



ELIZABETH ROTH, MEGAN S. SCHULER

National Survey on Drug Use and Health (NSDUH)

Veterans Data Hub Codebook

For more information on this publication, visit www.rand.org/t/TLA1363-2.

About RAND

RAND is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest. To learn more about RAND, visit www.rand.org.

Research Integrity

Our mission to help improve policy and decisionmaking through research and analysis is enabled through our core values of quality and objectivity and our unwavering commitment to the highest level of integrity and ethical behavior. To help ensure our research and analysis are rigorous, objective, and nonpartisan, we subject our research publications to a robust and exacting quality-assurance process; avoid both the appearance and reality of financial and other conflicts of interest through staff training, project screening, and a policy of mandatory disclosure; and pursue transparency in our research engagements through our commitment to the open publication of our research findings and recommendations, disclosure of the source of funding of published research, and policies to ensure intellectual independence. For more information, visit www.rand.org/about/research-integrity.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Published by the RAND Corporation, Santa Monica, Calif.

© 2024 RAND Corporation

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

This publication and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited; linking directly to its webpage on rand.org is encouraged. Permission is required from RAND to reproduce, or reuse in another form, any of its research products for commercial purposes. For information on reprint and reuse permissions, please visit www.rand.org/pubs/permissions.

About This Tool

This codebook was compiled as part of the RAND Veterans Policy Research Institute’s effort to develop a repository of datasets that assess veteran status. This codebook for the National Survey on Drug Use and Health (NSDUH) presents an overview of the data, how to access them, which variable(s) assess veteran status, other veteran-related questions, and additional resources for using the NSDUH. In this codebook, we cite directly from the original documentation of the NSDUH.

RAND Epstein Family Veterans Policy Research Institute

The RAND Epstein Family Veterans Policy Research Institute is dedicated to conducting innovative, evidence-based research and analysis to improve the lives of those who have served in the U.S. military. Building on decades of interdisciplinary expertise at RAND, the institute prioritizes creative, equitable, and inclusive solutions and interventions that meet the needs of diverse veteran populations while engaging and empowering those who support them.

For more information about the RAND Epstein Family Veterans Policy Research Institute, visit veterans.rand.org. Questions about this report or about the RAND Epstein Family Veterans Policy Institute should be directed to veteranspolicy@rand.org.

Funding

Funding for this publication was made possible by a generous gift from Daniel J. Epstein through the Epstein Family Foundation, which established the RAND Epstein Family Veterans Policy Research Institute in 2021.

Acknowledgments

The authors would like to thank Charles Goldman and Jeffrey Wenger of RAND for their peer reviews of this report. All errors or omissions remain the responsibility of the authors.

Contents

About This Tool	iii
Tables	vii
National Survey on Drug Use and Health (NSDUH): Veterans Data Hub Codebook	1
Overview to the National Survey on Drug Use and Health	1
Data Access	1
Key Variables	3
Additional Resources	9
Abbreviations	11
References	13

Tables

1.	Veteran Sample Size per Survey Wave, Relative to Total Sample Size.....	2
2.	Respondent Identifiers.....	3
3.	Panel Information.....	3
4.	Weights.....	4
5.	Veteran Status.....	4
6.	Sex and Gender.....	4
7.	Age.....	5
8.	Race and Ethnicity.....	5
9.	Geographic Locations.....	6
10.	Employment Status.....	6
11.	Education Status.....	6
12.	Income.....	7
13.	Health Insurance Coverage.....	8
14.	Period of Service.....	8
15.	U.S. Department of Veterans Affairs Care Recipient.....	8
16.	Other Relevant Variables.....	8

National Survey on Drug Use and Health (NSDUH): Veterans Data Hub Codebook

Overview to the National Survey on Drug Use and Health

The National Survey on Drug Use and Health (NSDUH) is an annual nationally representative sample survey of civilian, noninstitutionalized U.S. residents ages 12 and older that asks questions on substance use, substance use disorders, mental health, and treatment. The sample includes individuals living in households, sheltered people experiencing homelessness, and civilians living on military bases; it excludes active military members, people living in nursing homes or treatment centers, people in prison, and people experiencing homelessness who are not living in a shelter.

The NSDUH collects data on about 67,000 people annually with face-to-face field interviews complemented by computer-assisted self-interviewing (CASI) for sensitive questions. The sample is completely refreshed annually, so NSDUH surveys are cross-sectional and can be combined over time to increase sample sizes.

Toward the end of the survey, as part of the demographics data collection, all respondents ages 18 and older are asked whether they have ever been in the United States Armed Forces. Those who have are then asked whether they are currently on active duty (these respondents are released from the survey), in a reserve component, or “now separated or retired from the military.” Table 1 shows the total sample size of veterans in each wave of this dataset from 2001 to 2022, relative to the total sample size (using the definition of *veteran* provided in the “Demographic Characteristics” section).

Data Access

Public-Use Files

Publicly released NSDUH data can be downloaded from the Substance Abuse and Mental Health Services Administration (SAMHSA) website (SAMHSA, undated-b). NSDUH provides public datasets as year-specific files and in harmonized datasets that combine measures over time.

Restricted-Use Files

The U.S. Centers for Disease Control and Prevention’s (CDC’s) National Center for Health Statistics (NCHS) provides access to restricted NSDUH measures, including finer detail on geographic identifiers, continuous age, finer income detail, and other sensitive demographic variables, via Federal Statistical Research Data Centers (RDCs) (CDC, 2022). To obtain access

TABLE 1
Veteran Sample Size per Survey Wave, Relative to Total Sample Size

Year	Sample Size (Total)		Sample Size (Veteran)	
	Unweighted	Weighted	Unweighted	Weighted
2022	47,100	256,281,676	2,338	19,977,811
2021	47,291	253,824,662	2,501	19,226,679
2020	27,170	251,929,273	1,476	20,301,423
2019	42,739	250,316,210	2,184	20,173,652
2018	43,026	248,857,430	2,395	21,429,092
2017	42,554	247,160,541	2,375	21,085,263
2016	42,625	244,533,608	2,432	22,103,467
2015	43,561	242,801,072	2,489	21,429,107
2014	41,671	240,248,111	2,587	23,222,895
2013	37,424	237,498,836	1,887	23,012,506
2012	37,869	235,124,274	2,039	24,597,270
2011	39,133	232,625,299	2,128	23,401,841
2010	38,919	229,272,579	2,007	22,929,688
2009	37,707	227,206,546	2,049	23,799,078
2008	37,504	224,922,762	2,068	25,115,635
2007	37,708	222,604,121	2,166	24,229,965
2006	36,801	220,629,724	2,327	25,403,753
2005	37,227	217,865,412	2,275	26,217,073
2004	37,308	215,300,425	2,316	25,785,221
2003	37,026	212,686,651	2,372	27,094,201
2002	36,370	210,389,660	2,432	26,852,775
2001	38,132	202,035,837	2,700	26,274,594

to restricted data, one must submit a proposal to NCHS. Anyone working with these files must apply for Special Sworn Status and physically work inside an RDC facility.

Quick Tables/Data Downloads

SAMHSA provides access to many measures from both the public-use and the restricted-use datasets via its public data analysis system (PDAS) and restricted data analysis system (RDAS), respectively (SAMHSA, undated-a). These tools facilitate online analysis, explora-

tion, and download of many NSDUH measures. Users can select up to three classification variables by which to view measures, and, in many cases, this can be an effective avenue to access restricted data without having to use an RDC.

Key Variables

The NSDUH survey has changed over time, including changes in how the survey is administered (verbal versus computer-assisted), how certain questions are phrased, and how certain derived variables are computed. When a variable changes, NSDUH changes the variable name (e.g., by adding or incrementing a numerical suffix). It is imperative that, when combining files from different survey years, the analyst fully understand any changes in variables over time.

In many cases, NSDUH releases a variable's raw value plus a recoded and/or imputed variable. Recoded variables can be derived from source variables, and imputed variables employ statistical imputation to replace missing data with valid values. The names of imputed variables begin with "IR." Users are encouraged to use imputed or recoded variables when they exist.

Key variables for IDs and weights, demographic characteristics, economic characteristics, and veteran characteristics are presented in Table 2 through Table 16. Please note that the variable descriptions presented in these tables use 2019 variable names, and all variable definitions are sourced from the *2019 National Survey on Drug Use and Health Public Use File Codebook*, unless otherwise noted (Center for Behavioral Health Statistics and Quality, 2020).

IDs and Weights

NSDUH's individual respondent ID variable was named QUESTID in survey years up to 2003 (inclusive) and again in 2005; the individual respondent ID variable was named QUESTID2 in 2004 and from 2006 to present. As noted in Table 2, there is variation across survey years as to whether the QUESTID2 variable is coded as a character string or numeric value.

TABLE 2
Respondent Identifiers

Variable Name	Description
QUESTID	Individual respondent identifier (numeric)
QUESTID2	QUESTID: prior to 2003, and 2005 QUESTID2 (numeric): 2007–2013 QUESTID2 (character): 2004, 2006, and 2014–2019

TABLE 3
Panel Information

Variable Name	Description
YEAR	Year of interview

NSDUH provides sample weights with important documentation about their derivation and use in the codebook of each survey wave; anybody using NSDUH data should read this documentation carefully. Use ANALWT_C to generate year-specific weighted estimates. When generating a pooled estimate across multiple years, adjust the weight by dividing the final weight by the number of years of combined data. It is necessary to account for NSDUH's stratum (VESTR) and cluster replicate (VEREP) variables using the survey packages of standard statistical software to obtain correct variance estimates and inferential statistics (Table 4).

TABLE 4
Weights

Variable Name	Description	Values
ANALWT_C	Final person-level sample weight	
VESTR	Variance estimation stratum	20001–40050, depending on the survey year
VEREP	Variance estimation cluster replicates	01–02

Demographic Characteristics

Respondents who answer “yes” to ACTDEVER (ever on active duty) AND “now separated/retired from reserves/active duty” to MILSTAT (shown in bold in Table 5) should be classified as veterans.

TABLE 5
Veteran Status

Variable Name	Description	Values
SERVICE	Ever been in the U.S. armed forces	1 = Yes 2 = No
ACTDEVER	Ever on active duty in U.S. military/reserves; available from 2013 to present	1 = Yes 2 = No
MILSTAT	Current military status (current respondents on active duty are explicitly excluded from the sample, so there is no “1 = Active Duty” code for this variable)	2 = In a reserve component 3 = Now separated/retired from reserves/active duty
COMBATPY	Ever in combat zone on active duty	1 = Yes 2 = No

TABLE 6
Sex and Gender

Variable Name	Description	Values
IRSEX	Imputation revised gender (interviewers record respondents' gender)	1 = Male 2 = Female

TABLE 7**Age**

Variable Name	Description	Values
AGE2	Recoded final edited age	1 = Respondent is 12 years old 2 = Respondent is 13 years old 3 = Respondent is 14 years old 4 = Respondent is 15 years old 5 = Respondent is 16 years old 6 = Respondent is 17 years old 7 = Respondent is 18 years old 8 = Respondent is 19 years old 9 = Respondent is 20 years old 10 = Respondent is 21 years old 11 = Respondent is 22 or 23 years old 12 = Respondent is 24 or 25 years old 13 = Respondent is between 26 and 29 years old 14 = Respondent is between 30 and 34 years old 15 = Respondent is between 35 and 49 years old 16 = Respondent is between 50 and 64 years old 17 = Respondent is 65 years old or older

TABLE 8**Race and Ethnicity**

Variable Name	Description	Values
NEWRACE2	Race/Ethnicity	1 = NonHisp White 2 = NonHisp Black/Afr Am 3 = NonHisp Native Am/AK Native 4 = NonHisp Native HI/Other Pac Isl 5 = NonHisp Asian 6 = NonHisp more than one race 7 = Hispanic

NOTE: Afr Am = African American; AK = Alaskan; HI = Hawaiian; Native Am = Native American; NonHisp = non-Hispanic; Pac Isl = Pacific Islander.

To protect respondent privacy, NSDUH's public release files include very limited detail on geography. Detailed Region, State, County, Segment, and Block variables at the individual level can be requested with the restricted NSDUH data. A subset of summary tables can be generated at State and County levels from NSDUH's online PDAS system. Table 9 lists the only geographic elements released with the NSDUH public-release data.

TABLE 9
Geographic Locations

Variable Name	Description	Values
PDEN90 PDEN00 PDEN10	Population density (the different measures are based on different census years and are not fully comparable)	1 = Segment in a CBSA with 1 million or more persons 2 = Segment in a CBSA with <1 million persons 3 = Segment not in a CBSA
COUTYP2 COUTYP4	County metro/nonmetro status COUTYP2: 2004–2015 COUTYP4: 2016–present COUTYP2 and COUTYP4 are based on different rural/urban continuum codes and are not fully comparable.	1 = Large Metro 2 = Small Metro 3 = Nonmetro

NOTE: CBSA = core-based statistical area.

Economic Characteristics

TABLE 10
Employment Status

Variable Name	Description	Values
IRWRKSTAT18	Employment status for ages 18+	1 = Employed full time 2 = Employed part time 3 = Unemployed 4 = Other (incl. not in labor force)

EDUCCAT2 and EDUHIGHCAT are almost identical, except that EDUCCAT2 (used 2001–2014) had a response option for “some college,” and EDUHIGHCAT (used 2015 to present) specifies “some college/associate’s degree” (Table 11).

TABLE 11
Education Status

Variable Name	Description	Values
EDUCCAT2 EDUHIGHCAT	Education categories EDUCCAT2: 2001–2014 EDUHIGHCAT: 2015–present	1 = Less high school 2 = High school grad 3 = Some college (2001–2014) 3 = Some coll/Assoc Dg (2015–present) 4 = College graduate 5 = 12 to 17 year olds
COLLENR2 COLLENRST COLLENRLST	College enrollment. See notes in NSDUH codebook for details. COLLENR2: 2002–2014 COLLENRST: 2015 COLLENRLST: 2016–present	1 = Full Time College Student Aged 18–22 2 = Part Time College Student Aged 18–22 3 = Not Enrolled Aged 18–22 4 = Other Persons Aged 18–22 5 = Persons Aged 12–17, 23 or Older

In addition to measures for respondent total income, total family income, and poverty level, the NSDUH also includes indicators for receipt of public benefits—such as social security, Supplemental Security Income, and food stamps—and number of months on welfare (Table 12).

TABLE 12
Income

Variable Name	Description	Values
IRPINC3	Respondent total income	1 = Less than \$10,000 (Including Loss) 2 = \$10,000–\$19,999 3 = \$20,000–\$29,999 4 = \$30,000–\$39,999 5 = \$40,000–\$49,999 6 = \$50,000–\$74,999 7 = \$75,000 or more
IRFAMIN3	Total family income	1 = Less than \$10,000 (Including Loss) 2 = \$10,000–\$19,999 3 = \$20,000–\$29,999 4 = \$30,000–\$39,999 5 = \$40,000–\$49,999 6 = \$50,000–\$74,999 7 = \$75,000 or more
INCOME	Total family income recode	1 = Less than \$20,000 2 = \$20,000–\$49,999 3 = \$50,000–\$74,999 4 = \$75,000 or more
GOVTPROG	Participated in one or more government assistance programs	1 = Yes 2 = No
POVERTY2 POVERTY3	Poverty level (% of U.S. Census poverty threshold) POVERTY: 2003–2004 POVERTY2: 2004–2014 POVERTY3: 2015–present	. = Persons Aged 18–22 in a College Dorm 1 = Living in Poverty 2 = Income Up to 2X Fed Pov Thresh 3 = Income More Than 2X Fed Pov Thresh

NOTE: Fed Pov Thresh = federal poverty threshold.

The NSDUH collects a complete set of health insurance coverage variables (shown in Table 13), including Medicare, Medicaid/CHIP, TRICARE/CHAMPUS/CHAMPVA, and private health insurance. It collects number of months covered and uncovered in the past year and how long uncovered respondents have been without health insurance.

TABLE 13
Health Insurance Coverage

Variable Name	Description	Values
IRINSUR4	Imputed health insurance variables	1 = Yes, respondent does have [insurance type]
IRCHMPUS	IRINSUR4: Any health insurance	
IRMEDICR	IRCHMPUS: Military health TRICARE/VA/CHAMPUS/CHAMPVA	2 = No, respondent does not have [insurance type]
IRMCDCHP	IRMEDICR: Medicare	
IRPRVHLT	IRMCDCHP: Medicaid/CHIP	
IROTHHLT	IRPRVHLT: Private insurer	
	IROTHHLT: Other health insurance	

Veteran Characteristics

TABLE 14
Period of Service

Variable Name	Description	Values
ACTD2001	When did you serve on active duty in the United States Armed Forces or Reserve components? (respondents can reply 1=Yes to multiple military eras); these questions are available from 2013 to the present	ACTD2001: September 2001 or later
ACTD9001		ACTD9001: Aug 1990 to Aug 2001 (including Persian Gulf War)
ACTD7590		ACTD7590: May 1975 to July 1990
ACTDVIET		ACTDVIET: Vietnam Era (March 1961 to April 1975)
ACTDPRIV		ACTDPRIV: Any time before Vietnam Era (March 1961)

TABLE 15
U.S. Department of Veterans Affairs Care Recipient

Variable Name	Description	Values
TXYROTHER	During the past 12 months, have you received treatment for your [alcohol/drug] use in some other place besides these that have been listed?	1 = Yes 2 = No
TXYROTHSP2	Other place received treatment, past 12 months	52 = Military/VA Hospital

NOTE: VA = U.S. Department of Veterans Affairs.

TABLE 16
Other Relevant Variables

Variable Name	Description	Values
MILTFAMLY	Any immediate family currently in the U.S. military, whether or not they live with you 2016–present (Spouse or partner, mother or father, son or daughter, sister or brother)	1 = Yes
MILTSSPAR		2 = No
MILTPARNT		
MILTCHLDR		
MILTSIBLN		

Service-Connected Disability

The NSDUH has a number of disability measures, including mental health disability (World Health Organization Disability Assessment Schedule measure of functional impairment, Kessler Screening Scale for Psychological Distress) and drug and alcohol dependence measures, but none is specifically connected to a respondent's military service.

Additional Resources

Documentation, including codebooks, instruments, methodology notes, and public-use data downloads can be found at SAMHSA's "Data Sources" webpage under "National Survey on Drug Use and Health" (SAMHSA, undated-c).

Restricted-use files can be found at the CDC's "Data Hosting: National Survey on Drug Use and Health (NSDUH)" webpage (CDC, 2022). SAMHSA's "Data Tools" webpage offers a PDAS and RDAS online data explorer (SAMHSA, undated-a).

Abbreviations

NSDUH	National Survey on Drug Use and Health
SAMHSA	Substance Abuse and Mental Health Services Administration
CDC	U.S. Centers for Disease Control and Prevention
RDC	Research Data Center
PDAS	public data analysis system
RDAS	restricted data analysis system

References

CDC—*See* U.S. Centers for Disease Control and Prevention.

Center for Behavioral Health Statistics and Quality, *2019 National Survey on Drug Use and Health: Public Use File Codebook*, Substance Abuse and Mental Health Services Administration, October 20, 2020. As of May 28, 2024:

<https://www.datafiles.samhsa.gov/sites/default/files/field-uploads-protected/studies/NSDUH-2019/NSDUH-2019-datasets/NSDUH-2019-DS0001/NSDUH-2019-DS0001-info/NSDUH-2019-DS0001-info-codebook.pdf>

SAMHSA—*See* Substance Abuse and Mental Health Services Administration.

Substance Abuse and Mental Health Services Administration, “Data Tools,” webpage, undated-a. As of May 28, 2024:

<https://datatools.samhsa.gov/#/>

Substance Abuse and Mental Health Services Administration, “National Survey on Drug Use and Health (NSDUH),” webpage, undated-b. As of May 28, 2024:

<https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health>

Substance Abuse and Mental Health Services Administration, “Substance Abuse and Mental Health Data Archive,” webpage, undated-c. As of May 28, 2024:

<https://www.datafiles.samhsa.gov/>

U.S. Centers for Disease Control and Prevention, “Data Hosting: National Survey on Drug Use and Health (NSDUH),” webpage, December 15, 2022. As of May 28, 2024:

<https://www.cdc.gov/rdc/b1datatype/nsduh.htm>



This codebook was compiled as part of the RAND Veterans Policy Research Institute’s effort to develop a repository of datasets that assess veteran status. This codebook for the National Survey on Drug Use and Health (NSDUH) presents an overview of the data, how to access them, which variable(s) assess veteran status, other veteran-related questions, and additional resources for using the NSDUH. The authors cite directly from the original documentation of the NSDUH.

www.rand.org