

60080079 Introduction to Statistical Methods
Semester 2 2023-2024
Handout 1_1

High School and Beyond Data Set

The High School and Beyond data set is from a large-scale longitudinal study conducted by the National Opinion Research Center (1980). The students in the study are a nationally representative sample of $n = 600$ high school seniors with observations on 15 variables. The data are in SPSS format and can be downloaded from the class website. Below are brief descriptions of the variables.

Variable Number	Variable Name	Variable Label	Coded Responses
1	ID	Student identification number	3-digit numbers
2	SEX	Sex	1 = Male 2 = Female
3	RACE	Race or ethnicity	1 = Hispanic 2 = Asian 3 = Black 4 = White
4	SES	Socioeconomic status	1 = Low 2 = Medium 3 = High
5	SCH_TYPE	School type	1 = Public 2 = Private
6	HSP	High school program	1 = General 2 = Academic preparatory 3 = Vocational/technical
7	LOC_CON	Locus of control	Standardized score (mean 0, SD 1)
8	SEL_CON	Self control	Standardized score (mean 0, SD 1)
9	MOTIV	Motivation	Mean of 3 items scored 0 or 1
10	C_CHOICE	Career choice	1 = Clerical 2 = Craftsman 3 = Farmer 4 = Homemaker 5 = Laborer 6 = Manager 7 = Military 8 = Operative 9 = Professional I 10 = Professional II 11 = Proprietor 12 = Protective 13 = Sales 14 = School 15 = Service 16 = Technical 17 = Not working
11	READING	Reading score	Standardized score (mean 50, SD 10)
12	WRITING	Writing score	Standardized score (mean 50, SD 10)
13	MATH	Math score	Standardized score (mean 50, SD 10)
14	SCIENCE	Science score	Standardized score (mean 50, SD 10)
15	CIVICS	Civics score	Standardized score (mean 50, SD 10)