NHANES 2009–2012 dataset

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).

The NHANES **target population** is "the non-institutionalized civilian resident population of the United States".

Our dataset is derived from the data available for the 2009-2010 and 2011-2012 sample years and has approximately 10,000 observations (people) and about 70 variables.

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 oversample certain subpopulations like racial minorities. Naïve 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.

We have resampled the NHANES data to undo these oversampling effects. Our data is to be treated, for educational purposes, as a simple random sample from the American population.

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

[Disclaimer: Please note that this data set is underived from the NHANES database. It has been adapted for educational purposes. As such, it is NOT suitable to 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 NHANES websites:
http://www.cdc.gov/nchs/nhanes.htm, http://wwwn.cdc.gov/nchs/nhanes/search/nhanes11_12.aspx
http://wwwn.cdc.gov/nchs/nhanes/search/nhanes09_10.aspx*]*

VARIABLES:

The **NHANES2009-2012** dataset contains the following variables:

DEMOGRAPHIC VARIABLES					
Variable name	**Variable type**	**Code**	**Description**	**Study years**	
				2009-10	**2011-12**
ID	ID number				
Gender	Categorical	male female	Gender (sex) of study participant	Link	Link
Age	Numerical		Age in years at screening of study participant	Link	Link
AgeDecade	Categorical	0-9 10-19 …. 70+	Age decade. Derived from Age		
AgeMonths	Numerical		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	Link	Link
Race1	Categorical	Mexican, Hispanic White, Black, Other	Reported race of study participant	Link	Link
Race3	Categorical	Mexican, Hispanic White, Black Asian, Other	Reported race of study participant, including non-Hispanic Asian category	*Not available*	Link
Education	Categorical	8thGrade 9-11thGrade HighSchool SomeCollege CollegeGrad	Educational level of study participant Reported for participants aged 20 years or older	Link	Link
MaritalStatus	Categorical	Married Widowed Divorced Separated NeverMarried LivePartner	Marital status of study participant Reported for participants aged 20 years or older *LivePartner – living with partner*	Link	Link

DEMOGRAPHIC VARIABLES (continued)

Variable name	Variable type	Code	Description	Study years	
				2009-10	2011-12
HHIncome	Categorical	$0 - $4,999 $5,000 - $9,999 $10,000 - $14,999 $15,000 - $19,999 $20,000 - $24,999 $25,000 - $34,999 $35,000 - $44,999 $45,000 - $54,999 $55,000 - $64,999 $65,000 - $74,999 $75,000 - $99,999 $100,000 or More	Total annual gross income for the household	Link	Link
HHIncomeMid	Numerical		Total annual gross income for the household (taken from middle of income categories) *Income category* *Income value* Refused/don't know → NA $0 - $4,999 → $2,500 $5,000 - $9,999 → $7,500 $10,000 - $14,999 → $12,500 $15,000 - $19,999 → $17,500 $20,000 - $24,999 → $22,500 $25,000 - $34,999 → $30,000 $35,000 - $44,999 → $40,000 $45,000 - $54,999 → $50,000 $55,000 - $64,999 → $60,000 $65,000 - $74,999 → $70,000 $75,000 - $99,999 → $87,500 $100,000 or More → $100,000	Link	Link
Poverty	Numerical		A ratio of family income to poverty guidelines *Smaller numbers → more poverty*	Link	Link
HomeRooms	Numerical		How many rooms are in home of study participant (counting kitchen but not bathroom) *13 rooms = 13 or more rooms*	link	Link
HomeOwn	Categorical	Home Rent Other	Home of study participant or someone in their family is owned, rented or occupied by some other arrangement	Link	Link
Work	Categorical	Working Looking NotWorking	Type of work done by study participant last week Reported for participants aged 16 years or older	Link	Link

MEASUREMENT VARIABLES

Variable name	Variable type	Code	Description	Study years 2009-10	Study years 2011-12
Weight	Numerical		Weight in kg	[Link]	[Link]
Length	Numerical		Recumbent length in cm Reported for participants aged 0 – 3 years	[Link]	[Link]
HeadCirc	Numerical		Head circumference in cm Reported for participants aged 0 years (0 – 6 months)	[Link]	[Link]
Height	Numerical		Standing height in cm Reported for participants aged 2 years or older	[Link]	[Link]
BMI	Numerical		Body mass index (weight/height^2 in kg/m^2) Reported for participants aged 2 years or older	[Link]	[Link]
BMICatUnder20yrs	Categorical	UnderWeight NormWeight OverWeight Obese	Body mass index category Reported for participants aged 2 to 19 years $BMI < 5^{th}$ percentile $BMI\ 5^{th}\ to < 85^{th}$ percentile $BMI\ 85^{th}\ to < 95^{th}$ percentile $BMI \geq 95^{th}$ percentile	*Not available*	[Link]
BMI_WHO	Categorical	12.0_18.4 18.5_24.9 25.0_29.9 30.0_plus	Body mass index category Reported for participants aged 2 years or older	[Link]	[Link]
Pulse	Numerical		60 second pulse rate	[Link]	[Link]
BPSysAve	Numerical		Combined systolic blood pressure reading, following the procedure outlined for [BPXSAR]	[Link]	[Link]
BPDiaAve	Numerical		Combined diastolic blood pressure reading, following the procedure outlined for [BPXDAR]	[Link]	[Link]

MEASUREMENT VARIABLES (continued)

Variable name	Variable type	Code	Description	Study years	
				2009-10	**2011-12**
BPSys1	Numerical		Systolic blood pressure in mm Hg – first reading	Link	Link
BPDia1	Numerical		Diastolic blood pressure in mm Hg – first reading (consecutive readings)	Link	Link
BPSys2	Numerical		Systolic blood pressure in mm Hg – second reading (consecutive readings)	Link	Link
BPDia2	Numerical		Diastolic blood pressure in mm Hg – second reading	Link	Link
BPSys3	Numerical		Systolic blood pressure in mm Hg – third reading (consecutive readings)	Link	Link
BPDia3	Numerical		Diastolic blood pressure in mm Hg – third reading (consecutive readings)	Link	Link
Testosterone	Numerical		Testerone total (ng/dL) Reported for participants aged 6 years or older	*Not available*	Link

HEALTH VARIABLES

Variable name	Variable type	Code	Description	Study years 2009-10	2011-12
DirectChol	Numerical		Direct HDL cholesterol in mmol/L Reported for participants aged 6 years or older	Link	Link
TotChol	Numerical		Total cholesterol in mmol/L Reported for participants aged 6 years or older	Link	Link
UrineVol1	Numerical		Urine volume in mL – first test Reported for participants aged 6 years or older	Link	Link
UrineFlow1	Numerical		Urine flow rate in mL/min – first test Reported for participants aged 6 years or older (urine volume/time since last urination)	Link	Link
UrineVol2	Numerical		Urine volume in mL – second test Reported for participants aged 6 years or older	Link	Link
UrineFlow2	Numerical		Urine flow rate in mL/min – second test Reported for participants aged 6 years or older (urine volume/time since last urination)	Link	Link
Diabetes	Categorical	Yes No	Study participant told by a doctor or health professional that they have diabetes Reported for participants aged 1 year or older	Link	Link
DiabetesAge	Numerical		Age of study participant when first told they had diabetes Reported for participants aged 1 year or older	Link	Link
HealthGen	Categorical	Excellent Vgood Good Fair Poor	Self-reported rating of participant's health in general Reported for participants aged 12 years or older	Link	Link
DaysPhysHlthBad	Numerical		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	Link	Link
DaysMentHlthBad	Numerical		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	Link	Link

HEALTH VARIABLES (continued)

Variable name	Variable type	Code	Description	Study years 2009-10	Study years 2011-12
LittleInterest	Categorical	None Several Majority AlmostAll	Self-reported number of days where participant had little interest in doing things Reported for participants aged 18 years or older *Majority – more than half the days*	Link	Link
Depressed	Categorical	None Several Majority AlmostAll	Self-reported number of days where participant felt down, depressed or hopeless Reported for participants aged 18 years or older *Majority – more than half the days*	Link	Link
nPregnancies	Numerical		How many times participant has been pregnant Reported for female participants aged 20 years or older	Link	Link
nBabies	Numerical		How many of participants deliveries resulted in live births Reported for female participants aged 20 years or older	Link	Link
Age1stBaby	Numerical		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	Link	Link
SleepHrsNight	Numerical		Self-reported number of hours study participant usually gets at night on weekdays or workdays Reported for participants aged 16 years and older	Link	Link
SleepTrouble	Categorical	Yes No	Participant has told a doctor or other health professional that they had trouble sleeping Reported for participants aged 16 years and older	Link	Link

LIFESTYLE VARIABLES

Variable name	Variable type	Code	Description	Study years 2009-10	Study years 2011-12
PhysActive	Categorical	Yes No	Participant does moderate or vigorous-intensity sports, fitness or recreational activities Reported for participants 12 years or older	Link Link	Link Link
PhysActiveDays	Numerical		Number of days in a typical week that participant does moderate or vigorous-intensity activity Reported for participants 12 years or older	Link Link	Link Link
TVHrsDay	Categorical	0_hrs 0_to_1hr 1_hr 2_hr 3_hr 4_hr More_4_hr	Number of hours per day on average participant watched TV over the past 30 days Reported for participants 2 years or older	*Not available*	Link
CompHrsDay	Categorical	0_hrs 0_to_1hr 1_hr 2_hr 3_hr 4_hr More_4_hr	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	*Not available*	Link
TVHrsDayChild	Numerical		Number of hours per day on average participant watched TV over the past 30 days Reported for participants 2 to 11 years old	Link	*Not available*
CompHrsDayChild	Numerical		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	Link	*Not available*
Alcohol12PlusYr	Categorical	Yes No	Participant has consumed at least 12 drinks of any type of alcoholic beverage in any one year Reported for participants 18 years or older	Link	Link
AlcoholDay	Numerical		Average number of drinks consumed on days that participant drank alcoholic beverages Reported for participants aged 18 years or older	Link	Link
AlcoholYear	Numerical		Estimated number of days over the past year that participant drank alcoholic beverages Reported for participants aged 18 years or older	Link Link	Link Link

LIFESTYLE VARIABLES (continued)					
Variable name	**Variable type**	**Code**	**Description**	**Study years**	
				2009-10	**2011-12**
SmokeNow	Categorical	Yes No	Study participant currently smokes cigarettes regularly Reported for participants aged 20 years or older	Link	Link
Smoke100	Categorical	Yes No	Study participant has smoked at least 100 cigarettes in their entire life Reported for participants aged 20 years or older	Link	Link
SmokeAge	Numerical		Age study participant first started to smoke cigarettes fairly regularly Reported for participants aged 20 years or older	Link	Link
Marijuana	Categorical	Yes No	Participant has tried marijuana Reported for participants aged 18 to 59 years	Link	Link
AgeFirstMarij	Numerical		Age participant first tried marijuana Reported for participants aged 18 to 59 years	Link	Link
RegularMarij	Categorical	Yes No	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	Link	Link
AgeRegMarij	Numerical		Age of participant when first started regularly using marijuana Reported for participants aged 18 to 59 years	Link	Link
HardDrugs	Categorical	Yes No	Participant has tried cocaine, crack cocaine, heroin or methamphetamine Reported for participants aged 18 to 69 years	Link	Link
SexEver	Categorical	Yes No	Participant had had vaginal, anal, or oral sex Reported for participants aged 18 to 69 years	Link	Link
SexAge	Numerical		Age of participant when had sex for the first time Reported for participants aged 18 to 69 years	Link	Link
SexNumPartnLife	Numerical		Number of opposite sex partners participant has had any kind of sex with over their lifetime Reported for participants aged 18 to 69 years	Link Link	Link Link
SexNumPartYear	Numerical		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	Link Link	Link Link
SameSex	Categorical	Yes No	Participant has had any kind of sex with a same sex partner Reported for participants aged 18 to 69 years	Link Link	Link Link
SexOrientation	Categorical	Heterosexual Homosexual Bisexual	Participant's sexual orientation (self-described) Reported for participants aged 18 to 59 years	Link Link	Link Link
WTINT2YR **WTMEC2YR** **SDMVPSU** **SDMVSTRA**			Sample weighting variables	Link	Link