

## INFORMATION CONTAINED IN THIS DOCUMENT:

- DESCRIPTION OF FILES
- HOW TO GET STARTED
- DESCRIPTION OF VARIANCE ESTIMATION VARIABLES
- DESCRIPTION OF ID VARIABLES FOR MERGING FILES

## DESCRIPTION OF FILES

All of the files on the PIRLS NCES website contain data for the United States only. Files with data from all countries can be found at the PIRLS international website at:

[http://timssandpirls.bc.edu/pirls2006/user\\_guide.html](http://timssandpirls.bc.edu/pirls2006/user_guide.html).

The U.S. files available on this NCES website are:

- Student data
  - The data are contained in ASGUSAR2.DAT. This file contains questionnaire items and derived variables and index scores based on the student questionnaire; plausible values for all scales from the reading assessment; and student sampling weights and replicate weights. TOTWGT is the primary student-level weight and sums to the student population size in the U.S.
  - There are 5,190 cases in this file (9 students completed the assessment but have no questionnaire data). Since the data are hierarchical (students are clustered with classrooms and schools), each student record contains identification variables that enable the user to merge the student, teacher, and school data.
  - an SPSS syntax file, ASGUSAR2.SPS
  - a SAS syntax file, ASGUSAR2.SAS
  - a codebook file (ASGUSAR2\_CODEBOOK\_PUD.PDF) that includes variable names, variable location and format information, variable labels, question text, values, and unweighted frequencies.
- School data
  - The data are contained in ACGUSAR2.DAT. This file contains items from the school questionnaire, derived variables and index scores based on the school questionnaire, and the school weight, SCHWGT which is designed for school-level analysis. There are 183 cases in this file (1 school did not complete the questionnaire and therefore has missing data).
  - an SPSS syntax file, ACGUSAR2.SPS
  - a SAS syntax file, ACGUSAR2.SAS

- a codebook file (ACGUSAR2\_CODEBOOK\_PUD.PDF) that includes variable names, variable location and format information, variable labels, question text, values, and unweighted frequencies.
- Teacher data
  - The data are contained in ATGUSAR2.DAT. This file contains items from the teacher questionnaire and derived variables based on the teacher questionnaire. There are 253 cases in this file; 4 of those cases have missing data because the teachers did not complete the questionnaire). There is no weight variable on the teacher questionnaire data file because these variables are not designed to be analyzed independently. The teacher data can be merged with the student data using the student-teacher link file described below.
  - an SPSS syntax file, ATGUSAR2.SPS
  - a SAS syntax file, ATGUSAR2.SAS
  - a codebook file (ATGUSAR2\_CODEBOOK\_PUD.PDF) that includes variable names, variable location and format information, variable labels, question text, values, and unweighted frequencies.
- Student-Teacher link file
  - The data are contained in ASTUSAR2.DAT. This file contains items needed to link students and teachers. There are 5190 cases in this file. This file contains a special weight variable: the TCHWGT variable is specifically designed for using teacher background data in student level analysis. The sole purpose of this file is to link teacher-level data with student-level data to perform appropriate student-level analysis.
  - an SPSS syntax file, ASTUSAR2.SPS
  - a SAS syntax file, ASTUSAR2.SAS
  - a codebook file (ASTUSAR2\_CODEBOOK\_PUD.PDF) that includes variable names, variable location and format information, variable labels, question text, values, and unweighted frequencies
- Cognitive (assessment) item data
  - The data are contained in ASAUSAR2.DAT. This file contains student responses to each item in the assessment. Note that many of the items have not been released, so there is little descriptive information about them. There are 5190 cases in this file.
  - an SPSS syntax file, ASAUSAR2.SPS
  - a SAS syntax file, ASAUSAR2.SAS

- a codebook file (ASAUSAR2\_CODEBOOK\_PUD.PDF) that includes variable names, variable location and format information, variable labels, question text, values, and unweighted frequencies.
- Questionnaires
  - The U.S. version of the student questionnaire is in the file P06\_STQ\_USA.PDF.
  - The U.S. version of the teacher questionnaire is in the file P06\_TQ\_USA.PDF.
  - The U.S. version of the school questionnaire is in the file P06\_SQ\_USA.PDF
- Released items used in PIRLS 2006
  - Passages that were released after the 2006 administration of PIRLS can be found in P06\_UG\_SamplePassages1\_Items.pdf.
- PIRLS\_2006 data user's guide
  - This document, PIRLS\_2006\_Data\_Analysis\_Users\_Guide.PDF, contains information on the conduct of PIRLS in the United States. For complete, detailed information, on using PIRLS data, users should consult the **PIRLS 2006 User Guide for the International Database** (edited by Pierre Foy and Ann M. Kennedy) which can be found at the link below: [http://timssandpirls.bc.edu/pirls2006/user\\_guide.html](http://timssandpirls.bc.edu/pirls2006/user_guide.html).

## GETTING STARTED

Analysis of PIRLS data requires the use of specialized software or macros because of the use of plausible values. PIRLS students do not have just one score for a given scale; rather each student has five “plausible values” for each scale. Special software must be used that can estimate scores and correct standard errors. The TIMSS and PIRLS International Study Center has developed easy-to-use software called the IDB Analyzer that can be downloaded without charge from the website listed below:

[http://timssandpirls.bc.edu/pirls2006/user\\_guide.html](http://timssandpirls.bc.edu/pirls2006/user_guide.html).

The IDB Analyzer is designed to merge data across files (student, teacher, school) with ease, use the correct weight variables for each level of analysis, and produce variance estimates that take into account the plausible values and the sampling design of the study. The version available at the time of this posting requires that SPSS be installed on the computer as well as the IDB Analyzer. Users will also need to download the .dat and .sps files provided on this website to read into the IDB Analyzer.

Directions for using the IDB Analyzer can be found in the **PIRLS 2006 User Guide for the International Database** which can be downloaded from the same website.

Note that naming conventions for the U.S. data files are the same as for the international files, so that they can be read into the IDB Analyzer with ease. Users should download the U.S. files into a U.S. folder so as not to overwrite any international files they may have downloaded previously.

For users who do not wish to use the IDB Analyzer, the SPSS and SAS syntax files provided on the NCES website can be used to construct student, teacher, school, and assessment data files. Information on constructing macros for use with SPSS or SAS can be found in the PIRLS 2006 User Guide for the International Database, chapters 4 (SPSS) and 5 (SAS).

#### DESCRIPTION OF VARIANCE ESTIMATION VARIABLES

The student file (ASG), assessment file (ASA), and the student-teacher link file (AST) each contain JKZONE and JKREP variables. JKZONE is the jackknife sampling zone, or stratum, to which the student's school is assigned. JKREP is the jackknife replicate, or primary sampling unit, to which the student's school is assigned. These variables are to be used in student-level analysis.

The school file (ACG) contains JKCZONE and JKCREP variables which can be used for school-level analysis. JKCZONE is the jackknife sampling zone, or stratum, to which the school is assigned. JKCREP is the jackknife replicate, or primary sampling unit, to which the school is assigned.

Complete information on the weights and variance estimation variables is available and can be found beginning on page 36 of the **PIRLS 2006 User Guide for the International Database**. Please note that the IDB Analyzer is programmed to use these variables correctly.

#### DESCRIPTION OF ID VARIABLES USED TO MERGE FILES

The IDB Analyzer is programmed to use the correct ID variables to merge files.

For those electing to use SPSS, SAS, or other software, the major IDs are as follows:

IDSCHOOL is a four-digit identification code that uniquely identifies the participating schools within each country. The school codes are not unique across countries.

IDCLASS is a six-digit identification code that uniquely identifies the sampled classrooms within a country. The variable IDCLASS has a hierarchical structure and is formed by concatenating the IDSCHOOL variable and a two-digit sequential number identifying the sampled classrooms within a school.

IDSTUD is an eight-digit identification code that uniquely identifies each sampled student in a country. The variable IDSTUD also has a hierarchical structure and is formed by concatenating IDCLASS variable and a two digit sequential number identifying all students within each classroom.

IDTEACH is a six-digit identification code that uniquely identifies a teacher within a school. It has a hierarchical structure and is formed by the concatenation of IDSCHOOL and a two-digit sequential number within each school. This variable is on the AST and ATG files only.

IDLINK uniquely identifies the class for which a teacher answered a questionnaire. The combination of linking variables IDTEACH and IDLINK uniquely identifies all teacher-class combinations in the database. This variable is on the AST and ATG files only.