

**Analytic subsample for use with Suggested Course Extensions
for *The Chicago Guide to Writing about Multivariate Analysis, 2nd Edition***

This data set is a subset of data on children aged 0 to 10 years from the 1988–1994 National Health and Nutrition Examination Survey III (NHANES III)—a cross-sectional, population-based sample survey of the non-institutionalized population of the United States (US DHHS 1997). Blacks and Mexican Americans were oversampled in the NHANES III.

NOTE: The analytic subsample data set includes a randomly selected 90% subsample of the 9,813 cases used for the analyses in *The Chicago Guide to Writing about Multivariate Analysis, 2nd Edition* so the Ns and statistics based on this subsample will not match those in the book.

This subsample comprised a total of 8,832 children for whom family income and race/ethnicity were known. It includes 3,360 non-Hispanic white, 2,671 non-Hispanic black, and 2,801 Mexican American children. Children of other racial/ethnic backgrounds (mostly Asian) were excluded because there were not enough cases to study them as a separate group. All variables used in this analysis were based on information from the interview of the reporting adult in the household. To reduce recall bias, birth weight questions were asked only about children aged 0 to 10 years at the time of the survey.

Variable names have been changed to more substantively interpretable names; the question number or name from the original NHANES III data based is shown in parenthesis in the variable names used here.

Sampling weights provided for the NHANES III by the National Center for Health Statistics (US DHHS 1997) can be used to weight the statistics back to the national level.

NHANES III documentation

US Centers for Disease Control and Prevention. NHANES III Series 11 Data Files. Available online at <http://www.cdc.gov/nchs/nhanes/nh3data.htm>. Accessed January 2013.

US DHHS (Department of Health and Human Services). 1997. *National Health and Nutrition Examination Survey, III, 1988–1994*. CD-ROM Series 11, no. 1. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention.

National Center for Health Statistics. 1996. NHANES III HOUSEHOLD ADULT DATA FILE DOCUMENTATION. Available online at <http://archive.nlm.nih.gov/proj/dxpnet/nhanes/docs/doc/nhanes3/adult/adult.html> Accessed January 2013.

Variables in alphabetical order by variable name

Variable name Variable label

agemomb Age of mother at time of child's birth, years (HYA2)

range: 11 - 49

Variable name Variable label

birthwt Child's birth weight in grams (HYA7S)

range: 397 - 5896

. Missing

Variable name Variable label

educmom Highest grade completed for family reference person (HFHEDUCR)

range: 0 - 17

(continued)

Variable name Variable label

fincome Annual family income, unimputed (HFF19R)

Value	Value label
0	No income
1	Less than \$1000 (including loss)
2	\$ 1,000 - 1,999
3	\$ 2,000 - 2,999
4	\$ 3,000 - 3,999
5	\$ 4,000 - 4,999
6	\$ 5,000 - 5,999
7	\$ 6,000 - 6,999
8	\$ 7,000 - 7,999
9	\$ 8,000 - 8,999
10	\$ 9,000 - 9,999
11	\$10,000 - 10,999
12	\$11,000 - 11,999
13	\$12,000 - 12,999
14	\$13,000 - 13,999
15	\$14,000 - 14,999
16	\$15,000 - 15,999
17	\$16,000 - 16,999
18	\$17,000 - 17,999
19	\$18,000 - 18,999
20	\$19,000 - 19,999
21	\$20,000 - 24,999
22	\$25,000 - 29,999
23	\$30,000 - 34,999
24	\$35,000 - 39,999
25	\$40,000 - 44,999
26	\$45,000 - 49,999
27	\$50,000 and over
88*	Blank but applicable
99*	Don't know

Variable name Variable label

gender Gender of child (HSSEX)

Value	Value label
1	Male
2	Female

NOTE: Values marked with a "*" appear in the original data file but not in this subsample. They are included for the sake of completeness.

Variable name Variable label

hhsiz Number of persons in the household/dwelling (HSHSIZER)

range: 2 - 10
10 10+ persons

Variable name Variable label (original question number from the NHANES III)

idnum Family sequence ID number (DMPFSEQ)

range: 1 - 20073

Variable name Variable label

momsmoke Did mother smoke while pregnant with child (HYA3)

Value	Value label
1	Yes
2	No
8*	Blank but applicable
9*	Don't know

Variable name Variable label

povcat Poverty/income ratio, categorized

Value	Value label
1	<0.5
2	0.5-1.0
3	1.0-1.85
4	1.85-3.0
5	3.0+

Variable name Variable label

racethn Race/ethnicity of child (DMARETHN)

Value	Value label
1	Non-Hispanic white
2	Non-Hispanic black
3	Mexican-American
4*	Other

NOTE: Values marked with a "*" appear in the original data file but not in this subsample. They are included for the sake of completeness.

Variable name Variable label

sweight Final sampling weight (WTPFQX6)

range: 215.53 - 67904.26