IPUMS

User Extract nhis_00003.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_00003.dat'
Identification Number:	ddi2-41070110-c062-013b-7abd-0242c0a8b005-nhis_00003.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:	April 14, 2025
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_00003.dat
Responsibility Statement	
Authoring Entity:	IPUMS
Affiliation:	University of Minnesota
Production Statement	
Producer:	IPUMS
Affiliation:	University of Minnesota
Role:	Documentation
Date of Production:	April 14, 2025
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:	2025-04-14

Study Scope

Subject Information

Topic Classification:	Technical Household Variables -- HOUSEHOLD
	Geography Variables -- HOUSEHOLD
	Technical Person Variables -- PERSON
	Core Demographic Variables -- PERSON
	Education Variables -- PERSON
	Work Variables -- PERSON
	Material Hardship Variables -- PERSON
	General Health Variables -- PERSON
	General Coverage Variables -- PERSON
	Condition Variables -- PERSON
	Alcohol Consumption Variables -- PERSON
	Smoking Variables -- PERSON
	Physical Activity Variables -- PERSON
	Fruits and Vegetables (frequency) Variables -- PERSON
	Sugars and Fats (frequency) Variables -- PERSON
	Other Foods (frequency) Variables -- PERSON
	Sleep Variables -- PERSON
	COVID-19 Variables -- PERSON
Summary Data Description	
Time Period:	2022
Country:	United States
Notes	
Note:	Additional notes on a sample that is part of this study: 2022 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, Daniel Backman, Annie Chen, and Stephanie Richards. IPUMS Health Surveys: National Health Interview Survey, Version 7.4 [dataset]. Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D070.V7.4

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: Health including alcohol and other habits and demographics
	This extract is a revision of the user's previous extract, ID 44831950.

§ 3. File Description

File

File Name:	nhis_00003.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

39. TOMSAUCEMNO (Frequency eating tomato sauce, past month: Number of units)
40. SODAPNO (Frequency drinking soft drinks, number of time units)
41. FRIESPNO (Frequency eating fried potatoes, number of time units)
42. SPORDRMNO (Frequency drinking sports and energy drinks, past month: Number of units)
43. FRTDRINKMNO (Frequency drinking fruit-flavored drinks, past month: Number of units)
44. COFETEAMNO (Frequency drinking coffee and tea, past month: Number of units)
45. POTATONO (Frequency eating non-fried potatoes, number of time units)
46. PIZZANO (Frequency eating pizza, number of times units)
47. HRSLEEP (Usual hours sleep per day)
48. CVDSHT (COVID-19 vaccination)

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 Position:	14
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: "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:	32
End Position:	33
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:	34
End Position:	35
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:	36
End Position:	51
Width:	16
Variable Format:	character
Implied Decimal	0

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:	52
End Position:	58
Width:	7
Variable Format:	character
Implied Decimal Places:	0
Coder Instructions:	This is a 7-digit numeric variable with 0 implied decimal places

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:	59
End Position:	70
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: "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:	71
End Position:	71
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:	72
End Position:	72
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:	73
End Position:	75
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:	76
End Position:	76
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: "MARSTCUR"

Name:	MARSTCUR
Label:	Current marital status
Variable Text:	For sample adults (and, prior to 2019, for all persons), MARSTCUR describes the person's self-reported marital status. Beginning in 1997, the NCHS expanded the available responses to include â€ living with a partner,â€ defined as â€ unmarried opposite-sex and same-sex couples living together as if they were married.â€ Prior to this change, instructions to the Field Representative required that same-sex couples who reported being married be reclassified as unmarried and interviewed as different family units. Information on spousal location was collected from all persons who reported themselves as married, either through household composition and family relationship components (prior to 2019) or directly from respondents (SPLIVHERE and SPNOTHERSEP). Some respondents were asked a series of follow-up questions designed to gather information about their legal marital status. From 1997-2018, all persons who reported that they were cohabitating with a partner were asked if they had ever been married

	COHABEVMAR and what their legal status was COHABMARST. After the 2019 survey redesign, all persons who did not report themselves as married were asked similar questions about having been married (MARRIEDEV) and legal status (not available in public-use data). In all cases, a respondentâ€ s legal marital status was recoded by NCHS and is reported in the IPUMS variable MARSTLEG. This variable replaces MARSTCOHAB, which was deprecated in 2024.
Concept:	Core Demographic Variables -- PERSON
Start Position:	77
End Position:	77
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	Married, spouse present
2	Married, spouse absent
3	Married, spouse in household unknown
4	Separated
5	Divorced
6	Widowed
7	Living with partner
8	Never married
9	Unknown marital status

Variable: "EDUC"

Name:	EDUC
Label:	Educational attainment
Variable Text:	For sample adults (and, prior to 2019, for all persons aged 5 and older), EDUC reports the highest level of schooling an individual had completed, in terms of completed grades for persons with less than a high school degree, and in terms of degrees attained for high school graduates and those with higher education. Respondents were handed a card listing recognized categories and asked to identify the correct category. According to the Field Representative's Manuals for 1997-2000, only "regular schooling which advances a person toward an elementary or high school diploma, or a college/university/professional school (such as law, medicine, dentistry)" counted as education. Schooling in settings other than "regular schools" counted only if the credits obtained were acceptable in a regular school system. Thus, for example, "adult education" classes not taken for credit or for the attainment of a college degree did not count. For the end of high school, EDUC distinguishes between those who completed twelfth grade but did not attain a diploma (e.g., those who failed to pass state-mandated high school exit examinations), those who graduated from high school, and those passed the GED (General Equivalence Degree) Test (which certifies that the test taker has attained high school-level academic skills). Persons who attained a high school diploma in less than twelve years were to be recorded as "high school graduates," a guideline consistent with EDUC's emphasis on degrees attained rather than years spent in the classroom. The Field Representative's Manual for 1997-2000 directed interviewers to probe to determine whether persons who attended "post-graduate" high school classes but did not attend college had received a high school diploma. For nurses, interviewers were to determine whether training was received in a college or in a nursing school (with the grade completed at the last regular--non-nursing--school to be entered in the latter case). For a person attending special education classes or a school for persons with mental, physical, or developmental difficulties, interviewers were to ask which grade in a regular school most closely matched the individual's education level. Persons currently in school were generally assumed to have completed the previous grade.
Concept:	Education Variables -- PERSON
Start Position:	78
End Position:	80
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label

000	NIU
100	Grade 12 or less, no high school diploma or equivalent
101	Grade 8 or less (no further detail)
102	Never attended/kindergarten only
103	Grades 1-11 (no further detail)
104	Grade 1
105	Grade 2
106	Grade 3
107	Grade 4
108	Grade 5
109	Grade 6
110	Grade 7
111	Grade 8
112	Grade 9-12, no diploma (no further detail)
113	Grade 9
114	Grade 10
115	Grade 11
116	12th grade, no diploma
200	High school diploma or GED
201	High school graduate
202	GED or equivalent
300	Some college, no 4yr degree
301	Some college, no degree
302	AA degree: technical/vocational/occupational
303	AA degree: academic program

400	Bachelor's degree (BA,AB,BS,BBA)
500	Master's, Professional, or Doctoral Degree
501	Master's degree (MA,MS,Med,MBA)
502	Professional (MD,DDS,DVM,JD)
503	Doctoral degree (PhD, EdD)
504	Other degree
505	Professional School or Doctoral degree, topcoded (MD, DDS, DVM, JD, PhD, EdD)
996	No degree, years of education unknown
997	Unknown--refused
998	Unknown--not ascertained
999	Unknown--don't know

Variable: "HOURSWRK"

Name:	HOURSWRK
Label:	Total hours worked last week or usually
Variable Text:	For sample adults aged 18+ (and all persons age 18+ in 2004-2018) who worked at a paid or unpaid job or business last week, held a seasonal or contract position last week, or were not working last week because they were on vacation, on family/maternity leave, unable to work for health reasons, HOURSWRK reports the total hours that the sample person worked in the past week. or in the past two weeks at all jobs/ businesses or at the main job/business. The universe and time period for HOURSWRK were somewhat different for 1979 and 1991. HOURSWORK reports total hours worked in the past two weeks for sample persons aged 17+ in 1979 who, during the past 2 weeks, worked, had a job or business from which they were temporarily absent, or looked for work or were on layoff; in 1991, HOURSWORK reports total work hours worked in the past two weeks for sample persons aged 18+ in 1991 who were employed in the past 2 weeks and worked for a private company or the federal, state, or local government. Please refer to the Universe and Comparability tabs for more information about changes in the universe, time period, and the types of jobs included in HOURSWRK. Beginning in 2019, sample adults aged 18+ who held a seasonal or contract position last week continued to be identified by NCHS as included in the universe for HOURSWORK, but were omitted from the HOURSWRK question and had their responses set to "Unknown - Not Ascertained."
Concept:	Work Variables -- PERSON
Start Position:	81

End Position:	82
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
01	1
02	2
03	3
04	4
05	5
06	6
07	7
08	8
09	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16

17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41

42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66

67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91

92	92
93	93
94	94
95	95+ hours
97	Unknown--refused
98	Unknown--not ascertained
99	Unknown--don't know

Variable: "POVERTY"

Name:	POVERTY
Label:	Ratio of family income to poverty threshold
Variable Text:	For sample adults and sample children, POVERTY reports the ratio of family income to the U.S. Census Bureau's poverty thresholds for the year in question. Prior to 2019, POVERTY was reported for all persons. The poverty status of a family group is assigned to each member of the family, thus making POVERTY a person-level variable. The ratio of income to the poverty threshold is also calculated for adults who live alone or with persons they are not related to; in such cases, POVERTY is based on the individual's income. Values for POVERTY range from less than half of the poverty threshold for a family of a given size and age structure (i.e., with a given number of children under 18) to five or more times the poverty threshold for a family of a given size and age structure. For families where the number of children under age 18 in the family is equivalent to the number of family members, family income is presumed to be nonexistent, and the ratio of such nonexistent income to the poverty line is undefinable (with a code of 98 in POVERTY). The ratio of family income to the official poverty threshold in POVERTY is available only for 1997 forward. The dichotomous variable POORYN, which distinguishes persons with family income at or above the poverty threshold from persons with family income below the poverty threshold, is available for 1982 forward. For detailed discussion of the basis of the Census Bureau's official poverty thresholds, users should consult the variable description for POORYN.
Concept:	Material Hardship Variables -- PERSON
Start Position:	83
End Position:	84
Width:	2
Variable Format:	numeric

Implied Decimal Places:	0

Categories

Value	Label
10	Less than 1.0
11	Under 0.50
12	0.50 to 0.74
13	0.75 to 0.99
14	Less than 1.0 (no other detail)
20	1.00 to 1.99
21	1.00 to 1.24
22	1.25 to 1.49
23	1.50 to 1.74
24	1.75 to 1.99
25	1.00 - 1.99 (no further detail)
30	2.00 and over
31	2.00 to 2.49
32	2.50 to 2.99
33	3.00 to 3.49
34	3.50 to 3.99
35	4.00 to 4.49
36	4.50 to 4.99
37	5.00 and over
38	2.00 and over (no other detail)
98	Undefinable

99	Unknown

Variable: "HEIGHT"

Name:	HEIGHT
Label:	Height in inches without shoes
Variable Text:	HEIGHT reports the approximate height of sample adults and sample children ages 10-17, in inches. Prior to 2020, HEIGHT was only asked of adults. The values for this variable are based on self-reports or proxy reports/estimates by respondents for other household members, supplied in response to the question, "About how tall is [person] without shoes?" Such data are less accurate than, and thus not fully comparable to, data on height gleaned via direct measurement in other surveys, such as NHANES (the National Health and Nutrition Examination Survey fielded by the National Center for Health Statistics). In 1974 only, the question does not specify "without shoes." According to the 2020 Survey Description, for some variables, including HEIGHT, 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:	General Health Variables -- PERSON
Start Position:	85
End Position:	86
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
12	12
13	13

14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38

39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58 (1996: 58 inches or less)
59	59
60	60
61	61
62	62
63	63

64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77 (1996: 77+ inches)
78	78
79	79
80	80
81	81
82	82
83	83
84	84 (1976 - 1981: 84+ inches)
85	85
86	86
87	87
88	88

89	89
90	90
91	91
92	92
93	93
94	94+
95	Unknown-all causes
96	Exceptionally short or tall (2019-forward: or sex is unknown)
97	Unknown-refused
98	Unknown-not ascertained
99	Unknown-don't know

Variable: "WEIGHT"

Name:	WEIGHT
Label:	Weight in pounds without shoes
Variable Text:	WEIGHT reports the approximate weight of sample adults and sample children ages 10-17, in pounds. Prior to 2020, WEIGHT is only asked of adults. The values for this variable are based on self-reports or proxy reports/estimates by respondents for other household members, supplied in response to the question, "About how much does [person] weigh without shoes?" In 1974 only, the question does not specify "without shoes" and in 1976 and 1977 the question specifies "without clothes or shoes." Such verbally reported data are less accurate than, and thus not fully comparable to, data on weight gleaned via direct measurement in other surveys, such as NHANES (the National Health and Nutrition Examination Survey fielded by the National Center for Health Statistics). Indeed, the Field Representative's Manual for 1991 specified that one rationale for collecting height and weight data in the NHIS was to compare respondents' answers to the distribution of actual body measurements obtained in the Health and Nutrition Examination Survey, "to determine the reliability of self-reported or proxy-reported heights and weights."
Concept:	General Health Variables -- PERSON
Start Position:	87
End Position:	89

Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesWEIGHT is a 3-digit-numeric variable. The upper and lower weight thresholds for public data suppression vary over time. User should carefully review the codes below before making cross-year comparisons or interpreting results. 2019-forward codes 000: Not in Universe 100-299: 100-299 pounds 996: Exceptionally low or high weight or sex is missing 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know 1997-2018 codes 000: Not in Universe 99: 99 pounds or less (FEMALES only, 1997-2005) 100: 100 pounds or less (FEMALES only, 2006-2018) 126: 126 pounds or less (MALES only, 1997-2018) 259: 259 pounds or more (FEMALES only, 1997-2005) 274: 274 pounds or more (FEMALES only, 2006-2018) 285: 285 pounds or more (MALES only, 1997-2005) 299: 299 pounds or more (MALES only, 2006-2018) 996: Exceptionally low or high weight 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know In 1997-2018, data are NOT top and bottom coded, but instead cases outside of these upper and lower bounds are assigned to a code of "996: Exceptionally low or high weight". Additionally, these thresholds differ by sex. Pre-1997 codes: 000: Not in Universe (includes 2 reported "0" values in 1974, which has no bottom-code) 50: 50 pounds or less (1976-1995) 97: 97 pounds or less (1996) 290: 290 pounds or more (1996) 300: 300 pounds or more (1976 and 1978) 400: 400 pounds or more (1977, 1979-1981) 500: 500 pounds or more (1982-1995) 997: 997 pounds or more (1974) 995: Unknown - all causes

Variable: "BMICALC"

Name:	BMICALC
Label:	Body Mass Index, calculated from publicly released height and weight variables
Variable Text:	BMICALC reports the Body Mass Index, a measure of body fat based on height and weight, as calculated by IPUMS NHIS from the public use file data on height and weight. BMICALC was calculated using the following formula: [Weight in pounds/(Height in

	inches, squared)] multiplied by 703 and rounded to one digit past the implied decimal point. Individuals not asked their weight and height are coded "000" (for "Not in universe") for BMICALC. Individuals whose weight or height was topcoded, bottomcoded, or, for 1997 forward, an outlying value that was suppressed for confidentiality reasons, receive a code of "996" for BMICALC. The website of the Centers for Disease Control and Prevention on BMI reports the following "standard weight status categories associated with BMI ranges for adults." BMI below 18.5 is associated with "underweight" weight status; BMI 18.5 to 24.9 is associated with "normal" weight status; BMI 25.0 to 29.9 is associated with "overweight" weight status; BMI 30.0 and above is associated with "obese" weight status. The CDC website also notes, "the correlation between the BMI number and body fatness is fairly strong; however, the correlation varies by sex, race, and age." For example, women tend to have more body fat than men at the same BMI; older people, on average, tend to have more body fat than young adults; and highly trained athletes may have a high BMI because of increased muscularity rather than increased body fatness. The CDC adds, "It is important to remember that BMI is only one factor related to risk for disease," and that other predictors, such as waist circumference and levels of blood pressure and physical activity, are also important.
Concept:	General Health Variables -- PERSON
Start Position:	90
End Position:	93
Width:	4
Variable Format:	numeric
Implied Decimal Places:	1
Coder Instructions:	CodesBMICALC is a 4-digit numeric variable with one implied decimal place. That is, a value of 0123 should be interpreted as 12.3 The command files delivered with IPUMS extracts automatically divide BMICALC by 10, so no further adjustment is needed. 0.0 = NIU 996.0 = Not calculable

Variable: "HINOTCOVE"

Name:	HINOTCOVE
Label:	Health Insurance coverage status
Variable Text:	For sample adults and sample children, HINOTCOVE indicates whether the person currently lacks health insurance coverage. Prior to 2019, this variable is available for all persons. HINOTCOVE is a recoded variable created by the National Center for Health Statistics (NCHS) and included in the original NHIS public use data. HINOTCOVE, like other recoded health insurance variables in these data, is based on responses to a series of questions and on back editing carried out by NCHS staff. For the 1997 sample only, HINOTCOVE is constructed by IPUMS NHIS staff using back edited variables from the original NHIS public

	use data. The component variables used by IPUMS NHIS staff are available in IPUMS NHIS as HIPRIVATEE, HIMILITE, HIMCAIDE, HIMCAREE, HICHIPE, HISTATEE, and HIOTHGOVE.
Concept:	General Coverage Variables -- PERSON
Start Position:	94
End Position:	94
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	No, has coverage
2	Yes, has no coverage
7	Unknown-refused
8	Unknown-not ascertained
9	Unknown-don't know

Variable: "CANCEREV"

Name:	CANCEREV
Label:	Ever told had cancer
Variable Text:	CANCEREV identifies sample adults/persons who were ever diagnosed by a doctor or other health professional as having cancer. The Field Representative's Manuals for 1997-2000 directed the interviewer, "Do NOT include self-diagnosed conditions or conditions reported by a person who is not a doctor and not working with or for a doctor." The Manual for 2005 defined cancer as "diseases in which abnormal cells divide without control. Cancer cells can invade nearby tissue and can spread through the bloodstream and lymphatic system to other parts of the body." This definition, which was not routinely shared with respondents, was taken from the Healthy People 2010 Objectives, a comprehensive set

	of national health goals for the decade, which are coordinated by the Office of Disease Prevention and Health Promotion of the U.S. Department of Health and Human Services. Available in 2007, CANCERYR identifies sample children who were diagnosed by a doctor or other health professional as having cancer during the past 12 months.
Concept:	Condition Variables -- PERSON
Start Position:	95
End Position:	95
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: "CHEARTDIEV"

Name:	CHEARTDIEV
Label:	Ever told had coronary heart disease
Variable Text:	CHEARTDIEV identifies sample adults whom a doctor or other health professional ever diagnosed as having coronary heart disease. In 2002 and 2007, individuals who acknowledged ever being diagnosed with coronary heart disease were asked the follow-up question, "During the past 12 months, have you had coronary heart disease?" (CHEARTDIYR).

	Coronary heart disease was not defined in the survey documentation. The Medline Plus Medical Encyclopedia defines coronary heart disease as follows: "Coronary heart disease (or coronary artery disease) is a narrowing of the small blood vessels that supply blood and oxygen to the heart (coronary arteries). Coronary disease usually results from the build-up of fatty material and plaque (atherosclerosis). As the coronary arteries narrow, the flow of blood to the heart can slow or stop. The disease can cause chest pain (stable angina), shortness of breath, heart attack, or other symptoms. Coronary heart disease (CHD) is the leading cause of death in the United States for men and women." Related Variables In 1997 and subsequent years, the survey collected information from sample adults about a number of other conditions related to coronary heart disease. In 1997 forward, sample adults were asked whether they were ever told by a doctor or other health professional that they had angina pectoris (ANGIPECEV). Sample adults were also consistently asked whether they had ever been diagnosed with a heart attack or myocardial infarction (HEARTATTEV) or any other heart condition or heart disease (other than heart attack, angina pectoris, and coronary heart disease) (HEARTCONEV). In 2002 and 2007, follow-up questions for those giving affirmative answers covered whether sample adults had a heart attack (HEARTATYR), angina pectoris (ANGIPECYR), or any other heart condition or heart disease (HEARTCONYR) during the past 12 months. The 1999 survey collected information on age when first diagnosed with a heart attack (HEARTATAGE). In 1999 and 2002, the survey collected information on whether sample adults were ever diagnosed with congestive heart failure (CONHFAILEV) and, in 2002, about whether they had congestive heart failure during the past 12 months (CONHFAILYR). The 1977 survey included a single, broad question covering all types of heart disease or heart trouble. Specifically, persons age 20 and older were asked, "During the past 12 months, did you have heart disease or heart trouble?" (HEARTDISYR). For sample children, from 1997 forward, information was collected on whether the child was ever diagnosed with congenital heart disease (CONGHARTEV) or with some other (non-congenital) heart condition (NONCONHARTEV).
Concept:	Condition Variables -- PERSON
Start Position:	96
End Position:	96
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: "DIABETICEV"

Name:	DIABETICEV
Label:	Ever told had diabetes
Variable Text:	DIABETICEV identifies sample children and sample adults who had ever been diagnosed with "diabetes or sugar diabetes" by a doctor or other health professional. The format of the questions about diabetes differed for children and adults. For children, respondents were shown a flashcard listing several chronic conditions (including diabetes) and asked, "Looking at this list, has a doctor or health professional ever told you that [person] had any of these conditions?" For adults, the question was, "Have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?" For women, the question was preceded by the phrase, "Other than during pregnancy" (to exclude cases of gestational diabetes present only during pregnancy). The range of responses recognized also differed for children and adults. For children, responses were a dichotomous yes and no (or mentioned and not mentioned). For adults from 1997-2018, there was a third option, "Borderline." Borderline diabetes was not defined in the survey documentation, but outside sources indicate this term refers to "impaired glucose tolerance" or "impaired fasting glucose." In recent years, the term "pre-diabetes" has increasingly displaced the term "borderline diabetes." Beginning in 2019, a separate question was asked to report if a doctor or other health professional had ever told the respondent they have prediabetes or borderline diabetes (DIAPRE). Definitions of Diabetes The Field Representative's Manual for 1997-2000 states, "Do not include self-diagnosed diabetes, pre-diabetes, high sugar, or any condition other than 'diabetes' or 'sugar diabetes.' Do NOT include a doctor's diagnosis of 'gestational diabetes' or diabetes present only when a woman is pregnant." These directions were not routinely read to respondents and were contradicted by the survey form itself, with its inclusion of the "borderline" option (which overlaps with "pre-diabetes" and "high sugar"). Beginning in 2001, the Manual included the following definition of diabetes: "Diabetes is a chronic disorder of carbohydrate metabolism involving insulin. Symptoms include elevated sugar in the urine and the blood, excessive urination, thirst, hunger, weight loss, and itching." This definition was not routinely shared with respondents.Other Diabetes-Related Variables Some diabetes-related variables, such as DIABETICEV, were included in the National Health Interview Survey (NHIS) every year, beginning in 1997. From 1997 forward, the NHIS included questions about age at first diagnosis of diabetes (DIABETICAGE) for adult respondents who reported a diabetes diagnosis (other than "borderline.") Similarly, beginning in 1997, the survey consistently asked whether persons ever diagnosed with diabetes were now taking insulin (INSULIN) or were now taking diabetic pills (DIAPILLS). Other diabetes-related questions were periodically included in the survey, particularly in 1999 and 2003. The Manual for 2003 reported that the additional diabetes questions in that year were sponsored by the National Institute(s) of Health.

One group of the "periodic" diabetes questions collect further information on the time elapsed since first diagnosis of the illness. These include:

Years since first diagnosed with diabetes (DIAYRSAGO, in 1999-2005)

Diabetes diagnosed in past 12 months (DIAGNOSYR, in 1999)

Months ago, within year, diabetes diagnosed (DIAGNOSYRMO, in 1999)

A second group of "periodic" diabetes variables relate to the hemoglobin A1C test, a blood test which provides information about how high an individual's blood sugar level has been over the past 3 months. These variables include:

Ever heard of hemoglobin A1C (DIA1CKNOW in 1999, 2003)

Number exams for A1C hemoglobin, past 12 months (DIA1CEXAMYR, in 1999, 2003)

Last A1C hemoglobin level (DIA1CLEVEL, in 2003)

Safe A1C hemoglobin level, according to doctor (DIA1CSAFE, in 2003)

A third group of "periodic" diabetes-related variables focus on the frequency of checking blood glucose levels. These variables include:

Times per day check blood for glucose (DIAGLUCDAY, in 2003)

How often check blood for glucose: Time period (DIAGLUCTP, in 2003)

How often check blood for glucose: Number of units (DIAGLUCNO, in 2003)

A fourth group of "periodic" diabetes-related variables are concerned with the frequency of checking for foot sores. These variables include:

Number exams for foot sores, past 12 months (DIAFTEXAMYR, in 1999, 2003)

Times per week check feet for sores (DIAFTEXAMWK, in 2003)

How often check feet for sores: Time period (DIAFTEXAMTP, in 2003)

How often check feet for sores: Number of units (DIAFTEXAMNO, in 2003)

A fifth group of "periodic" diabetes-related variables relate to eye examinations and eye diseases linked to diabetes. These variables include:

Months since last eye exam with pupils dilated (DIAEYEXAMO, in 1999, 2002-2003)

Ever told had diabetic retinopathy (DIARETINEV, in 2002)

Had diabetic retinopathy, past 12 months (DIARETINYR, in 2002)
The last group of "periodic" diabetes-related variables concern contact with health professionals for the treatment and management of diabetes. These variables include:

Seeing one doctor for diabetes (DIAONEDR, in 2003)

Number times saw doctor for diabetes, past 12 months (DIADRYRNO, in 2003)

	Number times saw nurse/dietician for diabetes, past 12 months(DIANURSYRNO, in 2003) Ever took class in diabetes management (DIACLASS, in 1999) Starting in 2021, the following diabetes related variables were introduced: Time since A1C level checked (DIALASTA1C) How often individual felt overwhelmed from having diabetes (DIADISTRESS) Level of distress due to diabetes compared to before coronavirus pandemic (DIACVDSTRESS) See INSULIN for insulin related variables
Concept:	Condition Variables -- PERSON
Start Position:	97
End Position:	97
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
0	NIU
1	No or not mentioned
2	Yes or mentioned
3	Borderline
7	Unknown-refused
8	Unknown-not ascertained
9	Unknown-don't know

Variable: "HEARTATTEV"

Name:	HEARTATTEV
Label:	Ever told had heart attack
Variable Text:	HEARTATTEV indicates whether sample adults were ever told by a doctor or other health professional that they had "a heart attack (also called myocardial infarction)." Heart attack was not defined in the survey documentation. The Medline Plus Medical Encyclopedia defines heart attack as follows: "A heart attack (myocardial infarction) occurs when an area of heart muscle dies or is permanently damaged because of an inadequate supply of oxygen to that area." Related Variables In the 2002 and 2007 surveys, sample adults ever diagnosed with heart attack were asked if they had a heart attack during the past twelve months (HEARTATYR). In the 1999 survey, sample adults who were ever told by a doctor or other health professional that they had a heart attack were asked their age when first diagnosed with a heart attack (HEARTATAGE). Heart attacks are generally associated with coronary heart disease. In 1997 forward, sample adults were asked whether they were ever told by a doctor or other health professional that they had coronary heart disease (CHEARTDIEV). Sample adults were also consistently asked whether they had ever been diagnosed with angina pectoris (ANGIPECEV) or any other heart condition or heart disease (other than heart attack, angina pectoris, and coronary heart disease) (HEARTCONEV). In 2002 and 2007, follow-up questions for those giving affirmative answers covered whether sample adults had angina pectoris (ANGIPECYR), coronary heart disease (CHEARTDIYR), or any other heart condition or heart disease (HEARTCONYR) during the past 12 months. In 1999 and 2002, the survey collected information on whether sample adults were ever diagnosed with congestive heart failure (CONHFAILEV) and, in 2002, about whether they had congestive heart failure during the past 12 months (CONHFAILYR). The 1977 survey included a single, broad question covering all types of heart disease or heart trouble. Specifically, persons age 20 and older were asked, "During the past 12 months, did you have heart disease or heart trouble?" (HEARTDISYR). For sample children, from 1997 forward, information was collected on whether the child was ever diagnosed with congenital heart disease (CONGHARTEV) or with some other (non-congenital) heart condition (NONCONHARTEV).
Concept:	Condition Variables -- PERSON
Start Position:	98
End Position:	98
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: "STROKEV"

Name:	STROKEV
Label:	Ever told had a stroke
Variable Text:	STROKEV identifies adults whom a doctor or other health professional ever diagnosed as having had a stroke. In 2002 and 2007, persons who had ever been diagnosed with a stroke were asked whether they had a stroke during the past 12 months (STROKEYR). Please see Comparability and Universe tabs for complete information on universe changes between samples. The Field Representative's Manual for 1997-2000 directed interviewers to "include strokes, cerebrovascular accidents, and brain hemorrhages," but this directive was not stated during the interview. Stroke was defined in the Field Representative's Manual for 2001 forward as "a cerebral hemorrhage or embolism of the cerebral blood vessels," but this definition was not routinely shared with respondents. Related Variables Information on whether persons age 20 and older ever had a stroke is also available in 1977 (STROKEWOC). The information included in STROKEWOC was collected via two separate lines of questioning. Interviewers asked, "Has [name] ever had a stroke?" and "Has a doctor ever told [name] he had a stroke?" In addition, the person was recorded as having had a stroke and asked about a doctor's diagnosis of the condition if "stroke" was reported as the reason why an individual was limited in daily activities (i.e., "needs the help of another person in getting around," "stays in bed all or most of the day," "needs help in bathing/dressing/eating/using the toilet"). Only cases coded as "2" (Yes, medically confirmed) in STROKEWOC are roughly comparable to cases with an affirmative response in STROKEV. Along with the notable difference in question wording for the two variables, there is a difference in the range of responses that were accepted. In the 1977 inquiry, interviewers were directed to reask the question if a term other than "stroke" (such as "brain hemorrhage") was reported, and to consider as a "No" response any term other than "stroke." By contrast, as noted above, for STROKEV, other terms such as "cerebral hemorrhage" and "cerebrovascular accident" were accepted as "Yes" responses.
Concept:	Condition Variables -- PERSON

Start Position:	99
End Position:	99
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: "ALCANYNO"

Name:	ALCANYNO
Label:	Frequency drank alcohol in past year: Number of units
Variable Text:	For sample adults who have had at least one alcoholic drink in their entire life (ALCDRINKEV) and, before 2019, sample adults ages 18 and older who had at least 12 alcoholic beverages in their entire life (ALCLIFE), ALCANYNO reports the number of days (every day or the number of days per week, month, or year) for the frequency with which the person drinks alcohol. ALCANYNO should be paired with ALCANYTP (Frequency drank alcohol in past year: Time period) to obtain the complete frequency of alcohol consumption. Please see Comparability and Universe tabs for information on changes in the wording and universe between samples. Beginning in 2020, ALCANYNO is collected as part of the rotating core of NHIS alcohol use questions, fielded every other year. For more information about the NHIS alcohol use rotating core, please see ALCDRINKEV.

Concept:	Alcohol Consumption Variables -- PERSON
Start Position:	100
End Position:	102
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
000	Never
001	1 day
002	2 days
003	3 days
004	4 days
005	5 days
006	6 days
007	7 days
008	8 days
009	9 days
010	10 days
011	11 days
012	12 days
013	13 days

014	14 days
015	15 days
016	16 days
017	17 days
018	18 days
019	19 days
020	20 days
021	21 days
022	22 days
023	23 days
024	24 days
025	25 days
026	26 days
027	27 days
028	28 days
029	29 days
030	30 days
031	31 days
032	32 days
033	33 days
034	34 days
035	35 days
036	36 days
037	37 days
038	38 days

039	39 days
040	40 days
041	41 days
042	42 days
043	43 days
044	44 days
045	45 days
046	46 days
047	47 days
048	48 days
049	49 days
050	50 days
051	51 days
052	52 days
053	53 days
054	54 days
055	55 days
056	56 days
057	57 days
058	58 days
059	59 days
060	60 days
061	61 days
062	62 days
063	63 days

064	64 days
065	65 days
066	66 days
067	67 days
068	68 days
069	69 days
070	70 days
071	71 days
072	72 days
073	73 days
074	74 days
075	75 days
076	76 days
077	77 days
078	78 days
079	79 days
080	80 days
081	81 days
082	82 days
083	83 days
084	84 days
085	85 days
086	86 days
087	87 days
088	88 days

089	89 days
090	90 days
091	91 days
092	92 days
093	93 days
094	94 days
095	95 days
096	96 days
097	97 days
098	98 days
099	99 days
100	100 days
101	101 days
102	102 days
103	103 days
104	104 days
105	105 days
106	106 days
107	107 days
108	108 days
109	109 days
110	110 days
111	111 days
112	112 days
113	113 days

114	114 days
115	115 days
116	116 days
117	117 days
118	118 days
119	119 days
120	120 days
121	121 days
122	122 days
123	123 days
124	124 days
125	125 days
126	126 days
127	127 days
128	128 days
129	129 days
130	130 days
131	131 days
132	132 days
133	133 days
134	134 days
135	135 days
136	136 days
137	137 days
138	138 days

139	139 days
140	140 days
141	141 days
142	142 days
143	143 days
144	144 days
145	145 days
146	146 days
147	147 days
148	148 days
149	149 days
150	150 days
151	151 days
152	152 days
153	153 days
154	154 days
155	155 days
156	156 days
157	157 days
158	158 days
159	159 days
160	160 days
161	161 days
162	162 days
163	163 days

164	164 days
165	165 days
166	166 days
167	167 days
168	168 days
169	169 days
170	170 days
171	171 days
172	172 days
173	173 days
174	174 days
175	175 days
176	176 days
177	177 days
178	178 days
179	179 days
180	180 days
181	181 days
182	182 days
183	183 days
184	184 days
185	185 days
186	186 days
187	187 days
188	188 days

189	189 days
190	190 days
191	191 days
192	192 days
193	193 days
194	194 days
195	195 days
196	196 days
197	197 days
198	198 days
199	199 days
200	200 days
201	201 days
202	202 days
203	203 days
204	204 days
205	205 days
206	206 days
207	207 days
208	208 days
209	209 days
210	210 days
211	211 days
212	212 days
213	213 days

214	214 days
215	215 days
216	216 days
217	217 days
218	218 days
219	219 days
220	220 days
221	221 days
222	222 days
223	223 days
224	224 days
225	225 days
226	226 days
227	227 days
228	228 days
229	229 days
230	230 days
231	231 days
232	232 days
233	233 days
234	234 days
235	235 days
236	236 days
237	237 days
238	238 days

239	239 days
240	240 days
241	241 days
242	242 days
243	243 days
244	244 days
245	245 days
246	246 days
247	247 days
248	248 days
249	249 days
250	250 days
251	251 days
252	252 days
253	253 days
254	254 days
255	255 days
256	256 days
257	257 days
258	258 days
259	259 days
260	260 days
261	261 days
262	262 days
263	263 days

264	264 days
265	265 days
266	266 days
267	267 days
268	268 days
269	269 days
270	270 days
271	271 days
272	272 days
273	273 days
274	274 days
275	275 days
276	276 days
277	277 days
278	278 days
279	279 days
280	280 days
281	281 days
282	282 days
283	283 days
284	284 days
285	285 days
286	286 days
287	287 days
288	288 days

289	289 days
290	290 days
291	291 days
292	292 days
293	293 days
294	294 days
295	295 days
296	296 days
297	297 days
298	298 days
299	299 days
300	300 days
301	301 days
302	302 days
303	303 days
304	304 days
305	305 days
306	306 days
307	307 days
308	308 days
309	309 days
310	310 days
311	311 days
312	312 days
313	313 days

314	314 days
315	315 days
316	316 days
317	317 days
318	318 days
319	319 days
320	320 days
321	321 days
322	322 days
323	323 days
324	324 days
325	325 days
326	326 days
327	327 days
328	328 days
329	329 days
330	330 days
331	331 days
332	332 days
333	333 days
334	334 days
335	335 days
336	336 days
337	337 days
338	338 days

339	339 days
340	340 days
341	341 days
342	342 days
343	343 days
344	344 days
345	345 days
346	346 days
347	347 days
348	348 days
349	349 days
350	350 days
351	351 days
352	352 days
353	353 days
354	354 days
355	355 days
356	356 days
357	357 days
358	358 days
359	359 days
360	360 days
361	361 days
362	362 days
363	363 days

364	364 days
365	365 days
500	Every day
996	NIU
997	Unknown-refused
998	Unknown-not ascertained
999	Unknown-don't know

Variable: "ALCDAYSYR"

Name:	ALCDAYSYR
Label:	Frequency drank alcohol in past year: Days in past year
Variable Text:	For sample adults who have had at least one drink in their entire life (ALCDRINKEV), ALCDAYSYR reports the number of days per year during the past year that they drank alcoholic beverages. Please refer to the universe tab for more information on the universe for ALCDAYSWK before 2020. ALCDAYSWK draws on information from ALCANYNOE and ALCANYTPE to create a recode of the number of days per week. Beginning in 2020, ALCDAYSYR is collected as part of the rotating core of NHIS alcohol use questions, fielded every other year. For more information about the NHIS alcohol use rotating core, please see ALCDRINKEV. Data Collection and Recoding Interviewers began the section of the survey on alcohol by stating, "These next questions are about drinking alcoholic beverages. Alcohol includes beer, wine, wine coolers, liquors such as vodka, whiskey or rum, mixed drinks or cocktails with alcohol, and any other type of alcoholic drink." After asking whether the person has had at least one drink of any kind of alcohol in their entire life, they then asked, "During the past 12 months, how many days per week, per month or per year did you drink any type of alcoholic beverage?" Respondents could frame an answer using various time units (e.g., "5 times a week," "once a month," "10 times a year"), and the interviewer recorded the number stated and the time period that fit the answer (i.e., "Never/None," "Week," "Month," or "Year"). For the original NHIS public use files, the National Center for Health Statistics recoded these data in terms of days per week (for ALCDAYSWK), in terms of days per month (for ALCDAYSMO), and in terms of days per year (for ALCDAYSYR). The effect of reporting in time periods shorter than a year (i.e., times per week or per month) is evident in digit clustering in the frequencies for ALCDAYSYR. There is strong clustering of responses in codes that are multiples of 12, corresponding to responses given in terms of times per month. Users should, therefore, be wary of treating ALCDAYSYR as a simple interval variable. The Field Representative's Manual for 1997-2018 directed interviewers to "always probe for an exact number. If the sample adult reports a range or interval, assist the sample adult in making an estimate by probing. For example, you might ask, 'Could you give me a more exact number?'" To the same end, in 2000-2018, the survey instrument instructed interviewers, "If necessary, prompt with 'How many days per week, per month, or per year did you drink?'"

Concept:	Alcohol Consumption Variables -- PERSON
Start Position:	103
End Position:	105
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
000	0 days
001	1 day
002	2 days
003	3 days
004	4 days
005	5 days
006	6 days
007	7 days
008	8 days
009	9 days
010	10 days
011	11 days
012	12 days
013	13 days

014	14 days
015	15 days
016	16 days
017	17 days
018	18 days
019	19 days
020	20 days
021	21 days
022	22 days
023	23 days
024	24 days
025	25 days
026	26 days
027	27 days
028	28 days
029	29 days
030	30 days
031	31 days
032	32 days
033	33 days
034	34 days
035	35 days
036	36 days
037	37 days
038	38 days

039	39 days
040	40 days
041	41 days
042	42 days
043	43 days
044	44 days
045	45 days
046	46 days
047	47 days
048	48 days
049	49 days
050	50 days
051	51 days
052	52 days
053	53 days
054	54 days
055	55 days
056	56 days
057	57 days
058	58 days
059	59 days
060	60 days
061	61 days
062	62 days
063	63 days

064	64 days
065	65 days
066	66 days
067	67 days
068	68 days
069	69 days
070	70 days
071	71 days
072	72 days
073	73 days
074	74 days
075	75 days
076	76 days
077	77 days
078	78 days
079	79 days
080	80 days
081	81 days
082	82 days
083	83 days
084	84 days
085	85 days
086	86 days
087	87 days
088	88 days

089	89 days
090	90 days
091	91 days
092	92 days
093	93 days
094	94 days
095	95 days
096	96 days
097	97 days
098	98 days
099	99 days
100	100 days
101	101 days
102	102 days
103	103 days
104	104 days
105	105 days
106	106 days
107	107 days
108	108 days
109	109 days
110	110 days
111	111 days
112	112 days
113	113 days

114	114 days
115	115 days
116	116 days
117	117 days
118	118 days
119	119 days
120	120 days
121	121 days
122	122 days
123	123 days
124	124 days
125	125 days
126	126 days
127	127 days
128	128 days
129	129 days
130	130 days
131	131 days
132	132 days
133	133 days
134	134 days
135	135 days
136	136 days
137	137 days
138	138 days

139	139 days
140	140 days
141	141 days
142	142 days
143	143 days
144	144 days
145	145 days
146	146 days
147	147 days
148	148 days
149	149 days
150	150 days
151	151 days
152	152 days
153	153 days
154	154 days
155	155 days
156	156 days
157	157 days
158	158 days
159	159 days
160	160 days
161	161 days
162	162 days
163	163 days

164	164 days
165	165 days
166	166 days
167	167 days
168	168 days
169	169 days
170	170 days
171	171 days
172	172 days
173	173 days
174	174 days
175	175 days
176	176 days
177	177 days
178	178 days
179	179 days
180	180 days
181	181 days
182	182 days
183	183 days
184	184 days
185	185 days
186	186 days
187	187 days
188	188 days

189	189 days
190	190 days
191	191 days
192	192 days
193	193 days
194	194 days
195	195 days
196	196 days
197	197 days
198	198 days
199	199 days
200	200 days
201	201 days
202	202 days
203	203 days
204	204 days
205	205 days
206	206 days
207	207 days
208	208 days
209	209 days
210	210 days
211	211 days
212	212 days
213	213 days

214	214 days
215	215 days
216	216 days
217	217 days
218	218 days
219	219 days
220	220 days
221	221 days
222	222 days
223	223 days
224	224 days
225	225 days
226	226 days
227	227 days
228	228 days
229	229 days
230	230 days
231	231 days
232	232 days
233	233 days
234	234 days
235	235 days
236	236 days
237	237 days
238	238 days

239	239 days
240	240 days
241	241 days
242	242 days
243	243 days
244	244 days
245	245 days
246	246 days
247	247 days
248	248 days
249	249 days
250	250 days
251	251 days
252	252 days
253	253 days
254	254 days
255	255 days
256	256 days
257	257 days
258	258 days
259	259 days
260	260 days
261	261 days
262	262 days
263	263 days

264	264 days
265	265 days
266	266 days
267	267 days
268	268 days
269	269 days
270	270 days
271	271 days
272	272 days
273	273 days
274	274 days
275	275 days
276	276 days
277	277 days
278	278 days
279	279 days
280	280 days
281	281 days
282	282 days
283	283 days
284	284 days
285	285 days
286	286 days
287	287 days
288	288 days

289	289 days
290	290 days
291	291 days
292	292 days
293	293 days
294	294 day
295	295 days
296	296 days
297	297 days
298	298 days
299	299 days
300	300 days
301	301 days
302	302 days
303	303 days
304	304 days
305	305 days
306	306 days
307	307 days
308	308 days
309	309 days
310	310 days
311	311 days
312	312 days
313	313 days

314	314 days
315	315 days
316	316 days
317	317 days
318	318 days
319	319 days
320	320 days
321	321 days
322	322 days
323	323 days
324	324 days
325	325 days
326	326 days
327	327 days
328	328 days
329	329 days
330	330 days
331	331 days
332	332 days
333	333 days
334	334 days
335	335 days
336	336 days
337	337 days
338	338 days

339	339 days
340	340 days
341	341 days
342	342 days
343	343 days
344	344 days
345	345 days
346	346 days
347	347 days
348	348 days
349	349 days
350	350 days
351	351 days
352	352 days
353	353 days
354	354 days
355	355 days
356	356 days
357	357 days
358	358 days
359	359 days
360	360 days
361	361 days
362	362 days
363	363 days

364	364 days
365	365 days
995	Inconsistent
996	NIU
997	Unknown-refused
998	Unknown-not ascertained
999	Unknown-don't know

Variable: "CIGDAYMO"

Name:	CIGDAYMO
Label:	Number days smoked in past 30 days (some day smokers)
Variable Text:	For sample adults who have ever smoked 100 cigarettes and currently smoke some days (or whose current smoking status us unknown in 1997-2003), CIGDAYMO reports the number of days in the past 30 days respondents smoked one or more cigarettes. Definitions and Instruction The 1997-2000 Field Representative's Manuals stated, "The questions in this section concern cigarette smoking only. Accept whatever the Sample Adult reports, except if it is volunteered that he/she smoked a pipe, cigars of any kind, marijuana, hashish, 'crack', or the like." The Manuals for 2001 forward define a cigarette as "anything the respondent reports except cigars or any kind of marijuana." When asking the question associated with CIGADYMO, the 1997-2000 Manuals instructed interviewers to "[a]lways probe for an exact number" and to assist respondents in making an estimate if they reported a range or interval. The 1997-2000 Manuals also instructed interviewers, "If asked, explain that 'past 30 days' includes up to the day before the interview and not the day of the interview."Related Variables "Some day" smokers who smoked 1+ days or an unknown number of days during the past 30 days (according to CIGDAYMO) were asked the follow-up question, "On the average, when you smoked during the past 30 days, about how many cigarettes did you smoke a day?" Responses to this question are reported in CIGSDAY2. "Every day" smokers (IPUMS NHIS code 3 in SMOKFREQNOW) were asked, "On the average, how many cigarettes do you now smoke a day?" (CIGSDAY1). The variable CIGSDAY combines responses given to the questions associated with CIGSDAY1 and CIGSDAY2, reporting the average number of cigarettes smoked per day for "current smokers" (i.e., all "every day" smokers and "some day" smokers who smoked 1+ days or an unknown number of days in the past 30 days). The universe for CIGSDAY also includes (in 1997-2003) respondents whose current smoking status was unknown (IPUMS NHIS code 7-9 in SMOKFREQNOW). Respondents whose smoking status was unknown were all coded as "Unknown-not ascertained" in CIGSDAY (IPUMS NHIS code 98). CIGSDAY also reports the average number of cigarettes smoked per day for persons age 17+ who were current smokers (1970) and persons age 20+ who were part of a 1/3 subsample and who were current smokers (1976-1977).
Concept:	Smoking Variables -- PERSON

Start Position:	106
End Position:	107
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	None
01	1 day
02	2 days
03	3 days
04	4 days
05	5 days
06	6 days
07	7 days
08	8 days
09	9 days
10	10 days
11	11 days
12	12 days
13	13 days
14	14 days

15	15 days
16	16 days
17	17 days
18	18 days
19	19 days
20	20 days
21	21 days
22	22 days
23	23 days
24	24 days
25	25 days
26	26 days
27	27 days
28	28 days
29	29 days
30	30 days
96	NIU
97	Unknown-refused
98	Unknown-not ascertained
99	Unknown-don't know

Variable: "MOD10DMIN"

Name:	MOD10DMIN
Label:	Duration of moderate activity 10+ minutes: Minutes
Variable Text:	For sample adults who do moderate physical activities (MOD10FTP), MOD10DMIN is a recoded variable created by the staff at the National Center for Health Statistics that draws upon the information in MOD10DNO (Duration of moderate activity 10+ minutes: Number of units) and MOD10DTP (Duration of moderate activity 10+ minutes: Time

	period) and reports the duration of light or moderate leisure-time physical activities in terms of a standardized time unit, minutes. Thus, a respondent who engaged in moderate exercise for an hour and half would have a MOD10DMIN value of 90 (to represent 90 minutes). Light or moderate leisure-time activities are described in the survey questionnaire as activities that "cause only light sweating or a slight to moderate increase in breathing or heart rate." The Field Representative's Manual provides examples of light or moderate activities, although these examples were not routinely shared with respondents. These examples include leisurely walking or bicycling, slow swimming or dancing, and simple gardening. Beginning in 2020, individuals with a duration value (ascertained through combining MOD10DNO and MOD10DTP) equivalent to 720 minutes or 12 hours or more were coded as "996 - Extreme value." Additionally, the 2020 Survey Description notes that "an instrument error in 2020 allowed for entries of more than 24 hours in duration" (85). These entries were recoded to "Extreme value." For general discussion of how and why the NHIS collected information on adult physical activity for 1997 forward, see VIG10FNO.
Concept:	Physical Activity Variables -- PERSON
Start Position:	108
End Position:	110
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesMOD10DMIN is a 3-digit-numeric variable. 000: Not in Universe 996: Extreme value 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "VIG10DMIN"

Name:	VIG10DMIN
Label:	Duration of vigorous activity 10+ minutes: Minutes
Variable Text:	For sample adults who do vigorous physical activities (VIG10FTP) VIG10DMIN is a recoded variable created by the staff at the National Center for Health Statistics that draws upon the information in VIG10DNO (Duration of vigorous activity 10+ minutes: Number of units) and VIG10DTP (Duration of vigorous activity 10+ minutes: Time period) and reports the duration of vigorous leisure-time physical activities in terms of a standardized time unit, minutes. Thus, a respondent who engaged in vigorous exercise for an hour and half would have a VIG10DMIN value of 90 (to represent 90 minutes).

	Vigorous activities are described in the survey questionnaire as activities that "cause heavy sweating or large increases in breathing or heart rate." The Field Representative's Manual provides examples of vigorous activities, although these examples were not routinely shared with respondents. These examples include fast walking, fast bicycling, jogging, strenuous swimming or sports play, vigorous aerobic dance, and strenuous gardening. Beginning in 2020, individuals with a duration value (ascertained through combining VIG10DNO and VIG10DTP) equivalent to 720 minutes or 12 hours or more were coded as "996 - Extreme value." Additionally, the 2020 Survey Description notes that "an instrument error in 2020 allowed for entries of more than 24 hours in duration" (85). These entries were recoded to "Extreme value." For further discussion of data collection on leisure-time physical activity in general for 1997 forward, see VIG10FNO (Frequency of vigorous activity 10+ minutes: Number of units).
Concept:	Physical Activity Variables -- PERSON
Start Position:	111
End Position:	113
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesVIG10DMIN is a 3-digit-numeric variable. 000: Not in Universe 996: Extreme value 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "FRUTNO"

Name:	FRUTNO
Label:	Frequency eating fruit, number of time units
Variable Text:	For sample adults 18 and over, FRUTNO reports how many times the respondent ate fruit in a specified time period. This includes fresh, frozen, or canned fruit, but does not include fruit juice. FRUTNO should be used in conjunction with FRUTTP, which reports the corresponding time units stated by the respondent. FRUTNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.

Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	114
End Position:	116
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesFRUTNO is a 3-digit-numeric variable. 0: None 995: 995+ 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "VEGENO"

Name:	VEGENO
Label:	Frequency eating vegetables, number of time units
Variable Text:	For sample adults 18 and over, VEGENO reports how many times the respondent ate vegetables other than salads, white potatoes, and beans in a specified time period. VEGNO should be used in conjunction with VEGETP, which reports the corresponding time unit stated by the respondent. VEGENO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	117
End Position:	119
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0

Coder Instructions:	CodesVEGENO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 995: 995+ times 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "JUICEMNO"

Name:	JUICEMNO
Label:	Frequency drinking 100 percent fruit juice, past month: Number of units
Variable Text:	For sample adults 18 and over, JUICEMNO reports how many times the respondent consumed juice in a specified time period. This only includes 100% fruit juice such as orange, mango, apple, grape or pineapple; it does not include fruit drinks such as Kool-Aid, lemonade, cranberry juice cocktail, Hi-C or Tang. JUICEMNO should be used in conjunction with JUICEMTP, which reports the corresponding time unit stated by the respondent. JUICEMNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	120
End Position:	122
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesJUICEMNO is a 3-digit-numeric variable. 00: Never 95: 95+ 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "SALADSNO"

Name:	SALADSNO
Label:	Frequency eating green salad, number of time units

Variable Text:	For sample adults 18 and over, SALADSNO reports how many times the respondent ate green leafy or lettuce salad in a specified time period. This includes all tossed salads with lettuce or other greens. SALADSNO should be used in conjunction with SALADSTP, which reports the corresponding time unit stated by the respondent. For information about the portion size, see SALADSIZ. SALADSNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	123
End Position:	125
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesSALADSNO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 95: 95+ times 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "BEANNO"

Name:	BEANNO
Label:	Frequency eating beans: number of time units
Variable Text:	For sample adults 18 and over, BEANNO reports how many times the respondent ate cooked dried beans in a specified time period. This does not include green beans; examples provided of cooked dried beans include baked, pinto, or kidney beans, including beans in soup or chili. BEANNO should be used in conjunction with BEANTP, which reports the corresponding time unit stated by the respondent. For information about the portion size, see BEANSIZ. BEANNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer. For related variables, please use the IPUMS NHIS search function and drop-down menus.
Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	126

End Position:	128
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesBEANNO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "SALSAMNO"

Name:	SALSAMNO
Label:	Frequency eating salsa, past month: Number of units
Variable Text:	For sample adults 18 and over, SALSAMNO reports how many times the respondent ate Mexican-type salsa made with tomatoes in a specified time period. SALSAMNO should be used in conjunction with SALSAMTP, which reports the corresponding time unit (day, week, month, or year) stated by the respondent. SALSAMNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	129
End Position:	131
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesSALSAMNO is a 3-digit-numeric variable. 0: Never 996: Not in Universe 997: Unknown-refused

	998: Unknown-not ascertained 999: Unknown-don't know

Variable: "TOMSAUCEMNO"

Name:	TOMSAUCEMNO
Label:	Frequency eating tomato sauce, past month: Number of units
Variable Text:	For sample adults 18 and over, TOMSAUCEMNO reports how many times the respondent ate tomato sauce in a specified time period. This includes tomato sauce with spaghetti or noodles, or mixed into foods, such as lasagna. Tomato sauce on pizza was not included. TOMSAUCEMNO should be used in conjunction with TOMSAUCEMTP, which reports the corresponding time unit (per day, per week or per month) stated by the respondent. TOMSAUCEMNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Fruits and Vegetables (frequency) Variables -- PERSON
Start Position:	132
End Position:	134
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesTOMSAUCEMNO is a 3-digit-numeric variable. 0: Never 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "SODAPNO"

Name:	SODAPNO
Label:	Frequency drinking soft drinks, number of time units
Variable Text:	For sample adults 18 and over, SODAPNO reports how many times the respondent drank soda or soft drinks in a specified time period. SODAPNO should be used in conjunction with SODAPTP, which reports the corresponding time unit stated by the respondent. For information about the portion size, see SODASIZ. SODAPNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.

Concept:	Sugars and Fats (frequency) Variables -- PERSON
Start Position:	135
End Position:	137
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesSODAPNO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "FRIESPNO"

Name:	FRIESPNO
Label:	Frequency eating fried potatoes, number of time units
Variable Text:	For sample adults 18 and over, FRIESPNO reports how many times the respondent ate fried potatoes (including French fries, home fries, or hash brown potatoes), in a specified time period. FRIESPNO should be used in conjunction with FRIESPTP, which reports the corresponding time unit stated by the respondent. For information about the portion size, see FRIESSIZ. FRIESPNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Sugars and Fats (frequency) Variables -- PERSON
Start Position:	138
End Position:	140
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0

Coder Instructions:	CodesFRIESPNO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 995: 995+ 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "SPORDRMNO"

Name:	SPORDRMNO
Label:	Frequency drinking sports and energy drinks, past month: Number of units
Variable Text:	For sample adults ages 18 and older, SPORDRMNO reports the frequency with which the respondent drank sports and energy drinks, such as Gatorade, Red Bull, and Vitamin Water in the past month. SPORDRMNO should be used in conjunction with SPORDRMTP, which reports the corresponding time unit stated by the respondent. SPORDRMNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Sugars and Fats (frequency) Variables -- PERSON
Start Position:	141
End Position:	143
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesSPORDRMNO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "FRTDRINKMNO"

Name:	FRTDRINKMNO
Label:	Frequency drinking fruit-flavored drinks, past month: Number of units
Variable Text:	For sample adults 18 and over, FRTDRINKMNO reports how many times per month the respondent drank fruit-flavored drinks with sugar (such as Kool-aid, lemonade,

	cranberry cocktail, or fruit drinks they made at home and added sugar to) in a specified time period. FRTDRINKMNO should be used in conjunction with FRTDRINKMTP, which reports the corresponding time unit stated by the respondent. FRTDRINKMNO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Sugars and Fats (frequency) Variables -- PERSON
Start Position:	144
End Position:	146
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesFRTDRINKMNO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "COFETEAMNO"

Name:	COFETEAMNO
Label:	Frequency drinking coffee and tea, past month: Number of units
Variable Text:	For sample adults 18 and over, COFETEAMNO reports how many times the respondent drank coffee or tea sweetened with sugar or honey in a specified time period. Respondents were told to include drinks presweetened or sweetened by respondent, lattes and bottled tea, but not to include drinks artificially sweetened with things like Splenda or Equal. COFETEAMNO should be used in conjunction with COFETEAMTP, which reports the corresponding time unit stated by the respondent. COFETEAMNO is part of a series of variables related to food, food knowledge, and cancer.
Concept:	Other Foods (frequency) Variables -- PERSON
Start Position:	147
End Position:	149
Width:	3

Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesCOFETEAMNO is a 3-digit-numeric variable. 0: Never 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "POTATONO"

Name:	POTATONO
Label:	Frequency eating non-fried potatoes, number of time units
Variable Text:	For sample adults 18 and over, POTATONO reports how many times the respondent ate other potatoes (i.e., not fried) in a specified time period. Other potatoes are described as baked, boiled, or mashed potatoes as well as potato salad. POTATONO should be used in conjunction with POTATOTP, which reports the corresponding time unit stated by the respondent. For information about the portion size, see POTATSIZ. POTATONO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Other Foods (frequency) Variables -- PERSON
Start Position:	150
End Position:	152
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesPOTATONO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 995: 995+ times 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "PIZZANO"

Name:	PIZZANO
Label:	Frequency eating pizza, number of times units
Variable Text:	For sample adults 18 and over, PIZZANO reports how many times the respondent ate pizza in a specified time period. PIZZANO should be used in conjunction with PIZZATP, which reports the corresponding time unit stated by the respondent. For information about the portion size, see PIZSIZ. PIZZANO is part of a series of variables initiated in 1987 related to food, food knowledge, and cancer.
Concept:	Other Foods (frequency) Variables -- PERSON
Start Position:	153
End Position:	155
Width:	3
Variable Format:	numeric
Implied Decimal Places:	0
Coder Instructions:	CodesPIZZANO is a 3-digit-numeric variable. 0: Never or less than 6 times per year 996: Not in Universe 997: Unknown-refused 998: Unknown-not ascertained 999: Unknown-don't know

Variable: "HRSLEEP"

Name:	HRSLEEP
Label:	Usual hours sleep per day
Variable Text:	For sample adults, HRSLEEP reports how many hours, on average, the respondent sleeps per day. For surveys prior to 2004, interviewers were instructed to write down responses verbatim. From 2004 forward, interviewers were instructed to report the hours of sleep in whole numbers, rounding values of 30 minutes or more UP to the nearest hour or otherwise rounding down. IPUMS NHIS recodes the values from the earlier surveys into whole number values, following the same rules as 2004 forward. Beginning in 2020, HRSLEEP is one of a set of rotating core variables on sleep duration and quality collected by the NHIS every other year. In addition to HRSLEEP, other variables in the sleep rotating core include: Sample Adults SLEEPRESTFRQ: How often woke up feeling well rested, past 30 days

	SLEEPFALLFRQ: How often had trouble falling asleep, past 30 days SLEEPSTAYFRQ: How often had trouble staying asleep, past 30 days SLEEPMEDFRQ: How often took medication for sleep, past 30 days Sample Children CSLEEPREST: How often wake up feeling well rested in a typical school week CSLEEPGETUP: How often difficult to get out of bed in the morning in a typical school week CSLEEPTIRED: How often complain of being tired during the day in a typical school week CSLEEPNAPS: How often nap or fall asleep during the day in a typical school week CSLEEPBEDTIME: How often have a regular bedtime on school nights in a typical school week CSLEEPWKTIME: How often wake up at same time in a typical school week Related Variables on Sleep A set of questions on sleep were asked annually of sample adults in 2004-2018 that are similar to those asked beginning in 2020 but differ in the reference period (past week vs. past 30 days) and the structure of the response categories (number of days in past week vs. how often in the past 30 days). These variables include: SLEEPREST: Days woke up feeling rested, past week SLEEPFALL: Number of times having trouble falling asleep, past week SLEEPSTAY: Number of times having trouble staying asleep, past week SLEEPMEDS: Number of times taking medication for sleep, past week
Concept:	Sleep Variables -- PERSON
Start Position:	156
End Position:	157
Width:	2
Variable Format:	numeric
Implied Decimal Places:	0

Categories

Value	Label
00	NIU
01	1 hour (1990: 0 or 1 hour)

02	2 hours
03	3 hours
04	4 hours
05	5 hours
06	6 hours
07	7 hours
08	8 hours
09	9 hours
10	10 hours
11	11 hours
12	12 hours
13	13 hours
14	14 hours
15	15 hours
16	16 hours
17	17 hours
18	18 hours
19	19 hours
20	20 hours
21	21 hours
22	22 hours
23	23 hours
24	24 hours
25	Less than 1 hour
97	Unknown-refused

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

Variable: "CVDSHT"

Name:	CVDSHT
Label:	COVID-19 vaccination
Variable Text:	CVDSHT reports whether the respondent has had a COVID-19 vaccination. NCHS added nearly 30 questions related to COVID starting in quarter 3 of the 2020 NHIS questionnaire, eight of which are asked of both sample adults and sample children. For the full list of variables, as well as information about COVID-related changes to the NHIS, please see the user note.
Concept:	COVID-19 Variables -- PERSON
Start Position:	158
End Position:	158
Width:	1
Variable Format:	numeric
Implied Decimal Places:	0
Categories	

Value	Label
0	NIU
1	No
2	Yes
7	Refused
8	Not Ascertained
9	Don't Know