

IPUMS International: Building a census data time machine

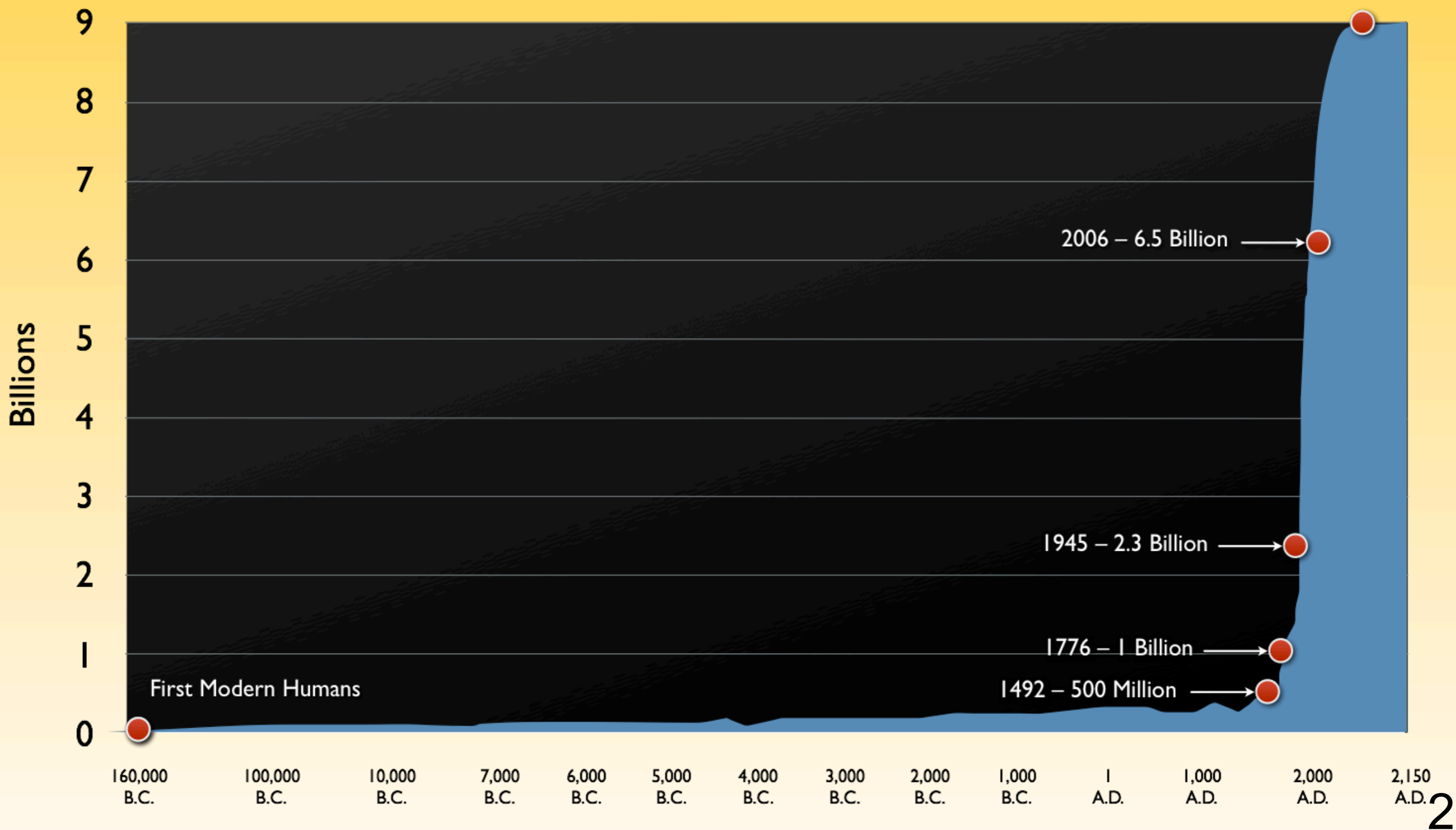
Peter Clark, Minnesota Population Center
<http://www.pop.umn.edu>



UNIVERSITY OF MINNESOTA

Driven to DiscoverSM

World Population Growth Over Time



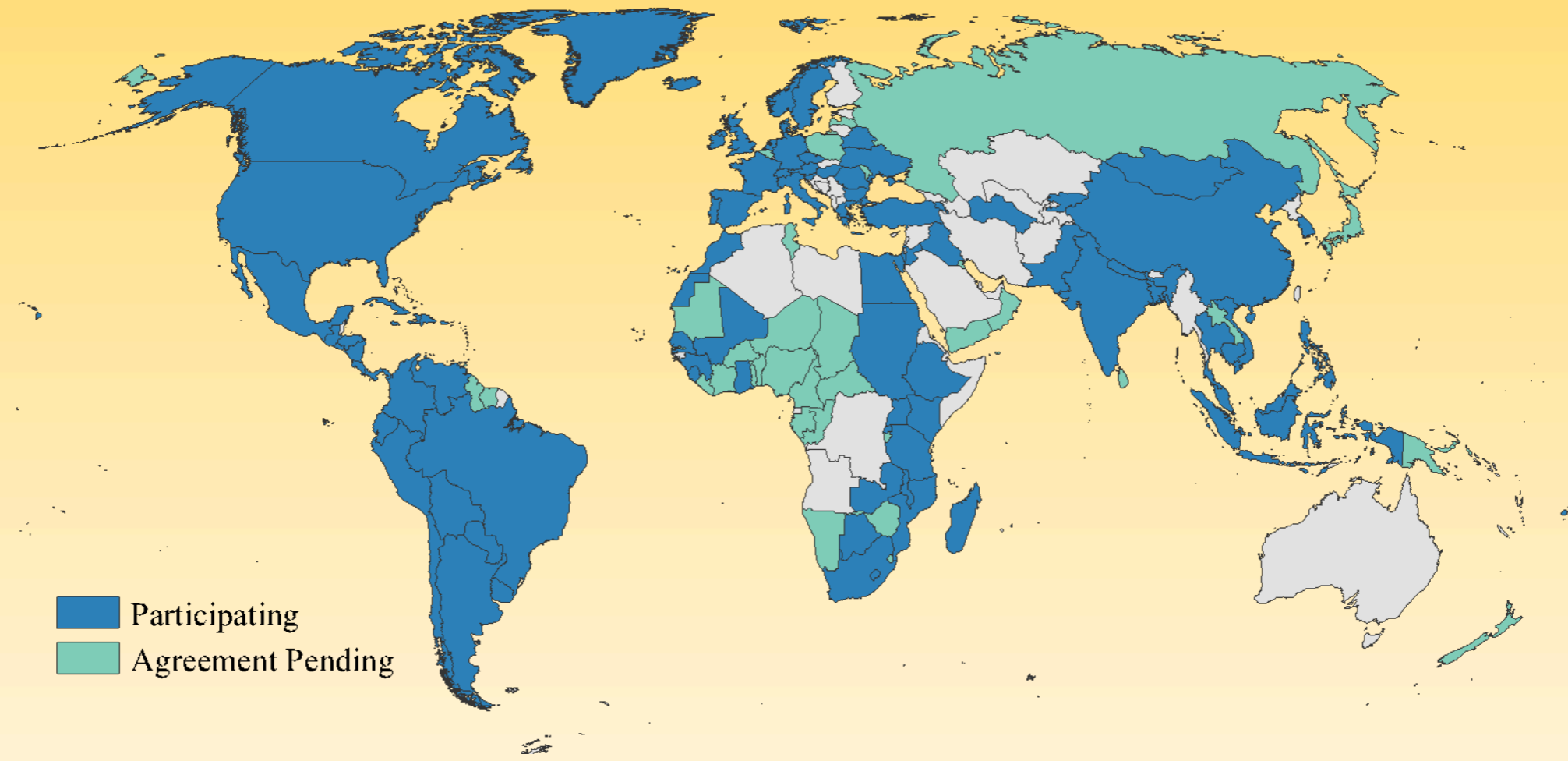
UNIVERSITY OF MINNESOTA
Driven to DiscoverSM

What's an IPUMS?

- Integrated Public Use Microdata Series
 - USA (1850-2008, not 1890)
 - International (machine-readable, post 1960)
 - Current Population Survey (CPS)
 - North Atlantic Population Project (NAPP)
 - Integrated Health Information Survey (IHIS)
- <http://www.ipums.org>



Worldwide Participation in IPUMS-I





Data Cart
 Your data extract
 0 variables
 11 samples
[VIEW CART](#)

[Home](#) [Select Data](#) [FAQ](#) [Contact Us](#) [Login](#)

Select Variables

Household ▾ Person ▾ A-Z ▾ Search

Change Samples

- Integrated variables [Help](#)
- Unharmonized variables [Options](#)
- [Country abbreviations](#)

Demographic Variables -- PERSON [\[top\]](#)

Add to cart	Variable	Variable Label	Type	Codes	CU 02	EC 01	GN 96	HU 01	IR 06	KG 99	RO 02	SD 08	TZ 88	US 00	VN 99
	RELATE	Relationship to household head	P	codes	X	X	X	X	X	X	X	X	X	X	X
	ERELATE	Relationship to head, Europe	P	codes	.	.	.	X	.	.	X
	AGE	Age	P	codes	X	X	X	X	X	X	X	X	X	X	X
	AGE2	Age, grouped into intervals	P	codes	X	X	X	X	X	X	X	X	X	X	X
	SEX	Sex	P	codes	X	X	X	X	X	X	X	X	X	X	X
	MARST	Marital status	P	codes	X	X	X	X	X	X	X	X	X	X	X
	EMARST	Marital status, Europe	P	codes	.	.	.	X	.	.	X
	CONSENS	Consensual union	P	codes	X	X	X	X	.	X	X
	POLYGAM	Polygamous union	P	codes	.	.	X
	AGEMARR	Age at first marriage or union	P	codes	X	X	.	.	.
	MARRYR	Year of first marriage	P	codes	X
	MARRNUM	Number of marriages or unions	P	codes	.	.	.	X	X
	SUBFREL	Relationship to head of subfamily	P	codes	.	.	.	X
	SUBFNUM	Subfamily membership number	P	codes	.	.	.	X
	BIRTHYR	Year of birth	P	codes	.	.	X	.	X	X	X	.	.	.	X
	BIRTHMO	Month of birth	P	codes	.	.	X	.	X	X	X	.	.	.	X

What's so hard about that?



What's so hard about that?

Censuses have
different structures:



What's so hard about that?

Censuses have
different structures:
Canada (2001)



What's so hard about that?

Censuses have different structures:
Canada (2001)

Ce questionnaire est disponible en français 1 800 591-2001

OFFICE USE ONLY

Prov. FED EA VN CCD **2B**

UD FR Forms 3

VR Line No. Map Block No. Form type No. of persons Questionnaire No. TD

3 of

NHS TR Coll.

1. 2.



FOR INFORMATION ONLY

CENSUS

Count yourself in! May 15, 2001

For over 300 years, the census has painted a picture of our people and the places in which we live. This census, on May 15, 2001, will continue the tradition by providing information needed by community groups, businesses and governments to develop plans for education and training, seniors' housing, day care, fire protection, public transport, and many other programs that are important to all of us.

As Canada's national statistics agency, Statistics Canada uses census data for producing statistical tables, analytical reports and for selecting samples or following up respondents for some of our surveys. These uses are strictly for statistical purposes and no one outside of the agency can have access to your identifiable information.

By law, Statistics Canada must take a census every five years, and every household must fill in a census form. Also, by law, Statistics Canada must protect the confidentiality of the personal information you provide. Our employees, including census takers, are personally liable to fines or imprisonment should they break the confidentiality of your information.

Please complete your census form and mail it back on May 15, 2001.

Thank you for your cooperation.

Ivan P. Fellegi
Chief Statistician of Canada

This information is collected under the authority of the *Statistics Act*, R.S.C. 1985, c. S-19, and must be provided by law. CONFIDENTIAL WHEN COMPLETED

STEP A **Begin here**
by printing your address

No. and street or lot and concession Apt. No.

City, town, village, Indian reserve Province / territory

Postal code Area code Telephone number

? For information and instructions see the Guide

Call us free of charge 1 800 591-2001

TTY/TDD users call: 1 877 881-8301

Statistics Canada Statistique Canada

Canada



UNIVERSITY OF MINNESOTA
Driven to DiscoverSM

What's so hard about that?

Censuses have different structures:
 Canada (2001)
 USA (2010)

Ce questionnaire est disponible en français 1 800 591-2001

OFFICE USE ONLY

Prov. FED EA VN CCD **2B**

UD FR Forms 3

VR Line No. Map Block No. Form type No. of persons Questionnaire No. TD 1 3 2.

2 4 Coll.

3



FOR INFORMATION ONLY

CENSUS

Count yourself in! May 15, 2001

For over 300 years, the census has painted a picture of our people and the places in which we live. This census, on May 15, 2001, will continue the tradition by providing information needed by community groups, businesses and governments to develop plans for education and training, seniors' housing, day care, fire protection, public transport, and many other programs that are important to all of us.

As Canada's national statistics agency, Statistics Canada uses census data for producing statistical tables, analytical reports and for selecting samples or following up respondents for some of our surveys. These uses are strictly for statistical purposes and no one outside of the agency can have access to your identifiable information.

By law, Statistics Canada must take a census every five years, and every household must fill in a census form. Also, by law, Statistics Canada must protect the confidentiality of the personal information you provide. Our employees, including census takers, are personally liable to fines or imprisonment should they break the confidentiality of your information.

Please complete your census form and mail it back on May 15, 2001.

Thank you for your cooperation.

Ivan P. Fellegi
 Chief Statistician of Canada

This information is collected under the authority of the *Statistics Act*, R.S.C. 1985, c. S-19, and must be provided by law.

CONFIDENTIAL WHEN COMPLETED

STEP A **Begin here**
 by printing your address

No. and street or lot and concession Apt. No.

City, town, village, Indian reserve Province / territory

Postal code Area code Telephone number

? For information and instructions see the Guide

Call us free of charge 1 800 591-2001

TTY/TDD users call: 1 877 881-8301

Statistics Canada Statistique Canada

Canada



UNIVERSITY OF MINNESOTA
 Driven to DiscoverSM

What's so hard about that?

Censuses have
different structures:
Canada (2001)
USA (2010)

United States
Census
2010

This is the official form for all the people at this address.
It is quick and easy, and your answers are protected by law.

U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU

Use a blue or black pen.
Start here

The Census must count every person living in the United States on April 1, 2010.
Before you answer Question 1, count the people living in this house, apartment, or mobile home using our guidelines.

- Count all people, including babies, who live and sleep here most of the time.

The Census Bureau also conducts counts in institutions and other places, so:

- Do not count anyone living away either at college or in the Armed Forces.
- Do not count anyone in a nursing home, jail, prison, detention facility, etc., on April 1, 2010.
- Leave these people off your form, even if they will return to live here after they leave college, the nursing home, the military, jail, etc. Otherwise, they may be counted twice.

The Census must also include people without a permanent place to stay, so:

- If someone who has no permanent place to stay is staying here on April 1, 2010, count that person. Otherwise, he or she may be missed in the census.

1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2010?
Number of people =

2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1?
Mark all that apply.

- Children, such as newborn babies or foster children
- Relatives, such as adult children, cousins, or in-laws
- Nonrelatives, such as roommates or live-in baby sitters
- People staying here temporarily
- No additional people

3. Is this house, apartment, or mobile home —
Mark ONE box.

- Owned by you or someone in this household with a mortgage or loan? *Include home equity loans.*
- Owned by you or someone in this household free and clear (without a mortgage or loan)?
- Rented?
- Occupied without payment of rent?

4. What is your telephone number? We may call if we don't understand an answer.
Area Code + Number
 - -

OMB No. 0607-0919-C; Approval Expires 12/31/2011.
Form **D-61** (1-15-2009)

5. Please provide information for each person living here. Start with a person living here who owns or rents this house, apartment, or mobile home. If the owner or renter lives somewhere else, start with any adult living here. This will be Person 1.
What is Person 1's name? *Print name below.*
Last Name
First Name MI

6. What is Person 1's sex? Mark ONE box.
 Male Female

7. What is Person 1's age and what is Person 1's date of birth?
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age on April 1, 2010 Month Day Year of birth
→ **NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.**

8. Is Person 1 of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — *Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.*

9. What is Person 1's race? Mark one or more boxes.

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — *Print name of enrolled or principal tribe.*
- Asian Indian Japanese Native Hawaiian
- Chinese Korean Guamanian or Chamorro
- Filipino Vietnamese Samoan
- Other Asian — *Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.*
- Other Pacific Islander — *Print race, for example, Fijian, Tongan, and so on.*
- Some other race — *Print race.*

10. Does Person 1 sometimes live or stay somewhere else?
 No Yes — Mark all that apply.

- In college housing For child custody
- In the military In jail or prison
- At a seasonal or second residence In a nursing home
- For another reason

→ If more people were counted in Question 1, continue with Person 2.

U.S. CENSUS BUREAU



What's so hard about that?

Censuses have different structures:
Canada (2001)
USA (2010)
Iraq (1997)

United States Census 2010

This is the official form for all the people at this address. It is quick and easy, and your answers are protected by law.

U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU

Use a blue or black pen.
Start here

The Census must count every person living in the United States on April 1, 2010.

Before you answer Question 1, count the people living in this house, apartment, or mobile home using our guidelines.

- Count all people, including babies, who live and sleep here most of the time.

The Census Bureau also conducts counts in institutions and other places, so:

- Do not count anyone living away either at college or in the Armed Forces.
- Do not count anyone in a nursing home, jail, prison, detention facility, etc., on April 1, 2010.
- Leave these people off your form, even if they will return to live here after they leave college, the nursing home, the military, jail, etc. Otherwise, they may be counted twice.

The Census must also include people without a permanent place to stay, so:

- If someone who has no permanent place to stay is staying here on April 1, 2010, count that person. Otherwise, he or she may be missed in the census.

1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2010?
Number of people =

2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1? Mark all that apply.

- Children, such as newborn babies or foster children
- Relatives, such as adult children, cousins, or in-laws
- Nonrelatives, such as roommates or live-in baby sitters
- People staying here temporarily
- No additional people

3. Is this house, apartment, or mobile home — Mark ONE box.

- Owned by you or someone in this household with a mortgage or loan? Include home equity loans.
- Owned by you or someone in this household free and clear (without a mortgage or loan)?
- Rented?
- Occupied without payment of rent?

4. What is your telephone number? We may call if we don't understand an answer.
Area Code + Number
 - -

5. Please provide information for each person living here. Start with a person living here who owns or rents this house, apartment, or mobile home. If the owner or renter lives somewhere else, start with any adult living here. This will be Person 1.
What is Person 1's name? Print name below.
Last Name
First Name MI

6. What is Person 1's sex? Mark ONE box.
 Male Female

7. What is Person 1's age and what is Person 1's date of birth? Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age on April 1, 2010 Month Day Year of birth

NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.

8. Is Person 1 of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

9. What is Person 1's race? Mark one or more boxes.

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — Print name of enrolled or principal tribe.
- Asian Indian
- Chinese
- Filipino
- Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.
- Japanese
- Korean
- Vietnamese
- Native Hawaiian
- Guamanian or Chamorro
- Samoan
- Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on.
- Some other race — Print race.

10. Does Person 1 sometimes live or stay somewhere else?
 No Yes — Mark all that apply.

- In college housing
- In the military
- At a seasonal or second residence
- For child custody
- In jail or prison
- In a nursing home
- For another reason

→ If more people were counted in Question 1, continue with Person 2.

OMB No. 0607-0919-C; Approval Expires 12/31/2011.
Form D-61 (1-15-2009)

U.S. CENSUS BUREAU



What's so hard about that?

Censuses have different structures:
 Canada (2001)
 USA (2010)
 Iraq (1997)

REPUBLIC OF IRAQ
 BOARD OF MINISTERS
 PLANNING COMMISSION
 CENTRAL STATISTICAL ORG.

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

أنا نعلم اليوم ، برغم الحصار
 مدمره المعتدون الاشرار ، وان
 جراح اقتصادنا ينتهي بعون الله
 وسيكون الازدهار فيه ازدهارا
 دائما وغير متذبذب ، ونحن يتوقف
 سعيه في الارتقاء .
 الرئيس القائد
 صدام حسين

1997 population census questionnaire

1 GOVERNORATE		2-QADAH		3 - NAHIYA	
---------------	--	---------	--	------------	--

4 - environment	5 - district name or mahalla name	6 - village name	7 - zone no.	8 - street no.	9 - house census no.	10 - household series no. in the zone	11 - no. of households in the residential unit
1 <input type="checkbox"/> urban 2 <input type="checkbox"/> rural	district no.	village no.					

put (x) in the correct square

12 - no. of bedrooms for the household	13 - ownership of the residential unit		14 - volume of the rented residential unit		15 - month of rent (I.D)
	private sector	gover. sector	1 <input type="checkbox"/> all unit 3 <input type="checkbox"/> two rooms	2 <input type="checkbox"/> one room 4 <input type="checkbox"/> three rooms and more	
	1 <input type="checkbox"/> possession 2 <input type="checkbox"/> rent	3 <input type="checkbox"/> other	4 <input type="checkbox"/> rent 5 <input type="checkbox"/> other		

(16-26) filled for each household living alone in the residential unit , if engaged with more than one households , one of the households fills these columns only .

16 - kind of residential unit	17 - the material in construction of the residential unit	18 - total no. of rooms in the residential unit	is there any				
1 <input type="checkbox"/> house 2 <input type="checkbox"/> flat 3 <input type="checkbox"/> hut	4 <input type="checkbox"/> tent 5 <input type="checkbox"/> caravan 6 <input type="checkbox"/> other	1 <input type="checkbox"/> bricks 2 <input type="checkbox"/> stone 3 <input type="checkbox"/> blocks	4 <input type="checkbox"/> concrete blocks 5 <input type="checkbox"/> mud 6 <input type="checkbox"/> other	19 - bathroom in the living unit	20 - water closet	21. bathroom and water closet in the same place	22 - kitchen
				1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no

23 - is the living unit provided with electricity	24 - is the living unit provided with telephon	25 - source of drinking water in the living unit		26 - drainage of the dirty water
1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	1 <input type="checkbox"/> water nets 2 <input type="checkbox"/> public tap	3 <input type="checkbox"/> well 4 <input type="checkbox"/> river or waterwheel 5 <input type="checkbox"/> other	1 <input type="checkbox"/> sewage system 2 <input type="checkbox"/> septic tank 3 <input type="checkbox"/> other

27 - filled for the people living in the establishments (hotels , clinics , prisons , etc)

1 <input type="checkbox"/> hotel or resthouse	2 <input type="checkbox"/> hostel (for students)	3 <input type="checkbox"/> building for social care	4 <input type="checkbox"/> medical establish
5 <input type="checkbox"/> mosques or churches , etc	6 <input type="checkbox"/> prisons , police stations	7 <input type="checkbox"/> residential units engaged with persons living outside households	8 <input type="checkbox"/> others
name of the establishment			



What's so hard about that?

Censuses have different structures:
 Canada (2001)
 USA (2010)
 Iraq (1997)
 Armenia (2001)

REPUBLIC OF IRAQ
 BOARD OF MINISTERS
 PLANNING COMMISSION
 CENTRAL STATISTICAL ORG.

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

أنا نعلم اليوم ، برغم الحصار
 مدمره المعتدون الاشرار ، وان
 جراح اقتصادنا ينتهي بعون الله
 وسيكون الازدهار فيه ازدهارا
 دائما وغير متذبذب ، ونحن يتوقف
 سعيه في الارتقاء .
 الرئيس القائد
 صدام حسين

1997 population census questionnaire

1 GOVERNORATE 2-QADAH 3- NAHIYA

4- environment 1 <input type="checkbox"/> urban 2 <input type="checkbox"/> rural	5- district name or mahalla name district no.	6- village name village no.	7- zone no.	8- street no.	9- house census no.	10- household series no. in the zone	11- no. of households in the residential unit
--	--	--------------------------------	-------------	---------------	---------------------	--------------------------------------	---

put (x) in the correct square

12- no. of bedrooms for the household	13- ownership of the residential unit		14- volume of the rented residential unit		15- month ly rent (I.D)
	private sector	gover. sector	1 <input type="checkbox"/> all unit 3 <input type="checkbox"/> two rooms	2 <input type="checkbox"/> one room 4 <input type="checkbox"/> three rooms and more	
	1 <input type="checkbox"/> possession 2 <input type="checkbox"/> rent	3 <input type="checkbox"/> other 4 <input type="checkbox"/> rent	5 <input type="checkbox"/> other		

(16-26) filled for each household living alone in the residential unit , if engaged with more than one households , one of the households fills these columns only.

16- kind of residential unit 1 <input type="checkbox"/> house 2 <input type="checkbox"/> flat 3 <input type="checkbox"/> hut	4 <input type="checkbox"/> tent 5 <input type="checkbox"/> caravan 6 <input type="checkbox"/> other	17- the material in construction of the residential unit 1 <input type="checkbox"/> bricks 2 <input type="checkbox"/> stone 3 <input type="checkbox"/> blocks	4 <input type="checkbox"/> concrete blocks 5 <input type="checkbox"/> mud 6 <input type="checkbox"/> other	18- total no. of rooms in the residential unit				is there any			
				19- bathroom in the living unit 1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	20- water closet 1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	21. bathroom and water closet in the same place 1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	22- kitchen 1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no				

23- is the living unit provided with electricity 1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	24- is the living unit provided with telephon 1 <input type="checkbox"/> yes 2 <input type="checkbox"/> no	25- source of drinking water in the living unit 1 <input type="checkbox"/> water nets 3 <input type="checkbox"/> well 5 <input type="checkbox"/> other 2 <input type="checkbox"/> public tap 4 <input type="checkbox"/> river or waterwheel	26- drainage of the dirty water 1 <input type="checkbox"/> sewage system 3 <input type="checkbox"/> other 2 <input type="checkbox"/> septic tank
--	---	---	--

27- filled for the people living in the establishments (hotels , clinics , prisons , etc)

1 <input type="checkbox"/> hotel or resthouse 5 <input type="checkbox"/> mosques or churches , etc	2 <input type="checkbox"/> hostel (for students) 6 <input type="checkbox"/> prisons , police stations	3 <input type="checkbox"/> building for social care 7 <input type="checkbox"/> residential units engaged with persons living outside households	4 <input type="checkbox"/> medical establish 8 <input type="checkbox"/> others
name of the establishment			



What's so hard about that?

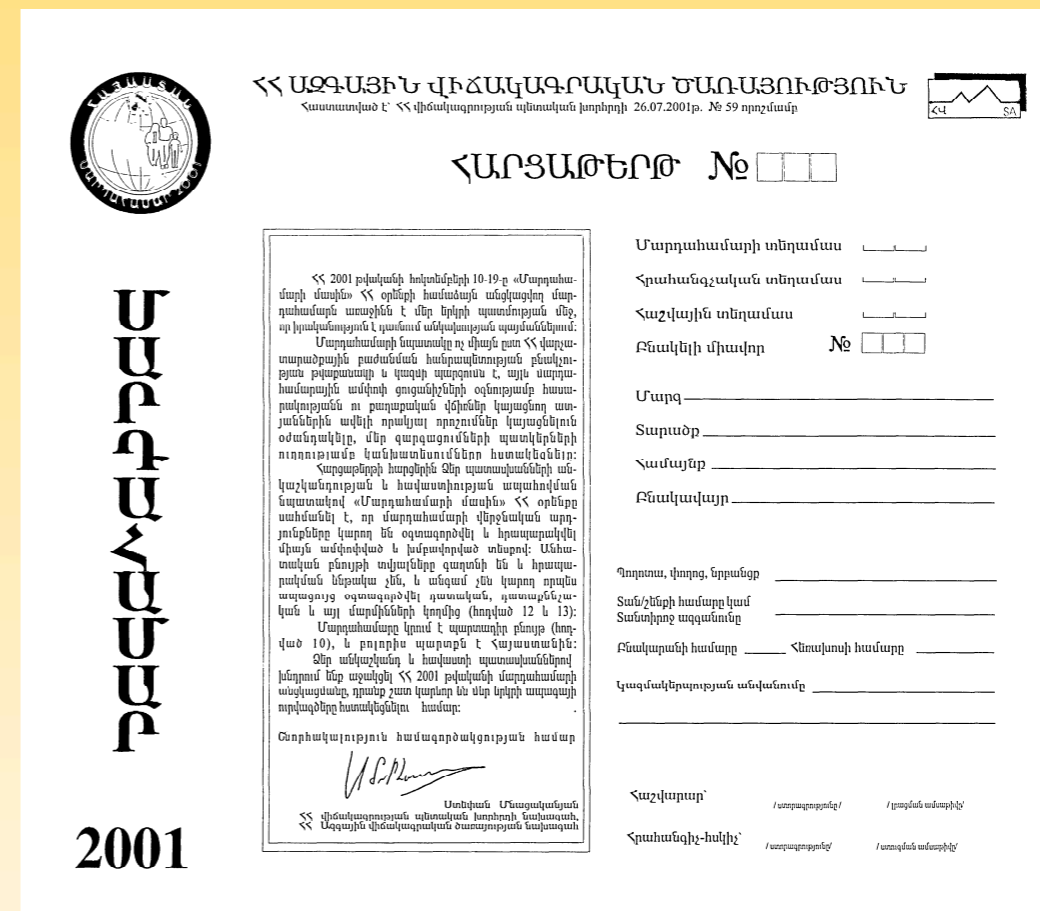
Censuses have different structures:

Canada (2001)

USA (2010)

Iraq (1997)

Armenia (2001)



The image shows a sample of an Armenian census form for the year 2001. At the top left is the coat of arms of Armenia. The title at the top center reads "ՀՀ ԱԶԳԱՅԻՆ ՎԻՃԱԿԱԳՐԱԿԱՆ ԾԱՌ-ԱՅՈՒԹՅՈՒՆ" (Republic of Armenia Statistical Service). Below the title is the date "Հաստատված է՝ ՀՀ վիճակագրության պետական խորհրդի 26.07.2001թ. № 59 որոշմամբ" (Approved by the State Statistical Committee of the RA on July 26, 2001, No. 59 decision). The form includes a section for "ՀԱՐՅԱԹԵՐԹ" (Household) with a box for the household number. It contains several lines for personal information: "Մարդահամարի տեղամաս" (Census district), "Հրահանգչական տեղամաս" (Municipality), "Հաշվային տեղամաս" (Settlement), "Բնակելի միավոր" (Housing unit) with a box for the unit number, "Մարզ" (Region), "Տարածք" (Territory), "Նամայնք" (Community), and "Բնակավայր" (Settlement). There are also lines for "Պոստա, փողոց, նրբանցք" (Post, street, alley), "Տան/չների համար կամ Տանտիրոջ ազգանունը" (House number or family name), "Բնակարանի համարը" (Apartment number) and "Նկատյալի համարը" (Tenement number), and "Կազմակերպության անվանումը" (Organization name). At the bottom, there are lines for "Հաշվարար" (Enumerator) and "Հրահանգիչ-հսկիչ" (Supervisor) with checkboxes for "Ատղազտորտեղ" (Present) and "Կրողին անտեղ" (Absent). The year "2001" is printed at the bottom left. The text is in Armenian.

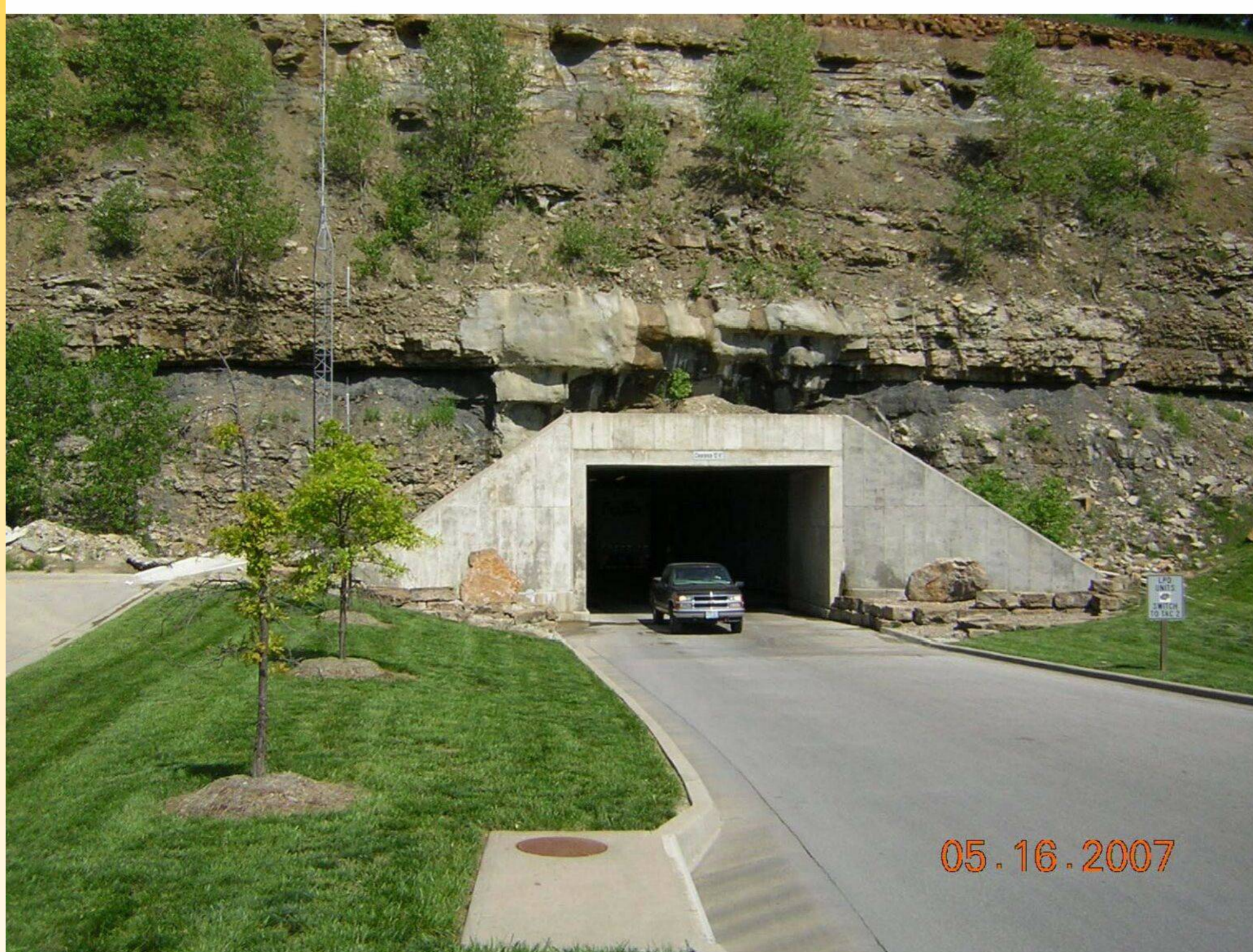


Adding data to IPUMS-I

- Data Acquisition
- Data Cleanup and Reformatting
- Data Integration
- Data Dissemination



The Problems: Acquisition



The Problems: Acquisition



The Problems: Cleanup (Physical)

1973 census tape #1: manual spacing required to recover data



The Problems: Cleanup (Logical)

- Reformatting - from whatever to Household/ Person hierarchical format
- Draw Sample (if necessary)
- Apply initial confidentiality measures (if necessary)
- Scan original documentation, and translate to English



The Problems: Cleanup (Logical)

- Recode variables to solely numeric values
- Check codes and marginal frequencies
- Document what we did

CR840018	label	cos1984
Marital status		P
		75
<tt>		
	0 NIU	B=Under age 10
	1 Consensual union	1=Consensual union
	2 Married	2=Married
	3 Separated	3=Separated
	4 Divorced	4=Divorced
	5 Widowed	5=Widowed
	6 Single	6=Single
	9 Missing	0=[undocumented]
	9 "	8=[undocumented]
</tt>		



The Problems: Data Integration

- Variable integration
- Constructed variables
- Variable documentation and descriptions



Integration - Marital Status?

China 1982	Columbia 1973
Never married Married Widowed Divorced	Consensual Union Married Separated or divorced Single Widowed
Kenya 1989	Mexico 1970
Single Married Monogamous Polygamous Widowed Divorced Separated	Married: Civil and Religious Married: Only Civil Married: Only Religious Consensual Union Widow Divorced Separated Single



Integration: Code

```
VAR: yrschl
// VAR: yrschl = Years of schooling.
switch(dataSet){
  case ke1989a:{
    int d = ke89a412(ln); // educ attainment
    int e = ke89a411(ln); // whether in school or ever attended
    int g = ke89a403(ln); // age
    if (g >= 20 && e != 1){
      // For persons age 20+ who are not currently in
      // school(?), reduce their years of schooling by 1 year
      // if their educational attainment was Form1
      // through Form4.
      switch(d){
        case 11: a = 8; break;
        case 12: a = 9; break;
        case 13: a = 10; break;
        case 14: a = 11; break;
      };
    }
  }break;
}
```



Integration: Documentation

```
<vardesc>
```

```
<var>  
MARST  
</var>
```

```
<desc>
```

MARST describes the person's current marital status according to law or custom. Individuals who remarried should report the status relevant to their most recent marriage. Census instructions rarely explicitly limit marital status to strictly legal unions.

```
</desc>
```

```
<comp>
```

The first digit of MARST is largely comparable across all samples. The second digit differentiates consensual unions from other forms of marriage and distinguishes among the categories separated, divorced, and married-spouse-absent. The final digit provides additional detail within the married and separated categories.

```
</comp>
```

```
<comp.am>
```

**The divorced category includes both legally registered and unregistered divorces.
All persons under age 15 are coded single.**

```
</comp.am>
```

```
<program><![CDATA[
```

```
Egypt 2006: if number of wives > 0, code marst as polygamous  
if (x=3 and (a408>1 and a408<8) )a=217;
```

```
]]></program>  
</vardesc>
```



Dissemination: Policy

- USA, CPS, IHIS - all US Domestic data
 - freely available to anyone
- IPUMS-International & NAPP
 - data usage governed by agreements with donor countries
 - only for academic use, must have an approved application to use
- No cost to use any MPC data project



Research Impact

- 5556 registered IPUMS-I users
- 36909 registered IPUMS USA users
- 526 bibliography citations for IPUMS-I

- 284 Google Scholar hits for IPUMS-I
- 4390 hits in Google Scholar for IPUMS
 - For comparison: 390K for Mathematica
 - and 1060 hits for “wolfram alpha”



Enabling Demographic Research

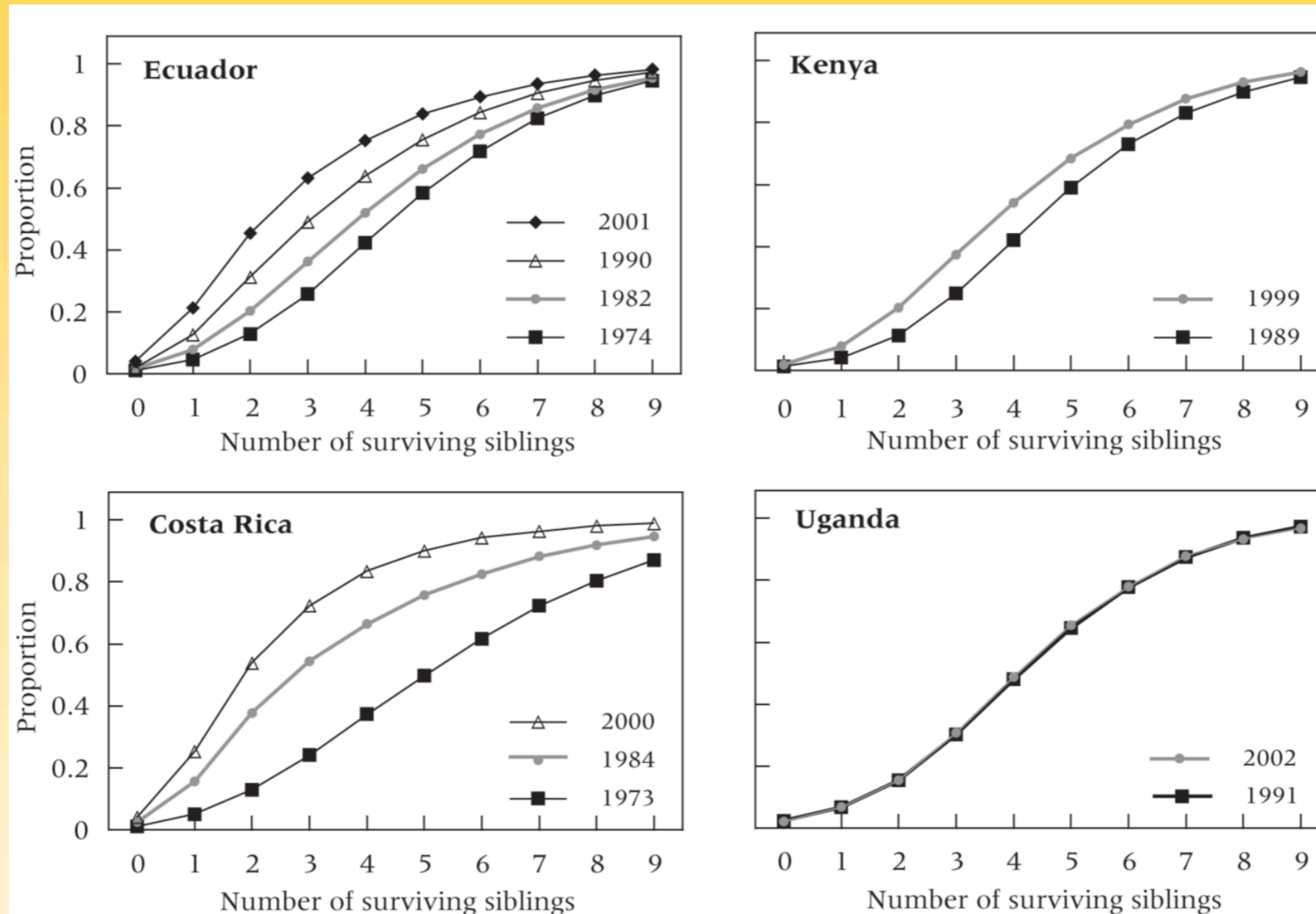


FIGURE 7. Cumulative distribution of siblings of children aged 9-11, eight countries: Proportion with specified number of surviving children or fewer

SOURCE: Estimates from IPUMS-International census microsamples (Minnesota Population Center 2007).

Enabling Demographic Research

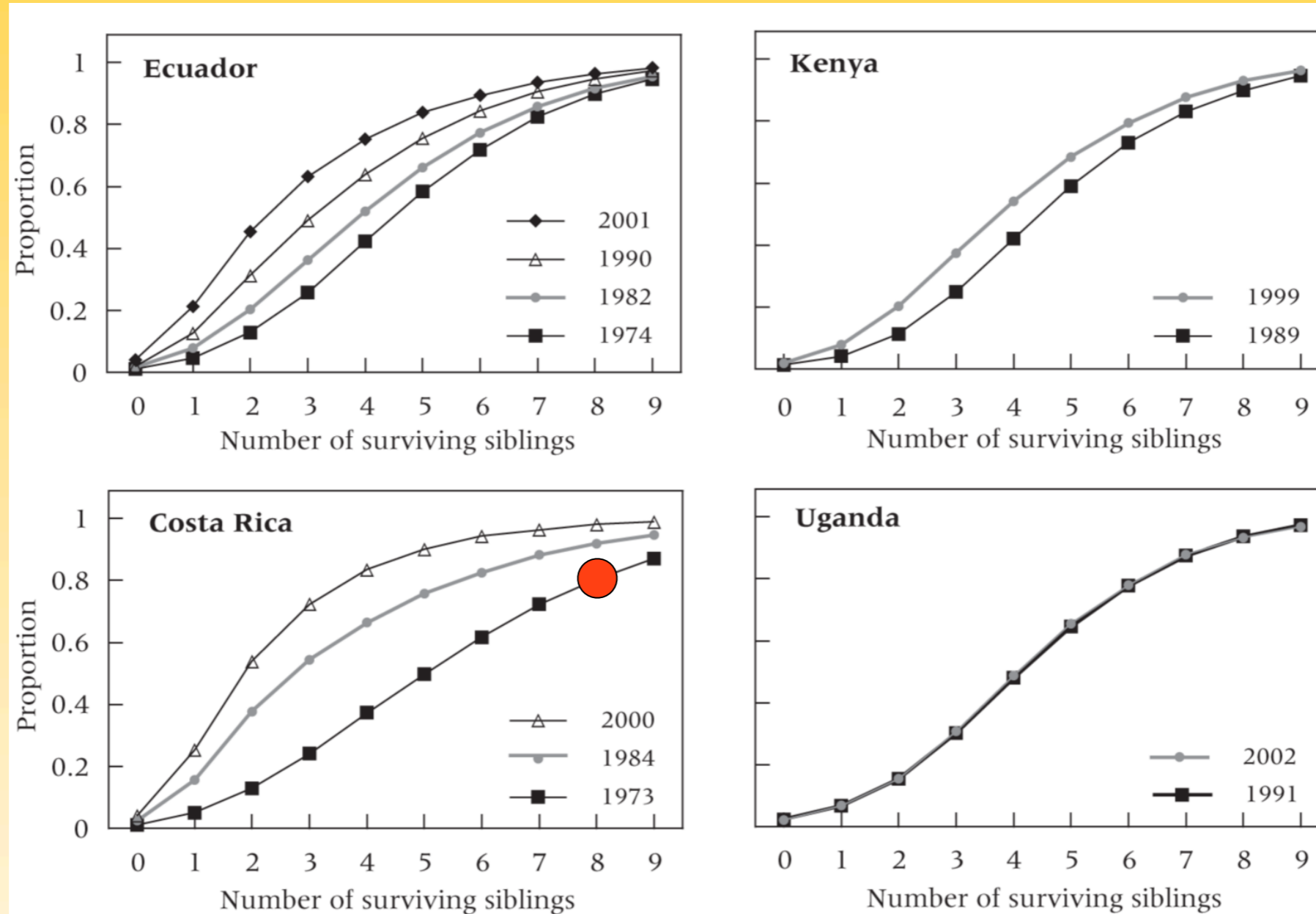


FIGURE 7. Cumulative distribution of siblings of children aged 9-11, eight countries: Proportion with specified number of surviving children or fewer

SOURCE: Estimates from IPUMS-International census microsamples (Minnesota Population Center 2007).

Enabling Demographic Research

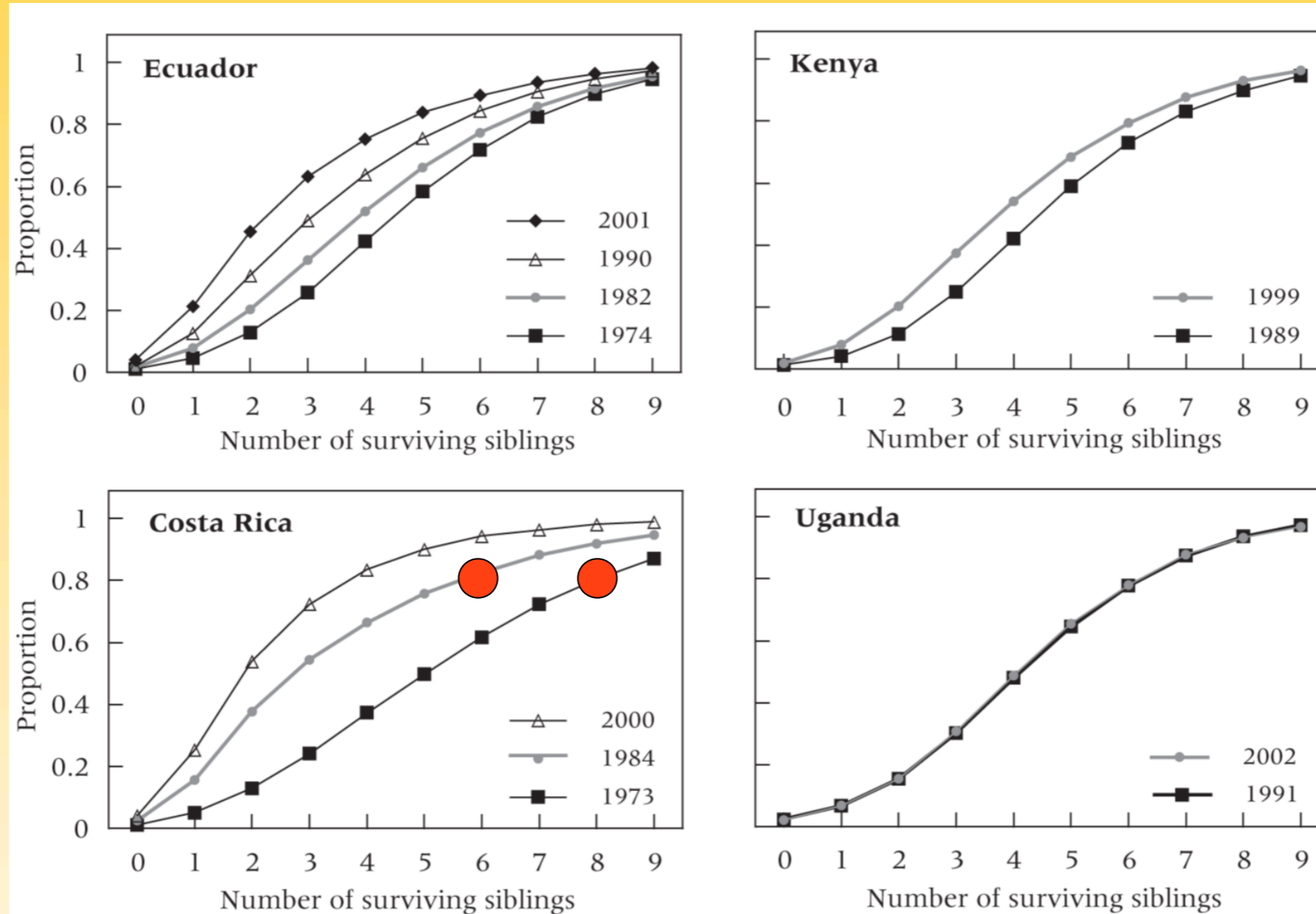


FIGURE 7. Cumulative distribution of siblings of children aged 9-11, eight countries: Proportion with specified number of surviving children or fewer

SOURCE: Estimates from IPUMS-International census microsamples (Minnesota Population Center 2007).

Enabling Demographic Research

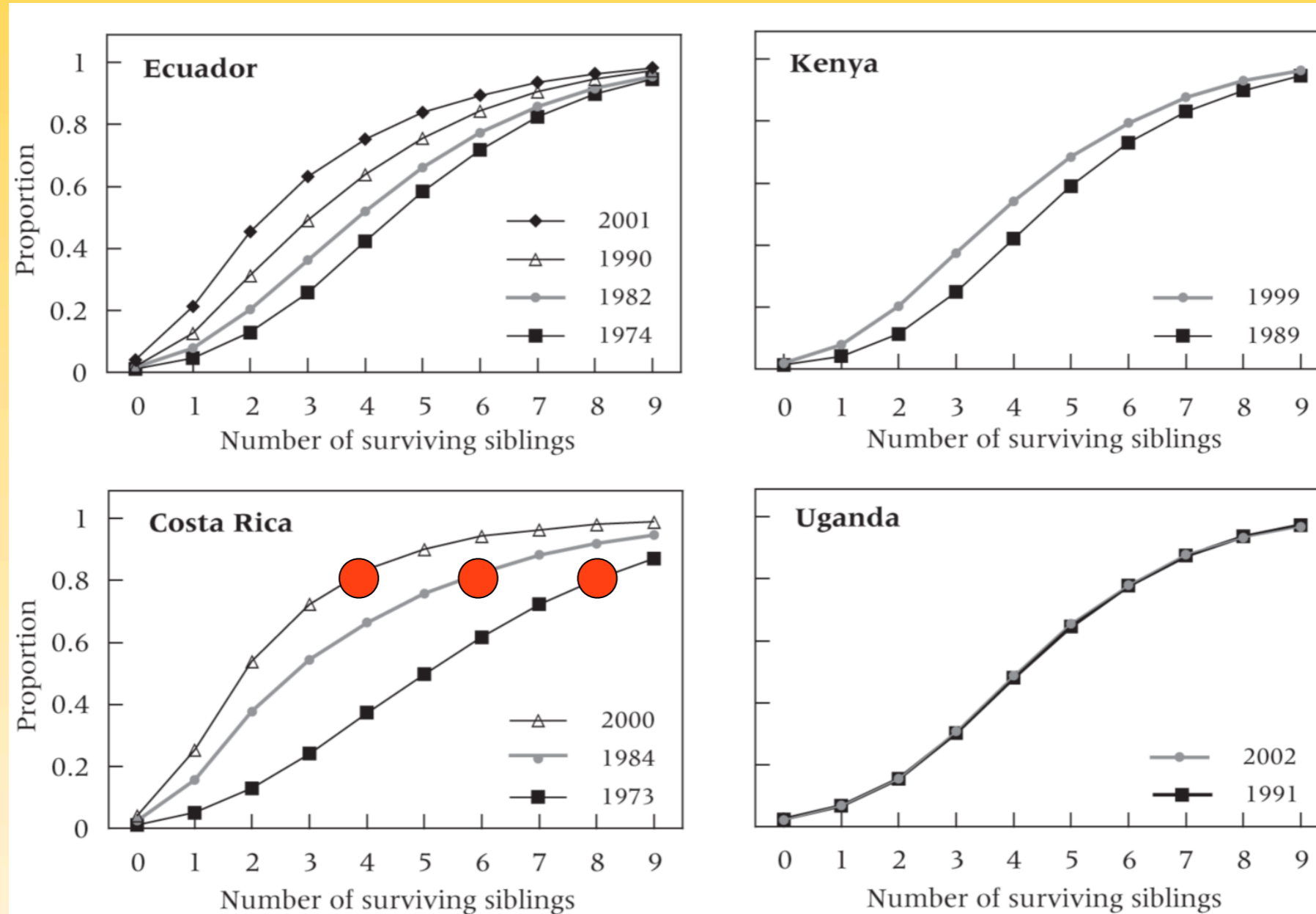


FIGURE 7. Cumulative distribution of siblings of children aged 9-11, eight countries: Proportion with specified number of surviving children or fewer
 SOURCE: Estimates from IPUMS-International census microsamples (Minnesota Population Center 2007).

Future

- Expecting to add another 100 samples over the next three years to IPUMS-International
- Word/Excel-based metadata system getting creaky
 - New metadata system (currently under development) will store metadata directly in database, allowing for checkin/checkout for editing.
- Use column-oriented db for actual data files
- Better leveraging of Big Data tools - hadoop?



Thank you!

Peter Clark
pclark@umn.edu

<http://www.pop.umn.edu>
<http://www.ipums.org>



Integration: Constructed Variables

(Simple household)

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

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

(Colombia 1985)

