National Health and Nutrition Examination Survey

August 2021-August 2023 Data Documentation, Codebook, and Frequencies

Demographic Variables and Sample Weights (DEMO_L)

Data File: DEMO_L.xpt

First Published: September 2024

Last Revised: NA

Component Description

After suspension of NHANES field operations in March 2020 due to the COVID-19 pandemic, data collection began for a new 2-year cycle in August 2021. The August 2021-August 2023 nationally representative data are based on an updated sample design and modified interview as well as examination procedures. See the Plan and Operations report of the NHANES August 2021-August 2023 for details on the survey content and procedures for the cycle. For more detail on changes to sample design and analytic guidelines specific to this cycle, please see the "Analytic Notes" section below.

This demographics file provides information on the following topics:

- Survey participant's household interview and examination status;
- · Interview and examination sample weights;
- Masked variance units;
- The six-month time period when the examination was performed;
- Pregnancy status;
- · The ratio of family income to poverty guidelines;
- · Total number of people in the Household;
- · Demographic information about the household reference person; and
- Other selected demographic information, such as gender, age, race/Hispanic origin, education, marital status, military service status, country of birth, and years of U.S. residence.

Eligible Sample

All participants in the NHANES August 2021-August 2023 sample are included in this dataset. The target age group for individual demographic variables in this file varies by the topic. Please review the codebook carefully.

Interview Setting and Mode of Administration

The family and sample person demographics questionnaires were asked by trained interviewers using the Computer-Assisted Personal Interview (CAPI) system in the participant's home or by telephone. The respondent selected the language of interview (English or Spanish) or requested that an interpreter be used. Hand cards, showing response choices or information that survey participants needed to answer the questions, were used for some questions. When

necessary, the interviewer further assisted the respondent by reading the response choices listed on the hand cards.

Participants 16 years and older and emancipated minors were interviewed directly. A proxy provided information for survey participants who were under 16 and for participants who could not answer the questions themselves.

The NHANES August 2021-August 2023 demographics questionnaires are available on the NHANES website.

Quality Assurance & Quality Control

The CAPI system is programmed with built-in consistency checks to reduce data entry errors and online help screens to assist interviewers in defining key terms.

After collection, interview data were reviewed by the NHANES field office staff for accuracy and completeness of selected items. The interviewers were required to audio record interviews and the recorded interviews were reviewed by NCHS staff and interviewer supervisors.

Data Processing and Editing

Frequency counts were checked, "skip" patterns were verified, and the plausibility of question responses was reviewed. Edits were made to some variables to ensure the completeness, consistency, and analytic usefulness of the data. Edits were also made to address data disclosure concerns.

SDDSRVYR: identifies the specific data release cycle. A value of "12" denotes NHANES August 2021-August 2023 data.

RIDSTATR: identifies whether a participant was both interviewed at home or by telephone and examined in the mobile examination center (MEC), or was only interviewed but never participated in the examination.

RIDAGEYR: Age in years at the time of the screening interview for survey participants between the ages of 1 and 79 years of age. Due to disclosure concerns, all responses of participants aged 80 years and older are coded as "80." In NHANES August 2021-August 2023, the weighted mean age for participants 80 years and older is 85 years.

RIDAGEYR was calculated based on the participant's date of birth. In rare cases, if the actual date of birth was missing but the participant's age in years was provided, the reported age was used.

RIDAGEMN: Age in months, at the time of the screening interview, is provided for participants who were less than 25 months of age at the time of examination (RIDEXAGM < 25). If the exact date of birth was not provided by the respondent, the age in months was calculated based on the imputed age in years at the time of the screening interview.

RIDEXAGM: Age in months, at the time of examination, for participants who were less than 240 months of age at the time of examination (RIDEXAGM < 240).

RIDRETH3: race-ethnicity categories used since the 2011-2012 survey cycle to accommodate the oversample of Asian Americans. It was derived from responses to the survey questions on race and Hispanic origin. Respondents who self-identified as "Mexican American" were coded as such (i.e., RIDRETH3=1) regardless of their other race-ethnicity identities. Otherwise,

participants who self-identified as "Hispanic" were coded as "2, Other Hispanic." Non-Hispanic participants were categorized based on their self-reported races: non-Hispanic White (RIDRETH3=3), non-Hispanic Black (RIDRETH3=4), non-Hispanic Asian (RIDRETH3=6), and other non-Hispanic races, including non-Hispanic multiracial (RIDRETH3=7). Code "5" was not used in RIDRETH3.

RIDRETH1: race-ethnicity variable that can be linked to the previous NHANES race-ethnicity variable in 1999-2010. Non-Hispanic Asian participants are grouped with other non-Hispanic races in code "5" (other non-Hispanic race including non-Hispanic multiracial) in RIDRETH1. Codes "6" and "7" were not used in RIDRETH1. Coding procedures for other categories in RIDRETH1 were compatible to RIDRETH3.

RIDEXMON: indicates the six-month time period when the examination was performed.

DMQMILIZ: provides information on whether the participant has ever served on active duty in the U.S. Armed Forces, military Reserves, or National Guard. Active duty does not include training for the Reserves or National Guard, but does include activation, for service in the U.S. or in a foreign country, in support of military or humanitarian operations.

This variable has been included in NHANES since 2011-2012. Prior to 2011, the information was collected using a question with different wording that asked if the participant had served in the U.S. Armed Forces, and released in variable DMQMILIT.

DMDBORN4: To address disclosure risk concerns, starting in 2011, country of birth was recoded into two categories: 1) Born in 50 U.S. states or Washington, DC; and 2) Born in other countries, including U.S. territories.

DMDYRUSR: the number of years the participant has lived in the United States. Participants born outside the U.S. were asked the month and year when they came to live in the U.S. (DMQ.160). Responses were used to calculate the length of U.S. residency and recoded into 6 categories: 1 = Less than 1 year, 2 = 1 to 4 years, 3 = 5 to 9 years, 4 = 10 to 14 years, 5 = 15 to 19 years, and 6 = 20 year or more. For participants who did not report the month of their arrival, a value of 7 (July) was used.

DMDEDUC2: the highest grade or level of education completed by adults 20 years and older.

DMDMARTZ: The marital status question was asked of people 14 years of age and older. Due to disclosure risk, marital status is only released for people 20 years of age and older, and recoded from the original 6 categories to 3 categories (1 = Married/Living with partner; 2 = Widowed/Divorced/Separated; 3 = Never married).

RIDEXPRG: Pregnancy status at the time of the health examination was ascertained for females 8–59 years of age. Due to disclosure risk, pregnancy status is released only for women 20-44 years of age. RIDEXPRG values are based on self-reported pregnancy status and urine pregnancy test results. People who reported they were pregnant at the time of exam were assumed to be pregnant (RIDEXPRG=1). Those who reported they were not pregnant or did not know their pregnancy status were further classified based on the results of the urine pregnancy test. If the respondent reported "no" or "don't know" and the urine test result was positive, the respondent was coded as pregnant (RIDEXPRG=1). If the respondent reported "no" and the urine test was negative, the respondent reported did not know her pregnancy status and the urine test was negative, the respondent was coded "could not be determined" (RIDEXPRG=3). People who were interviewed, but not examined also have an RIDEXPRG value = 3 (could not be determined).

DMDHHSIZ: the number of people in the participant's household, ranging from 1 to 7. Due to disclosure concerns, individuals living in households with 7 or more people are included in the

category "7 or more".

2/3/25. 12:23 PM

INDFMPIR: the ratio of family income to poverty level. The Department of Health and Human Services (HHS) poverty guidelines were used as the poverty measure to calculate this ratio. These guidelines are issued annually in the Federal Register and are used to determine financial eligibility for certain federal programs, such as Head Start, Supplemental Nutrition Assistance Program (SNAP), Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and the National School Lunch Program. The poverty guidelines vary by family size and geographic location.

INDFMPIR was calculated by dividing total annual family (or individual) income by the poverty guidelines specific to the survey year.

During the household interview, the respondent was asked to report total income for the entire family in the last calendar year in dollars. A family is defined as a group of two people or more related by birth, marriage, or adoption and residing together. Annual individual income was asked for households with one person, or households comprised of unrelated individuals.

If the respondent was not willing or able to provide an exact dollar figure, the interviewer asked an additional question to determine whether the income was < \$20,000 or $\ge $20,000$. Based on the respondent's answer to this question, he/she was asked to select a category of income from a hand card. For respondents who selected an income category, their family incomes were set as the midpoint of the selected range. INDFMPIR was not computed if the respondent only reported income as < \$20,000 or $\ge $20,000$ without additional details. INDFMPIR values at or above 5.00 were coded as 5.00 or more because of disclosure concerns. Values were not computed if income data were missing.

Household Reference Person: The household reference person is defined as the first household member 18 years of age or older listed on the household member roster who owns or rents the household residence. The household reference person is comparable to "family reference person" in NHANES programs prior to 1999. Household reference person data have often used in analysis as indicators characterizing the socioeconomic status for youth. This public use demographics file contains household reference person information for participants aged 0 to 19 years, including data on gender (DMDHRGND), age (DMDHRAGZ), education level (DMDHREDZ), and marital status (DMDHRMAZ), as well as the education level of the household reference person's spouse (DMDHSEDZ). Household reference person information for adult participants 20 years and older is available through NCHS Research Data Center.

Due to disclosure concerns, the household reference person's age was recoded into a categorical variable (<20, 20-39, 40-59, or 60+ years); the household reference person's education level and their spouse's education level variables were recoded from the original 5 categories to 3 categories (1 = Less than high school; 2 = High school grad/GED or some college/AA degree; 3 = College graduate or above); and the household reference person's marital status variable was recoded from the original 6 categories to 3 categories (1 = Married/Living with partner; 2 = Widowed/Divorced/Separated; 3 = Never married).

Analytic Notes

The August 2021–August 2023 NHANES sample design is similar to past cycles that drew a multi-year, stratified, clustered four-stage sample of the U.S. civilian noninstitutionalized population to provide national health estimates. However, notable design changes were made to adapt to the pandemic environment, including no person-level oversampling by race and Hispanic origin and income. As a result, the number of participants for certain demographic subgroups are noticeably lower than in previous cycles, and users should expect reduced statistical precision for these groups. Additionally, person-level oversampling based on age group was added. The potential increase in clustering by household was investigated with the finding that the mean number of participants from each household was not substantially different from previous cycles.

Although the August 2021-August 2023 cycle represents a two-year data collection period, there is a 15-month gap between it and the previous dataset (2017-March 2020). This gap in data collection presents an important analytic consideration given it represents the COVID-19 pandemic during which some health behaviors and outcomes were likely affected by disruptions that occurred in healthcare delivery, employment, and education. Data users are urged to proceed with caution before combining August 2021-August 2023 with previous cycles to calculate cross sectional health estimates or conducting trend analysis, given the 15-month period with unobserved data may include unusual patterns.

Age at screening: used to determine eligibility for an examination component and should be used for most analyses.

DMDMARTZ: Marital status is only released for people 20 years of age and older to reduce disclosure risk. Prior to 2007, marital status was released for participants aged 14 and older. In NHANES August 2021-August 2023, the percentage of people aged 14-19 who are married or living with a partner is less than 2%.

RIDEXPRG: To reduce disclosure risk, pregnancy status is released only for women aged 20-44 years. The percentage of pregnant women/girls aged 8-19 or 45-59 years is less than 1% in NHANES August 2021-August 2023.

Masked Variance Units (MVUs): Fifteen MVUs and 30 masked primary sampling units (PSUs) are included in the NHANES August 2021-August 2023. Each stratum has 2 PSUs. These MVUs are a collection of secondary sampling units that are aggregated into groups for the purpose of variance estimation. Variance estimates produced using the MVUs closely approximate variances that would have been estimated using the "true" sample design variance units. MVUs are used to protect confidentiality and to reduce disclosure risk and are described in NHANES Analytic Guidelines. Analysts should review the Guidelines carefully prior to analyzing survey data.

Sample Weights: The 2-year sample weights (WTINT2YR, WTMEC2YR, depending on the analysis) should be used for all NHANES August 2021-August 2023 analyses.

Please refer to the NHANES Analytic Guidelines and the on-line NHANES Tutorial for further details on the use of sample weights and other analytic issues.

Disclosure risks and issues pertaining to confidentiality protection prevent NCHS from releasing all the NHANES demographic variables publicly. Additional information and instructions for requesting use of these data are available through the NCHS RDC.

References

- Terry AL, Chiappa MM, McAllister J, Woodwell DA, Graber JE. Plan and operations of the National Health and Nutrition Examination Survey, August 2021–August 2023. National Center for Health Statistics. Vital Health Stat 1(66). 2024. DOI: https://stacks.cdc.gov/view/cdc/151927
- U.S. Department of Health & Human Services. Poverty Guidelines, Research, and Measurement. Washington, DC: U.S. Department of Health & Human Services. https://aspe.hhs.gov/topics/poverty-economic-mobility/poverty-guidelines/prior-hhs-poverty-guidelines-federal-register-references

Codebook and Frequencies

SEQN - Respondent sequence number

Variable Name: SEQN

SAS Label: Respondent sequence number

English Text: Respondent sequence number.

SDDSRVYR - Data release cycle

Variable Name: SDDSRVYR

SAS Label: Data release cycle

English Text: Data release cycle

Code or Value	Value Description	Count	Cumulative	Skip to Item
12	NHANES August 2021-August 2023 public release	11933	11933	
	Missing	0	11933	

RIDSTATR - Interview/Examination status

Variable Name: RIDSTATR

SAS Label: Interview/Examination status

English Text: Interview and examination status of the participant.

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Interviewed only	3073	3073	
2	Both interviewed and MEC examined	8860	11933	
	Missing	0	11933	

RIAGENDR - Gender

Variable Name: RIAGENDR

SAS Label: Gender

English Text: Gender of the participant.

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Male	5575	5575	
2	Female	6358	11933	
	Missing	0	11933	

RIDAGEYR - Age in years at screening

Variable Name: RIDAGEYR

SAS Label: Age in years at screening

English Text: Age in years of the participant at the time of screening. Individuals

80 and over are top-coded at 80 years of age.

Code or Value	Value Description	Count	Cumulative	Skip to Item
0 to 79	Range of Values	11408	11408	
80	80 years of age and over	525	11933	
	Missing	0	11933	

RIDAGEMN - Age in months at screening - 0 to 24 mos

Variable Name: RIDAGEMN

SAS Label: Age in months at screening - 0 to 24 mos

English Text: Age in months of the participant at the time of screening. Reported

for persons aged 24 months or younger at the time of exam (or

screening if not examined).

Code or Value	Value Description	Count	Cumulative	Skip to Item
0 to 24	Range of Values	377	377	
	Missing	11556	11933	

RIDRETH1 - Race/Hispanic origin

Variable Name: RIDRETH1

SAS Label: Race/Hispanic origin

English Text: Recode of reported race and Hispanic origin information

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Mexican American	1117	1117	
2	Other Hispanic	1373	2490	
3	Non-Hispanic White	6217	8707	
4	Non-Hispanic Black	1597	10304	
5	Other Race - Including Multi-Racial	1629	11933	
	Missing	0	11933	

RIDRETH3 - Race/Hispanic origin w/ NH Asian

Variable Name: RIDRETH3

SAS Label: Race/Hispanic origin w/ NH Asian

English Text: Recode of reported race and Hispanic origin information, with Non-

Hispanic Asian Category

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Mexican American	1117	1117	
2	Other Hispanic	1373	2490	
3	Non-Hispanic White	6217	8707	
4	Non-Hispanic Black	1597	10304	
6	Non-Hispanic Asian	681	10985	
7	Other Race - Including Multi-Racial	948	11933	
	Missing	0	11933	

RIDEXMON - Six-month time period

Variable Name: RIDEXMON

SAS Label: Six-month time period

English Text: Six month time period when the examination was performed - two

categories: November 1 through April 30, May 1 through October 31.

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	November 1 through April 30	4251	4251	
2	May 1 through October 31	4609	8860	
	Missing	3073	11933	

RIDEXAGM - Age in months at exam - 0 to 19 years

Variable Name: RIDEXAGM

SAS Label: Age in months at exam - 0 to 19 years

English Text: Age in months of the participant at the time of examination.

Reported for persons aged 19 years or younger at the time of

examination.

Code or Value	Value Description	Count	Cumulative	Skip to Item
0 to 239	Range of Values	2787	2787	
	Missing	9146	11933	

DMQMILIZ - Served active duty in US Armed Forces

Variable Name: DMQMILIZ

SAS Label: Served active duty in US Armed Forces

English Text: {Have you/Has SP} ever served on active duty in the U.S. Armed

Forces, military Reserves, or National Guard? (Active duty does not include training for the Reserves or National Guard, but does include activation, for service in the U.S. or in a foreign country, in support

of military or humanitarian operations.)

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Yes	694	694	
2	No	7606	8300	
7	Refused	1	8301	
9	Don't know	0	8301	
	Missing	3632	11933	

DMDBORN4 - Country of birth

Variable Name: DMDBORN4

SAS Label: Country of birth

English Text: Country of birth

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Born in 50 US states or Washington,	10039	10039	DMDEDUC2
2	Others	1875	11914	
77	Refused	0	11914	DMDEDUC2
99	Don't know	0	11914	DMDEDUC2
	Missing	19	11933	

DMDYRUSR - Length of time in US

Variable Name: DMDYRUSR

SAS Label: Length of time in US

English Text: Length of time the participant has been in the US.

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Less than 1 year	101	101	
2	1 to 4 years	260	361	
3	5 to 9 years	225	586	
4	10 to 14 years	160	746	
5	15 to 19 years	158	904	
6	20 years or more	912	1816	
77	Refused	13	1829	
99	Don't know	46	1875	
	Missing	10058	11933	

DMDEDUC2 - Education level - Adults 20+

Variable Name: DMDEDUC2

SAS Label: Education level - Adults 20+

English Text: What is the highest grade or level of school {you have/SP has}

completed or the highest degree {you have/s/he has} received?

English Instructions: HAND CARD DMQ1 READ HAND CARD CATEGORIES IF NECESSARY

ENTER HIGHEST LEVEL OF SCHOOL

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Less than 9th grade	373	373	
2	9-11th grade (Includes 12th grade with no diploma)	666	1039	
3	High school graduate/GED or equivalent	1749	2788	
4	Some college or AA degree	2370	5158	
5	College graduate or above	2625	7783	
7	Refused	0	7783	
9	Don't know	11	7794	
	Missing	4139	11933	

DMDMARTZ - Marital status

Variable Name: DMDMARTZ

SAS Label: Marital status

English Text: Marital status

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Married/Living with partner	4136	4136	
2	Widowed/Divorced/Separated	2022	6158	
3	Never married	1625	7783	
77	Refused	4	7787	
99	Don't know	5	7792	
	Missing	4141	11933	

RIDEXPRG - Pregnancy status at exam

Variable Name: RIDEXPRG

SAS Label: Pregnancy status at exam

English Text: Pregnancy status for females between 20 and 44 years of age at the

time of MEC exam.

Target: Females only 20 YEARS - 44 YEARS

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Yes, positive lab pregnancy test or self- reported pregnant at exam	41	41	
2	The participant was not pregnant at exam	1065	1106	
3	Cannot ascertain if the participant is pregnant at exam	397	1503	
	Missing	10430	11933	

DMDHHSIZ - Total number of people in the Household

Variable Name: DMDHHSIZ

SAS Label: Total number of people in the Household

English Text: Total number of people in the Household

Code or Value	Value Description	Count	Cumulative	Skip to Item
1 to 6	Range of Values	11309	11309	
7	7 or more people in the Household	624	11933	
	Missing	0	11933	

DMDHRGND - HH ref person's gender

Variable Name: DMDHRGND

SAS Label: HH ref person's gender

English Text: HH ref person's gender

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Male	1793	1793	
2	Female	2322	4115	
	Missing	7818	11933	

DMDHRAGZ - HH ref person's age in years

Variable Name: DMDHRAGZ

SAS Label: HH ref person's age in years

English Text: HH ref person's age in years

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	<20 years	83	83	
2	20-39 years	1985	2068	
3	40-59 years	1804	3872	
4	60+ years	252	4124	
	Missing	7809	11933	

DMDHREDZ - HH ref person's education level

Variable Name: DMDHREDZ

SAS Label: HH ref person's education level

English Text: HH ref person's education level

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Less than high school degree	550	550	
2	High school grad/GED or some college/AA degree	2004	2554	
3	College graduate or above	1192	3746	
	Missing	8187	11933	

DMDHRMAZ - HH ref person's marital status

Variable Name: DMDHRMAZ

SAS Label: HH ref person's marital status

English Text: HH ref person's marital status

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Married/Living with partner	2947	2947	
2	Widowed/Divorced/Separated	614	3561	
3	Never Married	459	4020	
	Missing	7913	11933	

DMDHSEDZ - HH ref person's spouse's education level

Variable Name: DMDHSEDZ

SAS Label: HH ref person's spouse's education level

English Text: HH ref person's spouse's education level

Code or Value	Value Description	Count	Cumulative	Skip to Item
1	Less than high school degree	300	300	
2	High school grad/GED or some college/AA degree	942	1242	
3	College graduate or above	885	2127	
	Missing	9806	11933	

WTINT2YR - Full sample 2-year interview weight

Variable Name: WTINT2YR

SAS Label: Full sample 2-year interview weight

English Text: Full sample 2 year interview weight.

Code or Value	Value Description	Count	Cumulative	Skip to Item
4584.463196 to 170968.343177	Range of Values	11933	11933	
	Missing	0	11933	

WTMEC2YR - Full sample 2-year MEC exam weight

Variable Name: WTMEC2YR

SAS Label: Full sample 2-year MEC exam weight

English Text: Full sample 2 year MEC exam weight.

Code or Value	Value Description	Count	Cumulative	Skip to Item
4581.595095 to 227108.296958	Range of Values	8860	8860	
0	Not MEC Examined	3073	11933	
	Missing	0	11933	

SDMVSTRA - Masked variance pseudo-stratum

Variable Name: SDMVSTRA

SAS Label: Masked variance pseudo-stratum

English Text: Masked variance unit pseudo-stratum variable for variance

estimation

Code or Value	Value Description	Count	Cumulative	Skip to Item
173 to 187	Range of Values	11933	11933	
	Missing	0	11933	

SDMVPSU - Masked variance pseudo-PSU

Variable Name: SDMVPSU

SAS Label: Masked variance pseudo-PSU

English Text: Masked variance unit pseudo-PSU variable for variance estimation

Code or Value	Value Description	Count	Cumulative	Skip to Item
1 to 2	Range of Values	11933	11933	
	Missing	0	11933	

INDFMPIR - Ratio of family income to poverty

Variable Name: INDFMPIR

SAS Label: Ratio of family income to poverty

English Text: Ratio of family income to poverty

Code or Value	Value Description	Count	Cumulative	Skip to Item
0 to 4.99	Range of Values	7743	7743	
5	Value greater than or equal to 5.00	2149	9892	
	Missing	2041	11933	