

# 1

## Description data

### *Description of the CTB Mathematics and Science data set*

In a study of the CTB - McGraw Hill, grade 8 students were evaluated regarding different objectives and subskills related to math and science. The data are found on the web site as a SAS data set, called MASC. The data set includes the results of 1500 students from 35 schools. Each of the students responded to 76 items, 56 items of the multiple choice type, and 22 items from the constructed response type. While the responses on the multiple choice items are coded 0 (wrong response) or 1 (correct response), the maximum score on the constructed response items varies from 1 to 4, depending on the item. The data set includes only data for the multiple choice items. Besides the indicator variables STUDENT, SCHOOL, and ITEM, the data set includes the following variables:

SCORE: the response of the students on the items,

GENDER: female = 1, male = 0,

TYPE: type of school, Catholic = 1, other private = 2, public = 3,

SIZE: the total enrollment rounded to the nearest hundred,

BACHELOR: transformed and standardized variable indicating the percentage of adults with Bachelors degree or higher, living in the area with the school zip code,

BORNINSTATE: transformed and standardized variable indicating the percentage of adults in the area with the school zip code, who were born in the state in which they currently reside,

MORTGAGE: transformed and standardized variable indicating the median of the monthly mortgage in the area of the school zip code,

OTHLANGUAGE: transformed and standardized variable indicating the percentage of households in the area of the school zip code, in which a language other than English is spoken,

MATH: items regarding mathematics = 1, items regarding science = 0, and finally

ITEM1-ITEM31 and ITEM42-ITEM66: 56 item indicators, one for each item.

Note: All zip code derived variables come from the 1990 US Census information.