

Raudenbush & Bryk Example Data Set: High School & Beyond

High School & Beyond is a nationally representative survey of U.S. public and Catholic high schools conducted by the National Center for Education Statistics (NCES). The data are a subsample of the 1982 HSB survey with 7,185 students from 160 schools. The average sample size per school is approximately 45 students. More information can be found at <http://nces.ed.gov/surveys/hsb/> and <http://pj.freefaculty.org/guides/stat/DataSets/HSB/NBER-High School and Beyond.htm>.

For use in HLM, there are two separate data sets: one for individual, student-level data (level 1) and one for school-level (level 2) data. The level-1 data file, *hsb1.sav*, has 7,185 cases and four variables (not including the school ID). The variables are:

- minority*, an indicator for student ethnicity (1 = minority, 0 = other)
- female*, an indicator for student gender (1 = female, 0 = male)
- ses*, (a standardized scale constructed from variables measuring parental education, occupation, and income)
- mathach*, a measure of mathematics achievement¹

LEVEL-1 DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
MINORITY	7185	0.27	0.45	0.00	1.00
FEMALE	7185	0.53	0.50	0.00	1.00
SES	7185	0.00	0.78	-3.76	2.69
MATHACH	7185	12.75	6.88	-2.83	24.99

At level 2, the data set *hsb2.sav* consists of 160 schools with 6 variables per school. The variables are:

- size* (school enrollment)
- sector* (1 = Catholic, 0 = public)
- pracad* (proportion of students in the academic track)
- disclim* (a scale measuring disciplinary climate)
- himnty* (1 = more than 40% minority enrollment, 0 = less than 40%)
- meanses* (mean of the SES values for the students in this school who are included in the level-1 file)

LEVEL-2 DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
SIZE	160	1097.82	629.51	100.00	2713.00
SECTOR	160	0.44	0.50	0.00	1.00
PRACAD	160	0.51	0.26	0.00	1.00
DISCLIM	160	-0.02	0.98	-2.42	2.76
HIMINTY	160	0.28	0.45	0.00	1.00
MEANSES	160	-0.00	0.41	-1.19	0.83

¹ This measure appears to be the 19-item *Mathematics I* measure designed for the HSB study, see Heyns, B., & Hilton, T. L. (1982). The cognitive tests for High School and Beyond: An assessment. *Sociology of Education*, 55, 89-102 for additional details on this measure and others in the HSB.

SPSS and R require only one data set, which I have called hsbmerged.sav.