IPUMS

User Extract nhis_00010.dat

Jump to Section

§ 1. Document Description

Citation

Title Statement	
Title:	Codebook for an Integrated Health Interview Series Data Extract
Subtitle:	DDI 2.5 metadata describing the extract file 'nhis_00010.dat'
Identification Number:	ddi2-e47d7570-e818-0134-9fb8-005056a35405-nhis_00010.dat-nhis.ipums.org
Responsibility Statement	
Authoring Entity:	IPUMS
Affiliation:	University of Minnesota
Production Statement	
Producer:	IPUMS
Affiliation:	University of Minnesota
Role:	Documentation
Date of Production:	August 22, 2022
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 nhis_00010.dat
Responsibility Statement	
Authoring Entity:	IPUMS
Affiliation:	University of Minnesota
Production Statement	
Producer:	IPUMS
Affiliation:	University of Minnesota
Role:	Documentation
Date of Production:	August 22, 2022
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:	2022-08-22

Study Scope

Subject Information	
Topic Classification:	Technical Household Variables -- HOUSEHOLD

	Geography Variables -- HOUSEHOLD
	Technical Person Variables -- PERSON
	Core Demographic Variables -- PERSON
	Ethnicity/Nativity Variables -- PERSON
	Education Variables -- PERSON
	Work Variables -- PERSON
	Total Income and Earnings Variables -- PERSON
	Limitation Variables -- PERSON
	Mortality Variables -- PERSON

Summary Data Description	
Time Period:	1986
Country:	United States

Summary Data Description	
Time Period:	1987
Country:	United States

Summary Data Description	
Time Period:	1988
Country:	United States

Summary Data Description	
Time Period:	1989
Country:	United States

Summary Data Description	
Time Period:	1990
Country:	United States

Summary Data Description

Time Period:	1991
Country:	United States

Summary Data Description

Time Period:	1992
Country:	United States

Summary Data Description

Time Period:	1993
Country:	United States

Summary Data Description

Time Period:	1994
Country:	United States

Summary Data Description

Time Period:	1995
Country:	United States

Summary Data Description

Time Period:	1996
Country:	United States

Summary Data Description

Time Period:	1997
Country:	United States

Summary Data Description

Time Period:	1998
Country:	United States

Summary Data Description

Time Period:	1999
Country:	United States

Summary Data Description

Time Period:	2000
Country:	United States

Summary Data Description

Time Period:	2001
Country:	United States

Summary Data Description

Time Period:	2002
Country:	United States

Summary Data Description

Time Period:	2003
Country:	United States

Summary Data Description

Time Period:	2004
Country:	United States

Summary Data Description

Time Period:	2005
Country:	United States

Summary Data Description

Time Period:	2006
Country:	United States

Summary Data Description

Time Period:	2007

Country:	United States

Summary Data Description

Time Period:	2008
Country:	United States

Summary Data Description

Time Period:	2009
Country:	United States

Summary Data Description

Time Period:	2010
Country:	United States

Summary Data Description

Time Period:	2011
Country:	United States

Summary Data Description

Time Period:	2012
Country:	United States

Summary Data Description

Time Period:	2013
Country:	United States

Summary Data Description

Time Period:	2014
Country:	United States

Summary Data Description

Time Period:	2015

| Country: | United States |

Summary Data Description

| Time Period: | 2016 |
| Country: | United States |

Summary Data Description

| Time Period: | 2017 |
| Country: | United States |

Summary Data Description

| Time Period: | 2018 |
| Country: | United States |

Notes

Note:	Additional notes on a sample that is part of this study: 1986 NHIS
	Additional notes on a sample that is part of this study: 1987 NHIS
	Additional notes on a sample that is part of this study: 1988 NHIS
	Additional notes on a sample that is part of this study: 1989 NHIS
	Additional notes on a sample that is part of this study: 1990 NHIS
	Additional notes on a sample that is part of this study: 1991 NHIS
	Additional notes on a sample that is part of this study: 1992 NHIS
	Additional notes on a sample that is part of this study: 1993 NHIS
	Additional notes on a sample that is part of this study: 1994 NHIS
	Additional notes on a sample that is part of this study: 1995 NHIS
	Additional notes on a sample that is part of this study: 1996 NHIS
	Additional notes on a sample that is part of this study: 1997 NHIS
	Additional notes on a sample that is part of this study: 1998 NHIS
	Additional notes on a sample that is part of this study: 1999 NHIS

	Additional notes on a sample that is part of this study: 2000 NHIS
	Additional notes on a sample that is part of this study: 2001 NHIS
	Additional notes on a sample that is part of this study: 2002 NHIS
	Additional notes on a sample that is part of this study: 2003 NHIS
	Additional notes on a sample that is part of this study: 2004 NHIS
	Additional notes on a sample that is part of this study: 2005 NHIS
	Additional notes on a sample that is part of this study: 2006 NHIS
	Additional notes on a sample that is part of this study: 2007 NHIS
	Additional notes on a sample that is part of this study: 2008 NHIS
	Additional notes on a sample that is part of this study: 2009 NHIS
	Additional notes on a sample that is part of this study: 2010 NHIS
	Additional notes on a sample that is part of this study: 2011 NHIS
	Additional notes on a sample that is part of this study: 2012 NHIS
	Additional notes on a sample that is part of this study: 2013 NHIS
	Additional notes on a sample that is part of this study: 2014 NHIS
	Additional notes on a sample that is part of this study: 2015 NHIS
	Additional notes on a sample that is part of this study: 2016 NHIS
	Additional notes on a sample that is part of this study: 2017 NHIS
	Additional notes on a sample that is part of this study: 2018 NHIS

Data Access - Use Statement

Confidentiality Declaration
The Public Health Service Act (Section 308 (d)) provides that the data collected by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC), may be used only for the purpose of health statistical reporting and analysis. Any effort to determine the identity of any reported case is prohibited by this law. NCHS does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted from the data files. Any intentional identification or disclosure of a person or establishment violates the assurances of confidentiality given to the providers of the information.

Therefore, users will:

Use the data in these data files for statistical reporting and analysis only.

Make no use of the identity of any person or establishment discovered inadvertently and advise the Director, NCHS, of any such discovery (301-458-4500).

Not link these data files with individually identifiable data from other NCHS or non-NCHS data files.

By using these data, you signify your agreement to comply with the above-stated statutorily-based requirements.

Contact Persons:	Integrated Health Interview Series
Affiliation:	IPUMS
URI:	http://www.nhis.ipums.org/

Citation Requirement

Publications and research reports based on the NHIS database must cite it appropriately. The citation is as follows:

Lynn A. Blewett, Julia A. Rivera Drew, Miriam L. King, Kari C.W. Williams, Natalie Del Ponte and Pat Convey. IPUMS Health Surveys: National Health Interview Survey, Version 7.1 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D070.V7.1

If possible, citations should also include the URL for the NHIS site: http://www.nhis.ipums.org.

Please see http://www.nhis.ipums.org/nhis/citation.shtml for precise formatting of the citation.

Conditions

The Public Health Service Act (Section 308 (d)) provides that the data collected by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC), may be used only for the purpose of health statistical reporting and analysis. Any effort to determine the identity of any reported case is prohibited by this law. NCHS does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted from the data files. Any intentional identification or disclosure of a person or establishment violates the assurances of confidentiality given to the providers of the information.

Therefore, users must:

Use the data in these data files for statistical reporting and analysis only.

Make no use of the identity of any person or establishment discovered inadvertently and advise the Director, NCHS, of any such discovery (301-458-4500).

Not link these data files with individually identifiable data from other NCHS or non-NCHS data files.

By using these data, you signify your agreement to comply with the above-stated statutorily-based requirements.

Furthermore, users of NHIS data must agree to abide by the conditions of use. Users must agree to the following conditions:

(1) Use the data in these data files for statistical reporting and analysis only

(2) Make no use of the identity of any person or establishment discovered inadvertently and advise the

Director of NCHS of any such discovery (301-458-4500)

(3) Do not link these data with individually-identifiable data from NCHS or non-NCHS data files

(4) No fees may be charged for use or distribution of the data. All persons are granted a limited license to use and distribute these data, but you may not charge a fee for the data if you distribute them to others.

(5) Cite the NHIS appropriately. Publications and research reports based on the database must cite it appropriately. Please see http://www.nhis.ipums.org/nhis/citation.shtml

(6) NHIS cannot be used to study small geographic areas. The smallest geographical areas identified in the NHIS are regions (groups of states) and a limited number of metropolitan areas.

(7) This system provides individual-level data only. The NHIS Data Extraction System will not produce tables. You will need to use a statistical software package, such as Stata, SAS, or SPSS, to analyze the downloaded data. Alternatively, you may use the NHIS-SDA tabulator to produce tables online, without making a data extract.

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:	User-provided description: NHIS class-mortality analyses with all variables, including 1986-2018 NHIS-NDI.
	This extract is a revision of the user's previous extract, ID 28175143.

§ 3. File Description

File

File Name:	nhis_00010.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 (Survey year)
2. SERIAL (Sequential Serial Number, Household Record)
3. STRATA (Stratum for variance estimation)
4. PSU (Primary sampling unit (PSU) for variance estimation)
5. NHISHID (NHIS Unique identifier, household)
6. HHWEIGHT (Household weight, final annual)
7. REGION (Region of residence)
8. PERNUM (Person number within family/household (from reformatting))
9. NHISPID (NHIS Unique Identifier, person)
10. HHX (Household number (from NHIS))
11. FMX (Family number (from NHIS))
12. PX (Person number of respondent (from NHIS).)
13. PERWEIGHT (Final basic annual weight)
14. SAMPWEIGHT (Sample Person Weight)
15. FWEIGHT (Final annual family weight)
16. INTERVWMO (Month of NHIS interview)
17. INTERVWYR (Year of NHIS interview)
18. ASTATFLG (Sample adult flag)
19. CSTATFLG (Sample child flag)
20. AGE (Age)
21. SEX (Sex)
22. MARSTAT (Legal marital status)
23. HISPETH (Hispanic ethnicity)
24. RACESR (Self-Reported Main Racial Background (Pre-1997 Revised OMB Standards))
25. EDUCREC1 (Educational attainment recode, nonintervalled)
26. EMPSTAT (Employment status in past 1 to 2 weeks)
27. OCC (Detailed occupation)
28. OCC1995 (Occupation, 1995 basis)
29. IND (Detailed industry recode, 1969-2003)
30. IND1995 (Industry, 1995 basis)
31. NUMEMPS (Number of employees at work)
32. CLASSWK2 (Class of worker, current or longest job)
33. JOBSECINCORP (Current self-employment job is incorporated business)
34. WHYNOWK2 (Corrected main reason not working last week (sample adults))
35. CLASSWK (Class of worker)
36. INCFAM97ON2 (Total combined family income (1997+ w. 2007 categories))
37. INCIMP1 (Imputed total combined family income (1997+ grouping))
38. INCIMP2 (Imputed total combined family income (1997+ grouping))
39. INCIMP3 (Imputed total combined family income (1997+ grouping))
40. INCIMP4 (Imputed total combined family income (1997+ grouping))
41. INCIMP5 (Imputed total combined family income (1997+ grouping))
42. IMPYFAMFLAG1 (Family income group imputation flag 1)
43. IMPYFAMFLAG2 (Family income group imputation flag 2)
44. IMPYFAMFLAG3 (Family income group imputation flag 3)
45. IMPYFAMFLAG4 (Family income group imputation flag 4)

46. IMPYFAMFLAG5 (Family income group imputation flag 5)

47. INCFAM8296 (Total combined family income (1982-96 grouping))

48. LAMTWRK (Limited in kind/amount of work (health))

49. MORTELIG (Eligibility status for mortality follow-up)

50. MORTSTAT (Final mortality status)

51. MORTDODQ (Quarter of death)

52. MORTDODY (Year of death)

53. MORTUCOD (Underlying cause of death (ICD-10))

54. MORTUCODLD (Leading underlying cause of death (ICD-10))

55. MORTWT (Weight adjusted for ineligible respondents in mortality analysis)

56. MORTDIAB (Diabetes flag from multiple cause of death (MCOD))

57. MORTHIPFX (Hip fracture flag from multiple cause of death (MCOD))

58. MORTHYPR (Hypertension flag from multiple cause of death (MCOD))

59. MORTCMS (Mortality status obtained from Centers for Medicaid and Medicare Services)

60. MORTNDI (Mortality match with National Death Index)

61. MORTSSA (Mortality status obtained from Social Security Administration)

62. MORTWTSA (Sample adult weight adjusted for ineligible respondents in mortality analysis)

Variable: "YEAR"

Name:	YEAR
Label:	Survey year
Variable Text:	YEAR is a four-digit variable reporting the calendar year (e.g., 2003) the survey was conducted and the data were collected. YEAR indicates the survey year reported on the household record.
Concept:	Technical Household Variables -- HOUSEHOLD
Start Position:	1
End Position:	4
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	This is a 4-digit numeric variable with 0 implied decimal places

Variable: "SERIAL"

Name:	SERIAL
Label:	Sequential Serial Number, Household Record
Variable Text:	SERIAL is an IPUMS NHIS-constructed value that is an identifying number unique to each household in a given survey year. The combination of YEAR and SERIAL provides a unique identifier for every household in the IPUMS NHIS database.
Concept:	Technical Household Variables -- HOUSEHOLD
Start Position:	5
End Position:	10
Width:	6
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesSERIAL is a 6-digit numeric variable.

Variable: "STRATA"

Name:	STRATA
Label:	Stratum for variance estimation
Variable Text:	STRATA is an IPUMS NHIS-constructed variable based on the NHIS sample design variables in the public use files concatenated with a sample design period indicator. STRATA represents the impact of the sample design stratification on the estimates of variance and standard errors. It is constant within a sample design period and changes between sample design periods. For analysis, researchers need to use STRATA in conjunction with PSU to account for stratification and clustering when computing variance estimates with IPUMS NHIS data. See the User Notes on variance estimation for additional information. After the 2019 redesign, there is no longer a sampling weight to produce household population estimates. See the user note on the calculation and use of sampling weights for additional information.
Concept:	Technical Household Variables -- HOUSEHOLD
Start Position:	11
End	14

Position:	
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesSTRATA is a 5-digit numeric variable. 000: Not in Universe

Variable: "PSU"

Name:	PSU
Label:	Primary sampling unit (PSU) for variance estimation
Variable Text:	PSU is the primary sampling unit variable that represents the impact of the sample design clustering on the estimates of variance and standard errors. It is constant within a sample design period and changes between sample design periods. For analysis, researchers need to use PSU in conjunction with STRATA to account for stratification and clustering when computing variance estimates with IPUMS NHIS data. See the User Notes on variance estimation for additional information. After the 2019 redesign, there is no longer a sampling weight to produce household population estimates. See the user note on the calculation and use of sampling weights for additional information.
Concept:	Technical Household Variables -- HOUSEHOLD
Start Position:	15
End Position:	17
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesPSU is a 3-digit numeric variable. 000: Not in Universe

Variable: "NHISHID"

Name:	NHISHID
Label:	NHIS Unique identifier, household
Variable Text:	NHISHID is an IPUMS NHIS-constructed value that is an identifying number unique to each household in a given survey year. Using NHISHID, analysts can link IPUMS NHIS data with additional household-level data elements, from the NHIS source public use data files, that are not currently available in IPUMS. NHISHID is a single variable constructed by concatenating multiple data elements from the NHIS source data files. Analysts wishing to link person-level NHIS data should use the person-level linking key NHISPID. See the user note on LINKING for instructions on how to create links between IPUMS NHIS and NHIS source data. After the 2019 redesign, there is no longer a sampling weight to produce household population estimates. See the user note on the calculation and use of sampling weights for additional information.
Concept:	Technical Household Variables -- HOUSEHOLD
Start Position:	18
End Position:	31
Width:	14
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	CodesNHISHID is a 14-character string variable.

Variable: "HHWEIGHT"

Name:	HHWEIGHT
Label:	Household weight, final annual
Variable Text:	HHWEIGHT is an IPUMS NHIS-constructed variable based on the Final Annual Household Weight for 1997 forward and the Final Basic Weight in 1969-1996 NHIS public use files. HHWEIGHT represents the inverse probability of household selection into the sample, adjusted for non-response. For analyses using the household as the unit of analysis (e.g., how many households contained a person who needed help with activities of daily living?), researchers should use the household weight. Rather than using HHWEIGHT, researchers should use PERWEIGHT or SAMPWEIGHT when analyzing person-level variables or variables from the sample adult/sample child

	supplements from 1997 forward. See the User Notes on the use of sampling weights for additional information.
Concept:	Technical Household Variables -- HOUSEHOLD
Start Position:	32
End Position:	37
Width:	6
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesHHWEIGHT is a 6-digit numeric variable.

Variable: "REGION"

Name:	REGION
Label:	Region of residence
Variable Text:	REGION reports the region of the U.S. where the housing unit containing survey participants was located. The geographic information included in REGION was added during processing, rather than ascertained via questioning. REGION is the smallest geographic unit identified in the IPUMS NHIS data for 1985 forward. In 1968-1984, a limited number (under 30) of metropolitan areas were also identified (METAREA).
	The four regions--Northeast, North Central/Midwest, South, and West--correspond to the U.S. regions recognized by the Census Bureau. Divisions and states included in the four regions are as follows:
	Northeast: New England Division (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, and Connecticut) and Middle Atlantic Division (New York, New Jersey, and Pennsylvania)
	North Central/Midwest: East North Central Division (Michigan, Ohio, Indiana, Illinois, Wisconsin) and West North Central Division (Minnesota, Iowa, Missouri, North Dakota, South Dakota, Kansas, and Nebraska)
	South: South Atlantic Division (Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, and Florida), East South Central Division (Kentucky, Tennessee, Mississippi, and Alabama), and West South Central Division (Texas, Arkansas, Oklahoma, and Louisiana)
	West: Pacific Division (Washington, Alaska, Oregon, California, and Hawaii) and Mountain Division (Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, and Nevada).
	After the 2019 redesign, there is no longer a sampling weight to produce household population estimates. See the user note on the calculation and use of sampling weights for additional information.

Concept:	Geography Variables -- HOUSEHOLD
Start Position:	38
End Position:	39
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	Northeast
02	North Central/Midwest
03	South
04	West
08	NO DATA IN ROUND
09	Unknown

Variable: "PERNUM"

Name:	PERNUM
Label:	Person number within family/household (from reformatting)
Variable Text:	PERNUM is an IPUMS NHIS-constructed variable that numbers all persons within each family or within each household consecutively (starting with "1") in the order in which they appear in the original NHIS data. PERNUM was created by IPUMS NHIS during the process of reformatting the original NHIS public use files.

Prior to 2019, PERNUM numbers all persons within each family. Beginning in 2019, PERNUM numbers all person within a household and has a maximum of "2," as only the sample child and sample adult are identified in the original NHIS data. For more information, see the Comparability section. |
| Concept: | Technical Person Variables -- PERSON |

Start Position:	40
End Position:	41
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesPERNUM is a 2-digit numeric variable.

Variable: "NHISPID"

Name:	NHISPID
Label:	NHIS Unique Identifier, person
Variable Text:	NHISPID is an IPUMS NHIS-constructed value that is an identifying number unique to each person in a given survey year. Using NHISPID, analysts can link IPUMS NHIS data with additional data elements, from the NHIS source public use data files, that are not currently available in IPUMS. For example, researchers can use NHISPID to link variables in IPUMS NHIS with other variables from the core NHIS files variables. Alternatively, users may use NHISPID to link variables from IPUMS with variables from NHIS supplements that are not yet part of IPUMS NHIS. NHISPID is a single variable constructed by concatenating multiple data elements from the NHIS source data files. Analysts wishing to link household-level NHIS data should use the household-level linking key NHISHID. See the user note on LINKING for instructions on how to create links between IPUMS NHIS and NHIS source data.
Concept:	Technical Person Variables -- PERSON
Start Position:	42
End Position:	57
Width:	16
Variable Format:	character
Implied	0

Decimal Places:	
Coder Instructions:	CodesNHISPID is a 16-character string variable.

Variable: "HHX"

Name:	HHX
Label:	Household number (from NHIS)
Variable Text:	For sample adults and sample children, HHX reports the person's household number on the original NHIS data. HHX is unique within sample but only unique across samples when combined with information on year. Prior to 2019, this variable was available for all persons.
Concept:	Technical Person Variables -- PERSON
Start Position:	58
End Position:	64
Width:	7
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	This is a 7-digit numeric variable with 0 implied decimal places

Variable: "FMX"

Name:	FMX
Label:	Family number (from NHIS)
Variable Text:	For all persons, FMX reports the person's family number within the household, as documented on the family record on the original NHIS data. FMX is unique only when combined with information on year and household.
Concept:	Technical Person Variables -- PERSON
Start Position:	65

End Position:	66
Width:	2
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	This is a 2-digit numeric variable with 0 implied decimal places

Variable: "PX"

Name:	PX
Label:	Person number of respondent (from NHIS).
Variable Text:	PX is the original person number assigned to each individual by the NHIS. PX is unique only when combined with information on year, household, and family.
Concept:	Technical Person Variables -- PERSON
Start Position:	67
End Position:	68
Width:	2
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	CodesPX is a 2-digit numeric variable.

Variable: "PERWEIGHT"

Name:	PERWEIGHT
Label:	Final basic annual weight
Variable Text:	PERWEIGHT is an IPUMS NHIS-constructed variable based on the Final Annual Weight in the original NHIS public use files. This weight should be used for many analyses at the

person level, particularly with variables for which information was collected about all family members. PERWEIGHT represents the inverse probability of selection into the sample, adjusted for non-response with post-stratification adjustments for age, race/ethnicity, and sex using the Census Bureau's population control totals. For each year, the sum of these weights is equal to that year's civilian, non-institutionalized U.S. population.

There are instances, however, when the researcher should use a different person-level weight, particularly with supplements where a random adult or child family member was selected for questioning. For example, researchers analyzing variables from the sample adult/sample child supplements from 1997 forward should use SAMPWEIGHT rather than PERWEIGHT.

Users should review the documentation for extracted variables--most notably the "Weights" section toward the top of each variable description--to ascertain which weight is the appropriate choice for a given survey year. See the User Notes on the use of sampling weights for additional information.

Concept:	Technical Person Variables -- PERSON
Start Position:	69
End Position:	80
Width:	12
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesPERWEIGHT is a 12-digit numeric variable.

Variable: "SAMPWEIGHT"

Name:	SAMPWEIGHT
Label:	Sample Person Weight
Variable Text:	SAMPWEIGHT is an IPUMS NHIS-constructed variable that represents, with a few exceptions, the random selection of a sample person in the household to complete a supplement survey. (The exceptions to this generalization are the 1983 health insurance supplement and 1993 health insurance, access to care, and family resources supplements, which were asked of all persons in the second half of the year.)

For survey years 1997 forward, SAMPWEIGHT is based on the Final Annual Sample Adult and Sample Child Weights in the original NHIS public use files. This weight should be used with variables taken from the sample adult or sample child supplements for 1997 on (as indicated by the supplement title in the "Source" section toward the top of each variable description).

SAMPWEIGHT also contains the sampling weights for a subset of the pre-1997 supplements, for which the definition of a sample person, and the rules governing the |

selection of sample persons, are less consistent. The universe statements for SAMPWEIGHT document who was included in the supplements requiring use of this weight on a year-by-year basis. SAMPWEIGHT consistently represents the inverse probability of selection into a supplement, adjusted for non-response with additional post-stratification adjustments for age, race/ethnicity, and sex using the Census Bureau's population control totals.

Users should review the documentation for extracted variables--most notably the "Weights" section toward the top of each variable description--to ascertain which weight is the appropriate choice for a given survey year. See the User Notes on the use of sampling weights for additional information.

Concept:	Technical Person Variables -- PERSON
Start Position:	81
End Position:	92
Width:	12
Variable Format:	numeric
Implied Decimal Places:	3
Coder Instructions:	CodesSAMPWEIGHT is a 12-digit numeric variable with three implied decimals. That is, values of 012345678912 should be interpreted as 12345678.912. The command files delivered with IPUMS extracts automatically divide SAMPWEIGHT by 1,000, so no further adjustment is needed.

Variable: "FWEIGHT"

Name:	FWEIGHT
Label:	Final annual family weight
Variable Text:	FWEIGHT is a variable based on the Final Annual Weight in the original NHIS public use files. This weight should be used for many analyses at the family level, particularly with variables for which information was collected about family characteristics. Creating FWEIGHT requires using independent estimates of the number of families from a reliable source such as the U.S. Census Bureau in order to perform post-stratification adjustments in a manner similar to what was done for the person-level weights. However, such independent estimates do not exist.

As a result, the NHIS used the "principle person" method to create family-level weights. The person weight with the smallest ratio adjustment within each family was used as a proxy for the family-level ratio adjustment. This method is believed to provide a more accurate estimate of the total number of U.S. families than either the use of other person weights within the family or the use of no ratio adjustments at all. Therefore, the Family weight contains factors for selection probabilities at the household level, household non-response adjustment, and several ratio adjustment factors that are applied to all person weights.Users should review the documentation for extracted variables--most notably the "Weights" section toward the top of each variable description--to ascertain which weight is |

	the appropriate choice for a given survey year. See the User Notes on the use of sampling weights for additional information.
Concept:	Technical Person Variables -- PERSON
Start Position:	93
End Position:	98
Width:	6
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesFWEIGHT is a 6-digit numeric variable.

Variable: "INTERVWMO"

Name:	INTERVWMO
Label:	Month of NHIS interview
Variable Text:	INTERVWMO indicates the month that the National Health Interview Survey (NHIS) was conducted.
Concept:	Technical Person Variables -- PERSON
Start Position:	99
End Position:	100
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0
Categories	

Value	Label
01	January

02	February
03	March
04	April
05	May
06	June
07	July
08	August
09	September
10	October
11	November
12	December
98	Unknown-not ascertained

Variable: "INTERVWYR"

Name:	INTERVWYR
Label:	Year of NHIS interview
Variable Text:	INTERVWYR indicates the year that the National Health Interview Survey (NHIS) was conducted.
Concept:	Technical Person Variables -- PERSON
Start Position:	101
End Position:	104
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	This is a 4-digit numeric variable with 0 implied decimal places

Variable: "ASTATFLG"

Name:	ASTATFLG
Label:	Sample adult flag
Variable Text:	ASTATFLG identifies the record of a sample adult. A sample adult is the one person over age 18 per household who is randomly selected to complete the sample adult questionnaire. For 1997-2018, a sample adult is the one adult per family who was selected at random by the computerized survey instrument to answer additional health-related questions, under the survey design instituted in 1997. All persons other than sample adults (those not having the value 1 in ASTATFLG) are coded as "not in universe" for variables derived from the additional questions asked only of sample adults. (Some additional questions were asked of both sample adults and sample children.) Every adult in each family, except for active duty armed forces members, was eligible to be selected as the sample adult.

Prior to 2019, ASTATFLG has the value "no one selected as sample adult" (code 4) for all adults in the family in the following two cases: 1) sample adult status was incorrectly assigned to someone in the armed forces, or 2) it could not be ascertained which adult in the family was chosen as sample adult.

For 2020 only, SALNGPRTFLG further distinguishes between sample adults included in the longitudinal sample and sample adults included in the partial sample. Please see the user note on COVID-related changes to the NHIS for more information on the longitudinal and partial samples.

A similar flag variable, CSTATFLG, identifies the record of a sample child.

This variable is automatically included in all extracts, but only has valid values for 1997 forward. |
Concept:	Technical Person Variables -- PERSON
Start Position:	105
End Position:	105
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	Sample adult, has record
2	Sample adult, no record

3	Not selected as sample adult
4	No one selected as sample adult
5	Armed forces member
6	AF member, selected as sample adult

Variable: "CSTATFLG"

Name:	CSTATFLG
Label:	Sample child flag
Variable Text:	CSTATFLG identifies the record of a sample child. A sample child is the one person under age 18 per household who is randomly selected to complete the sample child questionnaire. For 1997-2018, the sample child is the one person under age 18 per family who was selected at random by the computerized survey instrument for additional health-related questions, under the survey design instituted in 1997. All persons other than sample children (those not having the value 1 in CSTATFLG) are coded as "not in universe" for variables derived from these additional questions asked only of sample children. (Some additional questions were asked of both sample adults and sample children.) Prior to 2019, Every child in each family, except for active duty armed forces members and emancipated minors, was eligible to be selected as the sample child. Emancipated minors are persons age 14 to 17 who are married or living with a partner OR who live on their own without the supervision of a parent, other adult family member, or legal guardian. CSTATFLG has the value "no one selected as sample child" (code 4) for all children in the family in the following two cases prior to 2019: 1) if sample child status was incorrectly assigned to someone who was ineligible, or 2) if it was unknown which child in the family was chosen as sample child. A similar flag variable, ASTATFLG, identifies the record of a sample adult. This variable is automatically included in all extracts, but only has valid values for 1997 forward.
Concept:	Technical Person Variables -- PERSON
Start Position:	106
End Position:	106
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories	

Value	Label
0	NIU
1	Sample child-has record
2	Sample child-no record
3	Not selected as sample child
4	No one selected as sample child
5	Emancipated minor

Variable: "AGE"

Name:	AGE
Label:	Age
Variable Text:	AGE reports the individual's age, in years since their last birthday. Starting in 2019, "Unknown-refused" and "Unknown-don't know" are allowed responses. Prior to 2019 age is not coded as "unknown" for any persons included in the IPUMS NHIS data. As the public use file's codebooks for 1998-2003 state, "Because age is an important variable for instrument check items and in developing the weights, all respondents must have data on age."
Concept:	Core Demographic Variables -- PERSON
Start Position:	107
End Position:	109
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesAGE is a 3-digit-numeric variable. 085: Top code for 85 years or older (1963-1968 and 1997-forward) 090: Top code for 90 years or older (1996 only)

099: Top code for 99 years or older (1969-1995)
997: Unknown-refused
998: Unknown-not ascertained
999: Unknown-don't know

Variable: "SEX"

Name:	SEX
Label:	Sex
Variable Text:	SEX indicates whether the person was male or female. Starting in 2019, "Unknown-refused" and "Unknown-don't know" responses are allowed. Prior to 2019, SEX is not coded as "unknown" for any persons included in the IPUMS NHIS data. According to the 2020 Survey Description, for some variables, including SEX, the 2020 responses of sample adults that were part of the 2020 longitudinal sample were overwritten with their 2019 responses "to mitigate disclosure risks associated with differences in response from repeated measures among the same Sample Adults" (33). The sample adults' actual 2020 responses can be accessed through a Research Data Center (RDC). For more information on the 2020 longitudinal sample, please see SALNGPRTFLG.
Concept:	Core Demographic Variables -- PERSON
Start Position:	110
End Position:	110
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1	Male
2	Female
7	Unknown-refused
8	Unknown-not ascertained
9	Unknown-don't know

Variable: "MARSTAT"

Name:	MARSTAT
Label:	Legal marital status
Variable Text:	MARSTAT reports the person's legal marital status.
Concept:	Core Demographic Variables -- PERSON
Start Position:	111
End Position:	112
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
10	Married
11	Married - Spouse present
12	Married - Spouse not in household
13	Married - Spouse in household unknown
20	Widowed
30	Divorced
40	Separated
50	Never married
99	Unknown marital status

Variable: "HISPETH"

Name:	HISPETH
Label:	Hispanic ethnicity
Variable Text:	HISPETH identifies and classifies persons of Hispanic/Spanish/Latino origin or ancestry. According to the Field Representative's Manual for 1976-1981, "The term 'national origin or ancestry' refers to the national or cultural group from which the person is descended. A person may report his origin based on the origin of a parent, a grandparent, or some far-removed ancestor." The 1982 Field Representative's Manual noted, "There is no set rule as to how many generations are to be taken into account in determining origin." According to the 2020 Survey Description, for some variables, including HISPETH, the 2020 responses of sample adults that were part of the 2020 longitudinal sample were overwritten with their 2019 responses "to mitigate disclosure risks associated with differences in response from repeated measures among the same Sample Adults" (33). The sample adults' actual 2020 responses can be accessed through a Research Data Center (RDC). For more information on the 2020 longitudinal sample, please see SALNGPRTFLG.
Concept:	Ethnicity/Nativity Variables -- PERSON
Start Position:	113
End Position:	114
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
10	Not Hispanic/Spanish origin
20	Mexican
21	Mexican-Mexicano
22	Mexicano
23	Mexican-American
24	Chicano
30	Puerto Rican

40	Cuban/Cuban American
50	Dominican (Republic)
60	Other Hispanic
61	Central or South American
62	Other Latin American, type not specified
63	Other Spanish
64	Hispanic/Latino/Spanish, non-specific type
65	Hispanic/Latino/Spanish, type refused
66	Hispanic/Latino/Spanish, type not ascertained
67	Hispanic/Spanish, type don't know
70	Multiple Hispanic
90	Unknown
91	Unknown if Hispanic/Spanish origin
92	Two origins, unknown which is the main
93	Origin unknown, refused or not reported
99	NIU

Variable: "RACESR"

Name:	RACESR
Label:	Self-Reported Main Racial Background (Pre-1997 Revised OMB Standards)
Variable Text:	RACESR provides information on self-reported, main racial background of all persons, using the pre-1997 Office of Management and Budget's (OMB's) Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting.
Concept:	Ethnicity/Nativity Variables -- PERSON
Start Position:	115
End Position:	117
Width:	3

Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
100	White
200	Black/African American
300	Aleut, Alaskan Native, or American Indian
310	Alaskan Native or American Indian
320	Alaskan Native
330	Aleut
340	American Indian
400	Asian or Pacific Islander
410	Asian
411	Chinese
412	Filipino
413	Korean
414	Vietnamese
415	Japanese
416	Asian Indian
417	Other Asian (1999-2005)
420	Pacific Islander
421	Hawaiian
422	Samoan
423	Guamanian

430	Other Asian or Pacific Islander
431	Other Asian or Pacific Islander (1992-1995)
432	Other Asian or Pacific Islander (1996)
433	Other Asian or Pacific Islander (1997-1998)
500	Other race
510	Other race (1978)
520	Other race (1979-1991)
530	Other race (1992-1995)
540	Other race (1996)
550	Other race (1997-1998)
560	Other race (1999-2002)
570	Primary race not releasable
600	Multiple race, no primary race selected
900	Unknown
970	Unknown-refused
980	Unknown-not ascertained
990	Unknown-don't know

Variable: "EDUCREC1"

Name:	EDUCREC1
Label:	Educational attainment recode, nonintervalled
Variable Text:	EDUCREC1 is a recoded combination of two other variables, HIGRADE1 and EDUC, which measure educational attainment in different ways. HIGRADE1, available for 1982-1996, reports the respondent's highest grade of school or year of college completed, by single years. EDUC, available beginning in 1997, classifies high school graduates according to their highest degree or diploma obtained. In EDUCREC1, the HIGRADE1 categories are given the same codes as their approximate equivalents in EDUC.
Concept:	Education Variables -- PERSON
Start Position:	118

End Position:	119
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
01	Never attended/kindergarten only
02	Grade 1
03	Grade 2
04	Grade 3
05	Grade 4
06	Grade 5
07	Grade 6
08	Grade 7
09	Grade 8
10	Grade 9
11	Grade 10
12	Grade 11
13	Grade 12
14	1 to 3 years of college
15	4 years college/Bachelor's degree
16	5+ years of college
96	Unknown--all causes

97	Unknown--refused
98	Unknown--not ascertained
99	Unknown--not known

Variable: "EMPSTAT"

Name:	EMPSTAT
Label:	Employment status in past 1 to 2 weeks
Variable Text:	EMPSTAT reports whether sample adults were working last week--working for pay; temporarily absent due to illness, vacation, family or maternity leave, or another reason; performing seasonal or contract work; or working at a job or business but NOT for pay (WHYNOWK2). Responses of "no" include both sample adults who were part of the labor force but unemployed and sample adults who were out of the labor force. Additional detail about reasons for not working in the past week can be found in WHYNOWK2. Before 2019, EMPSTAT reports whether persons were part of the labor force--working or seeking work--and, if so, whether they worked, had a job or business from which they were temporarily absent, or were looking for work or on layoff during the preceding two weeks (for 1969-1996) or during the preceding week (for 1997-2018). In 1969-1996, EMPSTAT is a summary recode of responses to a series of questions. In 1969-1973 interviewers first asked, "Did [person] work at any time last week or the week before - (For females): not counting work around the house?" This question was also asked in 1974-1981, but the wording in these years omitted the instructions "For females" before the phrase "not counting work around the house." In 1982-1996 the question wording was changed to, "During those 2 weeks, did [person] work at any time at a job or business, not counting work around the house? (Include unpaid work in the family farm/business.) If the answer to the initial question asked in 1969-1996 was negative, interviewers asked whether the person had a job or business, even though he/she did not work during the reference period. The final questions used to categorize current employment status in 1969-1996 were, "Was he looking for work or on layoff from a job [during those 2 weeks]?" and, if yes, "Which--looking for work or on layoff from a job?"In 1997-2018, respondents were asked one general question to ascertain their employment status. Interviewers asked, "Which of the following was [person] doing last week?" In 1997-2000 interviewers read the following possible responses: "Working at a job or business," "With a job or business but not at work," "Looking for work," and "Not working at a job or business." Beginning in 2001, the response category "Working at a job or business" was split into two separate categories, "Working for pay at a job or business" and "Working, but not for pay, at a job or business." The rest of the response categories in 2001-2018 are the same as in 1997-2000.In 2019 forward, EMPSTAT only indicates whether someone was employed or not employed in the last week. Respondents are asked a series of questions that are used to code a final response to EMPSTAT: "LAST WEEK, did you work for pay at a job or business?" "Did you have a job or business LAST WEEK, but were temporarily absent due to illness, vacation, family or maternity leave, or some other reason?" "What is the MAIN reason you were not working for pay at a job or business last week?" "Does [sample adult] work for pay at a job or business?" NCHS uses these responses to code an individual as "Employed" or "Not employed" in EMPSTAT. Those who indicate they were performing seasonal or contract work or were

working at a job or business but not for pay are coded as "Employed."Definitions and Instructions

Concept:	Work Variables -- PERSON

Start Position:	120

End Position:	122

Width:	3

Variable Format:	numeric

Implied Decimal Places:	0

Categories

Value	Label
000	NIU
100	Employed
110	Working
111	Working for pay at job/business
112	Working, w/out pay, at job/business
120	With job, but not at work
121	With job, not at work: not laid-off, not looking
122	With job, not at work: looking
200	Not employed
210	Unemployed
211	Unemployed: On layoff
212	Unemployed: On layoff and looking
213	Unemployed: Unk if looking or laid off
214	Unemployed: Looking or on layoff

215	Unemployed: Have job to return to
216	Unemployed: Had job during the round
217	Unemployed: No job during reference period
220	Not in labor force
900	Unknown-all causes
997	Unknown-refused
998	Unknown-not ascertained
999	Unknown-don't know

Variable: "OCC"

Name:	OCC
Label:	Detailed occupation
Variable Text:	OCC reports the sample person's primary occupation, coded into a contemporary Classified Index of Occupation (from 1969 to 1982) and the Standard Occupational Classification (from 1983 and forward). Please see the Comparability and Universe tabs for more information on how this variable was constructed. To increase comparability over time, the IPUMS NHIS provides OCC1995, which applies a common occupational classification system over time. This recoded variable is also discussed at the "User Note on Occupation and Industry Variables".
Concept:	Work Variables -- PERSON
Start Position:	123
End Position:	124
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0
Categories	

Value	Label
00	

Variable: "OCC1995"

Name:	OCC1995
Label:	Occupation, 1995 basis
Variable Text:	OCC1995 is a modified version of the 1995 revised Standard Occupational Classification-based codes (SOC), offering researchers a consistent long-term classification of occupations from 1969 to present. All original occupation information is stored in OCC. For more information on all of the occupational schemes that were used in previous NHIS surveys, please refer to the "User Note on Occupation and Industry Variables." Please see the "Universe" tab for more detailed information on changes in the universe across survey samples.

Occupation coding in NHIS surveys underwent its two biggest changes in 1983, when the NHIS switched from using the Classified Index of Occupation to using the 1977 SOC, and in 2004, after the 2000 SOC was released. The introduction of the 1977 SOC and the 2000 SOC contributed to significant discontinuities in occupation coding in the NHIS history. The IPUMS NHIS staff chose the 1995 revised SOC as the standard because it falls in-between the years when the 1977 SOC and the 2000 SOC were implemented in the NHIS.

The construction of OCC1995 is based on the work done by the IPUMS-USA project, which harmonized Census Bureau occupation classifications from 1950 to present, using the 1990 Census Bureau occupational classification scheme as a standard. Based on a series of technical papers and crosswalks that provide detailed analyses on how the occupation coding scheme for each census year differed from the scheme used in the previous census year, the IPUMS staff traced the proportion of each occupation as it broke down into more specific occupations or as it was combined with others into a more aggregated occupation.

The IHIS followed procedures established by IPUMS-USA staff to create a version of the occupation codes that is comparable, to the extent possible, to the 1995 revised SOC that was used in NHIS surveys from 1995 to 2003. However, the IHIS staff was not able to follow the exact procedures employed by IPUMS-USA staff because the NHIS has released far less detail on occupation codes on the public use files than the Census or American Community Survey. For example, in the 2004 sample, the NHIS occupation code "59" (Communications equipment operators) combined multiple occupational classifications that were distinct in the Census occupational classifications, including code "5010" for switchboard operators, code "5020" for telephone operators, and code "5030" for all other communications equipment operators. |
Concept:	Work Variables -- PERSON
Start Position:	125
End Position:	128
Width:	4
Variable Format:	numeric

Implied
Decimal
Places:

0

Categories

Value	Label
0000	NIU
0100	Executive, Administrative, and Managerial Occupations
0101	Managers and administrators, except farm
0102	Officials and administrators, public administration
0103	Managers and administrators, except public administration
0104	Management related occupations
0200	Professional Specialty Occupations
0210	Engineers, Architechts, Surveyors, and Scientists:
0211	Engineers, scientists, and technicians
0212	Engineers and architects
0213	Engineers
0214	Architects and surveyors
0215	Scientists
0216	Natural mathematical and computer scientists
0220	Health Assessment, Diagnosing, and Treating Occupations:
0221	Health workers
0222	Health diagnosing occupations
0223	Health assessment and treating occupations
0230	Teachers, Librarians, Archivists, and Curators:
0231	Teachers, librarians and counselors
0232	Teachers, including professors and instructors
0240	All other professional specialty occupations

0241	All other professional, technical, and kindred workers
0242	Writers, artists, entertainers and athletes
0243	Other professional specialty occupations
0300	Technicians and Related Support Occupations
0301	Health technologists and technicians
0302	Engineering and science technicians
0303	Technologists, technicians except health
0400	Sales Occupations:
0401	Sales workers
0402	Supervisors and proprietors
0403	Sales representatives, commodities and finance
0404	Other sales
0500	Administrative Support Occupations, Including Clerical
0501	Mail carriers, baggagemen, teleg. messengers
0502	Stenos, typists, secretaries, receptionist
0503	Other clerical
0504	Financial records processing occupations
0505	Bookkeepers
0506	Office machine operators
0507	Computer equipment operators
0508	All other clerical workers
0509	Other administrative support
0600	Private Household Occupations:
0601	Private household workers
0602	Housekeepers, laundresses, baby sitters, etc.
0700	Protective Service Occupations

0701	Protective service
0702	Police and firefighters
0703	Other protective service occupations
0800	Service Occupations, Except Protective and Household
0801	All other service workers
0802	Food service
0803	Cleaning and building service
0804	Health and personal service
0805	Health service
0806	Personal service
0900	Farm Operators and Managers:
0901	Farmers and farm managers
0902	Farmers
0903	Farm managers
1000	Other Agricultural and Related Occupations:
1001	Farm workers and other agricultural workers
1002	Farm laborers and farm foremen
1003	Farm laborers
1004	Forestry and fishing occupations
1100	Mechanics and Repairers Occupation
1101	Mechanics and repairers
1102	Construction, including carpenters
1103	Machinists, mechanics, repairment, and metal craftsmen
1200	Construction and Extractive Trades Occupation
1201	Construction and extractive trades
1202	Apprentice, construction
1203	Carpenters

1204	Other construction craftsmen
1205	Mine operatives and laborers
1300	Precision Production Occupations:
1301	Precision production occupations
1302	Metal craftsmen, except mechanics
1303	All other craftsmen
1304	Apprentice, machinists, mechanics, metals
1400	Machine Operators, Assemblers, and Inspectors
1401	Operatives, except transport
1402	Machine operators and tenderers, except precision
1403	Fabricators, assemblers, inspectors, and samplers
1500	Transportation and Material Moving Occupations
1501	Transport equipment operatives
1502	Motor vehicle operators
1503	Other transportation, except motor vehicles
1504	Material moving equipment operators
1505	Construction laborers
1506	Freight, stock and material handlers
1507	All other operatives
1508	Laborers, except farm
1509	All other laborers except farm and mine
1600	Military
7777	Unknown-refused
8888	Unknown-not ascertained
9999	Unknown-don't know

Variable: "IND"

Name:	IND
Label:	Detailed industry recode, 1969-2003
Variable Text:	IND reports the sample person's primary industry, coded into a contemporary Standard Industrial Classification (from 1969 to 2003) and into the North American Industry Classification System (from 2004 and forward). Please see the Comparability and Universe tabs for more information on how this variable was constructed. To increase comparability over time, the IPUMS NHIS provides IND1995, which applies a common industrial classification system over time. This recoded variable is also discussed at the "User Note on Occupation and Industry Variables".
Concept:	Work Variables -- PERSON
Start Position:	129
End Position:	130
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU

Variable: "IND1995"

Name:	IND1995
Label:	Industry, 1995 basis
Variable Text:	IND1995 is a modified version of the 1995 revised Standard Industrial Classification-based codes (SIC), offering researchers a consistent long-term classification of industries from 1969 to present. All original industry codes are available in IND. For more information on all of the industrial schemes that were used in previous NHIS surveys, please refer to the "User Note on Occupation and Industry Variables". Please see the "Universe" tab for more detailed information on changes in the universe across survey samples.

Industry coding in NHIS surveys underwent its two biggest changes in 1983 when the NHIS switched from using the 1967 SIC to the 1972 SIC, and in 2004 after the 2002 North American Industry Classification System (NAICS) was released. The introduction of the 1972 SIC and the 2002 NAICS contributed to significant discontinuities in industry coding in the NHIS history. The IPUMS NHIS staff chose the 1995 revised SIC as the standard because it falls in-between the years when the 1972 SIC and the 2002 NAICS were implemented in the NHIS.

The construction of IND1995 is based on the work done by the IPUMS-USA project, which harmonized Census Bureau industry classifications from 1950 to present, using the 1990 Census Bureau industry classification scheme as a standard. Based on a series of technical papers and crosswalks that provide detailed analyses on how the industry coding scheme for each census year differed from the scheme used in the previous census year, the IPUMS staff traced the proportion of each industry as it broke down into more specific industries or as it was combined with others into a more aggregated industry.

IPUMS NHIS staff followed procedures established by IPUMS-USA staff to create a version of the industry codes that is comparable, to the extent possible, to the 1995 revised SIC that was used in NHIS surveys from 1995 to 2003. However, IPUMS NHIS staff were not able to follow the exact procedures employed by IPUMS-USA staff because the NHIS released far less detail on industry codes on the public use files than the Census or American Community Survey. For example, in the 2010 sample, the NHIS industry code "71" (amusement, gambling, and recreation industries) combined multiple industries that were distinct in the Census, including code "8580" (bowling centers) and code "8590" (other amusement, gambling, and recreation centers).

Concept:	Work Variables -- PERSON
Start Position:	131
End Position:	134
Width:	4
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0000	NIU
0100	Agriculture
0200	Forestry and Fisheries
0300	Mining
0400	Construction

0500	Nondurable products
0510	Food and kindred products
0520	Textile mill and finished textile products
0530	Printing, publishing, and allied industries
0540	Chemicals and allied products:
0600	Durable Goods
0610	Furniture, lumber, and wood
0620	Metal industries:
0621	Primary metal industries
0622	Fabricated metal industries, including ordnance
0630	Machinery, except electrical machinery
0640	Electrical machinery, equipment, and supplies
0650	Transportation equipment
0660	All other and unspecified products
0661	Other nondurable products
0662	Other durable products
0700	Transportation
0710	Railroads
0720	All other transportation
0721	Trucking service and warehousing
0722	All other transportation except trucking service and warehousing
0800	Communications
0900	Utilities and sanitary
1000	Wholesale trade
1100	Eating and drinking places
1200	Other retail trade

1210	General merchandise stores
1220	Food, bakery, and dairy stores
1230	Automotive dealers and gasoline stations
1240	Other and unspecified retail trade
1300	Finance, insurance, and real estate
1310	Banking and credit agencies
1320	Insurance, real estate, and other finance
1400	Business services
1500	Repair services
1600	Private households
1700	Other personal services
1800	Entertainment and recreation services
1900	Medical and other health services
1910	Hospitals
1920	Health services, except hospitals
2000	Educational services
2010	Elementary and secondary schools and colleges
2020	Other educational services
2100	Other miscellaneous services
2110	Social services, religious and membership organizations
2120	Legal, engineering and other professional services
2200	Public Aadministration
2300	Armed forces
7777	Unknown-refused
8888	Unknown-not ascertained
9999	Unknown-don't know

Variable: "NUMEMPS"

Name:	NUMEMPS
Label:	Number of employees at work
Variable Text:	For sample persons aged 18 and over from 1997 to 2003 who were working or were temporarily absent from a job or business last week; and sample persons aged 18 and over who were working or temporarily absent from a job or business last week, or who had ever worked, NUMEMPS reports the number of employees at the person's work location.
Concept:	Work Variables -- PERSON
Start Position:	135
End Position:	136
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
10	1 to 9 employees
20	10 to 24 employees
30	25 to 49 employees
40	50 to 99 employees
50	100 to 249 employees
60	250 to 499 employees
70	500 to 999 employees
80	1000+ employees
97	Unknown-refused

98	Unknown-not ascertained
99	Unknown-don't know

Variable: "CLASSWK2"

Name:	CLASSWK2
Label:	Class of worker, current or longest job
Variable Text:	For sample persons aged 18 and older who were working or temporarily absent from a job last week, or who ever worked, CLASSWK2 reports the type of job that the person held most recently. Please refer to the code tab for more detailed information on different types of jobs recorded.
Concept:	Work Variables -- PERSON
Start Position:	137
End Position:	137
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0
Categories	

Value	Label
0	NIU
1	Employee of private company for wages
2	Federal government employee
3	State government employee
4	Local government employee
5	Self-employed, own bus/prof/farm
6	Working w/out pay in family bus/farm

7	Refused
8	Not ascertained
9	Don't know

Variable: "JOBSECINCORP"

Name:	JOBSECINCORP
Label:	Current self-employment job is incorporated business
Variable Text:	For sample adults aged 18 and older who worked last week and were self-employed in their current job (or last or longest held job), JOBSECINCORP reports whether the business is incorporated. For more information about other related variables, please use the IPUMS NHIS search function and drop-down menus.
Concept:	Work Variables -- PERSON
Start Position:	138
End Position:	138
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0
Categories	

Value	Label
0	NIU
1	No
2	Yes
7	Unknown-refused
8	Unknown-not ascertained

9	Unknown-don't know

Variable: "WHYNOWK2"

Name:	WHYNOWK2
Label:	Corrected main reason not working last week (sample adults)
Variable Text:	For sample adults who were not working for pay last week and were not on temporary leave from a job or business, WHYNOWK2 reports the main reason why the person did not work last week. Prior to 2019, WHYNOWK2 is asked of sample adults whose corrected working status last week was not working at a job or business and was not looking for work (EMPSTATSA = 5) or who was with a job or business but not at work (EMPSTATSA = 2). Please refer to the code tab for more detailed information on why the person did not work last week.
Concept:	Work Variables -- PERSON
Start Position:	139
End Position:	140
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
10	Keeping house
20	Going to school
30	Retired
40	Unable to work for health reasons/disabled
41	Unable to work for health reasons
42	Disabled

50	Unemployed, laid off, looking for work
51	On layoff
52	On a planned vacation from work
53	On family or maternity leave
60	Have job/contract; off season
70	Working at job or business but not for pay
80	Other
97	Unknown-refused
98	Unknown-not ascertained
99	Unknown-don't know

Variable: "CLASSWK"

Name:	CLASSWK
Label:	Class of worker
Variable Text:	For sample adults who were working at a paying job last week, with a job or business last week but not at work, not working last week because they perform seasonal or contract work, or working at a non-paying job last week OR who did not have a job or business last week but did in the past 12 months, CLASSWK reports the type of main job they held. In 2003 and earlier years, CLASSWK reports whether persons were self-employed, worked as an employee in private industry or the public sector, or worked without pay in a family business or farm. If the individual worked more than one job, the question applied to the job where longer hours were worked; if the same hours were worked in two jobs, the question applied to the one the respondent considered most important or to the longest-held job. A related variable, CLASSWK2, collects similar information for the current or longest-held job.
Concept:	Work Variables -- PERSON
Start Position:	141
End Position:	142
Width:	2
Variable Format:	numeric

Implied Decimal Places:	0

Categories

Value	Label
00	NIU
10	Not employed
11	Not in labor force
12	Never worked
13	New worker
20	Private company, wage employee
30	Government employee
31	Federal government employee
32	Non-federal government employee
33	State government employee
34	Local government employee
40	Self-employed
41	Self-employed, incorp business
42	Self-employed, non-incorp business
50	Working w/o pay, family business/farm
96	Unknown--all causes
97	Unknown--refused
98	Unknown--not ascertained
99	Unknown--not known

Variable: "INCFAM97ON2"

Name:	INCFAM97ON2

Label:	Total combined family income (1997+ w. 2007 categories)
Variable Text:	For all persons, INCFAM97ON2 provides total grouped family income beginning in 1997, using income brackets introduced in 2007. Responses are reported in nominal dollars and have not been adjusted for inflation. The 2007 NHIS introduced a recoded income bracket variable to bridge the differences between income brackets from 1997 to 2006 with those from 2007 forward. Individuals whose 1997 to 2006 total family income brackets ranged from "$0 to $4,999" to "$25,000 to $34,999", and those who had incomplete income information but reported that their total family income was "Less than $20,000", are grouped together as "$0 to $34,999." Individuals whose 1997 to 2006 total family income brackets ranged from "$35,000 to $44,999" to "$65,000 to $74,999" are grouped together as "$35,000 to $74,999." See the Codes tab for changes to top-coded income. Those who had incomplete income information but reported that their total family income was $20,000 or more could not be incorporated into the new scheme and are coded separately. Data users wishing to use the original income codes for this variable may use INCFAM9706 (Total combined family income, 1997-2006 grouping) and/or INCFAM07ON (Total combined family income, 2007+). See the drop down menu for additional related variables.
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	143
End Position:	144
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
10	$0 - $34,999
20	$35,000-$74,999
30	$75,000 and over
31	$75,000-$99,999
32	$100,000 and over
96	$20,000 or more (no detail)

97	Refused
98	Not ascertained or undefined
99	Don't know

Variable: "INCIMP1"

Name:	INCIMP1
Label:	Imputed total combined family income (1997+ grouping)
Variable Text:	INCIMP1 is a variable that includes imputed values to replace missing data for the original variable INCFAM97ON2, a recoded variable of total combined family income (from all sources) in the previous calendar year. The complementary imputation flag variable IMPYFAMFLAG1 indicates whether responses in INCIMP1 were reported or imputed. Related Variables and Sources of Additional Information INCIMP1 is the first of five variables that contain imputed values for total family income. It was created as part of a set of variables that provide complete (i.e., without missing values) data on family income. One of the purposes of NHIS data is to study relationships between income and health and to monitor health and health care for persons at different income levels. However, as the technical documentation on "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples" describes, non-response rates are high for questions on total family income in the previous calendar year and personal earnings from employment in the previous calendar year. For more information on the imputation methodology, see EMPSTATIMP1. Before using the imputed income and earnings variables, researchers are strongly advised to read the NCHS documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files. This documentation cautions that each of the five datasets must be merged with other data from the survey to form a single completed dataset. For IPUMS NHIS data users, the imputed income files have already been merged with other data from each survey year for 1997 through the current year of data, as part of the process of adding these imputed income files and variables to the IPUMS NHIS database. The NCHS documentation for the imputed income files directs that analysis of the five versions of each imputed income variable should be done separately, using methods and software that are appropriate for such survey data (for example, SAS-callable SUDAAN or SAS-callable IVEware). Only then can estimates and standard errors be combined using the combining rules described in the aforementioned document on "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey." The 2018 imputed income file documentation further warns: The extra variability due to imputation CANNOT be incorporated by simply analyzing a SINGLE completed data set as if the imputed values were true values. Moreover, analysts SHOULD NOT create a single completed data set using the AVERAGE of the five sets of imputed values. Examples of correct data analyses and additional information about the procedures used to create the imputed data are provided in the technical documentation referred to above.
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	145

End Position:	146
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	$0-$4,999
02	$5,000-$9,999
03	$10,000-$14,999
04	$15,000-$19,999
05	$20,000-$24,999
10	$25,000-$34,999
11	$25,000-$29,999
12	$30,000-$34,999
20	$35,000-$44,999
21	$35,000-$39,999
22	$40,000-$44,999
30	$45,000-$54,999
31	$45,000-$49,999
32	$50,000-$54,999
40	$55,000-$64,999
41	$55,000-$59,999
42	$60,000-$64,999
50	$65,000-$74,999

51	$65,000-$69,999
52	$70,000-$74,999
60	$75,000 and over
61	$75,000-$79,999
62	$80,000-$84,999
63	$85,000-$89,999
64	$90,000-$94,999
65	$95,000-$99,999
66	$100,000 and over
67	$100,000-$104,999
68	$105,000-$109,999
69	$110,000-$114,999
70	$115,000 and over

Variable: "INCIMP2"

Name:	INCIMP2
Label:	Imputed total combined family income (1997+ grouping)
Variable Text:	INCIMP2 is a variable that includes imputed values to replace missing data for the original variable INCFAM97ON2, a recoded variable of total combined family income (from all sources) in the previous calendar year. The complementary imputation flag variable IMPYFAMFLAG2 indicates whether responses in INCIMP2 were reported or imputed.
	Related Variables and Sources of Additional Information
	INCIMP2 is the second of five variables that contain imputed values for total family income. It was created as part of a set of variables that provide complete (i.e., without missing values) data on family income. One of the purposes of NHIS data is to study relationships between income and health and to monitor health and health care for persons at different income levels. However, non-response rates are high for questions on total family income and personal earnings from employment in the previous calendar year. To obtain estimates of these variables for all survey participants, the National Center for Health Statistics created variables with values imputed for missing data for 1997 forward, using multiple-imputation methodology. The NHIS public use files with multiply imputed data consist of five files (and thus five versions of variables containing imputed values for missing data), one for each set of imputed values, to allow the assessment of variability due to imputation. For more information on the imputation methodology, see EMPSTATIMP1.
	Before using the imputed income and earnings variables, researchers are strongly advised to read the NCHS documentation on imputed income, such as 2018 Imputed Family/Personal

Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".

Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	147
End Position:	148
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	$0-$4,999
02	$5,000-$9,999
03	$10,000-$14,999
04	$15,000-$19,999
05	$20,000-$24,999
10	$25,000-$34,999
11	$25,000-$29,999
12	$30,000-$34,999
20	$35,000-$44,999
21	$35,000-$39,999
22	$40,000-$44,999
30	$45,000-$54,999
31	$45,000-$49,999
32	$50,000-$54,999

40	$55,000-$64,999
41	$55,000-$59,999
42	$60,000-$64,999
50	$65,000-$74,999
51	$65,000-$69,999
52	$70,000-$74,999
60	$75,000 and over
61	$75,000-$79,999
62	$80,000-$84,999
63	$85,000-$89,999
64	$90,000-$94,999
65	$95,000-$99,999
66	$100,000 and over
67	$100,000-$104,999
68	$105,000-$109,999
69	$110,000-$114,999
70	$115,000 and over

Variable: "INCIMP3"

Name:	INCIMP3
Label:	Imputed total combined family income (1997+ grouping)
Variable Text:	INCIMP3 is a variable that includes imputed values to replace missing data for the original variable INCFAM97ON2, a recoded variable of total combined family income (from all sources) in the previous calendar year. The complementary imputation flag variable IMPYFAMFLAG3 indicates whether responses in INCIMP3 were reported or imputed.
	Related Variables and Sources of Additional Information
	INCIMP3 is the third of five variables that contain imputed values for total family income. It was created as part of a set of variables that provide complete (i.e., without missing values) data on family income. One of the purposes of NHIS data is to study relationships between income and health and to monitor health and health care for persons at different income levels. However, non-response rates are high for questions on total family income and personal earnings from

employment in the previous calendar year. To obtain estimates of these variables for all survey participants, the National Center for Health Statistics created variables with values imputed for missing data for 1997 forward, using multiple-imputation methodology. The NHIS public use files with multiply imputed data consist of five files (and thus five versions of variables containing imputed values for missing data), one for each set of imputed values, to allow the assessment of variability due to imputation. For more information on the imputation methodology, see EMPSTATIMP1.

Before using the imputed income and earnings variables, researchers are strongly advised to read the NCHS documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".

Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	149
End Position:	150
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	$0-$4,999
02	$5,000-$9,999
03	$10,000-$14,999
04	$15,000-$19,999
05	$20,000-$24,999
10	$25,000-$34,999
11	$25,000-$29,999
12	$30,000-$34,999
20	$35,000-$44,999
21	$35,000-$39,999

22	$40,000-$44,999
30	$45,000-$54,999
31	$45,000-$49,999
32	$50,000-$54,999
40	$55,000-$64,999
41	$55,000-$59,999
42	$60,000-$64,999
50	$65,000-$74,999
51	$65,000-$69,999
52	$70,000-$74,999
60	$75,000 and over
61	$75,000-$79,999
62	$80,000-$84,999
63	$85,000-$89,999
64	$90,000-$94,999
65	$95,000-$99,999
66	$100,000 and over
67	$100,000-$104,999
68	$105,000-$109,999
69	$110,000-$114,999
70	$115,000 and over

Variable: "INCIMP4"

Name:	INCIMP4
Label:	Imputed total combined family income (1997+ grouping)
Variable Text:	INCIMP4 is a variable that includes imputed values to replace missing data for the original variable INCFAM97ON2, a recoded variable of total combined family income (from all sources) in the previous calendar year. The complementary imputation flag variable IMPYFAMFLAG4

indicates whether responses in INCIMP4 were reported or imputed.

Related Variables and Sources of Additional Information

INCIMP4 is the fourth of five variables that contain imputed values for total family income. It was created as part of a set of variables that provide complete (i.e., without missing values) data on family income.
 One of the purposes of NHIS data is to study relationships between income and health and to monitor health and health care for persons at different income levels. However, non-response rates are high for questions on total family income and personal earnings from employment in the previous calendar year. To obtain estimates of these variables for all survey participants, the National Center for Health Statistics created variables with values imputed for missing data for 1997 forward, using multiple-imputation methodology. The NHIS public use files with multiply imputed data consist of five files (and thus five versions of variables containing imputed values for missing data), one for each set of imputed values, to allow the assessment of variability due to imputation. For more information on the imputation methodology, see EMPSTATIMP1.

Before using the imputed income and earnings variables, researchers are strongly advised to read the NCHS documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".

Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	151
End Position:	152
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	$0-$4,999
02	$5,000-$9,999
03	$10,000-$14,999
04	$15,000-$19,999
05	$20,000-$24,999
10	$25,000-$34,999

11	$25,000-$29,999
12	$30,000-$34,999
20	$35,000-$44,999
21	$35,000-$39,999
22	$40,000-$44,999
30	$45,000-$54,999
31	$45,000-$49,999
32	$50,000-$54,999
40	$55,000-$64,999
41	$55,000-$59,999
42	$60,000-$64,999
50	$65,000-$74,999
51	$65,000-$69,999
52	$70,000-$74,999
60	$75,000 and over
61	$75,000-$79,999
62	$80,000-$84,999
63	$85,000-$89,999
64	$90,000-$94,999
65	$95,000-$99,999
66	$100,000 and over
67	$100,000-$104,999
68	$105,000-$109,999
69	$110,000-$114,999
70	$115,000 and over

Variable: "INCIMP5"

Name:	INCIMP5
Label:	Imputed total combined family income (1997+ grouping)
Variable Text:	INCIMP5 is a variable that includes imputed values to replace missing data for the original variable INCFAM97ON2, a recoded variable of total combined family income (from all sources) in the previous calendar year. The complementary imputation flag variable IMPYFAMFLAG5 indicates whether responses in INCIMP5 were reported or imputed. Related Variables and Sources of Additional Information INCIMP4 is the fifth of five variables that contain imputed values for total family income. It was created as part of a set of variables that provide complete (i.e., without missing values) data on family income. One of the purposes of NHIS data is to study relationships between income and health and to monitor health and health care for persons at different income levels. However, non-response rates are high for questions on total family income and personal earnings from employment in the previous calendar year. To obtain estimates of these variables for all survey participants, the National Center for Health Statistics created variables with values imputed for missing data for 1997 forward, using multiple-imputation methodology. The NHIS public use files with multiply imputed data consist of five files (and thus five versions of variables containing imputed values for missing data), one for each set of imputed values, to allow the assessment of variability due to imputation. For more information on the imputation methodology, see EMPSTATIMP1. Before using the imputed income and earnings variables, researchers are strongly advised to read the NCHS documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	153
End Position:	154
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	$0-$4,999
02	$5,000-$9,999

03	$10,000-$14,999
04	$15,000-$19,999
05	$20,000-$24,999
10	$25,000-$34,999
11	$25,000-$29,999
12	$30,000-$34,999
20	$35,000-$44,999
21	$35,000-$39,999
22	$40,000-$44,999
30	$45,000-$54,999
31	$45,000-$49,999
32	$50,000-$54,999
40	$55,000-$64,999
41	$55,000-$59,999
42	$60,000-$64,999
50	$65,000-$74,999
51	$65,000-$69,999
52	$70,000-$74,999
60	$75,000 and over
61	$75,000-$79,999
62	$80,000-$84,999
63	$85,000-$89,999
64	$90,000-$94,999
65	$95,000-$99,999
66	$100,000 and over
67	$100,000-$104,999
68	$105,000-$109,999

| 69 | $110,000-$114,999 |
| 70 | $115,000 and over |

Variable: "IMPYFAMFLAG1"

Name:	IMPYFAMFLAG1
Label:	Family income group imputation flag 1
Variable Text:	IMPYFAMFLAG1 is a generated variable that reports whether the value in the complementary variable INCIMP1, total combined family income, was directly reported (IPUMS NHIS code 0), was imputed without any other income information (IHIS code 1), or was imputed from some information on the level of income (IHIS code 2). Related Variables and Additional Sources of Information The complementary variable INCIMP1 was created as part of a set of variables that provide complete (i.e., without missing values) data on family income; the accompanying flag variables allow researchers to identify in which cases imputation was used. For more details on the purpose and methodology of imputation used in the NHIS, see EMPSTATIMP1. Before using the imputed income and earnings variables, researchers are strongly advised to read the National Center for Health Statistics documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	155
End Position:	155
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	Family income reported

1	Imputed, no income information reported
2	Imputed, some income information reported

Variable: "IMPYFAMFLAG2"

Name:	IMPYFAMFLAG2
Label:	Family income group imputation flag 2
Variable Text:	IMPYFAMFLAG2 is a generated variable that reports whether the value in the complementary variable INCIMP2, total combined family income, was directly reported (IPUMS NHIS code 0), was imputed without any other income information (IHIS code 1), or was imputed from some information on the level of income (IHIS code 2).

Related Variables and Additional Sources of Information

The complementary variable INCIMP2 was created as part of a set of variables that provide complete (i.e., without missing values) data on family income; the accompanying flag variables allow researchers to identify in which cases imputation was used. For more details on the purpose and methodology of imputation used in the NHIS, see EMPSTATIMP1.

Before using the imputed income and earnings variables, researchers are strongly advised to read the National Center for Health Statistics documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples". |
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	156
End Position:	156
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	Family income reported
1	Imputed, no income information reported

| 2 | Imputed, 2-category income reported |

Variable: "IMPYFAMFLAG3"

Name:	IMPYFAMFLAG3
Label:	Family income group imputation flag 3
Variable Text:	IMPYFAMFLAG3 is a generated variable that reports whether the value in the complementary variable INCIMP3, total combined family income, was directly reported (IPUMS NHIS code 0), was imputed without any other income information (IHIS code 1), or was imputed from some information on the level of income (IHIS code 2). Related Variables and Additional Sources of Information The complementary variable INCIMP3 was created as part of a set of variables that provide complete (i.e., without missing values) data on family income; the accompanying flag variables allow researchers to identify in which cases imputation was used. For more details on the purpose and methodology of imputation used in the NHIS, see EMPSTATIMP1. Before using the imputed income and earnings variables, researchers are strongly advised to read the National Center for Health Statistics documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	157
End Position:	157
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	Family income reported
1	Imputed, no income information reported
2	Imputed, 2-category income reported

Variable: "IMPYFAMFLAG4"

Name:	IMPYFAMFLAG4
Label:	Family income group imputation flag 4
Variable Text:	IMPYFAMFLAG4 is a generated variable that reports whether the value in the complementary variable INCIMP4, total combined family income, was directly reported (IPUMS NHIS code 0), was imputed without any other income information (IHIS code 1), or was imputed from some information on the level of income (IHIS code 2). Related Variables and Additional Sources of Information The complementary variable INCIMP4 was created as part of a set of variables that provide complete (i.e., without missing values) data on family income; the accompanying flag variables allow researchers to identify in which cases imputation was used. For more details on the purpose and methodology of imputation used in the NHIS, see EMPSTATIMP1. Before using the imputed income and earnings variables, researchers are strongly advised to read the National Center for Health Statistics documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	158
End Position:	158
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	Family income reported
1	Imputed, no income information reported
2	Imputed, 2-category income reported

Variable: "IMPYFAMFLAG5"

Name:	IMPYFAMFLAG5
Label:	Family income group imputation flag 5
Variable Text:	IMPYFAMFLAG5 is a generated variable that reports whether the value in the complementary variable INCIMP5, total combined family income, was directly reported (IPUMS NHIS code 0), was imputed without any other income information (IHIS code 1), or was imputed from some information on the level of income (IHIS code 2). Related Variables and Additional Sources of Information The complementary variable INCIMP5 was created as part of a set of variables that provide complete (i.e., without missing values) data on family income; the accompanying flag variables allow researchers to identify in which cases imputation was used. For more details on the purpose and methodology of imputation used in the NHIS, see EMPSTATIMP1. Before using the imputed income and earnings variables, researchers are strongly advised to read the National Center for Health Statistics documentation on imputed income, such as 2018 Imputed Family/Personal Earnings Files and "Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey: Methods and Examples".
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	159
End Position:	159
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	Family income reported
1	Imputed, no income information reported
2	Imputed, 2-category income reported

Variable: "INCFAM8296"

Name:	INCFAM8296
Label:	Total combined family income (1982-96 grouping)
Variable Text:	For all persons, INCFAM8296 provides total grouped family income from 1982 to 1996. Responses are reported in nominal dollars and have not been adjusted for inflation.

The field representative manual defined income as:

Wages and salaries including tips, commissions, Armed Forces pay and cash bonuses, as well as subsistence allowances.

Net income from unincorporated businesses, professional practices, or farms, or from rental property.

Social Security, or Supplemental Security Income.

Retirement, disability, and survivor pensions.

Interest and dividends.

Cash public assistance payments (welfare), excluding food stamps.

Veteran's payments.

Unemployment or workmen's compensation.

Alimony and child support.

Money regularly received from friends or relatives not living in the household.

Other periodic money income.
Income does NOT include:

Income "in kind," such as the value of room and board, free meals in restaurants, food stamps, free or reduced rent, value of crops produced by a farmer but consumed by his/her family, etc.

Lump sum payments of any kind, such as insurance payments, inheritances, or retirement.

Occasional gifts of money from persons not living in the household or any exchange of money between relatives living in the same household.

Money received from selling one's own house, car, or other personal property.

Withdrawals of savings from banks, retirement funds, or loans.

Tax refunds or any other refund or rebate.
See the drop down menu for additional related variables. |
Concept:	Total Income and Earnings Variables -- PERSON
Start Position:	160
End Position:	161
Width:	2

Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	Less than $1,000
02	$ 1,000 - 1,999
03	$2,000 - 2,999
04	$3,000 - 3,999
05	$4,000 - 4,999
06	$5,000 - 5,999
07	$6,000 - 6,999
08	$7,000 - 7,999
09	$8,000 - 8,999
10	$9,000 - 9,999
11	$10,000 - 10,999
12	$11,000 - 11,999
13	$12,000 - 12,999
14	$13,000 - 13,999
15	$14,000 - 14,999
16	$15,000 - 15,999
17	$16,000 - 16,999
18	$17,000 - 17,999
19	$18,000 - 18,999
20	$19,000 - 19,999
21	$20,000 - 24,999

22	$25,000 - 29,999
23	$30,000 - 34,999
24	$35,000 - 39,999
25	$40,000 - 44,999
26	$45,000 - 49,999
27	$50,000 and over
99	Unknown

Variable: "LAMTWRK"

Name:	LAMTWRK
Label:	Limited in kind/amount of work (health)
Variable Text:	For sample adults, LAMTWRK indicates whether they were currently unable to work at all, were limited in the amount or kind of work they could do, or were not limited in their work by a physical, mental, or emotional problem. Prior to 2019, the universe for LAMTWRK included all persons age 18 and older. Beginning in 2019, LAMTWRK is one of three functional limitation variables that the NHIS continues to collect but are not part of the Washington Group measures on disability. (See our user note on the Washington Group Disability Measures.) The three measures asked in addition to the Washington Group measures capture limitations in participation in everyday life activities by determining difficulties experienced by sample adults participating in work (LAMTWRK), running errands alone (DISERRANDP), and participating in social activities (DSOCIALP) due to a physical, mental, or emotional condition. The term "problem" was respondent-defined. Nonetheless, the Field Representative's Manual for 1997 forward informed interviewers what was intended by this term. The text states, A problem is defined as the respondent's perception of a chronic, perhaps permanent, departure from physical, mental or emotional well-being. Short-term disabilities (such as pregnancy or injury where full recovery is expected) should not be included as problems.The text for 2001 forward also states, A health problem is respondent defined. Generally speaking, though, it is any condition, physical, mental, or emotional, which causes limitation in activity . . . Do not include pregnancy or delivery as a health problem.The Manual for 1997 forward defined "limited" as referring to a person's ability to only partially perform a specific activity, perform that activity only part of the time, or not perform that activity at all. In practice, "limited" was respondent-defined. The Manual directed interviewers, "Do NOT define this term (limited) to respondents. If asked for a definition, emphasize that we are interested in whether the respondent thinks the person is limited in that specific activity or not." LANOWORK (Unable to work due to health problem) repeats information included in LAMTWRK (i.e., those with an affirmative response, code "2," in LANOWORK, are coded as "3," "Unable to work," in LAMTWRK).
Concept:	Limitation Variables -- PERSON
Start	162

Position:	
End Position:	162
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	Not limited in work
2	Limited in kind/amt of work
3	Unable to work
7	Unknown-refused
8	Unknown-not ascertained
9	Unknown-don't know

Variable: "MORTELIG"

Name:	MORTELIG
Label:	Eligibility status for mortality follow-up
Variable Text:	For all persons in the 1986-2014 NHIS samples and sample adults and sample children in the 2015-2018 NHIS samples, MORTELIG indicates whether they were eligible for follow-up through December 31, 2019 in the Linked Mortality File (LMF) by linkage to the National Death Index (NDI). Only persons (or sample adults, starting in 2015) who were 18 and older at the time of interview and who provided sufficient data were eligible for linkage to the NDI and inclusion in the public use files (IPUMS NHIS code 1 in MORTELIG). Linked mortality data for children (IPUMS NHIS code 2) are not available from the NHIS public use files; researchers interested in these data or in material with additional detail should review the documentation for the restricted access NHIS Linked Mortality Files. Adults with insufficient data for matching (IPUMS NHIS code 3) could not be linked to the NDI to create mortality variables. LMF Variables in the Public Use Update

The following variables were included in the 2019 LMF update:

MORTDIAB: Diabetes flag from multiple cause of death (MCOD)

MORTDODQ: Quarter of death

MORTDODY: Year of death

MORTELIG: Eligibility status for mortality follow-up

MORTHYPR: Hypertension flag from multiple cause of death (MCOD)

MORTSTAT: Final mortality status

MORTUCODLD: Leading underlying cause of death (ICD-10)

MORTWT: Weight adjusted for ineligible respondents in mortality analysis

MORTWTSA: Sample adult weight adjusted for ineligible respondents in mortality analysis Weights for Mortality Data Because adult survey participants who provided insufficient data for linking may differ from those who provided enough data for linking, ignoring the ineligible adult respondents (IPUMS NHIS code 3 in MORTELIG) could lead to biased mortality analysis. To correct this bias, NCHS developed eligibility-adjusted weights for use with the LMF update variables: MORTWT, available for 1986-2014, for use with variables from the NHIS person files, and MORTWTSA, available for 1997-2018, for use with variables from the sample adult files. When analyzing NHIS variables from the person files in conjunction with mortality variables, users should employ MORTWT (rather than the standard person weight, PERWEIGHT). When analyzing variables from the sample adult files in conjunction with mortality variables, users should employ MORTWTSA (rather than the standard sample weight, SAMPWEIGHT). In the 2015 and later NHIS samples, only sample adults are eligible for linkage to the NDI; analyses of the mortality data from these samples should employ MORTWTSA.

Concept:	Mortality Variables -- PERSON
Start Position:	163
End Position:	163
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1	Eligible
2	Under age 18

3	Ineligible
9	NIU

Variable: "MORTSTAT"

Name:	MORTSTAT
Label:	Final mortality status
Variable Text:	For persons aged 18 and older included in the 1986-2014 NHIS samples and sample adults 18 and older included in the 2015-2018 NHIS samples who provided sufficient data for linking (MORTELIG), MORTSTAT reports the final vital status (assumed alive or assumed deceased). For more information about the methodology to link NHIS records to the NDI, please see the NCHS report on the NHIS-NDI linkage methods and analytic considerations. MORTSTAT is one of a set of mortality variables released by the National Center for Health Statistics (NCHS) as part of the 2019 Linked Mortality File (LMF) update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG. For further information on the NDI and the creation of the mortality variables, see the variable description for MORTELIG and the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files. Related Variables For a list of all mortality variables from the LMF included in the IPUMS NHIS, please refer to the MORTELIG variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	164
End Position:	164
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0
Categories	

Value	Label

1	Assumed deceased
2	Assumed alive
9	NIU

Variable: "MORTDODQ"

Name:	MORTDODQ
Label:	Quarter of death
Variable Text:	For persons aged 18 and older included in the 1986-2014 NHIS samples and sample adults age 18 and older included in the 2015-2018 NHIS samples who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2019, MORTDODQ reports the quarter of death. For more information about the LMF methodology to link NHIS records to the NDI, please see the NCHS report on the NHIS-NDI linkage methods and analytic considerations. The related variable MORTDODY reports the year of death. For further information on the NDI and the creation of the mortality variables, see the variable description for MORTELIG and the NHIS Analytic Guidelines (2010) at NHIS Linked Mortality Files. MORTDODQ is one of a set of mortality variables released by the National Center for Health Statistics (NCHS) as part of the 2019 Linked Mortality File (LMF) update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG. Related Variables For a list of all mortality variables included in the IPUMS NHIS, please refer to the MORTELIG variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	165
End Position:	165
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1	January-March
2	April-June
3	July-September
4	October-December
9	NIU

Variable: "MORTDODY"

Name:	MORTDODY
Label:	Year of death
Variable Text:	For persons aged 18 and older included in the 1986-2014 NHIS samples and sample adults aged 18 and older included in the 2015-2018 NHIS samples who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2019, MORTDODY reports the year of death. For more information about the methodology to link NHIS records to the NDI, please see the NCHS report on the NHIS-NDI linkage methods and analytic considerations. The related variable MORTDODQ reports the quarter of death. MORTDODY is one of a set of mortality variables released by the National Center for Health Statistics (NCHS) as part of the 2019 Linked Mortality File (LMF) update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG. For further information on the NDI and the creation of the mortality variables, see the variable description for MORTELIG and the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files. Related Variables For a list of all mortality variables included in the IPUMS NHIS, please refer to the MORTELIG variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	166
End Position:	169
Width:	4
Variable Format:	numeric

Implied Decimal Places:	0

Categories

Value	Label
1986	1986
1987	1987
1988	1988
1989	1989
1990	1990
1991	1991
1992	1992
1993	1993
1994	1994
1995	1995
1996	1996
1997	1997
1998	1998
1999	1999
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
9999	NIU

Variable: "MORTUCOD"

Name:	MORTUCOD
Label:	Underlying cause of death (ICD-10)
Variable Text:	For persons age 18+ included in the NHIS during survey years 1986 to 2004 who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 2006, MORTUCOD reports the underlying cause of death. This information was determined by NCHS based on probabilistic matches of survey participants' NHIS records to National Death Index (NDI) records. MORTUCOD is a recode created by NCHS to classify underlying cause of death in 113 categories. Deaths that occurred prior to 1999 are coded based on the 9th revision of the International Statistical Classification of Diseases, Injuries, and Causes of Death (ICD-9). Deaths that occurred from 1999 forward were coded based on the 10th revision of the International Statistical Classification of Diseases, Injuries, and Causes of Death (ICD-10). This variable is not available for survey participants in the 2005 or later NHIS samples, or for decedents who died after December 2006. Weights for Mortality Data Because adult survey participants who provided insufficient data for linking may differ from those who provided enough data for linking, ignoring the ineligible adult respondents (IPUMS NHIS code 3 in MORTELIG) could lead to biased mortality analysis. To correct this bias, NCHS developed eligibility-adjusted weights: MORTWT, available for 1986-2014, for use with variables from the NHIS person files, and MORTWTSA, available for 1997-2014, for use with variables from the sample adult files. When analyzing NHIS variables from the person files in conjunction with MORTUCOD, users should employ MORTWT (rather than the standard person weight, PERWEIGHT). When analyzing variables from the sample adult files in conjunction with MORTUCOD, users should employ MORTWTSA (rather than the standard sample weight, SAMPWEIGHT).

For further information on the NDI and the creation of the mortality variables, see the variable description for MORTELIG and the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files.Related Variables
For other mortality variables included in IPUMS NHIS are the following, refer to the MORTSTAT variable description.

Concept:	Mortality Variables -- PERSON
Start Position:	170
End Position:	172
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
001	Salmonella infections
002	Shigellosis and amebiasis
003	Other intestinal infections
005	Respiratory tuberculosis
006	Other tuberculosis
007	Whooping cough
008	Scarlet fever and erysipelas
009	Meningococcal infection
010	Septicemia
011	Syphilis
012	Acute poliomyelitis
013	Arthropod-borne encephalitis

014	Measles
015	Viral hepatitis
016	Human immunodeficiency virus
017	Malaria
018	Other infectious parasitic disease
020	Cancer of lip, oral cavity, pharynx
021	Cancer of esophagus
022	Cancer of stomach
023	Cancer of colon, rectum, anus
024	Cancer of liver and bile ducts
025	Cancer of pancreas
026	Cancer of larynx
027	Cancer of trachea, bronchus, lung
028	Malignant melanoma of skin
029	Cancer of breast
030	Cancer of cervix uteri
031	Cancer of uterus
032	Cancer of ovary
033	Cancer of prostate
034	Cancer of kidney and renal pelvis
035	Cancer of bladder
036	Cancer of brain, nervous system
038	Hodgkin's disease
039	Non-Hodgkin's lymphoma
040	Leukemia
041	Multiple myeloma
042	Other lymph and blood cancer

043	Other, unspecified neoplasms
044	In situ and benign neoplasms
045	Anemias
046	Diabetes mellitus
048	Malnutrition
049	Other nutritional deficiencies
050	Meningitis
051	Parkinson's disease
052	Alzheimer's disease
055	Rheumatic fever/heart disease
056	Hypertensive heart disease
057	Hypertensive heart, renal disease
059	Acute myocardial infarction
060	Other acute ischemic heart disease
062	Atherosclerotic cardiovascular disease
063	Other chronic ischemic heart disease
065	Acute and subacute endocarditis
066	Pericardium disease, myocarditis
067	Heart failure
068	All other forms of heart disease
069	Primary hypertension, renal disease
070	Cerebrovascular diseases
071	Atherosclerosis
073	Aortic aneurysm and dissection
074	Other diseases of arteries, capillaries
075	Other disorders of circulatory system

077	Influenza
078	Pneumonia
080	Acute bronchitis and bronchiolitis
081	Acute lower resp. infection, unspec
083	Bronchitis, chronic and unspecified
084	Emphysema
085	Asthma
086	Other chronic lower resp. disease
087	Pneumoconioses
088	Pneumonitis from solids, liquids
089	Other respiratory system diseases
090	Peptic ulcer
091	Diseases of appendix
092	Hernia
094	Alcoholic liver disease
095	Other chronic liver disease
096	Cholelithiasis, gallbladder disease
098	Acute nephrotic syndrome
099	Chronic glomerulonephritis
100	Renal failure
101	Other disorders of kidney
102	Infections of kidney
103	Hyperplasia of prostate
104	Female pelvic inflammatory disease
106	Pregnancy with abortive outcome
107	Other pregnancy, birth complications
108	Certain perinatal conditions

109	Congenital mal or deformations
110	Abnormal clinical, lab findings, nec
111	All other diseases (Residual)
114	Motor vehicle accidents
115	Other land transport accidents
116	Water, air, space transport accidents
118	Falls
119	Accidental discharge of firearms
120	Accidental drowning
121	Accidental fire, smoke exposure
122	Accidental poisoning
123	Other nontransport accidents
125	Suicide by discharge of firearms
126	Suicide by other or means
128	Homicide by firearm discharge
129	Homicide by other means
130	Legal intervention
132	Firearm discharge, unknown intent
133	Other, unspecified events
134	War operations, their sequelae
135	Medical, surgical complications
999	NIU

Variable: "MORTUCODLD"

Name:	MORTUCODLD
Label:	Leading underlying cause of death (ICD-10)

Variable Text:	For persons aged 18 and older included in the 1986-2014 NHIS samples and sample adults aged 18 and older included in the 2015-2018 NHIS samples who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2019, MORTUCODLD reports the leading underlying cause of death. For more information about the methodology to link NHIS records to the NDI, please see the NCHS report on the NHIS-NDI linkage methods and analytic considerations. For survey respondents in the 2005 or later surveys and decedents who died after December 2006, MORTUCODLD is the most detailed information available on the linked NHIS-NDI public use files about underlying cause of death. For access to the detailed cause of death codes available on the linked NHIS-NDI restricted access files, see documentation on the restricted version of the NHIS Linked Mortality Files.

MORTUCODLD is a recode created by NCHS to classify underlying cause of death into 10 categories. Deaths that occurred prior to 1999 are coded based on the 9th revision of the International Statistical Classification of Diseases, Injuries, and Causes of Death (ICD-9). Deaths that occurred from 1999 forward were coded based on the 10th revision of the International Statistical Classification of Diseases, Injuries, and Causes of Death (ICD-10). In sample years 2015 forward, eligible respondents are limited to a single sample adult per household. Because of the smaller sample size, NCHS has grouped people with MORTUCODLD codes of 5-9 into code 10 (All other causes - residual) to lower the risk of respondent identification. Please see the Public-use Linked Mortality File Readme for information about these codes.

MORTUCODLD is one of a set of mortality variables released by the National Center for Health Statistics (NCHS) as part of the 2019 Linked Mortality File (LMF) update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG.

Related Variables
For a list of all mortality variables from the LMF included in the IPUMS NHIS, please refer to the MORTELIG variable description. |
Concept:	Mortality Variables -- PERSON
Start Position:	173
End Position:	174
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
01	Diseases of heart

02	Malignant neoplasms
03	Chronic lower respiratory diseases
04	Accidents (unintentional injuries)
05	Cerebrovascular diseases
06	Alzheimer's disease
07	Diabetes mellitus
08	Influenza and pneumonia
09	Nephritis, nephrotic syndrome and nephrosis
10	All other causes (residual)
96	NIU

Variable: "MORTWT"

Name:	MORTWT
Label:	Weight adjusted for ineligible respondents in mortality analysis
Variable Text:	For persons aged 18 and older included in the 1986 to 2014 NHIS samples, MORTWT reports the sampling weight that adjusts for ineligible respondents in analyses of data linking NHIS participants during survey years 1986 to 2014 to the National Death Index (NDI) with mortality follow up through December 31, 2019. Because of changes in the universe of NHIS participants linked to the NDI, that is, included in the Linked Mortality File (LMF) data, mortality analyses including participants in the 2015 and later NHIS samples should instead use MORTWTSA.
	Because adult survey participants who provided insufficient data for linking may differ from those who provided enough data for linking, ignoring the ineligible adult respondents could lead to biased mortality analysis. To correct this bias, the National Center for Health Statistics (NCHS) developed the eligibility-adjusted weights for use with variables in the 2019 LMF update: MORTWT and MORTWTSA.
	NCHS cautions that the MORTWT, rather than the standard person weight (PERWEIGHT), should be used when analyzing mortality variables in conjunction with variables from the person files for all years other than 1986. Similarly, MORTWTSA should be used when analyzing mortality variables in conjunction with variables from the sample adult files. In 1986, NCHS did not produce an eligibility-adjusted weight for use with the mortality data and instead recommends that analysts use values for PERWEIGHT; the IPUMS variable MORTWT in 1986 reports values of PERWEIGHT for persons eligible for mortality follow-up (MORTELIG values of 1). In the 2015 and later samples, only sample adults are eligible for linkage to the NDI and all analyses of the mortality data from the 2015 and later samples should employ MORTWTSA.
	MORTWT is one of a set of mortality variables released by NCHS as part of the 2019 LMF update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG.

	Related Variables For a list of all mortality variables from the LMF included in the IPUMS NHIS, please refer to the MORTELIG variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	175
End Position:	182
Width:	8
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesMORTWT is a variable with eight columns and no implied decimal places.

Variable: "MORTDIAB"

Name:	MORTDIAB
Label:	Diabetes flag from multiple cause of death (MCOD)
Variable Text:	For persons aged 18 and older included in the 1986-2014 NHIS samples and sample adults age 18 and older included in the 2015-2018 NHIS samples who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2019, MORTDIAB indicates whether, in addition to the leading underlying cause of death (MORTUCODLD), diabetes was listed as a contributing cause of death. For more information about the methodology to link NHIS records to the NDI, please see the National Center for Health Statistics (NCHS) report on the NHIS-NDI linkage methods and analytic considerations. Diabetes was classified based on an ICD-9 code of 250 for 1986-1998 and on an ICD-10 code of E10-E14 for 1999 forward. MORTDIAB is one of a set of mortality variables released by the NCHS as part of the 2019 Linked Mortality File (LMF) update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG. Related Variables MORTDIAB is one of three multiple cause of death flags that indicate whether specific health conditions were listed as contributing causes of death (in addition to the underlying cause of death). MORTHIPFX, not available for persons participating in the NHIS after 2004 or decedents who died after December 2006, indicates whether hip fracture was listed as a contributing cause of death; and MORTHYPR indicates whether hypertension was listed as a contributing cause of death. For a list of all mortality variables included in the IPUMS NHIS, please refer to the MORTELIG variable description.

Concept:	Mortality Variables -- PERSON
Start Position:	183
End Position:	183
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1	No
2	Yes
9	NIU

Variable: "MORTHIPFX"

Name:	MORTHIPFX
Label:	Hip fracture flag from multiple cause of death (MCOD)
Variable Text:	For persons age 18+ included in the NHIS during survey years 1986 to 2004 who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 2006, MORTHIPFX indicates whether, in addition to the underlying cause of death (MORTUCOD), hip fracture was listed as a contributing cause of death. This information was determined by NCHS based on probabilistic matches of survey participants' NHIS records to National Death Index (NDI) records. Hip fracture was classified based on an ICD-9 code of E820 for 1986-1998 and on an ICD-10 code of S70-S722 for 1999-2004. This variable is not available for NHIS participants in the 2005 or later surveys, or for decedents who died after December 2006.

Use of Weights with Mortality Data Because adult survey participants who provided insufficient data for linking may differ from those who provided enough data for linking, ignoring the ineligible adult respondents (IPUMS NHIS code 3 in MORTELIG) could lead to biased mortality analysis. To correct this bias, NCHS developed eligibility-adjusted weights: MORTWT, available for 1986-2009, for use with variables from the NHIS person files, and MORTWTSA, available for 1997-2009, for use with variables from the sample adult files. When analyzing NHIS variables from the person files in conjunction with mortality variables, users should employ MORTWT (rather than the standard person weight, PERWEIGHT). When analyzing variables from the sample adult files in conjunction with mortality variables, users should employ MORTWTSA (rather than the standard sample weight, SAMPWEIGHT).Related |

Variables MORTHIPFX was one of three multiple cause of death flags that indicate whether specific health conditions were listed as contributing causes of death (in addition to the underlying cause of death). MORTDIAB indicates whether diabetes was listed as a contributing cause of death; MORTHYPR indicates whether hypertension was listed as a contributing cause of death.

For other mortality variables included in IPUMS NHIS, please refer to the MORTSTAT variable description.

For further information on the NDI and the creation of the mortality variables, see the variable description for MORTELIG and the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files.

Concept:	Mortality Variables -- PERSON
Start Position:	184
End Position:	184
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1	No
2	Yes
9	NIU

Variable: "MORTHYPR"

Name:	MORTHYPR
Label:	Hypertension flag from multiple cause of death (MCOD)
Variable Text:	For persons ages 18 and older included in the 1986-2014 NHIS samples and sample adults 18 and older included in the 2015-2018 NHIS samples who provided sufficient data for linking (MORTELIG) and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2019. MORTHYPR indicates whether, in addition to the underlying cause of death (MORTUCODLD), hypertension was listed as a contributing cause of death. For more information about the methodology to link NHIS records to the NDI, please see the the National Center for Health Statistics (NCHS) report on the NHIS-NDI linkage methods and

analytic considerations. Hypertension was classified based on an ICD-9 code of 401 or 403 for 1986-1998 and on an ICD-10 code of I10 or I12 for 1999 forward.

MORTHYPR is one of a set of mortality variables released by the NCHS as part of the Linked Mortality File (LMF) update. The LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG. For more information on LMFs in general, see the IPUMS user note about the NHIS-LMFs.

Related Variables MORTHYPR was one of three multiple cause of death flags that indicate whether specific health conditions were listed as contributing causes of death (in addition to the underlying cause of death). MORTDIAB indicates whether diabetes was listed as a contributing cause of death; MORTHIPFX, not available for NHIS participants in the 2005-forward surveys or for decedents who died after December 2006, indicates whether hip fracture was listed as a contributing cause of death.

For a list of all mortality variables from the LMF included in the IPUMS NHIS, please refer to the MORTELIG variable description.

For further information on the NDI and the creation of the mortality variables, see the variable description for MORTELIG and the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files.

Concept:	Mortality Variables -- PERSON
Start Position:	185
End Position:	185
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
1	No
2	Yes
9	NIU

Variable: "MORTCMS"

Name:	MORTCMS

Label:	Mortality status obtained from Centers for Medicaid and Medicare Services
Variable Text:	For persons age 18+ included in the NHIS during survey years 1994-1998 who provided sufficient data for linking (MORTELIG and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2006, MORTCMS indicates that the source of the mortality data was the Centers for Medicaid and Medicare Services (CMS), rather than the National Death Index (NDI). Physicians and hospitals treating Medicaid and Medicare enrollees must report a preliminary assessment of the patient's cause of death (no further documentation is required). This creates a death record in the CMS database. Sources of Information on Mortality Status The National Health Interview Survey Linked Mortality Files Analytic Guidelines state, "The primary determination of mortality for eligible participants is based upon matching survey records to the NDI. However, NCHS also relies on additional sources of information to determine the mortality status of a survey participant. Other sources of mortality include the Social Security Administration, the Centers for Medicare and Medicaid Services, and death certificates. If a source of mortality, other than a NDI record was available, the participant was considered deceased." Other variables identifying the sources of mortality information are: MORTSSA (Mortality status obtained from Social Security Administration) MORTNDI (Mortality match with National Death Index) The NDI is a central database that collects and stores death records from State vital statistics offices. As noted above, mortality ascertainment is first attempted through a probabilistic match between NHIS and NDI death certificate records, if there is sufficient identifying data (Social Security number, date of birth, and name). For additional information, see the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files.Related Variables For other mortality variables included in the IPUMS NHIS, please refer to the MORTSTAT variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	186
End Position:	186
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU

1	No
2	Yes
9	Not applicable in this year

Variable: "MORTNDI"

Name:	MORTNDI
Label:	Mortality match with National Death Index
Variable Text:	For persons age 18+ included in the NHIS during survey years 1986 to 2009 who provided sufficient data for linking (MORTELIG and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2011, MORTNDI indicates that the source of mortality data was the National Death Index (NDI). The NDI is a central database that collects and stores death records from State vital statistics offices. Mortality ascertainment is first attempted through a probabilistic match between NHIS and NDI death certificate records, if there is sufficient identifying data (Social Security number, date of birth, and name). For additional information, see the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files. Sources of Information on Mortality Status The National Health Interview Survey Linked Mortality Files Analytic Guidelines state, "The primary determination of mortality for eligible participants is based upon matching survey records to the NDI. However, NCHS also relies on additional sources of information to determine the mortality status of a survey participant. Other sources include the Social Security Administration, the Centers for Medicare and Medicaid Services, and death certificates. If a source of mortality, other than a NDI record was available, the participant was considered deceased." (p. 4) Variables that identify these other sources of mortality data are: MORTCMS (Mortality status obtained from Centers for Medicaid and Medicare Services) MORTSSA (Mortality match with Social Security Administration records) Related Variables For a list of the other mortality variables included in the IPUMS NHIS, please refer to the MORTSTAT variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	187
End Position:	187
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	No
2	Yes

Variable: "MORTSSA"

Name:	MORTSSA
Label:	Mortality status obtained from Social Security Administration
Variable Text:	For persons age 18+ included in the NHIS during survey years 1986 to 2009 who provided sufficient data for linking (MORTELIG and whose final vital status was judged deceased (MORTSTAT) as of December 31, 2011, MORTSSA indicates that the mortality data were obtained from the Social Security Administration (SSA), rather than from the National Death Index (NDI). The SSA maintains a Death Master File (DMF), a database available to the public containing death notices for enrollees in the U.S. Social Security program. Sources of Information on Mortality Status The National Health Interview Survey Linked Mortality Files Analytic Guidelines state, "The primary determination of mortality for eligible participants is based upon matching survey records to the NDI. However, NCHS also relies on additional sources of information to determine the mortality status of a survey participant. Other sources of mortality include the Social Security Administration, the Centers for Medicare and Medicaid Services, and death certificates. If a source of mortality, other than a NDI record was available, the participant was considered deceased." (p. 4) Other variables identifying the sources of mortality information are: MORTCMS (Mortality status obtained from Centers for Medicaid and Medicare Services) MORTNDI (Mortality match with National Death Index) The NDI is a central database that collects and stores death records from State vital statistics offices. As noted above, mortality ascertainment is first attempted through a probabilistic match between NHIS and NDI death certificate records, if there is sufficient identifying data (Social Security number, date of birth, and name). For additional information, see the technical documents from NCHS on the creation and analysis of the NHIS Linked Mortality Files.Related Variables For a list of the other mortality variables included in the IPUMS NHIS, please refer to the MORTSTAT variable description.
Concept:	Mortality Variables -- PERSON
Start Position:	188
End Position:	188
Width:	1

Variable Format:	numeric

Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	No
2	Yes

Variable: "MORTWTSA"

Name:	MORTWTSA
Label:	Sample adult weight adjusted for ineligible respondents in mortality analysis
Variable Text:	For sample adults aged 18 and older included in the 1997-2018 samples, MORTWTSA reports the sampling weight that adjusts for ineligible respondents in analyses of data linking NHIS participants during survey years 1997 to 2018 to the National Death Index (NDI) with mortality follow up through December 31, 2019. Because of changes in the universe of NHIS participants linked to the NDI, that is, included in the Linked Mortality File (LMF) data, this is the only sampling weight available for use in mortality analyses including participants in the 2015 and later NHIS samples. Because adult survey participants who provided insufficient data for linking may differ from those who provided enough data for linking, ignoring the ineligible adult respondents could lead to biased mortality analysis. To correct this bias, the National Center for Health Statistics (NCHS) developed the eligibility-adjusted weights for use with variables in the 2019 LMF update. MORTWTSA is a mortality weight designed for use with sample adult variables and is available for the NHIS years 1997-2018. (A similar variable, MORTWT, is appropriate for weighting variables from the NHIS person files in conjunction with mortality variables and is available for the NHIS years 1986-2014). When analyzing NHIS variables from the sample adult files in conjunction with any of the mortality variables, users should employ MORTWTSA rather than the sample adult weight (SAMPWEIGHT). SAMPWEIGHT is based on the Final Annual Sample Adult Weights (for survey years 1997 forward) in the original NHIS public use files. MORTWTSA is one of a set of mortality variables released by the NCHS as part of the 2019 LMF update. The 2019 LMF includes mortality information for participants in the 1986-2018 NHIS samples with mortality follow up in the NDI through December 31, 2019. For more information on the LMF, including which variables it offers and the appropriate sampling weights to use when analyzing the mortality data, see the description tab of MORTELIG. Related Variables For a list of all mortality variables from the LMF included in the IPUMS NHIS, please refer to the MORTELIG variable description.

Concept:	Mortality Variables -- PERSON
Start Position:	189
End Position:	196
Width:	8
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesMORTWTSA is a variable with eight columns and no implied decimals.