1990 Census Names Recovery Project

Diane Cronkite and Trent Alexander
Center for Administrative Records Research and Applications
U.S. Census Bureau

FedCASIC Workshop
May 4, 2016

Project Team

U.S. Census Bureau

Trent Alexander, Diane Cronkite, Denise Flanagan-Doyle, Catherine Massey, Amy O'Hara

Stanford University

Jonathan Fisher, David Grusky, Matthew Snipp, Aliya Saperstein

National Academies of Sciences, Engineering, and Medicine

Robert Hauser, Carol House, Michael Hout

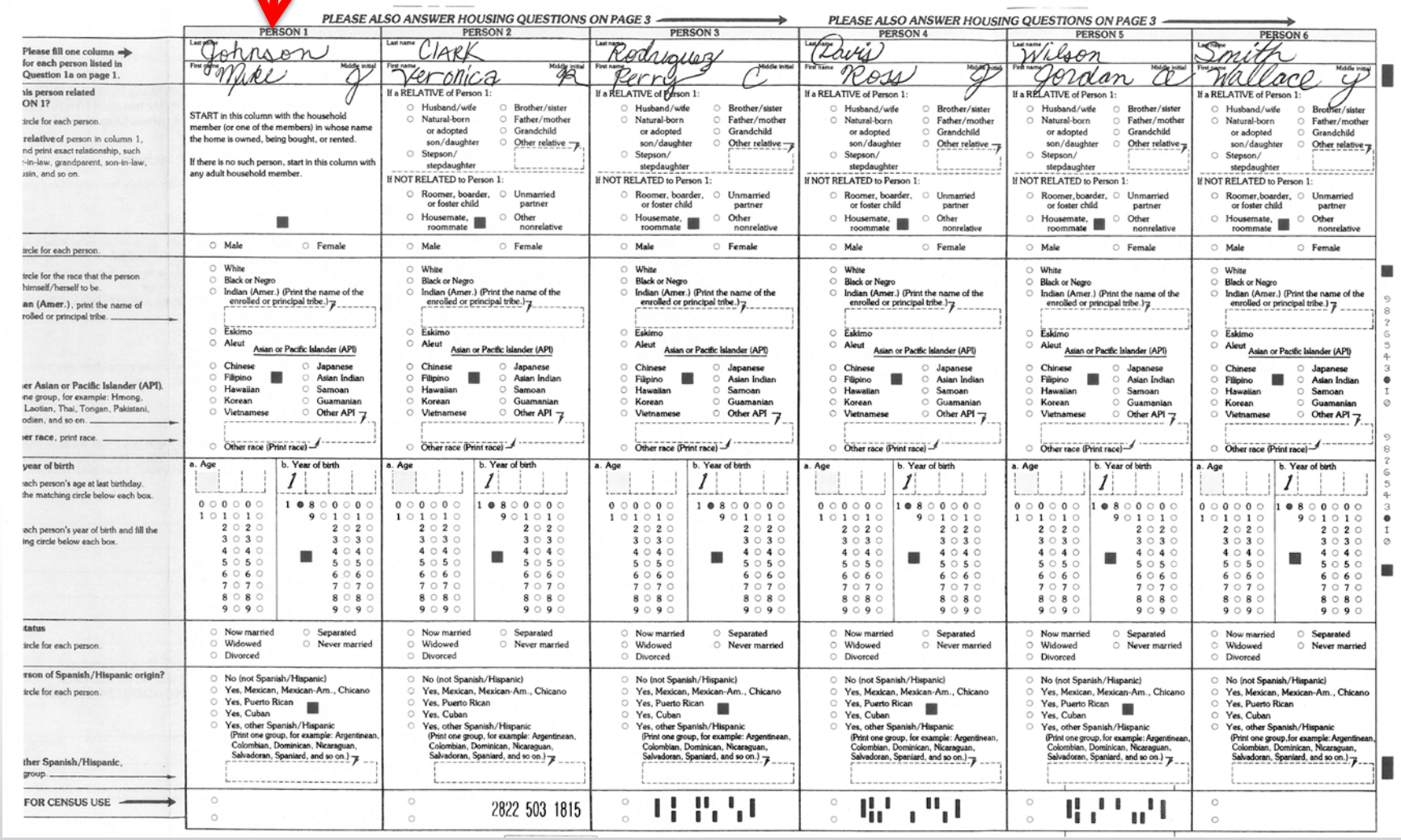
Background

- CARRA creates and uses person-level data linkages
  - Supports agency's use of administrative records
  - Increases utility of Census and Survey data
- Dedicated staff links administrative records, censuses, surveys
- Linkage keys available from 2000-present
- Key goal of this project
  - Demonstrate methods to extend infrastructure back in time
  - Start with 1990, then move to prior Censuses

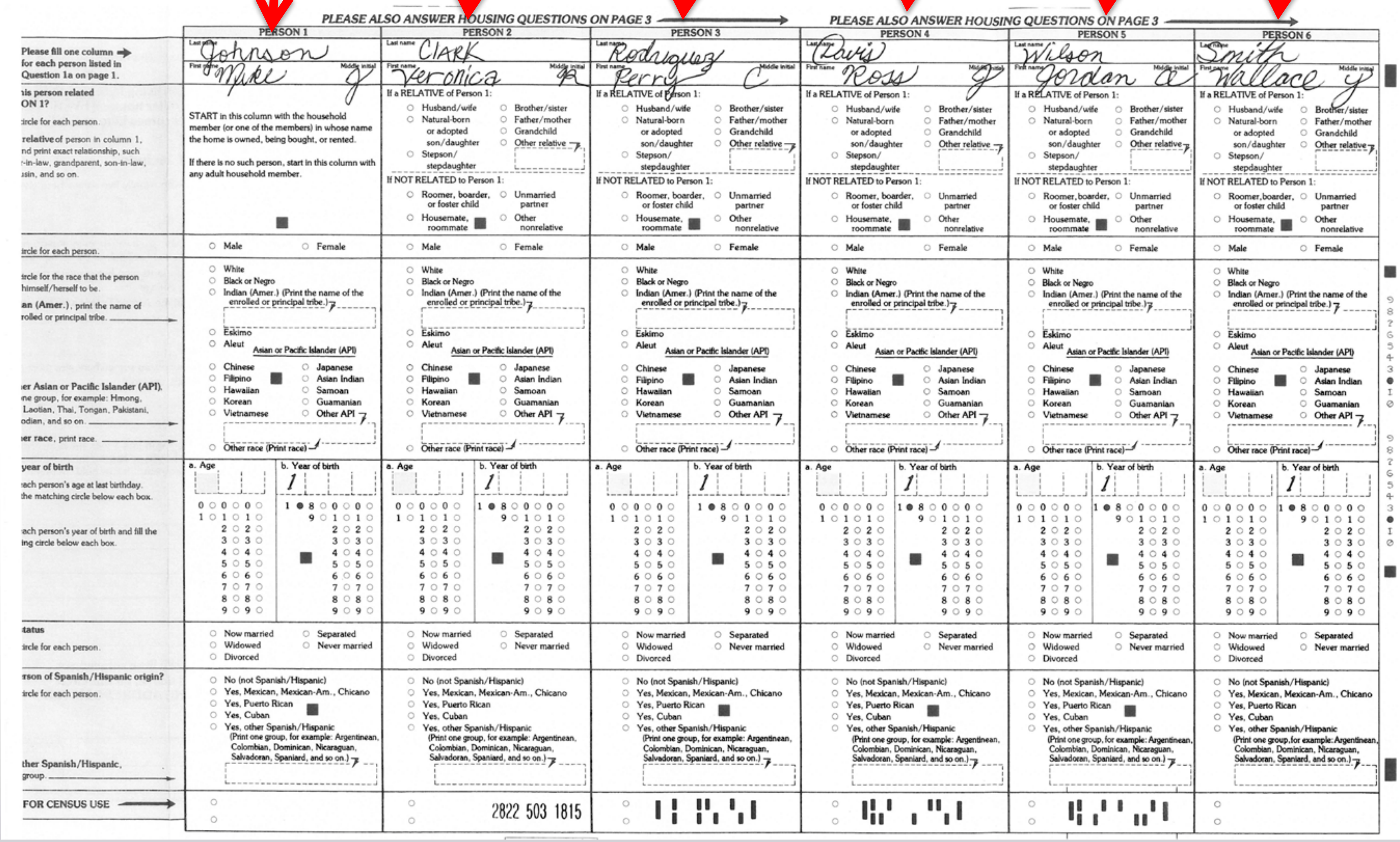
1990 Name Recovery Pilot

- 1990 Census
  - Names were handwritten on Census form but not captured electronically
  - Census forms exist on 130,000 microfilm reels
    - Census Bureau's National Processing Center
    - National Archives and Records Administration
  - Most other variables are already available in microdata file

Last name, first name, middle initial



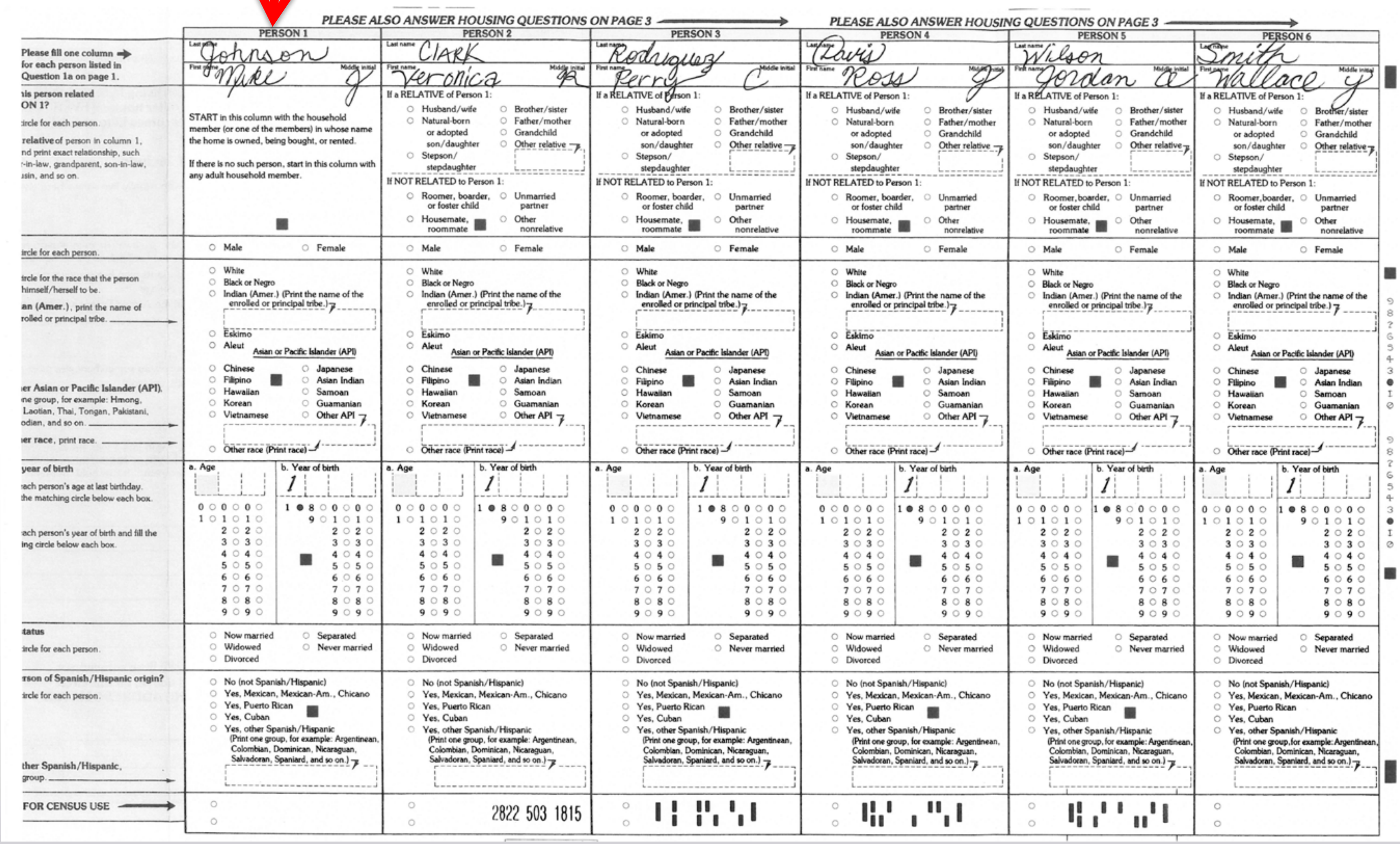
Up to six names on this page



PLEASE ALSO ANSWER HOUSING QUESTIONS ON PAGE 3 → PLEASE ALSO ANSWER HOUSING QUESTIONS ON PAGE 3 →

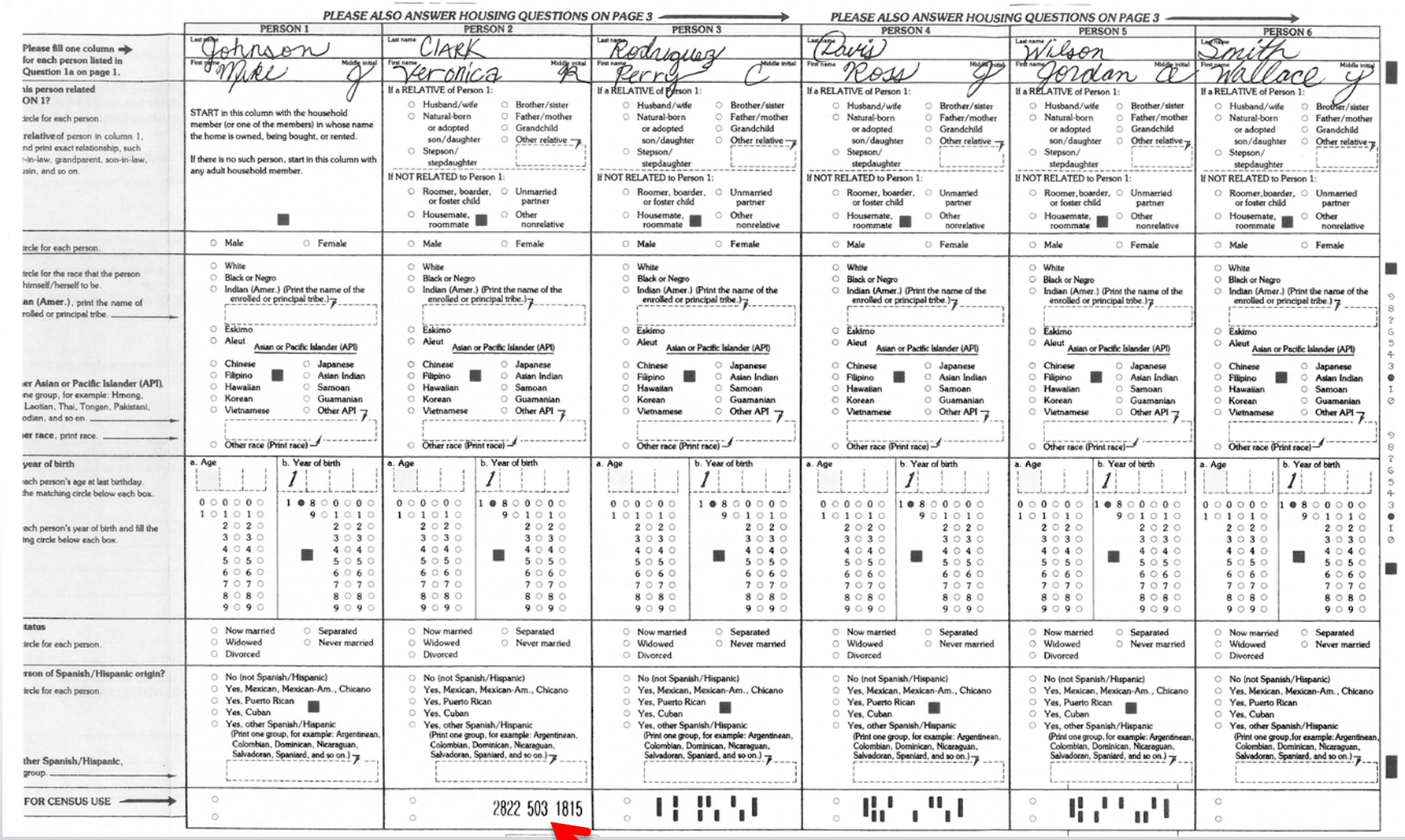
	PERSON 1	PERSON 2	PERSON 3	PERSON 4	PERSON 5	PERSON 6
Please fill one column → for each person listed in Question 1a on page 1.	Last name: Johnson; First name: Mike; Middle initial: J	Last name: CLARK; First name: Veronica; Middle initial: R	Last name: Rodriguez; First name: Perry; Middle initial: C	Last name: Davis; First name: Ross; Middle initial: J	Last name: Wilson; First name: Jordan; Middle initial: A	Last name: Smith; First name: Wallace; Middle initial: J
…is person related …ON 1? …ircle for each person. …relative of person in column 1, …nd print exact relationship, such …-in-law, grandparent, son-in-law, …sin, and so on.	START in this column with the household member (or one of the members) in whose name the home is owned, being bought, or rented. If there is no such person, start in this column with any adult household member.	If a RELATIVE of Person 1: Husband/wife; Natural-born or adopted son/daughter; Stepson/stepdaughter; Brother/sister; Father/mother; Grandchild; Other relative. If NOT RELATED to Person 1: Roomer, boarder, or foster child; Housemate, roommate; Unmarried partner; Other nonrelative	(same as Person 2)	(same as Person 2)	(same as Person 2)	(same as Person 2)
…ircle for each person.	Male / Female	Male / Female	Male / Female	Male / Female	Male / Female	Male / Female
…ircle for the race that the person himself/herself to be. …an (Amer.), print the name of …rolled or principal tribe. …er Asian or Pacific Islander (API), …ne group, for example: Hmong, Laotian, Thai, Tongan, Pakistani, …odian, and so on. …er race, print race.	White; Black or Negro; Indian (Amer.) (Print the name of the enrolled or principal tribe.); Eskimo; Aleut; Asian or Pacific Islander (API): Chinese, Filipino, Hawaiian, Korean, Vietnamese, Japanese, Asian Indian, Samoan, Guamanian, Other API; Other race (Print race)	(same)	(same)	(same)	(same)	(same)
…year of birth …ach person's age at last birthday. …the matching circle below each box. …ach person's year of birth and fill the …ing circle below each box.	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1
…tatus …ircle for each person.	Now married; Widowed; Divorced; Separated; Never married	(same)	(same)	(same)	(same)	(same)
…rson of Spanish/Hispanic origin? …ircle for each person. …ther Spanish/Hispanic, …group.	No (not Spanish/Hispanic); Yes, Mexican, Mexican-Am., Chicano; Yes, Puerto Rican; Yes, Cuban; Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)	(same)	(same)	(same)	(same)	(same)
FOR CENSUS USE →		2822 503 1815				

Person number



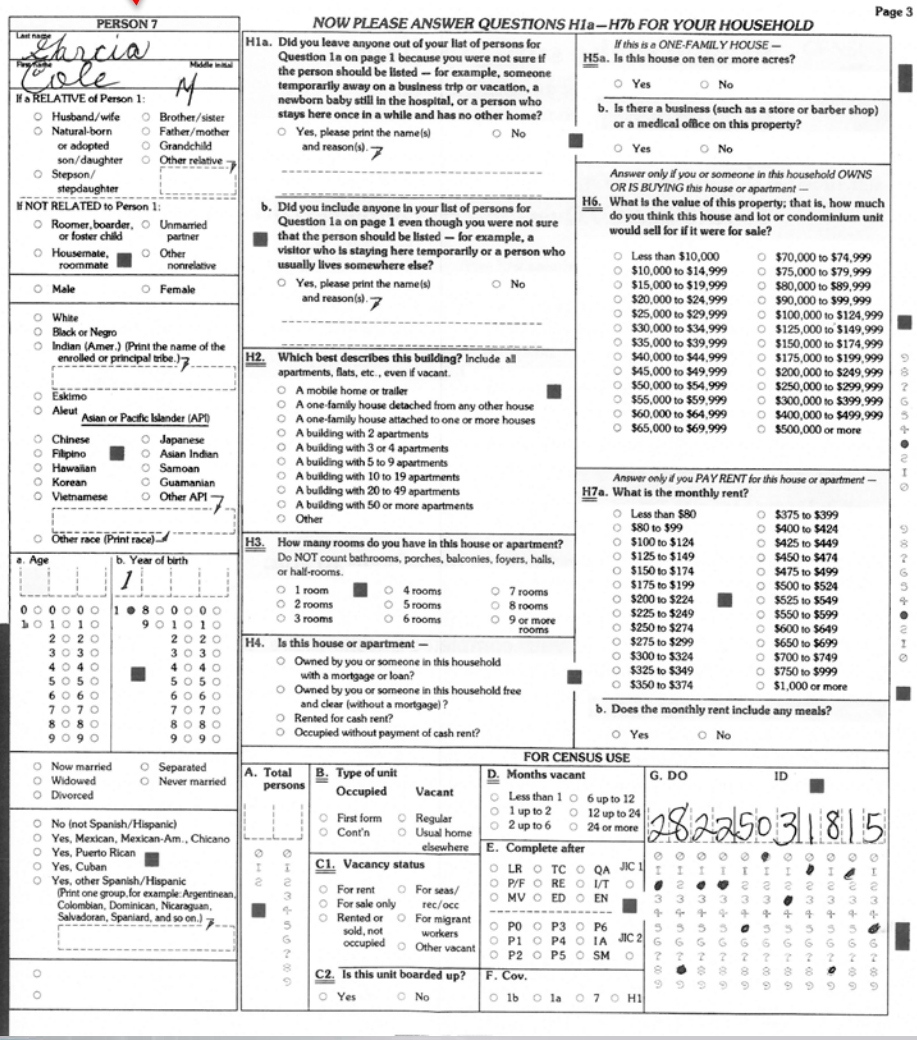
PLEASE ALSO ANSWER HOUSING QUESTIONS ON PAGE 3 → PLEASE ALSO ANSWER HOUSING QUESTIONS ON PAGE 3 →

	PERSON 1	PERSON 2	PERSON 3	PERSON 4	PERSON 5	PERSON 6
Please fill one column → for each person listed in Question 1a on page 1.	Last name: Johnson; First name: Mike; Middle initial: J	Last name: CLARK; First name: Veronica; Middle initial: R	Last name: Rodriguez; First name: Perry; Middle initial: C	Last name: Davis; First name: Ross; Middle initial: G	Last name: Wilson; First name: Jordan; Middle initial: A	Last name: Smith; First name: Wallace; Middle initial: J
…is person related …ON 1? …ircle for each person. …relative of person in column 1, …nd print exact relationship, such …-in-law, grandparent, son-in-law, …sin, and so on.	START in this column with the household member (or one of the members) in whose name the home is owned, being bought, or rented. If there is no such person, start in this column with any adult household member.	If a RELATIVE of Person 1: ○ Husband/wife ○ Natural-born or adopted son/daughter ○ Stepson/stepdaughter ○ Brother/sister ○ Father/mother ○ Grandchild ○ Other relative; If NOT RELATED to Person 1: ○ Roomer, boarder, or foster child ○ Housemate, roommate ○ Unmarried partner ○ Other nonrelative	If a RELATIVE of Person 1: ○ Husband/wife ○ Natural-born or adopted son/daughter ○ Stepson/stepdaughter ○ Brother/sister ○ Father/mother ○ Grandchild ○ Other relative; If NOT RELATED to Person 1: ○ Roomer, boarder, or foster child ○ Housemate, roommate ○ Unmarried partner ○ Other nonrelative	If a RELATIVE of Person 1: ○ Husband/wife ○ Natural-born or adopted son/daughter ○ Stepson/stepdaughter ○ Brother/sister ○ Father/mother ○ Grandchild ○ Other relative; If NOT RELATED to Person 1: ○ Roomer, boarder, or foster child ○ Housemate, roommate ○ Unmarried partner ○ Other nonrelative	If a RELATIVE of Person 1: ○ Husband/wife ○ Natural-born or adopted son/daughter ○ Stepson/stepdaughter ○ Brother/sister ○ Father/mother ○ Grandchild ○ Other relative; If NOT RELATED to Person 1: ○ Roomer, boarder, or foster child ○ Housemate, roommate ○ Unmarried partner ○ Other nonrelative	If a RELATIVE of Person 1: ○ Husband/wife ○ Natural-born or adopted son/daughter ○ Stepson/stepdaughter ○ Brother/sister ○ Father/mother ○ Grandchild ○ Other relative; If NOT RELATED to Person 1: ○ Roomer, boarder, or foster child ○ Housemate, roommate ○ Unmarried partner ○ Other nonrelative
…ircle for each person.	○ Male ○ Female	○ Male ○ Female	○ Male ○ Female	○ Male ○ Female	○ Male ○ Female	○ Male ○ Female
…ircle for the race that the person …himself/herself to be. …an (Amer.), print the name of …rolled or principal tribe. → …er Asian or Pacific Islander (API), …ne group, for example: Hmong, Laotian, Thai, Tongan, Pakistani, …odian, and so on. → …er race, print race. →	○ White ○ Black or Negro ○ Indian (Amer.) (Print the name of the enrolled or principal tribe.) ○ Eskimo ○ Aleut; Asian or Pacific Islander (API): ○ Chinese ○ Filipino ○ Hawaiian ○ Korean ○ Vietnamese ○ Japanese ○ Asian Indian ○ Samoan ○ Guamanian ○ Other API; ○ Other race (Print race)	○ White ○ Black or Negro ○ Indian (Amer.) (Print the name of the enrolled or principal tribe.) ○ Eskimo ○ Aleut; Asian or Pacific Islander (API): ○ Chinese ○ Filipino ○ Hawaiian ○ Korean ○ Vietnamese ○ Japanese ○ Asian Indian ○ Samoan ○ Guamanian ○ Other API; ○ Other race (Print race)	○ White ○ Black or Negro ○ Indian (Amer.) (Print the name of the enrolled or principal tribe.) ○ Eskimo ○ Aleut; Asian or Pacific Islander (API): ○ Chinese ○ Filipino ○ Hawaiian ○ Korean ○ Vietnamese ○ Japanese ○ Asian Indian ○ Samoan ○ Guamanian ○ Other API; ○ Other race (Print race)	○ White ○ Black or Negro ○ Indian (Amer.) (Print the name of the enrolled or principal tribe.) ○ Eskimo ○ Aleut; Asian or Pacific Islander (API): ○ Chinese ○ Filipino ○ Hawaiian ○ Korean ○ Vietnamese ○ Japanese ○ Asian Indian ○ Samoan ○ Guamanian ○ Other API; ○ Other race (Print race)	○ White ○ Black or Negro ○ Indian (Amer.) (Print the name of the enrolled or principal tribe.) ○ Eskimo ○ Aleut; Asian or Pacific Islander (API): ○ Chinese ○ Filipino ○ Hawaiian ○ Korean ○ Vietnamese ○ Japanese ○ Asian Indian ○ Samoan ○ Guamanian ○ Other API; ○ Other race (Print race)	○ White ○ Black or Negro ○ Indian (Amer.) (Print the name of the enrolled or principal tribe.) ○ Eskimo ○ Aleut; Asian or Pacific Islander (API): ○ Chinese ○ Filipino ○ Hawaiian ○ Korean ○ Vietnamese ○ Japanese ○ Asian Indian ○ Samoan ○ Guamanian ○ Other API; ○ Other race (Print race)
…year of birth …ach person's age at last birthday. …the matching circle below each box. …ach person's year of birth and fill the …ing circle below each box.	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1	a. Age; b. Year of birth 1
…tatus …ircle for each person.	○ Now married ○ Widowed ○ Divorced ○ Separated ○ Never married	○ Now married ○ Widowed ○ Divorced ○ Separated ○ Never married	○ Now married ○ Widowed ○ Divorced ○ Separated ○ Never married	○ Now married ○ Widowed ○ Divorced ○ Separated ○ Never married	○ Now married ○ Widowed ○ Divorced ○ Separated ○ Never married	○ Now married ○ Widowed ○ Divorced ○ Separated ○ Never married
…rson of Spanish/Hispanic origin? …ircle for each person. …ther Spanish/Hispanic, …group. →	○ No (not Spanish/Hispanic) ○ Yes, Mexican, Mexican-Am., Chicano ○ Yes, Puerto Rican ○ Yes, Cuban ○ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)	○ No (not Spanish/Hispanic) ○ Yes, Mexican, Mexican-Am., Chicano ○ Yes, Puerto Rican ○ Yes, Cuban ○ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)	○ No (not Spanish/Hispanic) ○ Yes, Mexican, Mexican-Am., Chicano ○ Yes, Puerto Rican ○ Yes, Cuban ○ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)	○ No (not Spanish/Hispanic) ○ Yes, Mexican, Mexican-Am., Chicano ○ Yes, Puerto Rican ○ Yes, Cuban ○ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)	○ No (not Spanish/Hispanic) ○ Yes, Mexican, Mexican-Am., Chicano ○ Yes, Puerto Rican ○ Yes, Cuban ○ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)	○ No (not Spanish/Hispanic) ○ Yes, Mexican, Mexican-Am., Chicano ○ Yes, Puerto Rican ○ Yes, Cuban ○ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)
FOR CENSUS USE →	○ ○	○ ○ 2822 503 1815	○ ○	○ ○	○ ○	○ ○



11-digit household ID

Person 7 information



PERSON 7

Last name: Garcia
First name: Cole
Middle initial: M

If a RELATIVE of Person 1:
- Husband/wife
- Natural-born or adopted son/daughter
- Stepson/stepdaughter
- Brother/sister
- Father/mother
- Grandchild
- Other relative

If NOT RELATED to Person 1:
- Roomer, boarder, or foster child
- Housemate, roommate
- Unmarried partner
- Other nonrelative

- Male
- Female

- White
- Black or Negro
- Indian (Amer.) (Print the name of the enrolled or principal tribe.)
- Eskimo
- Aleut

Asian or Pacific Islander (API)
- Chinese
- Filipino
- Hawaiian
- Korean
- Vietnamese
- Japanese
- Asian Indian
- Samoan
- Guamanian
- Other API

- Other race (Print race)

a. Age: 1
b. Year of birth: 1980

- Now married
- Widowed
- Divorced
- Separated
- Never married

- No (not Spanish/Hispanic)
- Yes, Mexican, Mexican-Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)

NOW PLEASE ANSWER QUESTIONS H1a–H7b FOR YOUR HOUSEHOLD

Page 3

H1a. Did you leave anyone out of your list of persons for Question 1a on page 1 because you were not sure if the person should be listed — for example, someone temporarily away on a business trip or vacation, a newborn baby still in the hospital, or a person who stays here once in a while and has no other home?
- Yes, please print the name(s) and reason(s).
- No

b. Did you include anyone in your list of persons for Question 1a on page 1 even though you were not sure that the person should be listed — for example, a visitor who is staying here temporarily or a person who usually lives somewhere else?
- Yes, please print the name(s) and reason(s).
- No

H2. Which best describes this building? Include all apartments, flats, etc., even if vacant.
- A mobile home or trailer
- A one-family house detached from any other house
- A one-family house attached to one or more houses
- A building with 2 apartments
- A building with 3 or 4 apartments
- A building with 5 to 9 apartments
- A building with 10 to 19 apartments
- A building with 20 to 49 apartments
- A building with 50 or more apartments
- Other

H3. How many rooms do you have in this house or apartment? Do NOT count bathrooms, porches, balconies, foyers, halls, or half-rooms.
- 1 room
- 2 rooms
- 3 rooms
- 4 rooms
- 5 rooms
- 6 rooms
- 7 rooms
- 8 rooms
- 9 or more rooms

H4. Is this house or apartment —
- Owned by you or someone in this household with a mortgage or loan?
- Owned by you or someone in this household free and clear (without a mortgage)?
- Rented for cash rent?
- Occupied without payment of cash rent?

If this is a ONE-FAMILY HOUSE —

H5a. Is this house on ten or more acres?
- Yes
- No

b. Is there a business (such as a store or barber shop) or a medical office on this property?
- Yes
- No

Answer only if you or someone in this household OWNS OR IS BUYING this house or apartment —

H6. What is the value of this property; that is, how much do you think this house and lot or condominium unit would sell for if it were for sale?
- Less than $10,000
- $10,000 to $14,999
- $15,000 to $19,999
- $20,000 to $24,999
- $25,000 to $29,999
- $30,000 to $34,999
- $35,000 to $39,999
- $40,000 to $44,999
- $45,000 to $49,999
- $50,000 to $54,999
- $55,000 to $59,999
- $60,000 to $64,999
- $65,000 to $69,999
- $70,000 to $74,999
- $75,000 to $79,999
- $80,000 to $89,999
- $90,000 to $99,999
- $100,000 to $124,999
- $125,000 to $149,999
- $150,000 to $174,999
- $175,000 to $199,999
- $200,000 to $249,999
- $250,000 to $299,999
- $300,000 to $399,999
- $400,000 to $499,999
- $500,000 or more

Answer only if you PAY RENT for this house or apartment —

H7a. What is the monthly rent?
- Less than $80
- $80 to $99
- $100 to $124
- $125 to $149
- $150 to $174
- $175 to $199
- $200 to $224
- $225 to $249
- $250 to $274
- $275 to $299
- $300 to $324
- $325 to $349
- $350 to $374
- $375 to $399
- $400 to $424
- $425 to $449
- $450 to $474
- $475 to $499
- $500 to $524
- $525 to $549
- $550 to $599
- $600 to $649
- $650 to $699
- $700 to $749
- $750 to $999
- $1,000 or more

b. Does the monthly rent include any meals?
- Yes
- No

FOR CENSUS USE

A. Total persons

B. Type of unit
Occupied: First form; Cont'n
Vacant: Regular; Usual home elsewhere

C1. Vacancy status
- For rent
- For sale only
- Rented or sold, not occupied
- For seas/rec/occ
- For migrant workers
- Other vacant

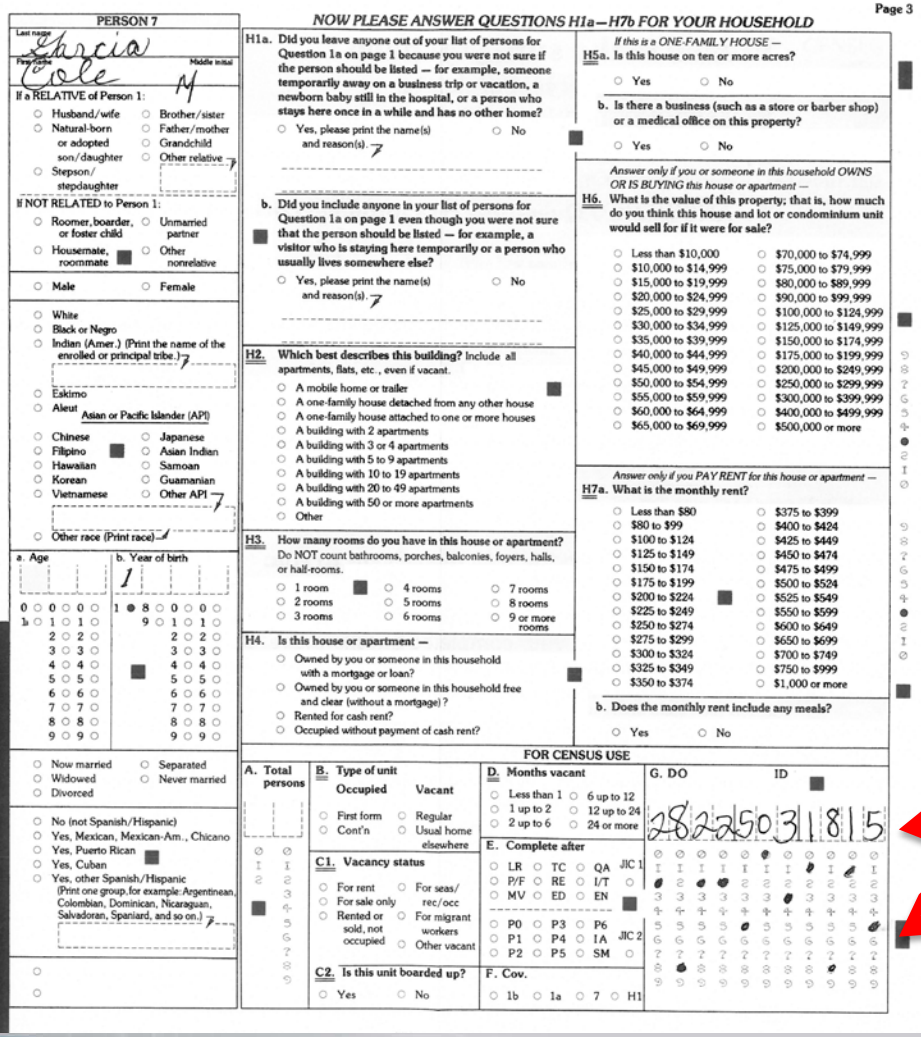
C2. Is this unit boarded up?
- Yes
- No

D. Months vacant
- Less than 1
- 1 up to 2
- 2 up to 6
- 6 up to 12
- 12 up to 24
- 24 or more

E. Complete after
LR TC QA JIC 1
P/F RE I/T
MV ED EN
P0 P3 P6 JIC 2
P1 P4 IA
P2 P5 SM

F. Cov.
1b 1a 7 H1

G. DO ID: 28225031815



PERSON 7

Last name: Garcia
First name: Cole
Middle initial: M

If a RELATIVE of Person 1:
- Husband/wife
- Natural-born or adopted son/daughter
- Stepson/stepdaughter
- Brother/sister
- Father/mother
- Grandchild
- Other relative — 7

If NOT RELATED to Person 1:
- Roomer, boarder, or foster child
- Housemate, roommate
- Unmarried partner
- Other nonrelative

- Male
- Female

- White
- Black or Negro
- Indian (Amer.) (Print the name of the enrolled or principal tribe.) — 7
- Eskimo
- Aleut

Asian or Pacific Islander (API)
- Chinese
- Filipino
- Hawaiian
- Korean
- Vietnamese
- Japanese
- Asian Indian
- Samoan
- Guamanian
- Other API — 7
- Other race (Print race) — 7

a. Age: 1

b. Year of birth: 1980

- Now married
- Widowed
- Divorced
- Separated
- Never married

- No (not Spanish/Hispanic)
- Yes, Mexican, Mexican-Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.) — 7

NOW PLEASE ANSWER QUESTIONS H1a–H7b FOR YOUR HOUSEHOLD

Page 3

H1a. Did you leave anyone out of your list of persons for Question 1a on page 1 because you were not sure if the person should be listed — for example, someone temporarily away on a business trip or vacation, a newborn baby still in the hospital, or a person who stays here once in a while and has no other home?
- Yes, please print the name(s) and reason(s). — 7
- No

b. Did you include anyone in your list of persons for Question 1a on page 1 even though you were not sure that the person should be listed — for example, a visitor who is staying here temporarily or a person who usually lives somewhere else?
- Yes, please print the name(s) and reason(s). — 7
- No

H2. Which best describes this building? Include all apartments, flats, etc., even if vacant.
- A mobile home or trailer
- A one-family house detached from any other house
- A one-family house attached to one or more houses
- A building with 2 apartments
- A building with 3 or 4 apartments
- A building with 5 to 9 apartments
- A building with 10 to 19 apartments
- A building with 20 to 49 apartments
- A building with 50 or more apartments
- Other

H3. How many rooms do you have in this house or apartment? Do NOT count bathrooms, porches, balconies, foyers, halls, or half-rooms.
- 1 room
- 2 rooms
- 3 rooms
- 4 rooms
- 5 rooms
- 6 rooms
- 7 rooms
- 8 rooms
- 9 or more rooms

H4. Is this house or apartment —
- Owned by you or someone in this household with a mortgage or loan?
- Owned by you or someone in this household free and clear (without a mortgage)?
- Rented for cash rent?
- Occupied without payment of cash rent?

If this is a ONE-FAMILY HOUSE —

H5a. Is this house on ten or more acres?
- Yes
- No

b. Is there a business (such as a store or barber shop) or a medical office on this property?
- Yes
- No

Answer only if you or someone in this household OWNS OR IS BUYING this house or apartment —

H6. What is the value of this property; that is, how much do you think this house and lot or condominium unit would sell for if it were for sale?
- Less than $10,000
- $10,000 to $14,999
- $15,000 to $19,999
- $20,000 to $24,999
- $25,000 to $29,999
- $30,000 to $34,999
- $35,000 to $39,999
- $40,000 to $44,999
- $45,000 to $49,999
- $50,000 to $54,999
- $55,000 to $59,999
- $60,000 to $64,999
- $65,000 to $69,999
- $70,000 to $74,999
- $75,000 to $79,999
- $80,000 to $89,999
- $90,000 to $99,999
- $100,000 to $124,999
- $125,000 to $149,999
- $150,000 to $174,999
- $175,000 to $199,999
- $200,000 to $249,999
- $250,000 to $299,999
- $300,000 to $399,999
- $400,000 to $499,999
- $500,000 or more

Answer only if you PAY RENT for this house or apartment —

H7a. What is the monthly rent?
- Less than $80
- $80 to $99
- $100 to $124
- $125 to $149
- $150 to $174
- $175 to $199
- $200 to $224
- $225 to $249
- $250 to $274
- $275 to $299
- $300 to $324
- $325 to $349
- $350 to $374
- $375 to $399
- $400 to $424
- $425 to $449
- $450 to $474
- $475 to $499
- $500 to $524
- $525 to $549
- $550 to $599
- $600 to $649
- $650 to $699
- $700 to $749
- $750 to $999
- $1,000 or more

b. Does the monthly rent include any meals?
- Yes
- No

FOR CENSUS USE

A. Total persons

B. Type of unit
Occupied: First form; Cont'n
Vacant: Regular; Usual home elsewhere

C1. Vacancy status
- For rent
- For sale only
- Rented or sold, not occupied
- For seas/rec/occ
- For migrant workers
- Other vacant

C2. Is this unit boarded up?
- Yes
- No

D. Months vacant
- Less than 1
- 1 up to 2
- 2 up to 6
- 6 up to 12
- 12 up to 24
- 24 or more

E. Complete after
- LR TC QA JIC 1
- P/F RE I/T
- MV ED EN
- P0 P3 P6
- P1 P4 IA JIC 2
- P2 P5 SM

F. Cov.
- 1b
- 1a
- 7
- H1

G. DO ID: 28225031815

Handwritten household ID with FOSDIC bubbles

Scope of Work

- **Scan Microfilm**
- Make hand-keyed “truth data”
- Do Optical Character Recognition (OCR)
- Evaluate Results

Scanning Microfilm

- Goals
  - Determine best settings
  - Determine best scanner
  - Estimate cost of scanning all 130,000 reels

Scanning Microfilm

MEKEL Technology

CROWLEY

MEET THE MACH MICROFILM SCANNERS

And Meet (or Exceed) True Optical DPI Requirements with Ease and Precision

For archivists, records managers and service bureaus there is no greater measure of success than accuracy to the original media. In today's business climate, that accuracy has to be paired with maximum efficiency in order to be cost-effective. The Mekel Technology ***MACH-series*** microfilm scanners combine the best of worlds, using NARA and Library of Congress preservation specifications as the gold standards and pairing them with innovative technology for an unmatched return on investment.

ALL MEKEL PRODUCTION MICROFILM SYSTEMS FEATURE:

- Robust work ethic; works 24/7 with minimal maintenance requirements
- 16mm and 35mm microfilm conversion capability
- Scanning up to 1,000' rollfilm and 3M/ANSI/Ektamate cartridges as standard
- Conversion to most electronic imaging formats, including PDF varieties and archival formats
- Fast throughput
- Highly-refined precision lenses
- Mekel's one-of-a-kind ***QuantuM Scan*** and ***QuantuM Process*** suite
- 100% accurate image capture; no re-loading necessary – scan once and done
- A sophisticated toolkit of automated and manual image processing features to guarantee the best image quality possible; eliminates the need for re-scans
- An external PC, allowing for maximum uptime and technology updates
- Calibrated/focused LED light source
- Sturdy case with built-in reel storage
- Stable and gentle transport system
- "Load and go" intuitive operation

MACH12 Key Features

- High optical resolution: 100-750 dpi true optical resolution
- 12-bit camera exceeds true optical dpi requirements for capture of books, newspapers, manuscripts, financial records and more
- Unrivaled image quality and capture accuracy – excellent tonality and depth of field
- Crystal clear scanning of small fonts and highly-reduced images
- Higher OCR accuracy as a result of better image quality

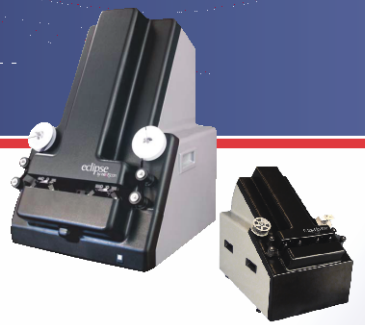
MACH10 Key Features

- Highest efficiency for the combination of speed and image quality
- Scans up to 700 images per minute/3 minutes per 100' roll (up to 1400 images per minute via speed mode)

MACH5 Key Features

- Offers speed and image quality when scanning a wide variety of microfilm in various states of composition
- Scans under 6 minutes per 100'roll

Eclipse with nextStar



The World's Premier scanning platform for film

The Eclipse 16/35mm roll film scanning solution provides a complete package of hardware, software, training and support to provide the fastest throughput and highest quality digital images in the market today.

The Eclipse 16/35mm roll film scanning solution includes the scanner, PC, high resolution lens, NextStar software, on-site training and long term support. Also offered with this solution is an optional expanded digital storage server (RSD), cassette adapter and/or 1000' reel adapter.

The system can handle all of the challenging properties of roll film scanning including:

- Variable size documents
- Variable position images
- Multiple film files
- Variable indexing to digital files
- Captures up to 3 levels of blips
- Enhancement of signatures and handwritten data on filmed images
- Multiple output formats including PDF/A

Unique scanner features:

- Scan 16 and 35mm roll film at speeds up to 1000 ppm and beyond
- Patented LuminTec illumination system runs cool with lower power consumption, longer lamp life and sharper images
- Ribbon scanning software for guaranteed image capture and detection
- Patented film transport to ensure linearity (eliminates image distortion due to environmental vibration)
- Special capabilities for handling 1000' reels and blipped film
- Choice of scanner chassis to accommodate internal or external expanded digital storage
- Scan system workflow is configured to minimize any impact on your internal network

nextScan The Next Generation in Film and Fiche Scanning Technology

Scanning Microfilm

- Worked at Census's National Processing Center
  - Census has copy of the archival original reels
  - 2 microfilm scanners
  - Scanned 600 reels
    - Mix of short-form and long-form census
    - >1,000,000 total images
  - October – December 2015

Scanning Microfilm

- National Archives master version
  - Census Bureau version is a copy of these
  - National Archives scanned 2 reels
    - Provided images to Census
    - Images looked better but achieved similar OCR results
    - Is performing OCR themselves on these reels

Scope of Work

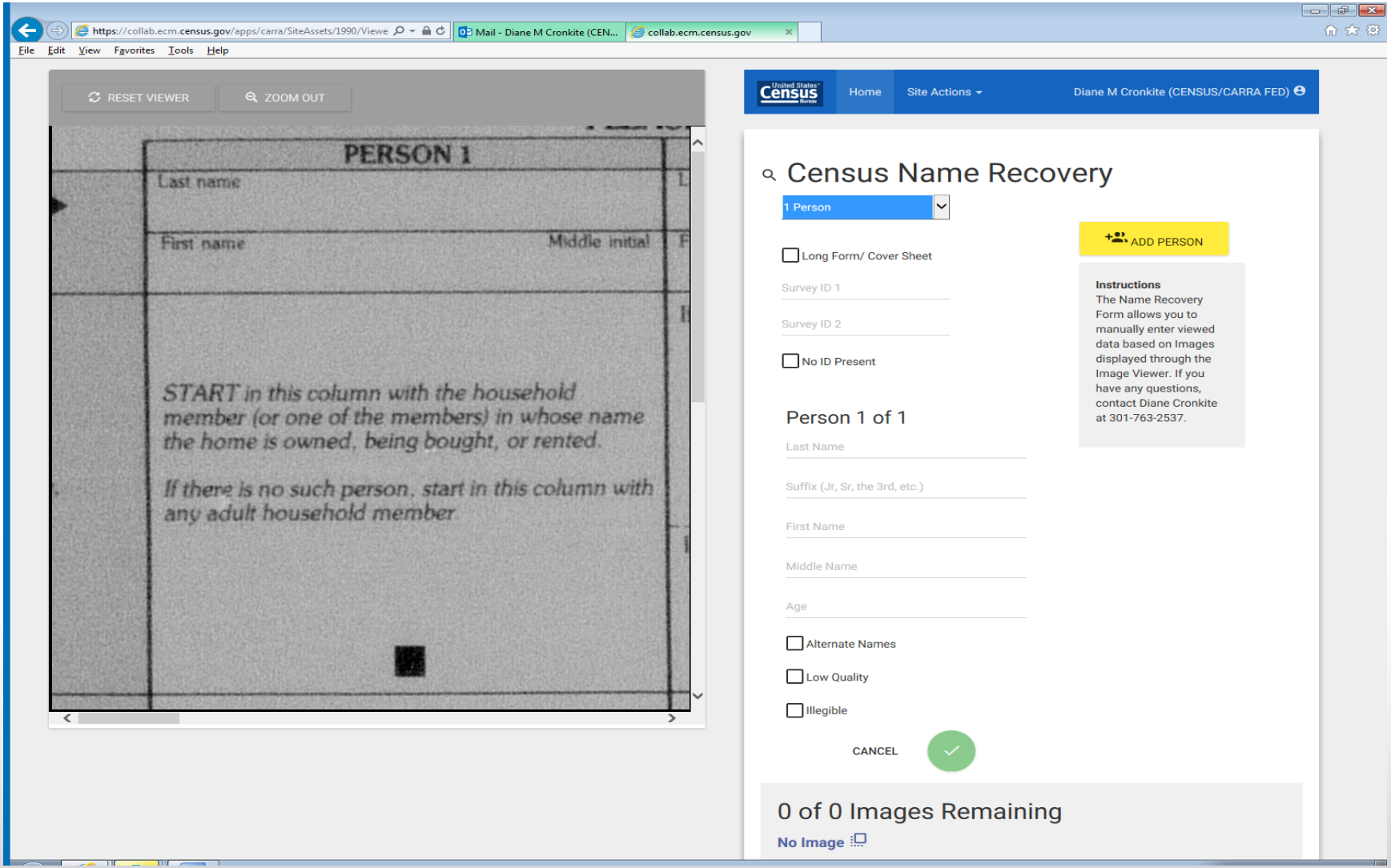
- Scan Microfilm
- **Make hand-keyed "truth data"**
- Do Optical Character Recognition (OCR)
- Evaluate Results

Hand-keyed “Truth Data”

- Goals
  - Create “truth data” to evaluate OCR results
  - Double-keyed data gives a measure of keying error
  - To measure proportion of hard-to-read names
    - If a person cannot read the name, this should not be counted as an OCR “error”

Hand-keyed "Truth Data"

- Used a key-from-image SharePoint application
  - Developed by Census's Center for Applied Technology
- Hand-keyed 44,000 names – double-keyed
  - double-keyed names matched 95% of the time
- December 2015 – March 2016


https://collab.ecm.census.gov/apps/carra/SiteAssets/1990/Viewe
Mail - Diane M Cronkite (CEN...
collab.ecm.census.gov
File Edit View Favorites Tools Help
RESET VIEWER
ZOOM OUT
PERSON 1
Last name
First name
Middle initial
START in this column with the household member (or one of the members) in whose name the home is owned, being bought, or rented.
If there is no such person, start in this column with any adult household member.
United States Census Bureau
Home
Site Actions
Diane M Cronkite (CENSUS/CARRA FED)
Census Name Recovery
1 Person
ADD PERSON
Long Form/ Cover Sheet
Survey ID 1
Survey ID 2
No ID Present
Instructions
The Name Recovery Form allows you to manually enter viewed data based on Images displayed through the Image Viewer. If you have any questions, contact Diane Cronkite at 301-763-2537.
Person 1 of 1
Last Name
Suffix (Jr, Sr, the 3rd, etc.)
First Name
Middle Name
Age
Alternate Names
Low Quality
Illegible
CANCEL
0 of 0 Images Remaining
No Image

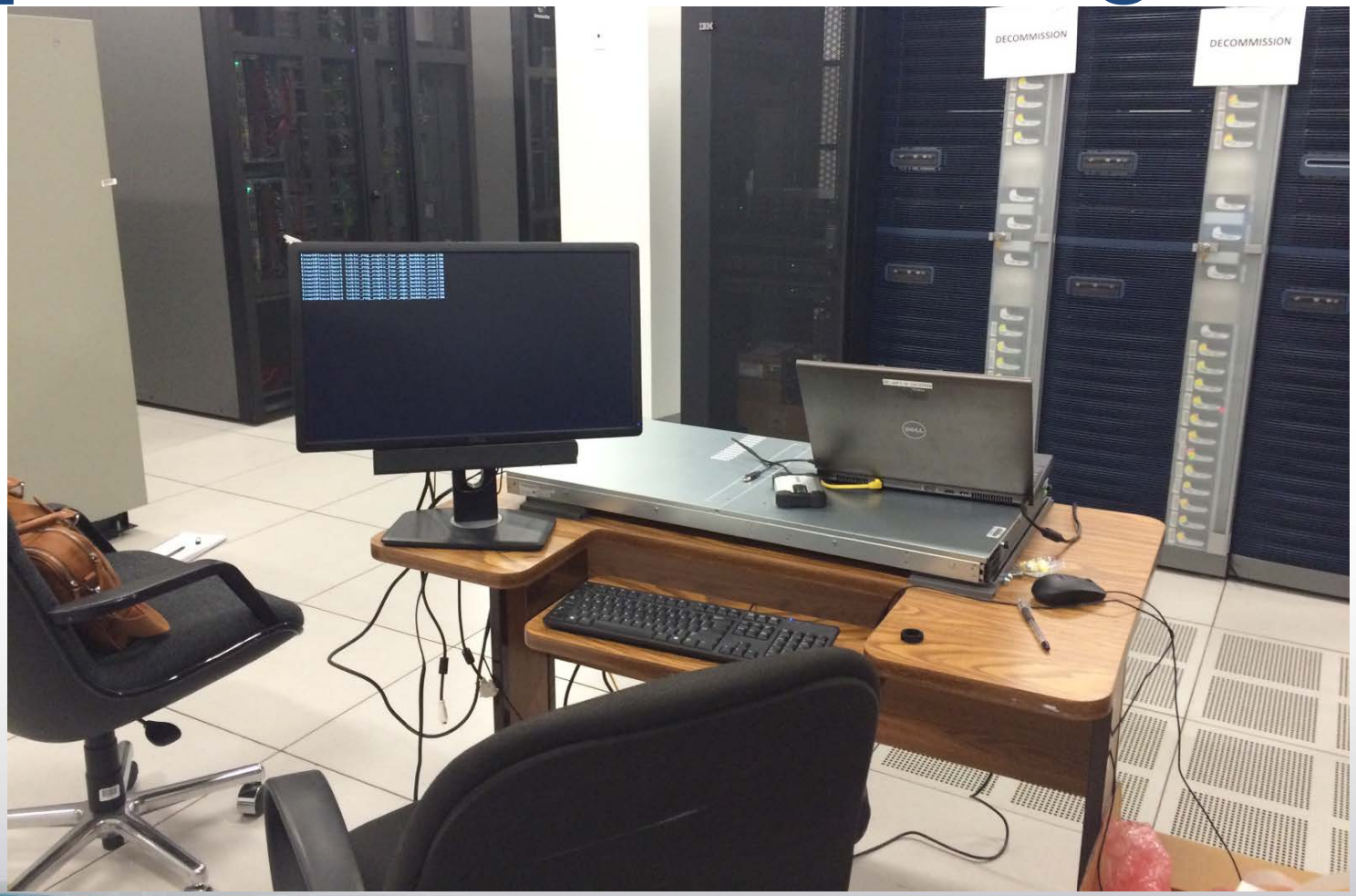
Scope of Work

- Scan Microfilm
- Make hand-keyed “truth data”
- **Do Optical Character Recognition (OCR)**
- Evaluate Results

Optical Character Recognition

- Currently happening at Census HQ
  - One academic institution and one company
  - Both providing own servers
    - Quarantined machines
    - Hard-drives will be destroyed

Optical Character Recognition



Optical Character Recognition

- Census created “dictionaries” of every name
  - Included every first and last name ever associated with a Social Security number
  - Initial dictionaries were too large
  - Provided dictionaries including names of 95% and 99% of the population (reduced size by >half)

Optical Character Recognition

- OCR began in March 2016
  - Both teams provided preliminary results in April
  - Will be complete in May 2016
  - National Archives is also doing limited OCR

Scope of Work

- Scan Microfilm
- Make hand-keyed “truth data”
- Do Optical Character Recognition (OCR)
- **Evaluate Results**

Preliminary Results: Scanning and Truth Data

- Scanning: identified optimal scanner and settings
- “truth data”: have measures of keying error rate and hard-to-read names
- OCR: will compare output of 2 teams and National Archives

Preliminary Results: Optical Character Recognition

- OCR quality when compared to truth data
  - Perfect Matches
    - Household ID 85%
    - First name 71%
    - Last name 67%

Preliminary Results: Optical Character Recognition

- OCR quality when compared to truth data
  - "Good" Matches (Jaro-Winkler distance >0.83)
    - First name 82%
      - Mary and Gary
      - Cora and Lora
      - Morgan and Megan
    - Last name 78%
      - Conners and Coppers
      - Leke and Lake
      - Boyd and Byrd

Preliminary Results: Unexpected Discovery

- Located file that will greatly simplify linkage
  - Addresses for every record in 1990
    - Street addresses, rural routes, apartment numbers, etc.
  - Includes householder name for 30% of units
    - 27 million names
    - Were collected in a pre-census address canvassing operation at a cost of $68 million in 1989 dollars

Preliminary Results: Unexpected Discovery

- Why is the address file a big deal?
  - Allows for address linkage to a same-year administrative records composite that has names
  - Composite contains demographic characteristics to validate address matches

- May enable good links without names

Next steps

- Assign linkage keys to 40,000 names
  - Will be conducted by CARRA's linkage staff
  - Can compare links made with OCR'ed names to links made with hand-keyed names
- Will answer question:
  - Is current OCR good enough for linkage?

Next steps

- Create linked file for entire 1990 Census
  - Will use address/name file that was discovered
  - Matched to IRS/SSA data from same period

- Seeking funding for similar pilots on Censuses of 1960, 1970, and 1980
  - No address file for those
  - OCR is the only option

Thank you

Diane Cronkite

diane.m.cronkite@census.gov