package ‘NHANES’

October 12, 2022

Type Package

Title Data from the US National Health and Nutrition Examination Study

Version 2.1.0

Date 2015-07-02

Author Randall Pruim <rpruim@calvin.edu>

Maintainer Randall Pruim <rpruim@calvin.edu>


Depends R (>= 3.0)

Suggests mosaic

License GPL (>= 2)

LazyLoad yes

LazyData yes

NeedsCompilation no

Repository CRAN

Date/Publication 2015-07-02 06:47:16

R topics documented:

NHANES ................................................................. 2

Index ................................................................. 8
Description

This is survey data collected by the US National Center for Health Statistics (NCHS) which has conducted a series of health and nutrition surveys since the early 1960’s. Since 1999 approximately 5,000 individuals of all ages are interviewed in their homes every year and complete the health examination component of the survey. The health examination is conducted in a mobile examination centre (MEC).

Usage

data(NHANES)

Format

data frames with raw and resampled versions of the NHANES data. See below for details and descriptions of the variables.

Details

The NHANES target population is "the non-institutionalized civilian resident population of the United States". NHANES, (American National Health and Nutrition Examination surveys), use complex survey designs (see http://www.cdc.gov/nchs/data/series/sr_02/sr02_162.pdf) that over-sample certain subpopulations like racial minorities. Naive analysis of the original NHANES data can lead to mistaken conclusions. The percentages of people from each racial group in the data, for example, are quite different from the way they are in the population.

NHANES and NHANESraw each include 75 variables available for the 2009-2010 and 2011-2012 sample years. NHANESraw has 20,293 observations of these variables plus four additional variables that describe that sample weighting scheme employed. NHANES contains 10,000 rows of data resampled from NHANESraw to undo these oversampling effects. NHANES can be treated, for educational purposes, as if it were a simple random sample from the American population.

A list of the variables in the data set follows appears below along with variable descriptions and links to the original NHANES documentation.

NHANES warning

The following warning comes directly from the NHANES web site:

For NHANES datasets, the use of sampling weights and sample design variables is recommended for all analyses because the sample design is a clustered design and incorporates differential probabilities of selection. If you fail to account for the sampling parameters, you may obtain biased estimates and overstate significance levels.
Disclaimer

Please note that the data sets provided in this package are derived from the NHANES database and have been adapted for educational purposes. As such, they are NOT suitable for use as a research database. For research purposes you should download original data files from the NHANES website and follow the analysis instructions given there. Further details and relevant documentation can be found on the following NHANES websites

- http://www.cdc.gov/nchs/nhanes.htm,

Study Variables

SurveyYr Which survey the participant participated in.
ID Participant identifier.

Demographic Variables


Gender Gender (sex) of study participant coded as male or female
Age Age in years at screening of study participant. Note: Subjects 80 years or older were recorded as 80.
AgeDecade Categorical variable derived from age with levels 0–9, 10–19, … 70+
AgeMonths Age in months at screening of study participant. Reported for participants aged 0 to 79 years for 2009 to 2010 data Reported for participants aged 0 to 2 years for 2011 to 2012 data.
Race1 Reported race of study participant: Mexican, Hispanic, White, Black, or Other.
Race3 Reported race of study participant, including non-Hispanic Asian category: Mexican, Hispanic, White, Black, Asian, or Other. Not available for 2009-10.
Education Educational level of study participant Reported for participants aged 20 years or older. One of 8thGrade, 9–11thGrade, HighSchool, SomeCollege, or CollegeGrad.
MaritalStatus Marital status of study participant. Reported for participants aged 20 years or older. One of Married, Widowed, Divorced, Separated, NeverMarried, or LivePartner (living with partner).
HHIncomeMid Numerical version of HHIIncome derived from the middle income in each category
Poverty A ratio of family income to poverty guidelines. Smaller numbers indicate more poverty
HomeRooms How many rooms are in home of study participant (counting kitchen but not bathroom). 13 rooms = 13 or more rooms.
HomeOwn One of Home, Rent, or Other indicating whether the home of study participant or someone in their family is owned, rented or occupied by some other arrangement.
Physical Measurements


Weight  Weight in kg
Length  Recumbent length in cm. Reported for participants aged 0 - 3 years.
HeadCirc  Head circumference in cm. Reported for participants aged 0 years (0 - 6 months).
Height  Standing height in cm. Reported for participants aged 2 years or older.
BMI  Body mass index (weight/height2 in kg/m2). Reported for participants aged 2 years or older.
BMICatUnder20yrs  Body mass index category. Reported for participants aged 2 to 19 years. One of UnderWeight (BMI < 5th percentile) NormWeight (BMI 5th to < 85th percentile), OverWeight (BMI 85th to < 95th percentile), Obese (BMI >= 95th percentile).
BMI_WHO  Body mass index category. Reported for participants aged 2 years or older. One of 12.0_18.4, 18.5_24.9, 25.0_29.9, or 30.0_plus.
Pulse  60 second pulse rate
BPSysAve  Combined systolic blood pressure reading, following the procedure outlined for BPXSAR.
BPDiaAve  Combined diastolic blood pressure reading, following the procedure outlined for BPXDAR.
BPSys1  Systolic blood pressure in mm Hg – first reading
BPDia1  Diastolic blood pressure in mm Hg – second reading (consecutive readings)
BPSys2  Systolic blood pressure in mm Hg – second reading (consecutive readings)
BPDia2  Diastolic blood pressure in mm Hg – second second
BPSys3  Systolic blood pressure in mm Hg third reading (consecutive readings)
BPDia3  Diastolic blood pressure in mm Hg – third reading (consecutive readings)
Testosterone  Testerone total (ng/dL). Reported for participants aged 6 years or older. Not available for 2009-2010.

Health Variables


DirectChol  Direct HDL cholesterol in mmol/L. Reported for participants aged 6 years or older.
TotChol  Total HDL cholesterol in mmol/L. Reported for participants aged 6 years or older.
UrineVol1  Urine volume in mL – first test. Reported for participants aged 6 years or older.
UrineFlow1  Urine flow rate (urine volume/time since last urination) in mL/min – first test. Reported for participants aged 6 years or older.
UrineVol2  Urine volume in mL – second test. Reported for participants aged 6 years or older.
UrineFlow2  Urine flow rate (urine volume/time since last urination) in mL/min – second test. Reported for participants aged 6 years or older.
Diabetes  Study participant told by a doctor or health professional that they have diabetes. Reported for participants aged 1 year or older as Yes or No.
**DiabetesAge**  Age of study participant when first told they had diabetes. Reported for participants aged 1 year or older.

**HealthGen**  Self-reported rating of participant’s health in general. Reported for participants aged 12 years or older. One of Excellent, Vgood, Good, Fair, or Poor.

**DaysPhysHlthBad**  Self-reported number of days participant’s physical health was not good out of the past 30 days. Reported for participants aged 12 years or older.

**DaysMentHlthBad**  Self-reported number of days participant’s mental health was not good out of the past 30 days. Reported for participants aged 12 years or older.

**LittleInterest**  Self-reported number of days where participant had little interest in doing things. Reported for participants aged 18 years or older. One of None, Several, Majority (more than half the days), or AlmostAll.

**Depressed**  Self-reported number of days where participant felt down, depressed or hopeless. Reported for participants aged 18 years or older. One of None, Several, Majority (more than half the days), or AlmostAll.

**nPregnancies**  How many times participant has been pregnant. Reported for female participants aged 20 years or older.

**nBabies**  How many of participants deliveries resulted in live births. Reported for female participants aged 20 years or older.

**PregnantNow**  Pregnancy status at the time of the health examination was ascertained for females 8-59 years of age. Due to disclosure risks pregnancy status was only be released for women 20-44 years of age. The information used included urine pregnancy test results and self-reported pregnancy status. Urine pregnancy tests were performed prior to the dual energy x-ray absorptiometry (DXA) exam. Persons who reported they were pregnant at the time of exam were assumed to be pregnant. As a result, if the urine test was negative, but the subject reported they were pregnant, the status was coded as "Yes". If the urine pregnancy results were negative and the respondent stated that they were not pregnant, the respondent was coded as "No". If the urine pregnancy results were negative and the respondent did not know her pregnancy status, the respondent was coded "unknown". Persons who were interviewed, but not examined also have a value of "unknown". In addition there are missing values.

**Age1stBaby**  Age of participant at time of first live birth. 14 years or under = 14, 45 years or older = 45. Reported for female participants aged 20 years or older.

**SleepHrsNight**  Self-reported number of hours study participant usually gets at night on weekdays or workdays. Reported for participants aged 16 years and older.

**SleepTrouble**  Participant has told a doctor or other health professional that they had trouble sleeping. Reported for participants aged 16 years and older. Coded as Yes or No.

**Lifestyle Variables**


**PhysActive**  Participant does moderate or vigorous-intensity sports, fitness or recreational activities (Yes or No). Reported for participants 12 years or older.

**PhysActiveDays**  Number of days in a typical week that participant does moderate or vigorous-intensity activity. Reported for participants 12 years or older.
TVHrsDay  Number of hours per day on average participant watched TV over the past 30 days. Reported for participants 2 years or older. One of 0_to_1hr, 1_hr, 2_hr, 3_hr, 4_hr, More_4_hr. Not available 2009-2010.

CompHrsDay Number of hours per day on average participant used a computer or gaming device over the past 30 days. Reported for participants 2 years or older. One of 0_hrs, 0_to_1hr, 1_hr, 2_hr, 3_hr, 4_hr, More_4_hr. Not available 2009-2010.

TVHrsDayChild Number of hours per day on average participant watched TV over the past 30 days. Reported for participants 2 to 11 years. Not available 2011-2012.

CompHrsDayChild Number of hours per day on average participant used a computer or gaming device over the past 30 days. Reported for participants 2 to 11 years old. Not available 2011-2012.

Alcohol12PlusYr Participant has consumed at least 12 drinks of any type of alcoholic beverage in any one year. Reported for participants 18 years or older as Yes or No.

AlcoholDay Average number of drinks consumed on days that participant drank alcoholic beverages. Reported for participants aged 18 years or older.

AlcoholYear Estimated number of days over the past year that participant drank alcoholic beverages. Reported for participants aged 18 years or older.

SmokeNow Study participant currently smokes cigarettes regularly. Reported for participants aged 20 years or older as Yes or No, provided they answered Yes to having smoked 100 or more cigarettes in their life time. All subjects who have not smoked 100 or more cigarettes are listed as NA here.

Smoke100 Study participant has smoked at least 100 cigarettes in their entire life. Reported for participants aged 20 years or older as Yes or No.

SmokeAge Age study participant first started to smoke cigarettes fairly regularly. Reported for participants aged 20 years or older.

Marijuana Participant has tried marijuana. Reported for participants aged 18 to 59 years as Yes or No. AgeFirstMarijAge participant first tried marijuana. Reported for participants aged 18 to 59 years.

RegularMarij Participant has been/is a regular marijuana user (used at least once a month for a year). Reported for participants aged 18 to 59 years as Yes or No.

AgeRegMarij Age of participant when first started regularly using marijuana. Reported for participants aged 18 to 59 years.

HardDrugs Participant has tried cocaine, crack cocaine, heroin or methamphetamine. Reported for participants aged 18 to 69 years as Yes or No.

SexEver Participant had had vaginal, anal, or oral sex. Reported for participants aged 18 to 69 years as Yes or No.

SexAge Age of participant when had sex for the first time. Reported for participants aged 18 to 69 years.

SexNumPartnLife Number of opposite sex partners participant has had any kind of sex with over their lifetime. Reported for participants aged 18 to 69 years.

SexNumPartYear Number of opposite sex partners participant has had any kind of sex with over the past 12 months. Reported for participants aged 18 to 59 years.

SameSex Participant has had any kind of sex with a same sex partner. Reported for participants aged 18 to 69 years ad Yes or No.
**SexOrientation** participant’s sexual orientation (self-described). Reported for participants aged 18 to 59 years. One of Heterosexual, Homosexual, Bisexual.

**Weighting Variables (NHANESraw only)**

WTINT2YR, WTMEC2YR, SDMVPSU, SDMVSTRA Sample weighting variables. For more details see one of the following.
- [http://www.cdc.gov/Nchs/tutorials/environmental/orientation/sample_design/index.htm](http://www.cdc.gov/Nchs/tutorials/environmental/orientation/sample_design/index.htm)
- [http://www.cdc.gov/nchs/nhanes/nhanes2009-2010/DEMO_F.htm#WTINT2YR](http://www.cdc.gov/nchs/nhanes/nhanes2009-2010/DEMO_F.htm#WTINT2YR) and

**Source**

These data were originally assembled by Michelle Dalrymple of Cashmere High School and Chris Wild of the University of Auckland, New Zealand for use in teaching statistics.

**Examples**

```r
# Due to the sampling design, some races were over/under-sampled.
rbind(  
  NHANES = table(NHANES$Race1) / nrow(NHANES),  
  NHANESraw = table(NHANESraw$Race1) / nrow(NHANESraw),  
  diff = (table(NHANES$Race1) - table(NHANESraw$Race1)) / nrow(NHANESraw)  
)

# SmokeNow is only asked of people who answer Yes to Smoke100
if (require(mosaic)) {
  nhanes <-
    NHANES %>%
    mutate(      
      SmokingStatus = derivedFactor(        
        Current = SmokeNow == "Yes",        
        Former = SmokeNow == "No",        
        Never = Smoke100 == "No"        
      )
    )
  tally(~SmokingStatus, data = nhanes)
}
```
Index

NHANES, 2
NHANESraw (NHANES), 2