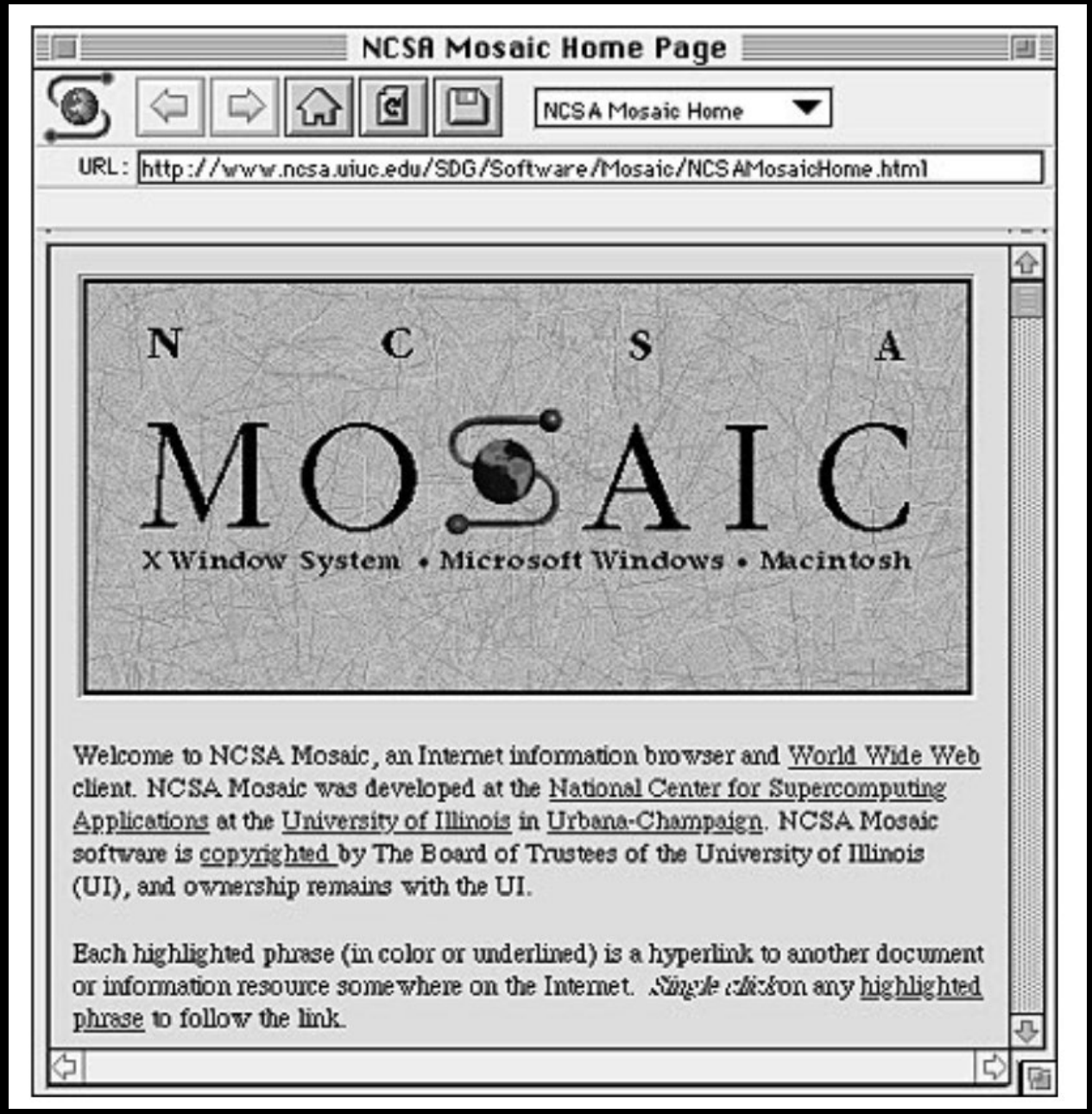
Materials and Supplies. The largest category of supplies is 9-track tapes. Despite their inefficiency for very large data sets, 9-track tapes are presently the only medium used by the Inter-University Consortium for importing and exporting data. Accordingly, we are requesting funds for 150 tapes in the first year and 300 tapes in the final year. For internal storage and backup we will rely on 2 Gbyte Exabyte tape cartridges.

Description of supplies	Year 1	Year 2	Year 3	Total
450 2400' 6250BPI 9-Track tapes	2,048	-	4,095	6,143
200 2 Gbyte Exabyte tape cartridges	600	360	240	1,200
300 1.44 Mbyte 3.5" diskettes	189	189	189	567
Printing supplies	260	260	260	780
General office supplies	500	525	550	1,575
Total	3,597	1,334	5,334	10,265

FTP log for November 19, 1993:
First IPUMS data dissemination via anonymous FTP

```
Mon Nov  8 10:01:38 1993 1 quads.uchicago.edu 742 /pub/gzip/README a
Mon Nov  8 10:21:00 1993 1116 quads.uchicago.edu 5552885 /raw1850/raw
Tue Nov  9 22:57:22 1993 1 ukanaix.cc.ukans.edu 872 /README a _ o a W
Tue Nov  9 22:58:56 1993 1 ukanaix.cc.ukans.edu 6455 /raw1850/coll850
Wed Nov 10 11:13:53 1993 1 ukanaix.cc.ukans.edu 872 /README a _ o a W
Mon Nov 15 07:43:06 1993 4 131.130.42.102 3460 /nfs/legohead/files0/h
Mon Nov 15 10:33:32 1993 1 ukanaix.cc.ukans.edu 872 /README a _ o a W
Mon Nov 15 13:09:51 1993 1 larry.hist.umn.edu 3460 /nfs/legohead/file
Mon Nov 15 19:56:26 1993 1 ukanaix.cc.ukans.edu 872 /README a _ o a W
Wed Nov 17 18:29:02 1993 1 legohead.hist.umn.edu 742 /pub/gzip/README
Fri Nov 19 11:12:12 1993 876 ucracl.ucr.edu 8909034 /pub/ipums/1880/ip
Fri Nov 19 11:47:55 1993 1 larry.hist.umn.edu 36434 /nfs/legohead/file
Fri Nov 19 14:52:40 1993 12866 ucracl.ucr.edu 10370306 /pub/ipums/188
Sat Nov 20 02:17:29 1993 1 ukanaix.cc.ukans.edu 28 /pub/.message a _ c
Sat Nov 20 02:18:39 1993 1 ukanaix.cc.ukans.edu 6455 /raw1850/coll850
Sat Nov 20 16:53:06 1993 1 ukanaix.cc.ukans.edu 872 /README a _ o a W
```

On November 11, 1993, eight days before the first IPUMS Internet download, an undergraduate from the University of Illinois released the first successful web browser for a PC:
NCSA Mosaic 1.0



Welcome to the Social History Research Lab!

University of Minnesota

This web site currently contains the data and documentation for the **Integrated Public Use Microdata Series (IPUMS)**. The **IPUMS** is a database consisting of 23 samples of the U.S. Census from 1850 to 1990. The **IPUMS** assigns the different samples consistent codes and integrates their documentation. At present, you may only download compressed **IPUMS** data files. To obtain a DOS decompression program click here.

If you need uncompressed data or have any other questions about the IPUMS database, contact us at **ipums@atlas.socsci.umn.edu**

If you have problems. questions, or suggestions about this page, send e-mail to **block@torgo.hist.umn.edu**

The User's Guide for the IPUMS is available online in Word 6.0 format (see below). This document is 800 pages long, and is contained in 61 separate files. A printed version is available from us for $30; contact **ipums@atlas.socsci.umn.edu**

IPUMS DATA FILES:

To download an IPUMS sample, just click on it:

1850 sample (6.9M)

1880 sample (21.1M)

1900 sample (3.6M)

1910 sample (14.0M)

1920 sample (14.4M)

1940 sample (62.4M)

1950 sample (72.8M)

1960 sample (77.4M)

March 10,1995: The first IPUMS website

Among the first 15,000

IPUMS Team, 1995

IPUMS Data Extract

Sample Selection

Reset Values | Continue To Variable Selection

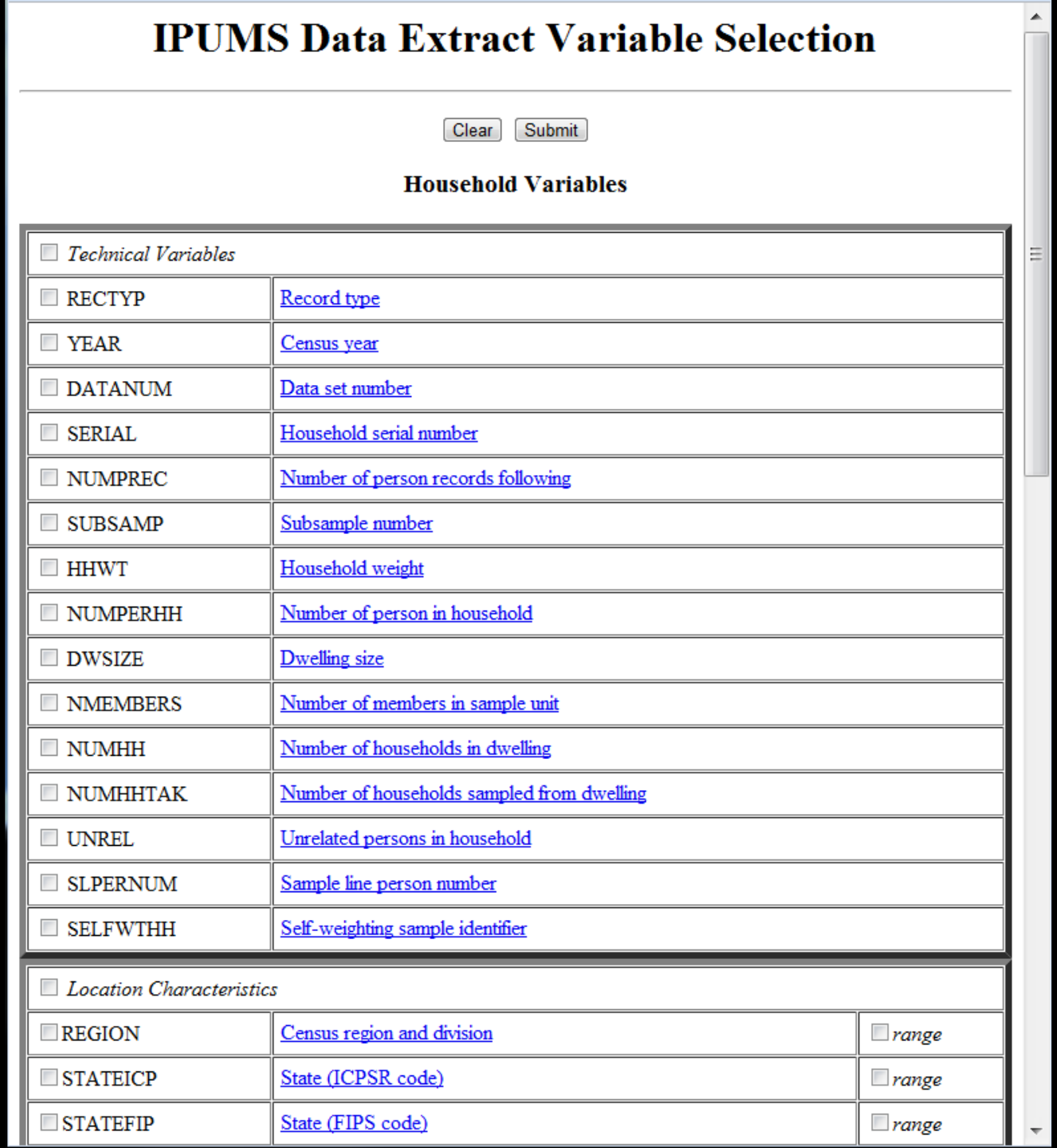
What is your email address?

Census Year NOTE: IPUMS-95 contains only the 1970 5% and 15% state and county group samples, the 1980 B sample, and the 1990 1% sample.	☐ 1850 Sample ☐ 1880 Sample ☐ 1900 Sample ☐ 1910 Sample ☐ 1920 Sample ☐ 1940 Sample ☐ 1950 Sample ☐ 1960 Sample ☐ 1970 5% State Sample ☐ 1970 15% State Sample ☐ 1980 B Sample ☐ 1990 1% Sample
Sample Density	◉ Tiny ○ Small ○ Regular
File Type	○ Flat ◉ Hierarchical
Data Quality Flags	☐ Include all data quality flags

(ip´-ums) IPUMS

November 5, 1995: Interactive data extract system

Step 1: select census years, sample density, and desired format



IPUMS Data Extract Variable Selection

Clear Submit

Household Variables

Technical Variables		
RECTYP	Record type	
YEAR	Census year	
DATANUM	Data set number	
SERIAL	Household serial number	
NUMPREC	Number of person records following	
SUBSAMP	Subsample number	
HHWT	Household weight	
NUMPERHH	Number of person in household	
DWSIZE	Dwelling size	
NMEMBERS	Number of members in sample unit	
NUMHH	Number of households in dwelling	
NUMHHTAK	Number of households sampled from dwelling	
UNREL	Unrelated persons in household	
SLPERNUM	Sample line person number	
SELFWTHH	Self-weighting sample identifier	

Location Characteristics		
REGION	Census region and division	*range*
STATEICP	State (ICPSR code)	*range*
STATEFIP	State (FIPS code)	*range*

Step 2:
select variables

IPUMS Data Extract Case Selection

Clear Submit

Household Variables

Region	☐ New England ☐ Middle Atlantic ☐ East North Central ☐ West North Central ☐ South Atlantic ☐ East South Central ☐ West South Central ☐ Mountain ☐ Pacific ☐ Military/Military reservations ☐ PUMS boundaries cross state lines ☐ State not identified
State	Alabama Alaska Arizona Arkansas California Colorado Connecticut Delaware District of Columbia Florida

Person Variables

Age	from 0 to All
Sex	☐ Male ☐ Female
Race	☐ White ☐ Black/Negro ☐ American Indian ☐ Chinese

Step 3:
Subset
population

RELATE - **P 48-49 General**

P 48-51 Detailed

Relationship to household head/householder

Availability:

1850	1860	1870	1880	1900	1910	1920	1940	1950	1960	1970	1980	1990
			X	X	X	X	X	X	X	X	X	X

Universe:

All persons.

Codes and Frequencies

Description:

RELATE describes an individual's relationship to the head of household or householder. Beginning in 1880, data on household relationship was asked of every person. The *general* relationship code is reasonably comparable across years. The *detailed* code makes distinctions that cannot be made in all years.

The relationship codes are divided into two categories—relatives (codes 1-10) and non-relatives (codes 11-12). The codes for relatives are self-explanatory; the non-relative codes are divided into three groups: "Partner, Friend, Visitor," roughly described as persons who do not pay or work for their accommodations (unless they share ownership), "Other Non-Relatives" including those persons paying or working for accommodations, and "Institutional Inmates." See the comparability discussion for further information about the coding scheme.

Comparability:

The general code is comparable across all years. Users should note, however, that there are some fundamental differences between the early period (before 1940) and the later period (1940-1990). Group quarters residence is a primary distinction in the relationship variable for the later period, but before 1940 relationship to head was recorded regardless of group quarters status. Persons classified as related to the head (codes 1 through 10) in the early period would have been classified in the 'Other non-relative" category based on their group quarters status in the later years. We decided not to impose consistency because it would have resulted in the loss of too much information in the early period.

Users may recode RELATE for higher comparability, but we recommend caution because a recode may lose considerable information, changing what were considered ordinary households in the early years into group quarters. To impose the 1970 group quarters definition on all years, create

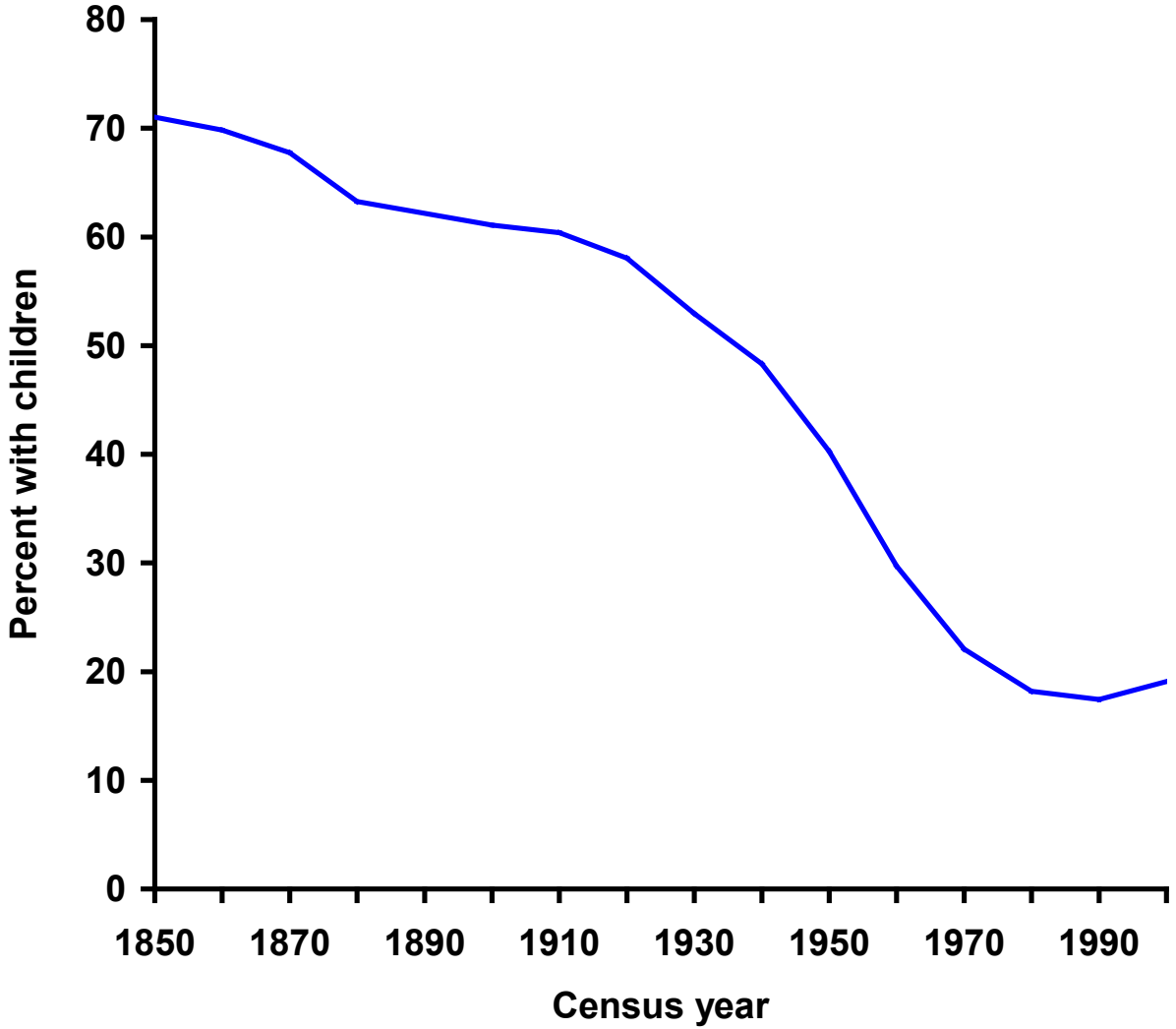
1997:

Hypertext variable-level documentation accessible from every stage of the extract system

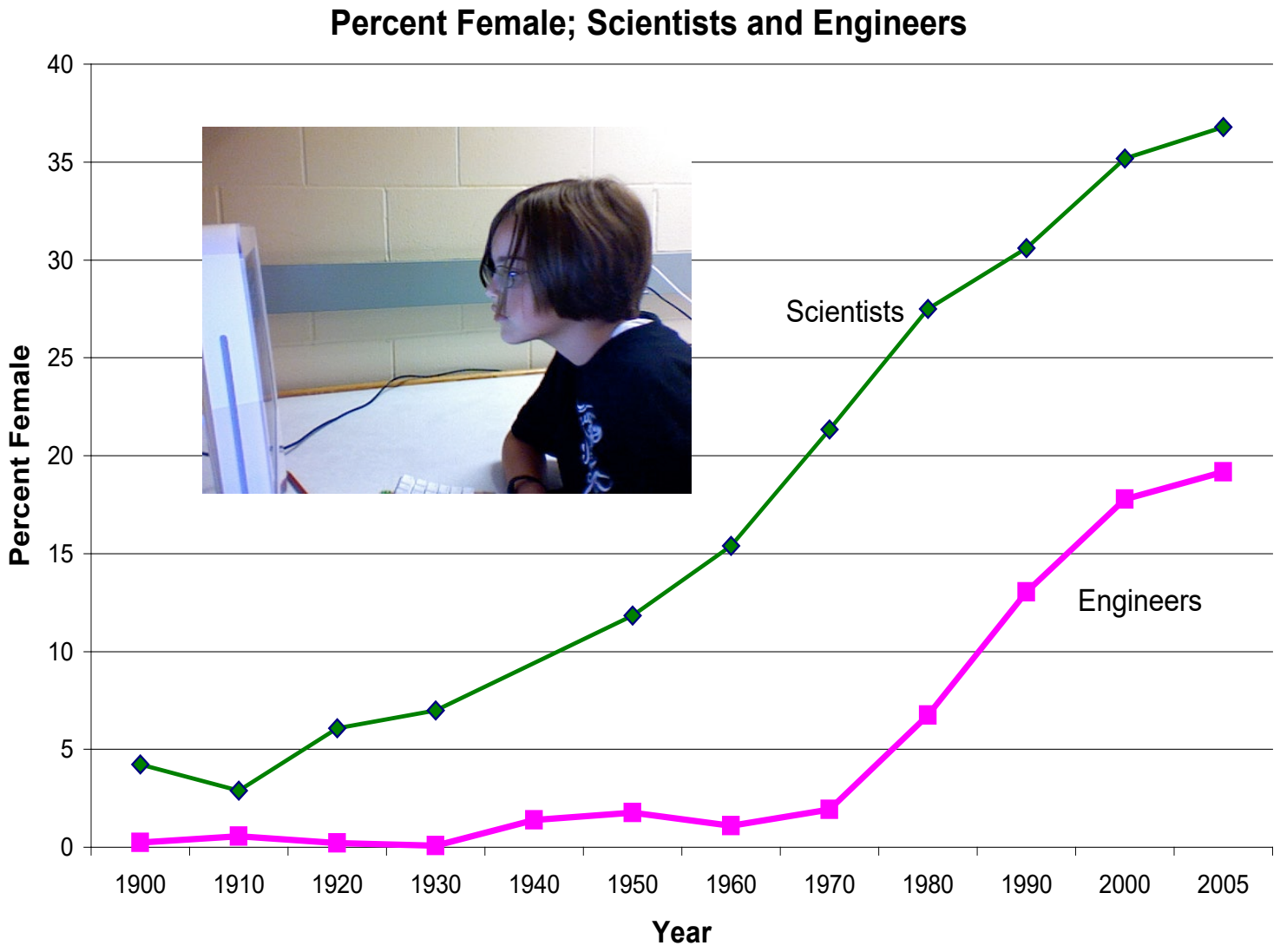
1992-2006

- Missing census years added (1850, 1860, 1870, 1920, 1930)
- Preston samples expanded (1900, 1910)

Percent of elderly (65+) residing with own children: United States, 1850-2000



IPUMS Graph from “A Century of Women in Science and Engineering,” 2006 History Day project by Abby Norling-Ruggles, age 12



IPUMS Big Data Developments of 1999

- IPUMS International
- Full-Count Microdata

1999: IPUMS International

IPUMS-*International*

International Integrated Microdata Access System

- Sun Microsystems Provides Equipment Grant (Press Release 4-26-00)
- Announcement of Funding (Press Release 9-15-99)
- Summary
- Project Proposal to NSF
- Guidelines for Collaboration
- Data Security Policy
- Contacts
- Microdata Inventory
- Principles of Harmonization
- Prototype Application to Use the Microdata
- Task Force on the 2000 PUMS

Major funding provided by the National Science Foundation

IPUMS-International is affiliated with the IPUMS project


Khartoum, CBS-Sudan

1973 Sudan Census Tapes arrive

Dhaka, Bangladesh
Bureau of Statistics


omputer·Link


Computer-Link

Participating Countries

Starting in 1999, NAPP added historical data 1703-1930 for Iceland, Ireland, Germany, Norway, Sweden, and the UK

NAPP

North Atlantic Population Project

Home | Create an Extract | Variables | FAQ | Contact Us

The Project

What is NAPP?
Progress Report
Release Dates
Revision History
FAQ

Data

NAPP Registration
Create Data Extract
Download Data Extract
Analyze Data Online
Citation and Use

Documentation

Samples
Variables
Census Questionnaires
Other Resources

Contact Us

Feedback
NAPP Participants

Welcome to NAPP

North Atlantic Population Project News: In May 2009, NAPP added a 2% sample of the **1851 Census of Great Britain.**

IPUMS linked representative samples: The IPUMS database links cases from the 1880 U.S. census to 1% samples of all other U.S. censuses from 1850 to 1930.

Nappsters in Grundarfjördur, Iceland

IPUMS Full-Count Census Projects

Census	Project start	Major contributors
1880	1999	Church of Latter-Day Saints
1960-2000	2001	Census Bureau
1850	2009	FamilySearch
1940	2012	Ancestry
1860-1870	2015	Ancestry
1900-1930	2015	Ancestry/FamilySearch
1950	2021	Ancestry/FamilySearch

Church News
THE CHURCH OF JESUS CHRIST OF LATTER-DAY SAINTS

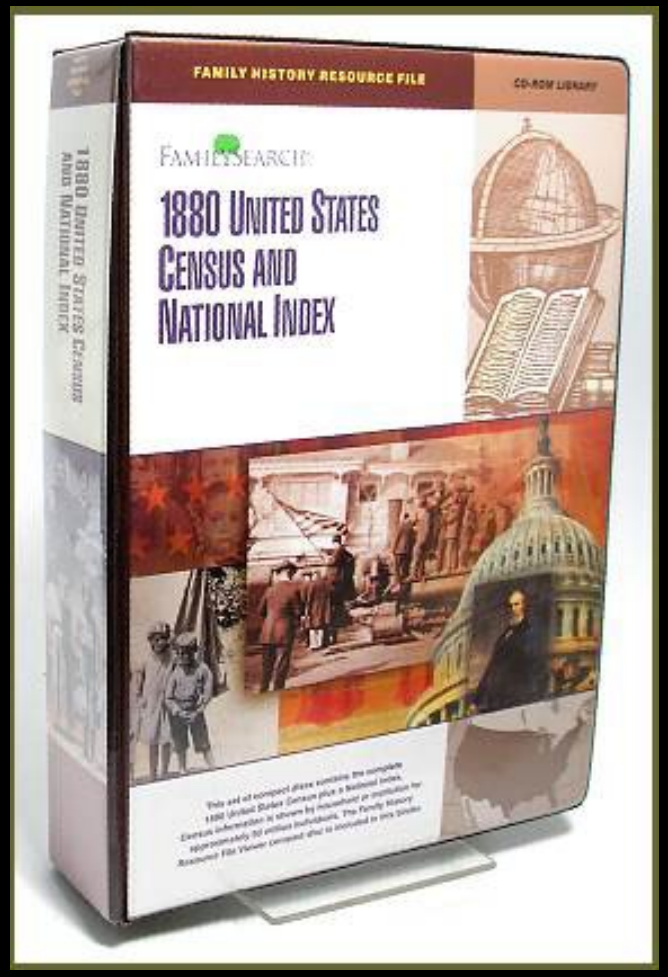
1880 U.S. Census to be available on CDs

Published: Saturday, May 26, 2001

The U.S. 1880 Census, the first ever U.S. census published in its complete format, has been automated and the 56-CD-ROM set will be available June 1.

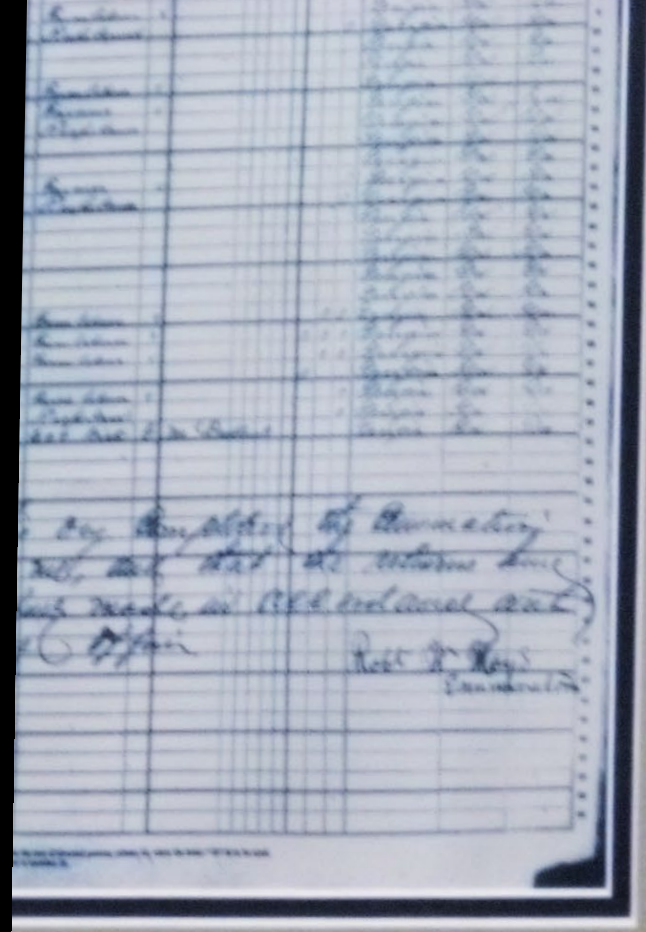
The census extraction project was started 17 years ago and survived through changing intentions and technologies. After 11.5 million hours of work by dedicated extractors, the huge database was submitted to the Minnesota Population Center of the University of Minnesota, which partnered with the Family and Church History Department in the considerable work of cleaning up and completing the project. The record is composed of 35 data CDs, 20 Index CDs and a viewer CD and will sell for $49.

With 50.5 million names, the record has twice as many names as the recently announced Ellis Island Records (24 million names), and has 12 times (6.5 million) as many names of African-Americans as the recently announced Freedman's Bank Records (480,000 names). It is a fully extracted record with every name indexed.


FAMILY HISTORY RESOURCE FILE
CD-ROM LIBRARY
FAMILYSEARCH
1880 UNITED STATES CENSUS AND NATIONAL INDEX

56 CD-ROM Set

Collaboration with the Latter-Day Saints



Presented to the

MINNESOTA
POPULATION CENTER
UNIVERSITY OF
MINNESOTA

For their exceptional
contribution to the
development of the

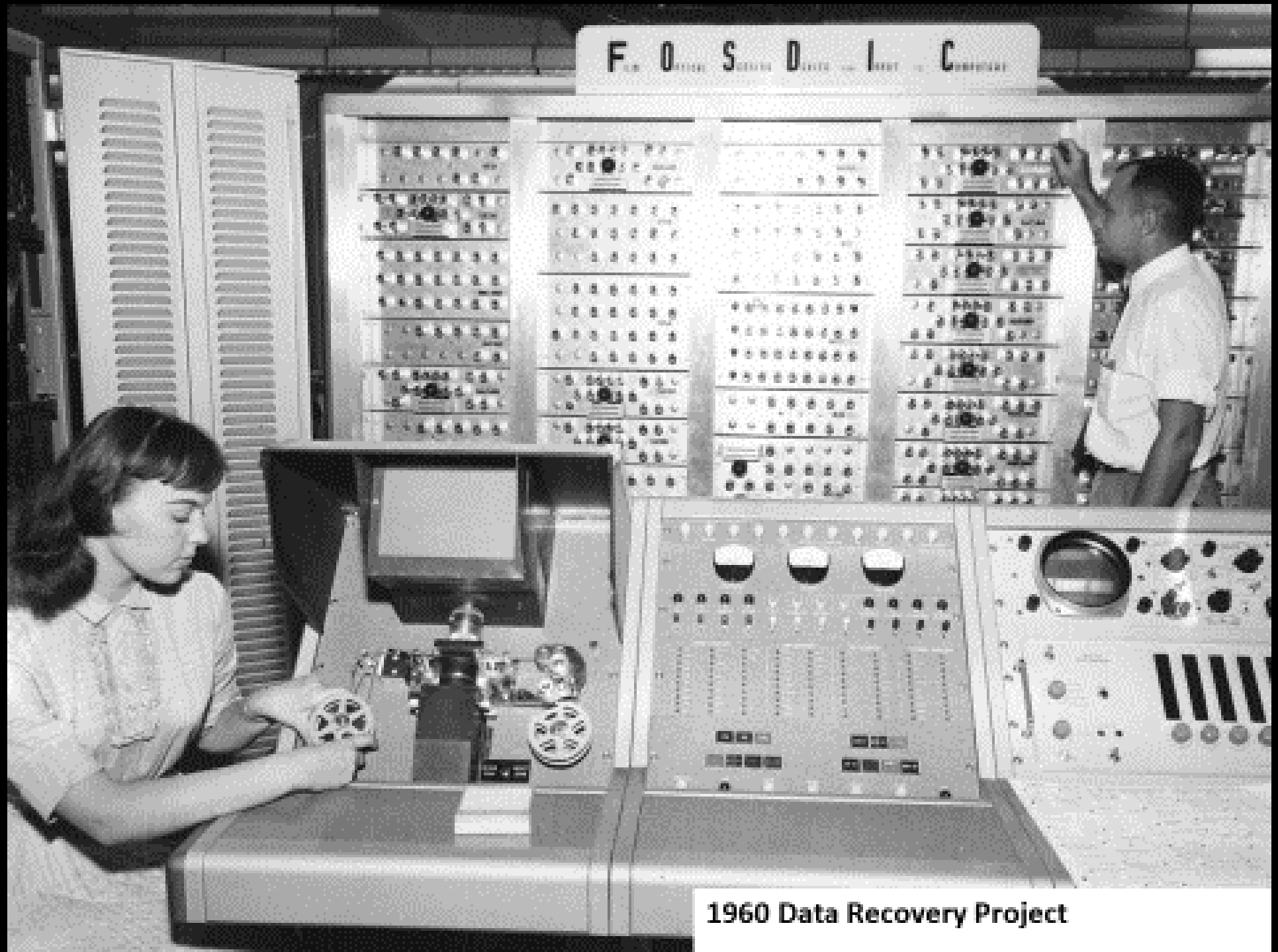
1880
UNITED STATES
CENSUS ON COMPACT
DISC

In Cooperation with
The Church of Jesus Christ
of Latter-day Saints

2002

2001: National Historical Census Files Project

1. Recover, clean, document, and verify the complete 1960, 1970, 1980, and 1990 long-form and short-form microdata
2. Create an IPUMS-compatible version of internal census microdata files from 1960 to the present



1960 Data Recovery Project

P22. Did this person work at any time last week?

Include part-time work such as a Saturday job, delivering papers, or helping without pay in a family business or farm. Do **not** count own housework.

Yes ☐ No ☐

P23. How many hours did he work last week (at all jobs)?

(If exact figure not known, give best estimate)

1 to 14 hours ☐	40 hours ☐
15 to 29 hours ☐	41 to 48 hours ☐
30 to 34 hours ☐	49 to 59 hours ☐
35 to 39 hours ☐	60 hours or more ☐

P24. Was this person looking for work, or on layoff from a job?

Yes ☐ No ☐

P25. Does he have a job or business from which he was temporarily absent all last week because of illness, vacation, or other reasons?

Yes ☐ No ☐

P26. When did he last work at all, even for a few days?

(Check one box)

Working now ☐	1949 or earlier ☐
In 1960 ☐	
In 1959 ☐	Never worked ☐
1955 to 1958 ☐	
1950 to 1954 ☐	

P27. Occupation *(Answer 1, 2, or 3)*

1. This person **last** worked in 1949 or earlier ... ☐
   This person has never worked ...

OR

2. On active duty in the Armed Forces now ... ☐

OR

If this person worked last week, answer questions P28 and P29.

P28. What city and county did he work in last week?

If he worked in more than one city or county, give place where he worked most last week.

a. City or town ...

b. If city or town — Did he work inside the city limits? ... Yes ☐ No ☐

c. County ... State ...

P29. How did he get to work last week?

(Check one box for principal means used last week)

Railroad ☐	Taxicab ☐	Walk only ☐
Subway or elevated ☐	Private auto or car pool ☐	Worked at home ☐
Bus or streetcar ☐	Other means — Write in:	

P30. Last year (1959), did this person work at all, even for a few days?

Yes ☐ No ☐

P31. How many weeks did he work in 1959, either full-time or part-time? Count paid vacation, paid sick leave, and military service as weeks worked.

(If exact figure not known, give best estimate)

13 weeks or less ☐	40 to 47 weeks ☐
14 to 26 weeks ☐	48 to 49 weeks ☐
27 to 39 weeks ☐	50 to 52 weeks ☐

P32. How much did this person earn in 1959 in wages, salary, commissions, or tips from all jobs?

Before deductions for taxes, bonds, dues, or other items. *(Enter amount or check "None." If exact figure not known,*

Form 60 **PH-3** (5)

U.S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS

1960 CENSUS OF POPULATION AND HOUSING

Approval expires 12-31-60

CONFIDENTIAL—The Census is required by the United States Constitution and further authorized by 13 U.S.C. 5, 9, 141, 221–4. The law requires that the inquiries be answered completely and accurately, and guarantees that the information furnished will be accorded confidential treatment. The Census report cannot be used for purposes of taxation, investigation, or regulation.

2. Page No. from PH–1 or PH–2

Mark below

	0	1	2	3	4	5	6	7	8	9
Hnd's	○	○	○	○						
Tens	○	○	○	○	○	○	○	○	○	○
Units	○	○	○	○	○	○	○	○	○	○

3. *If this is a continuation of a household started on a previous sheet, mark below and omit all "H" items on this sheet.*

Continuation ○

H1. Sample key

A ○
GQ ○
(If GQ, fill H17–H18 and omit remaining "H" items)

H3. Type

House, apt., flat ○
Trailer ○

H4. Access

Direct from outside or common hall ○
Through another unit ○

H5. Kitchen or cooking

For exclusive use ○
Shared or none ○

H6. Condition

Sound ○
Deteriorating ○
Dilapidated ○

H7. Occupancy

Occupied ○
Vacant:
Year-round ○
Migratory wkr ○
Seasonal ○

H8. Rooms

1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10+ ○

H9. Water

Running water in structure
Hot and cold ○
Cold only ○
Run. water outside ○
No running water ○

H10. Toilet

For exclusive use ○
Shared ○
None ○

H11. Bath

For exclusive use ○
Shared ○
None ○

H12. Owned or rented

Owned or being bought ○
Rented ○
No cash rent ○

H13. Vacancy status

For rent ○
For sale only ○
Rtd or sold not occ ○
For occasional use ○
Other vacant ○

FOR ALL UNITS AND GQ's (Omit in type Z ED's)

H17. Is this house—

On a city lot (or apt. bldg.)? ○ *(Omit H18)*

On a place of less than 10 acres? ○ *(Ask H18a)*

On a place of 10 or more acres? ○ *(Ask H18b)*

H18a. *If occupied—* **Last year, 1959, did sales of crops, livestock and other farm products from this place amount to—**

$250 or more? ○
Less than $250 (or none)? ○

H18b. *If occupied—* **Last year, 1959, did sales of crops, livestock and other farm products from this place amount to—**

$50 or more? ○
Less than $50 (or none)? ○

FOR ALL UNITS, OCCUPIED AND VACANT

In all following questions, use "house", "apartment", "flat", or "rooms", as appropriate, instead of "unit"

H19. How many bedrooms are in this *unit*?

No bedroom ○
1 ○
2 ○
3 ○
4 or more ○

H20. About when was this house originally built?

In 1959 or 1960 ○
1955 to 1958 ○
1950 to 1954 ○
1940 to 1949 ○
1930 to 1939 ○
1929 or earlier ○

H21. How is this *unit* heated?

Steam or hot water ○
Warm air furnace ○
Floor, wall, or pipeless furnace ○
Built-in electric units ○
Other means—with flue ○
Other means—no flue ○
Not heated ○

FOR OCCUPIED UNITS

H22. What fuel is used most for →

	a. Heating this *unit*?	b. Cooking?	c. Heating water?
Coal or coke	○	○	○
Wood	○	○	○
Utility gas	○	○	○
Bottled, tank, or LP gas	○	○	○
Electricity	○	○	○
Fuel oil, kerosene, etc.	○	○	○
Other	○	○	○
No fuel used	○	○	○

H23. Is there a clothes washing machine in this *unit*?

Wringer or spinner ○
Automatic or semi-automatic ○
Washer-dryer combination ○
No washing machine ○

H24. Is there a clothes dryer in this *unit*?

Electric ○
Gas ○
No clothes dryer ○

H25. Is there television in this *unit*?

1 set ○
2 or more ○
No television ○

H26. Are there any radios in this *unit*?

1 set ○
2 or more ○
No radio ○

H27. Is there any air conditioning in this *unit*?

Room units: 1 ○
2 or more ○
Central system ○
No air conditioning ○

H28. Is there a home food freezer separate from the refrigerator in this *unit*?

Yes ○
No ○

FOR OCCUPIED UNITS

H29–H34. *(Omit)*

H35. Is there a telephone on which people in this *unit* can be called?

Yes ○
No ○

H35a. *If "Yes" in H35—* **What is the telephone number?**

H36. How many automobiles are owned or regularly used by people who live in this *unit*?

Include company cars kept at home

No auto ○
1 ○
2 ○
3 or more ○

H37. *If this is a trailer—* **Is the trailer mobile or has it been put on a permanent foundation?**

Mobile ○
On permanent foundation ○

FOR VACANT UNITS

H38. How many months has this *unit* been vacant?

Up to 1 month ○
1 up to 2 ○
2 up to 4 ○
4 up to 6 ○
6 or more ○

If "OWNED OR BEING BOUGHT" or "VACANT—FOR SALE ONLY" (Omit if 10 acres or more)

H39. (H14) Description of property *(From back page of questionnaire)*

1 unit, no business ○
1 unit, with business ○
2 or more units ○

If transcribing from PH–2, copy from items H14, H15, H16 into H39, H40, H41.

H40. (H15) *If "1 unit, no business"—* **About how much do you think this property would sell for on today's market?** *(If vacant:* **What is the price asked for this property?)**

Under $5,000 ○	15,000–17,400 ○
5,000– 7,400 ○	17,500–19,900 ○
7,500– 9,900 ○	20,000–24,900 ○
10,000–12,400 ○	25,000–34,900 ○
12,500–14,900 ○	35,000 or over ○

If "RENTED" or "VACANT—FOR RENT"

H41. (H16) What is the monthly rent for this *unit*?

Enter to nearest dollar

	$	0	1	2	3	4	5	6	7	8	9
Hnd's		○	○	○	○	○	○	○	○	○	○
Tens		○	○	○	○	○	○	○	○	○	○
Units		○	○	○	○	○	○	○	○	○	○

H42. *(Omit in Z ED's)* **Does the rent include any land used for farming (or ranching)?** Yes ○ No ○

H43. In addition to rent, does renter pay for—

a. Electricity? Yes ○ No ○
b. Gas? Yes ○ No ○
c. Water? Yes ○ No ○

H44. *If "Yes" in H43 and if occupied—* **What is the average monthly cost for—**

	$	0	1	2	3	4	5	6	7	8	9
a. Electricity?	Tens	○	○	○	○	○	○	○	○	○	○
	Units	○	○	○	○	○	○	○	○	○	○
b. Gas?	$	0	1	2	3	4	5	6	7	8	9
	Tens	○	○	○	○	○	○	○	○	○	○
	Units	○	○	○	○	○	○	○	○	○	○
c. Water?	$	0	1	2	3	4	5	6	7	8	9
	Tens	○	○	○	○	○	○	○	○	○	○
	Units	○	○	○	○	○	○	○	○	○	○

H45. In addition to rent, does renter pay for—

Oil, coal, wood or kerosene? Yes ○ No ○

H46. *If "Yes" in H45 and if occupied—* **What is the total yearly cost for—**

Oil, coal, wood or kerosene?

	$	0	1	2	3	4	5	6	7	8	9
	Hnd's	○	○	○	○	○	○	○	○	○	○
	Tens	○	○	○	○	○	○	○	○	○	○
	Units	○	○	○	○	○	○	○	○	○	○

2 3 H1 H3 H4 H5 H6 H7 H8 H9 H10 H11 H12 H13 H14 H15 H16 H-) H17 H18 H19 H20 H21 H22 H23 H24 H25 H26 H27 H28 H35 H35a H36 H37 H38 H39 H40 H41 H42 H43 H44 H45 H46

or none)? ○

ales of crops,
arm products
to—

..... ○

r none)? ○

ACANT

ment", "flat",
unit"

..... ○

..... ○

..... ○

..... ○

..... ○

Bottled, tank, or LP gas	○	.....	○	.....	○
Electricity	○	.....	○	.....	○
Fuel oil, kerosene, etc.	○	.....	○	.....	○
Other	○	.....	○	.....	○
No fuel used	○	.....	○	.....	○

H23. Is there a clothes washing machine in this *unit*?

- Wringer or spinner ..... ○
- Automatic or semi-automatic ○
- Washer-dryer combination ○
- No washing machine ..... ○

H24. Is there a clothes dryer in this *unit*?

- Electric ..... ○
- Gas ..... ○
- No clothes dryer ○

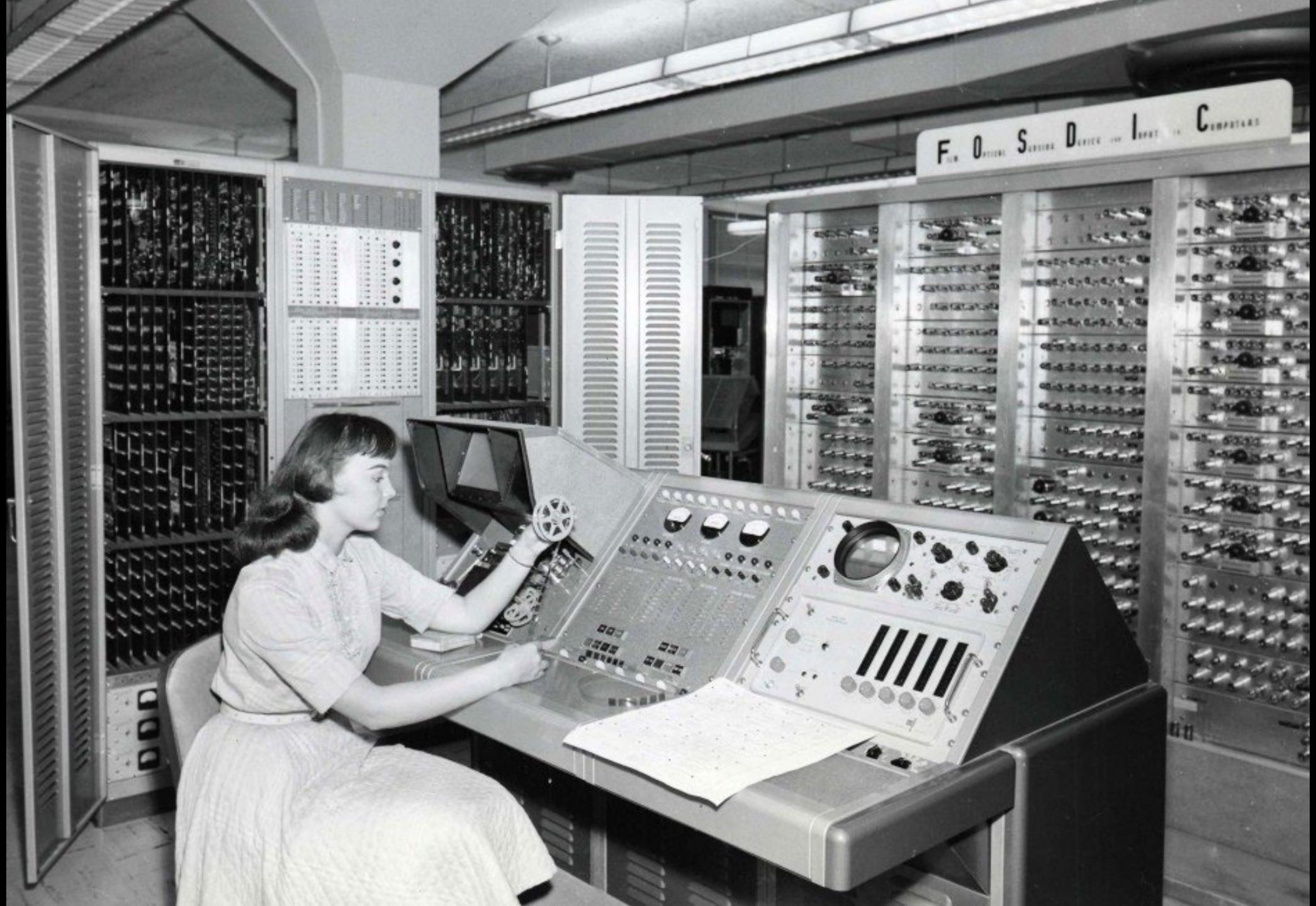
H25. Is there television

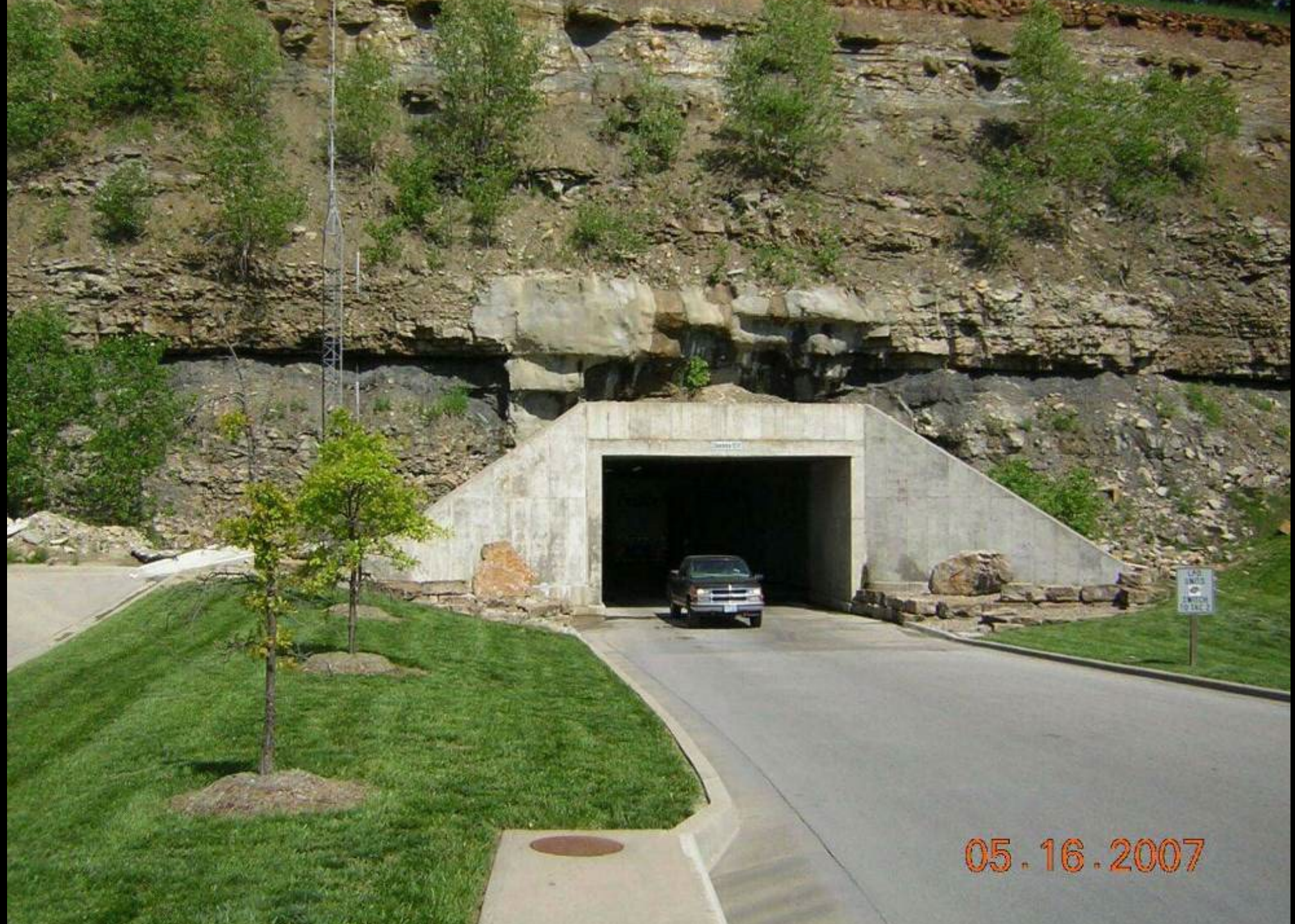
H35a.
What

.....

H36.
are ow
by pe
unit?

*Includ
kept a*




05.16.2007


DIRECTORY
05.16.2007

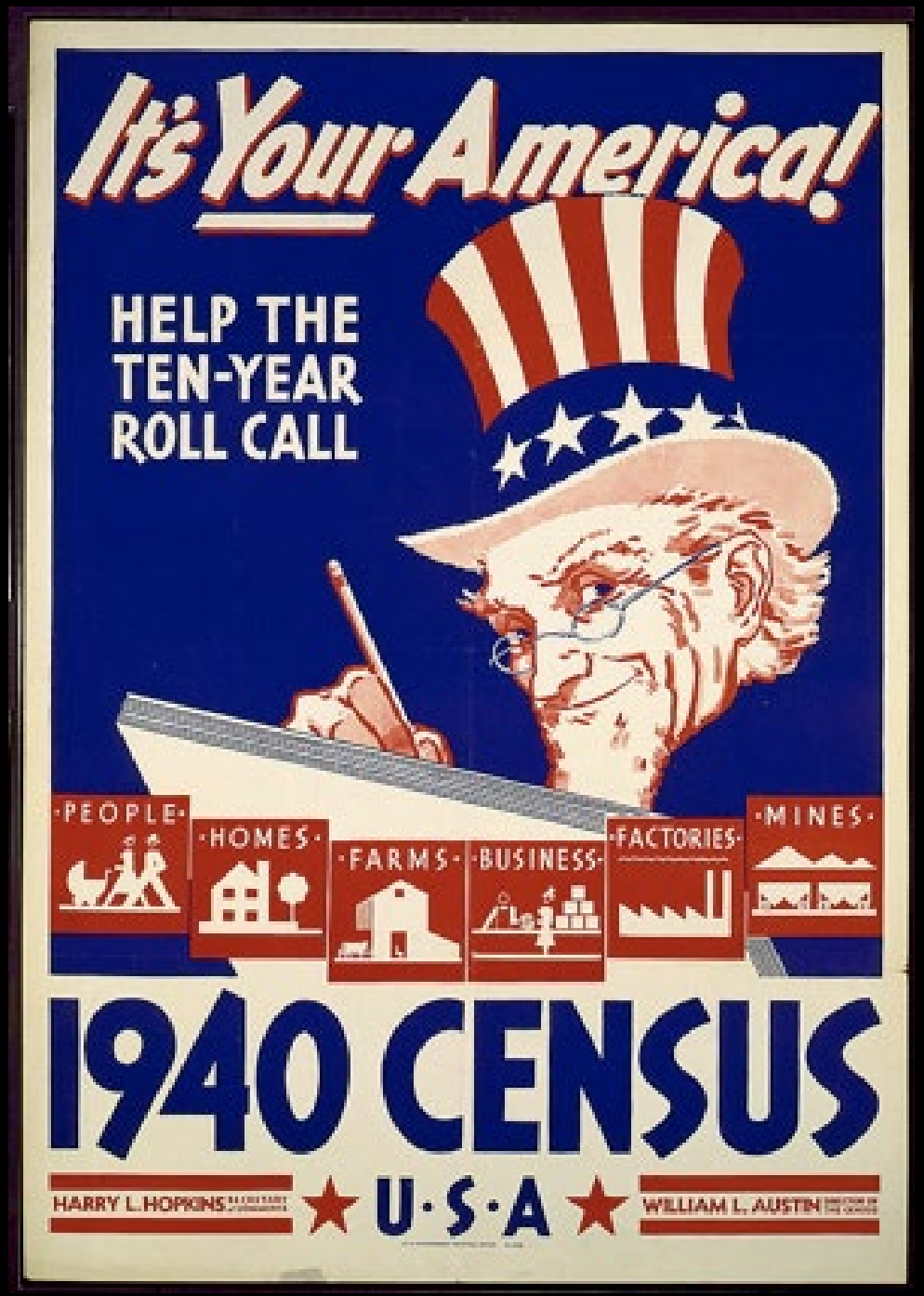

eclipse
09.20.2007

ancestry.com®

IPUMS

132 Million Persons
71 Fields
8.6 billion keystrokes

Funding from NIA, NICHD, NSF



2015: Big Microdata Expansion Project

- 1790-1840 Household data
- 1860-1870 Ancestry genealogical data
- 1900-1930 Ancestry/FamilySearch merged data
- 1860-1930 missing fields keyed

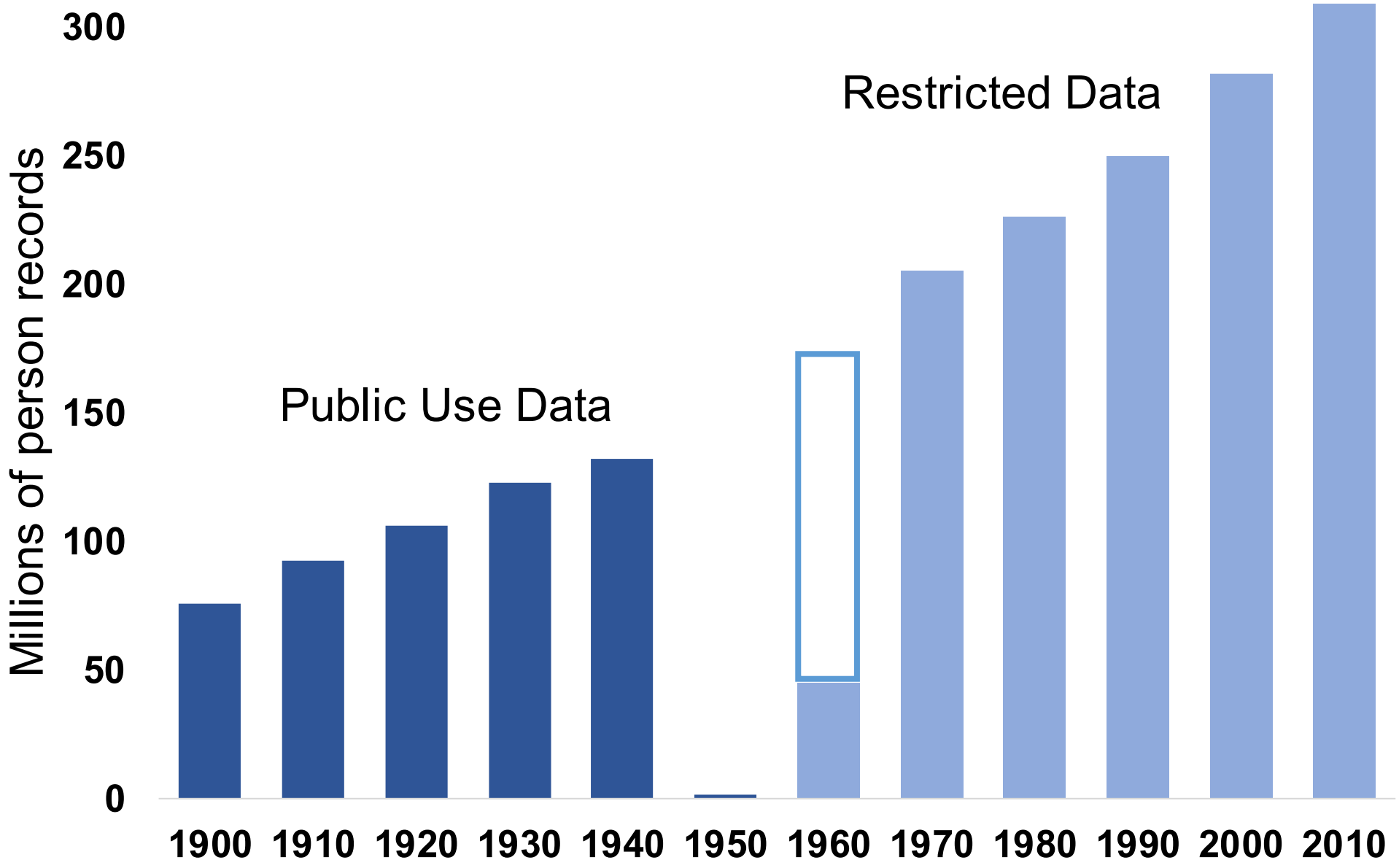


Figure 1. Available census microdata, 1900-2010

2001

NHGIS
National Historical Geographic Information System

Minnesota Population Center
University of Minesota

NHGIS Home | Search

NHGIS Home | Proposal | Downloads

View the full text of the project proposal submitted in Fall 2000 (222 KB; PDF)

There is a a postdoctoral associate position open for this project.

NHGIS Overview

This project makes the census accessible to researchers within the framework of a comprehensive National Historical Geographic Information System (NHGIS). United States summary census data are the primary source of statistical information about growth and change of the American population. The great bulk of these data exist in machine-readable form, but they are largely inaccessible. Approximately 670 gigabytes of data covering the period 1790 through 2000 exist or are in preparation, but they are scattered across dozens of archives and stored in incompatible formats on CD-ROM, magnetic tape, or paper. Only a small fraction of these data are available on the Internet, and even those offer only primitive documentation and extraction tools. Moreover, census summary data cannot be effectively exploited without clear definitions of each geographic unit, but high-quality electronic boundary files exist only for the 1990 census year.

Technological change presents an unprecedented opportunity to make these data readily available for social science research. Bringing the complete census

2003

ihis INTEGRATED HEALTH INTERVIEW SERIES

- Home
- IHIS Components
- Benefits
- Collaborators
- NHIS
- Staff
- Advisory Board

Welcome

The Integrated Health Interview Series (IHIS) is comprised of National Health Interview Survey (NHIS) data and documentation for the period 1963 to 2003. The IHIS will multiply the value of NHIS data by allowing researchers to make consistent comparisons across four decades of dramatic change in public health, and thus to study the health status of Americans as a dynamic process. It will enable a substantial body of new scientific and policy-relevant research into health behavior and disparities, access to and use of medical care, population aging, progress toward public health goals like Healthy People 2010, and many other topics.

PROJECT
About ATUS-X
Time Use Variables
Demographic Variables
Linked CPS Supplements

EXTRACT BUILDER
Build an Extract
Download/Revise Extracts

ATUS DOCUMENTATION
About ATUS
Samples
User's Guides
Questionnaires

USER RESOURCES
FAQ
Citation and Use
Bibliography
Revisions

American Time Use Survey Data Extract Builder

ATUS-X is a project dedicated to making it easy for researchers to use data from the American Time Use Survey (ATUS). The ATUS is an ongoing time diary study that is funded by the U.S. Bureau of Labor Statistics and fielded by the U.S. Census Bureau. The goals of the ATUS-X project are:

- To collect and preserve ATUS data and documentation
- To harmonize the data from different years of data collection
- To make it easy to create data files containing the ATUS variables a user needs

Time is our scarcest resource...use it wisely!

New Users

Create an ATUS-X Account
Navigating the Site

ATUS-X News

2012 data
Release Plans
Metabolic Equivalent Values for Activities

MPC Data Projects

IPUMS-CPS and others

Funding provided by:
National Institutes of Health
Economic Research Service

Supported by:
MPC Minnesota Population Center
Maryland Population Research Center

2011

IDHS

Integrated Demographic and Health Series

HOME | OVERVIEW | PARTNERS | STAFF

Documentation

- Surveys
- MEASURE DHS Links
- Sample Descriptions
- Source Documents

Email Updates

Sign up to receive updates about IDHS

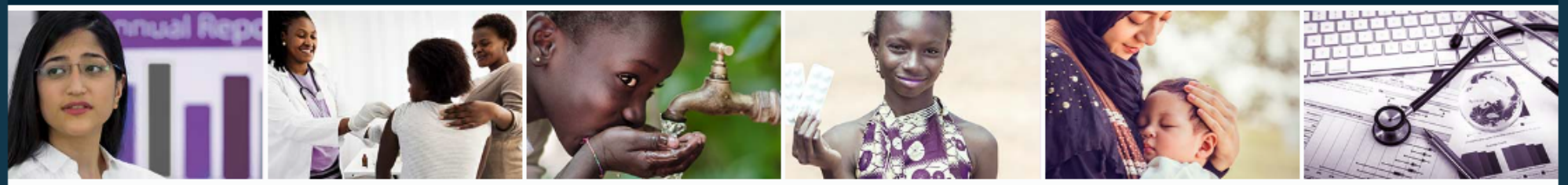
Welcome to IDHS

The Demographic and Health Surveys (DHS) are the main source of information on health in the developing world. IDHS is designed to facilitate analysis of DHS data across time and space.

IDHS will:

- Focus initally on Africa and India
- Display variable availability across surveys
- Code data consistently across years and countries
- Provide variable-specific documentation
- Create customized multi-country and multi-year datasets
- Release data beginning in 2014

2017



DATA

BROWSE AND SELECT DATA

MY DATA EXTRACTS

REGISTER TO USE IPUMS PMA

DOCUMENTATION

ERRATA AND REVISIONS

SAMPLE DESCRIPTIONS

QUESTIONNAIRES

USER GUIDE

PROJECT

ABOUT IPUMS PMA

PMA2020 HOMEPAGE

RESEARCH

MONITORING KEY FAMILY-PLANNING INDICATORS

IPUMS-PMA harmonizes the Performance Monitoring and Accountability 2020 (PMA2020) data series. It provides an interactive web dissemination system for PMA2020 data with variable documentation on hundreds of harmonized variables on family planning, water and sanitation, and menstrual hygiene management. PMA2020 is fielded by the Bill & Melinda Gates Foundation and Johns Hopkins University using streamlined and high-frequency data collection in 11 FP2020 pledging countries.

11 COUNTRIES ·· 96 SAMPLES ·· OVER 1800 VARIABLES ·· OVER 900 THOUSAND RECORDS

CREATE AN EXTRACT

Get Data

CREATE AN ACCOUNT

Register

2020

IPUMS IHGIS

HOME | SELECT DATA | MY DATA | SUPPORT

IPUMS IHGIS
ABOUT
SOURCE ORGANIZATIONS
REGISTER
DONATE TO IHGIS

DATA
BROWSE AND SELECT DATA
DOWNLOAD OR REVISE MY DATA

SUPPLEMENTAL DATA
GEOGRAPHY & GIS

DOCUMENTATION
DATASET DESCRIPTIONS

SUPPORT
FAQ
USER GUIDE

ANALYSIS-READY TABLES FROM INTERNATIONAL CENSUSES

The **International Historical Geographic Information System** (IHGIS) provides data tables from population and housing censuses as well as agricultural censuses from around the world, along with corresponding GIS boundary files. Data from diverse published documents are provided in a standardized, ready-to-use structure with thorough, consistently formatted metadata. *Read more*.

START HERE:

Get Data

IPUMS USA

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

VISIT SITE

IPUMS CPS

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

VISIT SITE

IPUMS INTERNATIONAL

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

VISIT SITE

IPUMS GLOBAL HEALTH

Health survey data for Africa and Asia, including harmonized data collections for DHS and PMA. Learn More

VISIT SITE

IPUMS NHGIS

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

VISIT SITE

IPUMS IHGIS

Tabular and GIS data from population, housing, and agricultural censuses around the world. Learn More

Find additional spatial population & environmental data in IPUMS Terra.

VISIT SITE

IPUMS TIME USE

IPUMS HEALTH SURVEYS

IPUMS HIGHER ED

INTRODUCING IPUMS DHS CALENDAR DATA

August 24
9:00-10:00am CT
Free Webinar

LEARN MORE AND REGISTER

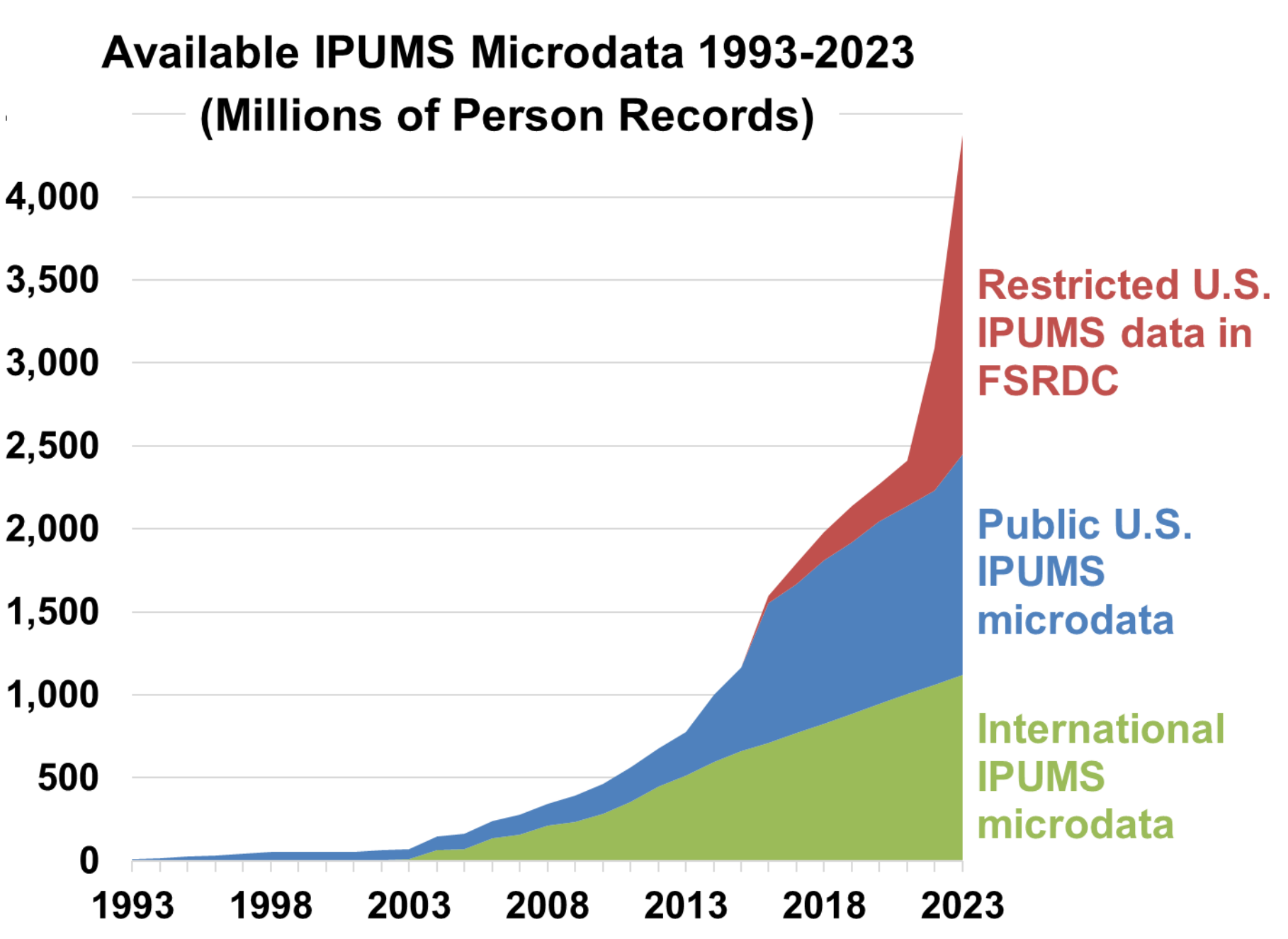
CENSUS BUREAU PLANS MAJOR CHANGES TO PUBLIC DATA

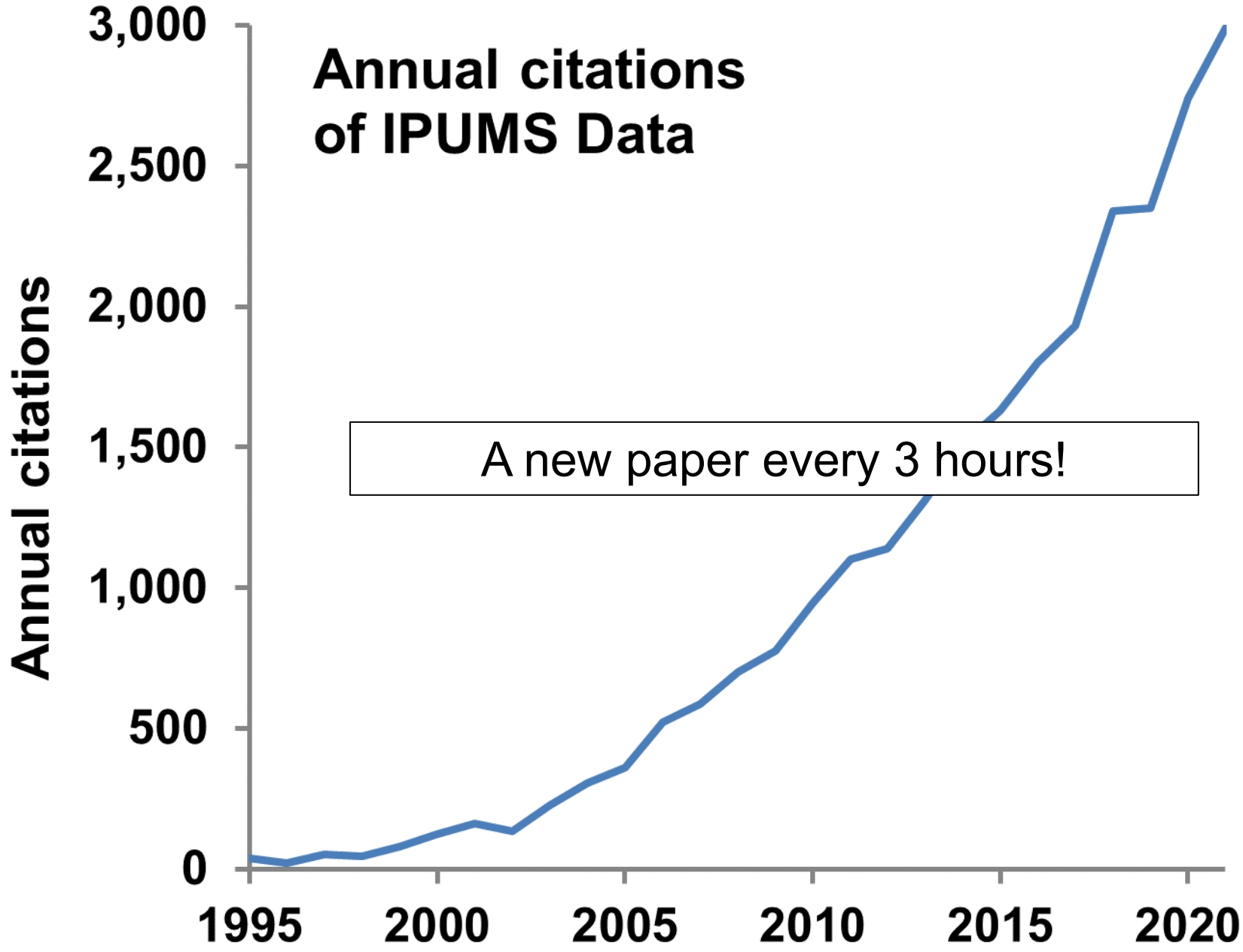
New data may be unusable for research and planning

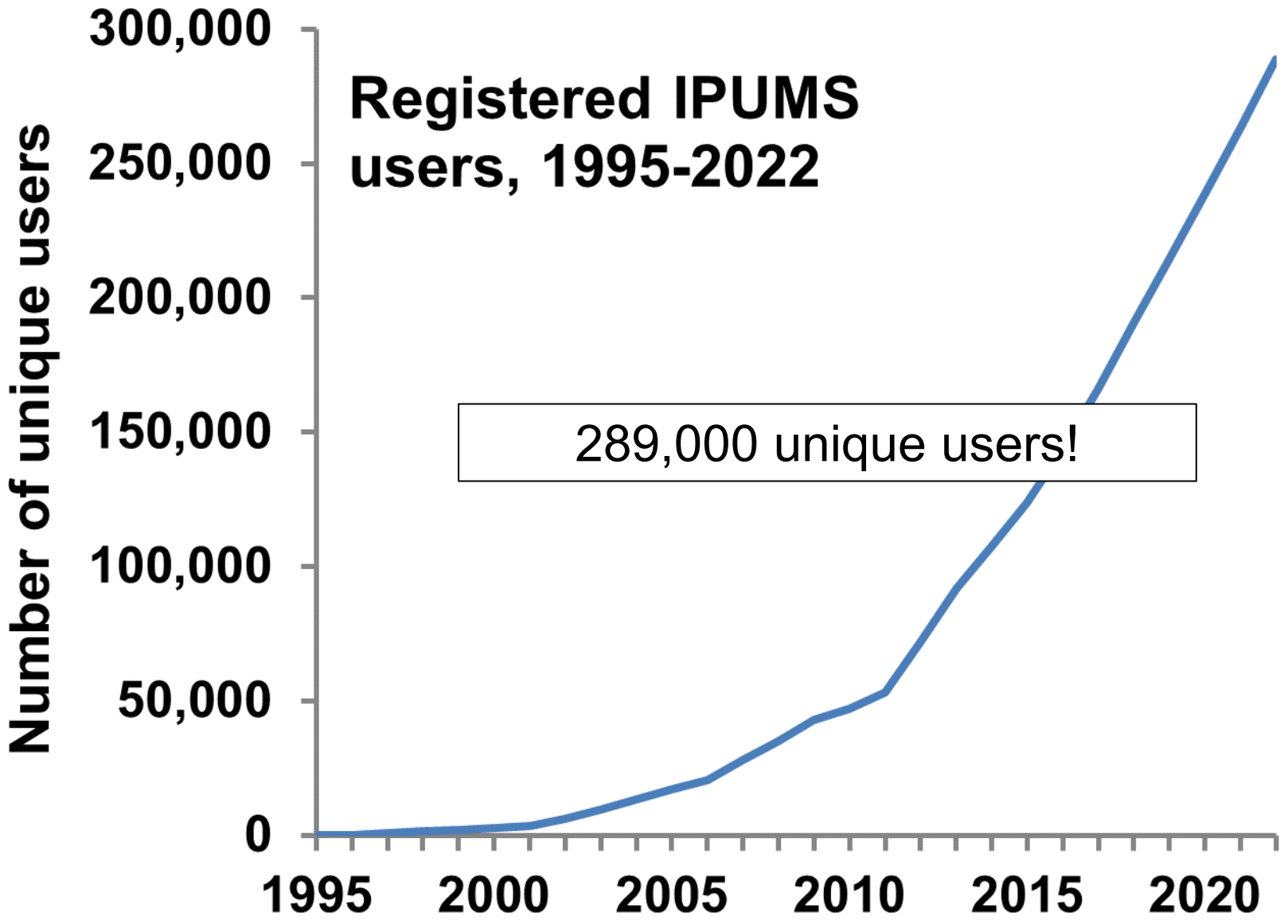
LEARN MORE

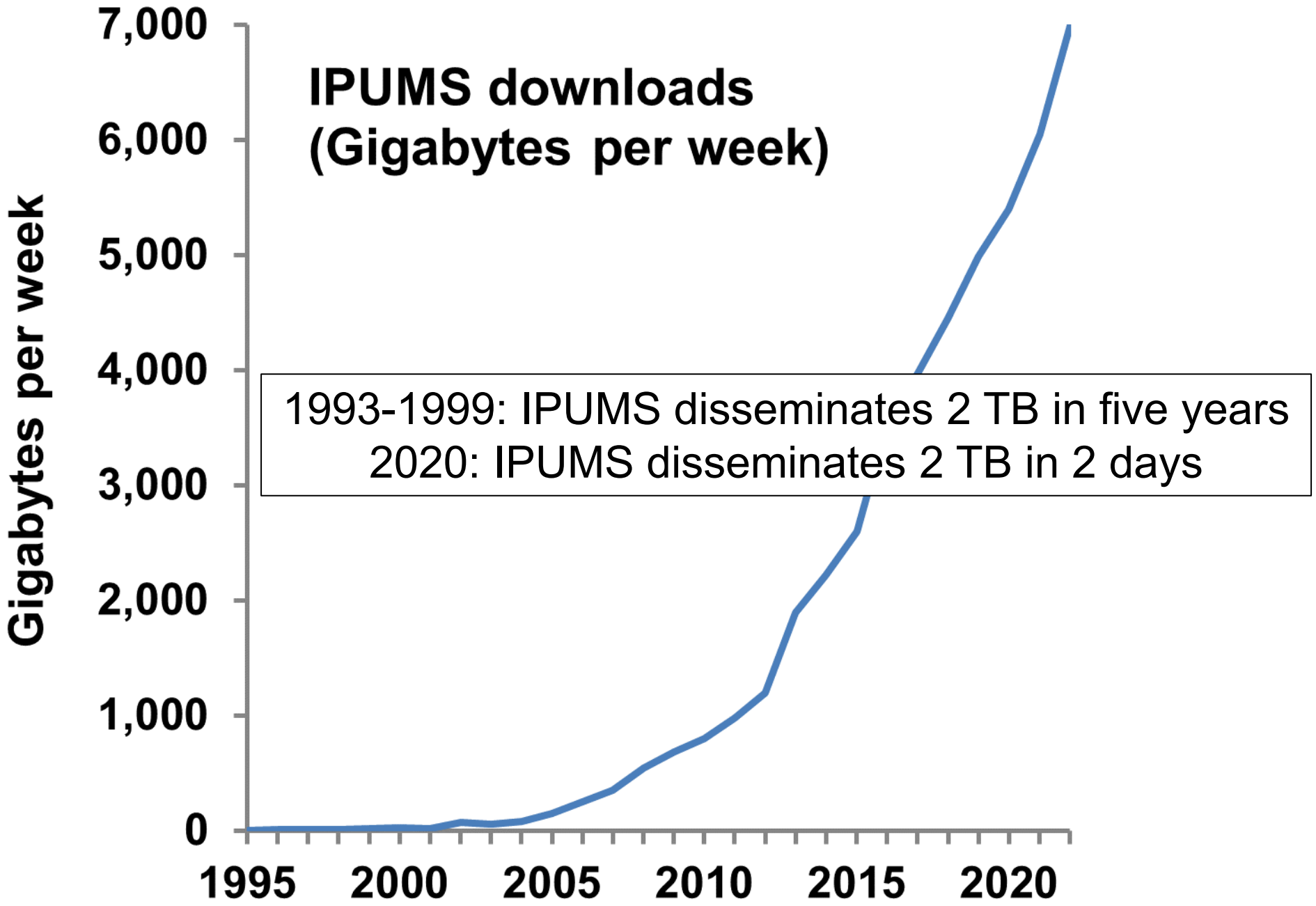
CALENDAR

2022 Data-Intensive Research Conference
Jul 20, 2022 to Jul 21, 2022
Minneapolis, MN


Available IPUMS Microdata 1993-2023
(Millions of Person Records)
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
1993
1998
2003
2008
2013
2018
2023
Restricted U.S. IPUMS data in FSRDC
Public U.S. IPUMS microdata
International IPUMS microdata


Annual citations of IPUMS Data
Annual citations
3,000
2,500
2,000
1,500
1,000
500
0
1995
2000
2005
2010
2015
2020
A new paper every 3 hours!


Registered IPUMS users, 1995-2022
Number of unique users
300,000
250,000
200,000
150,000
100,000
50,000
0
1995
2000
2005
2010
2015
2020
289,000 unique users!


IPUMS downloads
(Gigabytes per week)
Gigabytes per week
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
1995
2000
2005
2010
2015
2020
1993-1999: IPUMS disseminates 2 TB in five years
2020: IPUMS disseminates 2 TB in 2 days

Census Longitudinal Infrastructure Project (CLIP)

- Restricted data project, housed within the Federal Statistical Research Data Centers
- Link 1940 census to Social Security records, assigning an identity key to each person
- 1940 can then be linked to a wide range of later data that are already identified

CLIP Linking Strategy

IPUMS

Multigenerational Longitudinal Panel (MLP)

- Link censuses from 1850-1940 to each other and to administrative records
- First release 8/2020, new versions planned annually

1940 Census

Numident

Genealogies

1850-1930 Censuses

-2020 aths

MLP Linking Strategy

Why Big Microdata?

- You can’t link samples
- Broad geographic coverage mitigates the problem of out-migration, which is a big problem in some local studies
- Can study small subgroups (e.g. twins, indigenous, residents of public housing in NYC)
- Can link full-count data to small datasets, such as labor union data or recent surveys of aging
- Full-count data allows contextual analysis to understand environmental impacts and community effects (e.g., Ferrie’s lead exposure)

The Future

- Build on existing historical data infrastructure to create national-scale historical longitudinal data wherever possible, linking all available records (e.g., vital records, health surveys, population registers, censuses, administrative records)
- Make these data compatible across time and across national boundaries
- Develop compatible contextual data from governmental statistics, remote sensing, and other sources and make them easily interoperable with microdata

Lessons from IPUMS

- To make diverse datasets interoperable we need data integration metadata
- To make data derived from a broad range of sources widely used, we need powerful data access tools that make it easy to merge data across time and space
- When feasible, we should connect historical data with recent data to maximize salience and fundability

Thank You.