IPUMS

User Extract usa_00004.dat

Jump to Section

§ 1. Document Description

Citation

Title Statement	
Title:	Codebook for an IPUMS USA Data Extract
Subtitle:	DDI 2.5 metadata describing the extract file 'usa_00004.dat'
Identification Number:	ddi2-0f98ff10-e197-013c-b091-0242ac190004-usa_00004.dat-usa.ipums.org
Responsibility Statement	
Authoring Entity:	IPUMS
Affiliation:	University of Minnesota
Production Statement	
Producer:	IPUMS
Affiliation:	University of Minnesota
Role:	Documentation
Date of Production:	June 7, 2024
Place of Production:	IPUMS, 50 Willey Hall, 225 - 19th Avenue South, Minneapolis, MN 55455
Distribution Statement	
Contact Persons:	IPUMS
Affiliation:	University of Minnesota
URI:	https://ipums.org

§ 2. Study Description

Citation

Title Statement	
Title:	User Extract usa_00004.dat

Responsibility Statement	
Authoring Entity:	IPUMS
Affiliation:	University of Minnesota
Production Statement	
Producer:	IPUMS
Affiliation:	University of Minnesota
Role:	Documentation
Date of Production:	June 7, 2024
Place of Production:	IPUMS, 50 Willey Hall, 225 - 19th Avenue South, Minneapolis, MN 55455
Distribution Statement	
Contact Persons:	IPUMS
Affiliation:	University of Minnesota
URI:	https://ipums.org
Version Statement	
Date:	2024-06-07

Study Scope

Subject Information	
Topic Classification:	Technical Variables -- HOUSEHOLD
	Geographic Variables -- HOUSEHOLD
	Technical Variables -- PERSON
	Demographic Variables -- PERSON
	Education Variables -- PERSON
	Income Variables -- PERSON
Summary Data Description	
Time Period:	2016
Country:	United States
Notes	
Note:	Additional notes on a sample that is part of this study: 2012-2016, ACS 5-year Density of the full data file: 5.0% Density of this extract: 5.0%

Data Access - Use Statement

Confidentiality Declaration	

None	
Contact Persons:	IPUMS USA
Affiliation:	IPUMS
URI:	http://usa.ipums.org

Citation Requirement

Publications and research reports based on the IPUMS USA database must cite it appropriately. The citation should include the following:

Steven Ruggles, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler. IPUMS USA: Version 15.0 [dataset]. Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D010.V15.0

The licensing agreement for use of IPUMS USA data requires that users supply us with the title and full citation for any publications, research reports, or educational materials making use of the data or documentation. Please add your citation to the IPUMS bibliography at http://bibliography.ipums.org/.

Conditions

Users of IPUMS USA data must agree to abide by the conditions of use. A user's license is valid for one year and may be renewed. Users must agree to the following conditions:

(1) No fees may be charged for use or distribution of the data.

(2) Cite IPUMS appropriately. For information on proper citation, refer to the citation requirement section of this DDI document.

(3) Tell us about any work you do using the IPUMS. Publications, research reports, or presentations making use of IPUMS USA should be added to our Bibliography. Continued funding for the IPUMS depends on our ability to show our sponsor agencies that researchers are using the data for productive purposes.

(4) The IPUMS cannot be used for genealogical research

(5) It is difficult to use the IPUMS to study small geographic areas. In the IPUMS census samples for years 1940-present, no places having a population of fewer than 100,000 persons can be identified.

(6) Use it for GOOD -- never for EVIL.

(7) Please notify ipums@umn.edu regarding errors in the data or documentation.

Disclaimer

The user of the data acknowledges that the original collector of the data, the authorized distributor of the data, and the relevant funding agency bear no responsibility for use of the data or for interpretations or inferences based upon such uses.

Study Notes

Notes	
Note:	ACS_2012_2016
	This extract is a revision of the user's previous extract, ID 21648106.

§ 3. File Description

File

File Name:	usa_00004.dat
Contents of Files:	Microdata records
Type:	rectangular
File Type:	ISO-8859-1 data file

Data Format:	fixed length fields
Place of File Production:	IPUMS, 50 Willey Hall, 225 - 19th Avenue South, Minneapolis, MN 55455

§ 4. Variable Description

Jump to Variable

1. YEAR (Census year)
2. MULTYEAR (Actual year of survey, multi-year ACS/PRCS)
3. SAMPLE (IPUMS sample identifier)
4. US2016C_ST (State Code)
5. US2016C_ADJINC (Adjustment factor for income and earnings dollar amounts (6 implied decimal places))
6. PERNUM (Person number in sample unit)
7. PERWT (Person weight)
8. US2016C_PWGTP (Person's weight)
9. US2016C_AGEP (Age)
10. US2016C_SCHL (Educational attainment)
11. US2016C_SEX (Sex)
12. US2016C_PINCP (Total person's income (signed))

Variable: "YEAR"

Name:	YEAR
Label:	Census year
Variable Text:	YEAR reports the four-digit year when the household was enumerated or included in the census, the ACS, and the PRCS. For the multi-year ACS/PRCS samples, YEAR indicates the last year of data included (e.g., 2007 for the 2005-2007 3-year ACS/PRCS; 2008 for the 2006-2008 3-year ACS/PRCS; and so on). For the actual year of survey in these multi-year data, see MULTYEAR.
Concept:	Technical Variables -- HOUSEHOLD
Start Position:	1
End Position:	4
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1850	1850
1860	1860
1870	1870
1880	1880

1900	1900
1910	1910
1920	1920
1930	1930
1940	1940
1950	1950
1960	1960
1970	1970
1980	1980
1990	1990
2000	2000
2001	2001
2002	2002
2003	2003
2004	2004
2005	2005
2006	2006
2007	2007
2008	2008
2009	2009
2010	2010
2011	2011
2012	2012
2013	2013
2014	2014
2015	2015
2016	2016
2017	2017
2018	2018
2019	2019
2020	2020
2021	2021

2022	2022

Variable: "MULTYEAR"

Name:	MULTYEAR
Label:	Actual year of survey, multi-year ACS/PRCS
Variable Text:	MULTYEAR identifies the actual year of survey in multi-year ACS/PRCS samples. For example, the 3-year ACS and PRCS data files each include cases from three single-year files. For these multi-year samples, the YEAR variable identifies the last year of data (2007 for the 2005-2007 3-year data; 2008 for the 2006-2008 data; and so on). MULTYEAR gives the single-year sample from which the case was drawn (2005, 2006, or 2007 for the 2005-2007 3-year data; 2006, 2007, or 2008 for the 2006-2008 3-year data; and so on).
Concept:	Technical Variables -- HOUSEHOLD
Start Position:	5
End Position:	8
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesThis variable is valid only for multi-year ACS and PRCS samples. MULTYEAR is a 4-digit numeric variable which reports the actual year of survey in multi-year ACS/PRCS samples. MULTYEAR specific variable codes for missing, edited, or unidentified observations, observations not applicable (N/A), observations not in universe (NIU), top and bottom value coding, etc. are provided below if applicable by Census year (and data sample if specified).

Variable: "SAMPLE"

Name:	SAMPLE
Label:	IPUMS sample identifier
Variable Text:	SAMPLE identifies the IPUMS sample from which the case is drawn. Each sample receives a unique 6-digit code. The codes are structured as follows: The first four digits are the year of the census/survey. The next two digits identify the sample within the year. For most censuses, IPUMS has multiple datasets which were constructed using different sampling techniques (i.e. size/demographic of the sample population, geographic coverage level or location, or duration of the sampling period for the ACS/PRCS samples). The availability table for each variable indicates whether that variable is available in only certain samples for a given year. For further discussion of sample differences, see "Sample Designs.". Note: SAMPLE replaces DATANUM. Though the last two digits in SAMPLE do not correlate exactly with the now-deprecated DATANUM, the variable serves the same purpose of assigning a unique id to all cases that belong to the same dataset.
Concept:	Technical Variables -- HOUSEHOLD
Start Position:	9
End Position:	14
Width:	6

Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
202204	2018-2022, PRCS 5-year
202203	2018-2022, ACS 5-year
202202	2022 PRCS
202201	2022 ACS
202104	2017-2021, PRCS 5-year
202103	2017-2021, ACS 5-year
202102	2021 PRCS
202101	2021 ACS
202004	2016-2020, PRCS 5-year
202003	2016-2020, ACS 5-year
202001	2020 ACS
201904	2015-2019, PRCS 5-year
201903	2015-2019, ACS 5-year
201902	2019 PRCS
201901	2019 ACS
201804	2014-2018, PRCS 5-year
201803	2014-2018, ACS 5-year
201802	2018 PRCS
201801	2018 ACS
201704	2013-2017, PRCS 5-year
201703	2013-2017, ACS 5-year
201702	2017 PRCS
201701	2017 ACS
201604	2012-2016, PRCS 5-year
201603	2012-2016, ACS 5-year
201602	2016 PRCS

201601	2016 ACS
201504	2011-2015, PRCS 5-year
201503	2011-2015, ACS 5-year
201502	2015 PRCS
201501	2015 ACS
201404	2010-2014, PRCS 5-year
201403	2010-2014, ACS 5-year
201402	2014 PRCS
201401	2014 ACS
201306	2009-2013, PRCS 5-year
201305	2009-2013, ACS 5-year
201304	2011-2013, PRCS 3-year
201303	2011-2013, ACS 3-year
201302	2013 PRCS
201301	2013 ACS
201206	2008-2012, PRCS 5-year
201205	2008-2012, ACS 5-year
201204	2010-2012, PRCS 3-year
201203	2010-2012, ACS 3-year
201202	2012 PRCS
201201	2012 ACS
201106	2007-2011, PRCS 5-year
201105	2007-2011, ACS 5-year
201104	2009-2011, PRCS 3-year
201103	2009-2011, ACS 3-year
201102	2011 PRCS
201101	2011 ACS
201008	2010 Puerto Rico 10%
201007	2010 10%
201006	2006-2010, PRCS 5-year
201005	2006-2010, ACS 5-year
201004	2008-2010, PRCS 3-year

201003	2008-2010, ACS 3-year
201002	2010 PRCS
201001	2010 ACS
200906	2005-2009, PRCS 5-year
200905	2005-2009, ACS 5-year
200904	2007-2009, PRCS 3-year
200903	2007-2009, ACS 3-year
200902	2009 PRCS
200901	2009 ACS
200804	2006-2008, PRCS 3-year
200803	2006-2008, ACS 3-year
200802	2008 PRCS
200801	2008 ACS
200704	2005-2007, PRCS 3-year
200703	2005-2007, ACS 3-year
200702	2007 PRCS
200701	2007 ACS
200602	2006 PRCS
200601	2006 ACS
200502	2005 PRCS
200501	2005 ACS
200401	2004 ACS
200301	2003 ACS
200201	2002 ACS
200101	2001 ACS
200008	2000 Puerto Rico 1%
200007	2000 1%
200006	2000 Puerto Rico 1% sample (old version)
200005	2000 Puerto Rico 5%
200004	2000 ACS
200003	2000 Unweighted 1%
200002	2000 1% sample (old version)

200001	2000 5%
199007	1990 Puerto Rico 1%
199006	1990 Puerto Rico 5%
199005	1990 Labor Market Area
199004	1990 Elderly
199003	1990 Unweighted 1%
199002	1990 1%
199001	1990 5%
198007	1980 Puerto Rico 1%
198006	1980 Puerto Rico 5%
198005	1980 Detailed metro/non-metro
198004	1980 Labor Market Area
198003	1980 Urban/Rural
198002	1980 1%
198001	1980 5%
197009	1970 Puerto Rico Neighborhood
197008	1970 Puerto Rico Municipio
197007	1970 Puerto Rico State
197006	1970 Form 2 Neighborhood
197005	1970 Form 1 Neighborhood
197004	1970 Form 2 Metro
197003	1970 Form 1 Metro
197002	1970 Form 2 State
197001	1970 Form 1 State
196002	1960 5%
196001	1960 1%
195002	1950 100% database
195001	1950 1%
194002	1940 100% database
194001	1940 1%
193004	1930 100% database
193003	1930 Puerto Rico

193002	1930 5%
193001	1930 1%
192003	1920 100% database
192002	1920 Puerto Rico sample
192001	1920 1%
191004	1910 100% database
191003	1910 1.4% sample with oversamples
191002	1910 1%
191001	1910 Puerto Rico
190004	1900 100% database
190003	1900 1% sample with oversamples
190002	1900 1%
190001	1900 5%
188003	1880 100% database
188002	1880 10%
188001	1880 1%
187003	1870 100% database
187002	1870 1% sample with black oversample
187001	1870 1%
186003	1860 100% database
186002	1860 1% sample with black oversample
186001	1860 1%
185002	1850 100% database
185001	1850 1%

Variable: "US2016C_ST"

Name:	US2016C_ST
Label:	State Code
Variable Text:	
Universe:	2012-2016, ACS 5-year: All households and group quarter.
Concept:	Geographic Variables -- HOUSEHOLD
Start Position:	15

End Position:	16
Width:	2
Variable Format:	character
Implied Decimal Places:	0

Categories

Value	Label
1	Alabama/AL
2	Alaska/AK
4	Arizona/AZ
5	Arkansas/AR
6	California/CA
8	Colorado/CO
9	Connecticut/CT
10	Delaware/DE
11	District of Columbia/DC
12	Florida/FL
13	Georgia/GA
15	Hawaii/HI
16	Idaho/ID
17	Illinois/IL
18	Indiana/IN
19	Iowa/IA
20	Kansas/KS
21	Kentucky/KY
22	Louisiana/LA
23	Maine/ME
24	Maryland/MD
25	Massachusetts/MA
26	Michigan/MI
27	Minnesota/MN
28	Mississippi/MS

29	Missouri/MO	
30	Montana/MT	
31	Nebraska/NE	
32	Nevada/NV	
33	New Hampshire/NH	
34	New Jersey/NJ	
35	New Mexico/NM	
36	New York/NY	
37	North Carolina/NC	
38	North Dakota/ND	
39	Ohio/OH	
40	Oklahoma/OK	
41	Oregon/OR	
42	Pennsylvania/PA	
44	Rhode Island/RI	
45	South Carolina/SC	
46	South Dakota/SD	
47	Tennessee/TN	
48	Texas/TX	
49	Utah/UT	
50	Vermont/VT	
51	Virginia/VA	
53	Washington/WA	
54	West Virginia/WV	
55	Wisconsin/WI	
56	Wyoming/WY	

Variable: "US2016C_ADJINC"

Name:	US2016C_ADJINC
Label:	Adjustment factor for income and earnings dollar amounts (6 implied decimal places)
Variable Text:	
Universe:	2012-2016, ACS 5-year:

Concept:	Technical Variables -- HOUSEHOLD
Start Position:	17
End Position:	23
Width:	7
Variable Format:	character
Implied Decimal Places:	0

Categories

Value	Label
1007588	2016 factor (1.007588 * 1.00000000)
1013916	2015 factor (1.001264 * 1.01263642)
1022342	2014 factor (1.008425 * 1.01380104)
1038170	2013 factor (1.007549 * 1.03039158)
1056030	2012 factor (1.010207 * 1.04536021)

Variable: "PERNUM"

Name:	PERNUM
Label:	Person number in sample unit
Variable Text:	PERNUM numbers all persons within each household consecutively in the order in which they appear on the original census or survey form. When combined with SAMPLE and SERIAL, PERNUM uniquely identifies each person within the IPUMS.
Concept:	Technical Variables -- PERSON
Start Position:	24
End Position:	27
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesPERNUM is a 4-digit numeric variable which numbers all persons within each household consecutively in the order in which they appear on the original census or survey form. PERNUM specific variable codes for missing, edited, or unidentified observations, observations not applicable (N/A), observations not in universe (NIU), top and bottom value coding, etc. are provided below if applicable by Census year (and data sample if specified).

Variable: "PERWT"

Name:	PERWT

Label:	Person weight
Variable Text:	PERWT indicates how many persons in the U.S. population are represented by a given person in an IPUMS sample. It is generally a good idea to use PERWT when conducting a person-level analysis of any IPUMS sample. The use of PERWT is optional when analyzing one of the "flat" or unweighted IPUMS samples. Flat IPUMS samples include the 1% samples from 1850-1930, all samples from 1960, 1970, and 1980, the 1% unweighted samples from 1990 and 2000, the 10% 2010 sample, and any of the full count 100% census datasets. PERWT must be used to obtain nationally representative statistics for person-level analyses of any sample other than those. For further explanation of the sample weights, see "Sample Designs" and "Sample Weights". See also HHWT for a corresponding variable at the household level, and SLWT for a weight variable used with sample-line records in 1940 and 1950.
Concept:	Technical Variables -- PERSON
Start Position:	28
End Position:	37
Width:	10
Variable Format:	numeric
Implied Decimal Places:	2
Coder Instructions:	CodesPERWT is a 6-digit numeric variable which indicates how many persons in the U.S. population are represented by a given person in an IPUMS sample and has two implied decimals. For example, a PERWT value of 010461 should be interpreted as 104.61. PERWT specific variable codes for missing, edited, or unidentified observations, observations not applicable (N/A), observations not in universe (NIU), top and bottom value coding, etc. are provided below if applicable by Census year (and data sample if specified). PERWT Specific Variable Codes

Variable: "US2016C_PWGTP"

Name:	US2016C_PWGTP
Label:	Person's weight
Variable Text:	
Universe:	2012-2016, ACS 5-year: All persons.
Concept:	Technical Variables -- PERSON
Start Position:	38
End Position:	42
Width:	5
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	This is a 5-digit numeric variable with 0 implied decimal places

Variable: "US2016C_AGEP"

Name:	US2016C_AGEP

Label:	Age
Variable Text:	
Universe:	2012-2016, ACS 5-year: All persons.
Concept:	Demographic Variables -- PERSON
Start Position:	43
End Position:	44
Width:	2
Variable Format:	character
Implied Decimal Places:	0

Categories

Value	Label
0	Under 1 year
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19

20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51

52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83

84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97

Variable: "US2016C_SCHL"

Name:	US2016C_SCHL
Label:	Educational attainment
Variable Text:	
Universe:	2012-2016, ACS 5-year: All persons.
Concept:	Education Variables -- PERSON
Start Position:	45
End Position:	46
Width:	2
Variable Format:	character
Implied Decimal Places:	0

Categories

Value	Label
1	No schooling completed
2	Nursery school, preschool
3	Kindergarten

4	Grade 1
5	Grade 2
6	Grade 3
7	Grade 4
8	Grade 5
9	Grade 6
10	Grade 7
11	Grade 8
12	Grade 9
13	Grade 10
14	Grade 11
15	12th grade - no diploma
16	Regular high school diploma
17	GED or alternative credential
18	Some college, but less than 1 year
19	1 or more years of college credit, no degree
20	Associate's degree
21	Bachelor's degree
22	Master's degree
23	Professional degree beyond a bachelor's degree
24	Doctorate degree
BB	N/A (less than 3 years old)

Variable: "US2016C_SEX"

Name:	US2016C_SEX
Label:	Sex
Variable Text:	
Universe:	2012-2016, ACS 5-year: All persons.
Concept:	Demographic Variables -- PERSON
Start Position:	47
End Position:	47
Width:	1

Variable Format:	character
Implied Decimal Places:	0
Categories	

Value	Label
1	Male
2	Female

Variable: "US2016C_PINCP"

Name:	US2016C_PINCP
Label:	Total person's income (signed)
Variable Text:	
Universe:	2012-2016, ACS 5-year: Persons age 15+.
Concept:	Income Variables -- PERSON
Start Position:	48
End Position:	56
Width:	9
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	This is a 9-digit numeric variable with 0 implied decimal places