# SHPPS 2000 Data Files and Documentation Data User's Guide

SHPPS 2000 data include 23 questionnaires, 27 data files, and 27 codebooks. Files are grouped by administration level (state, district, school, and classroom) and by component.

## File Names

File names include three parts. The first part identifies the component, the second part identifies the level, and the third part indicates the survey year.

| Component | Part |
|---|---|
| Health education | hlthed |
| Physical education and activity | physed |
| Health services | hlthser |
| Mental health and social services | mental |
| Healthy and safe school environment | env |
| Nutrition services | foodser |
| Faculty and staff health promotion | faculty |
| Status | status |

| Level | Part |
|---|---|
| State | s |
| District | d |
| School | l |
| Classroom | c |

The questionnaire and codebook files use the same naming convention but add the words "questionnaire" or "codebook" to the file name.

For example, the files for health education at the state level would be:

| Health education file | Name |
|---|---|
| SPSS data | hltheds2000.sav |
| Questionnaire | hltheds2000questionnaire.txt |
| Codebook | hltheds2000codebook.rtf |

# Data Files

The data files are available in:
- SAS
- SPSS
- Microsoft Access formats

The sections below provide information specific to each data format.

**SAS** Format:
SAS version 7.0 data files. The variables have been formatted and missing values have been defined. A format library has been provided and must be referenced by SAS programs for the datasets to be used. The library is named formats.sas7bcat. For detailed information on using SAS datasets and format libraries, please refer to the documentation provided with your SAS system.

**SPSS** Format:
SPSS version 9.0 data files. The data have been formatted and missing values have been defined. For detailed information on using SPSS datasets, please refer to the documentation provided with your SPSS system.

**Access** Format:
Access data files. In some cases, there are too many variables for one Access table. In these instances, the dataset is split into multiple tables. For detailed information on using Access databases, please refer to the documentation provided with your Access software.

# Status Files
The status files contain additional information for the component-specific datasets. There are four status files, one for each level (state, district, school, and classroom). The status files contain
- Enrollment information for district, school, and classroom data
- School type for school and classroom data
- QED data

# Questionnaires
The questionnaires are available in either PDF or txt formats.

# Codebooks
The codebooks are available in rich text format (rtf). The codebooks contain the
- Variable names*
- Variable labels
- Value range and codes
- Frequencies

*The first three letters of the variable name identifies the questionnaire for the state and district files (the school [l] and classroom [c] files are identified by single unique letters). The number at the end of the variable names represents the question number from the associated questionnaire.

# SHPPS 2000 Identification Variables

The SHPPS 2000 datasets include many identification variables.  The following is documentation of the origin, structure, and purposes of these variables.

## State Variables

There are two identification variables on the state status file:

UNIQ_KEY This is a 6-digit variable (e.g., 010000)
     Digits 1-2 indicate the state.
     Digits 3-6 are set to zero.

ST_CODE This is a 2-digit variable (e.g., 14).
     Each state in the sample has a unique 2-digit code; these are used as the first two digits of UNIQ_KEY.

There are three identification variables on each of the seven state questionnaire files. Each of the state questionnaire files contains a similar variable; Mental Health and Social Services (MHSS) will be used in the examples below.

MHSSSID This is an 8-digit variable (e.g., 01000005)
     Digits 1-6 are equivalent to UNIQ_KEY contained in the state status file.
     Digits 7-8 indicate the component.  For state files:
       HED 01
       PED 11
       HSV 03
       MHSS 05
       FSV 09
       ENV 13
       FSHP 07

MHSKEY This variable is equivalent to UNIQ_KEY contained in the state status file.

MHSCODE This variable is equivalent to ST_CODE contained in the state status file.

### District Variables

There are four identification variables on the district status file:

UNIQ_KEY    This is a 6-digit variable (e.g., 010003)
                   Digits 1-2 indicate the state.
                   Digits 3-6 are a sequential indicator of the district within the state.

ST_CODE    This is a 2-digit variable (e.g., 14).
                   This is equivalent to ST_CODE contained in the state status file.

DSDS_ID    This is a 4-digit variable (e.g., 0067)
                   This variable is equivalent to the last four digits of UNIQ_KEY contained in this file.

DISTRICT    This is a 6-digit variable produced by QED. It should be used in conjunction with ST_CODE when merging district data sets to a school or classroom data file. It is possible that QED used the same district code in more than one state; using it together with the ST_CODE will prevent error.

There is one identification variable on each of the seven district questionnaire files. Each of the district questionnaire files contain a similar variable; Mental Health and Social Services (MHSS) will be used in the example below.

MHSSDID    This is an 8-digit variable (e.g., 01000106)
                   Digits 1-6 are equivalent to UNIQ_KEY contained in the district status file.
                   Digits 7-8 indicate the component. For district files:
                           HED 02
                           PED 12
                           HSV 04
                           MHSS 06
                           FSV 10
                           ENV 14
                           FSHP 08
                   To merge district questionnaire and status files, link the first 6 digits of this identification variable (e.g., MHSSDID) with UNIQ_KEY contained in the district status file.

## School Variables

There are five identification variables on the school status file:

UNIQ_KEY    This is a 6-digit variable (e.g., 011009)
                Digits 1-2 indicate the state.
                Digits 3-6 are a sequential indicator of the school within the state.

ST_CODE    This is a 2-digit variable (e.g., 14).
                This is equivalent to ST_CODE contained in the state status file.

DSDS_ID    This variable is mis-named/labeled. It actually has nothing to do with districts. It is a 4-digit variable (e.g., 1067) generated internally, and is equivalent to the last four digits of UNIQ_KEY contained in <u>this file</u>.

DISTRICT    This is a 6-digit variable produced by QED. It should be used in conjunction with ST_CODE when merging school data sets to a district data file. It is possible that QED used the same district code in more than one state; using it together with the ST_CODE will prevent error.

SCH_ID    This is a 6-digit variable and is equivalent to UNIQ_KEY in this file.

There is one identification variable on each of the seven school-level questionnaire files. Each of the school-level questionnaire files contains a similar variable; Mental Health and Social Services (MHSS) will be used in the example below.

MHSSID    This is an 8-digit variable (e.g., 01100103)
                Digits 1-6 are equivalent to UNIQ_KEY contained in the school status file.
                Digits 7-8 indicate the component. For school files:
                       HED 01
                       PED 06
                       HSV 02
                       MHSS 03
                       FSV 05
                       ENV 07
                       FSHP 04

                To merge school-level questionnaire and status files, link the first 6 digits of this identification variable (e.g., MHSSID) with UNIQ_KEY contained in the school status file.

## Classroom Variables

There are six identification variables on the classroom status file:

UNIQ_KEY   This is a 6-digit variable (e.g., 011009) equivalent to UNIQ_KEY on the school files.
Digits 1-2 indicate the state.
Digits 3-6 are a sequential indicator of the school within the state.

CLASSID   This is an 8-character variable (e.g., 01100919 or 011009S9).
Digits 1-6 are equivalent to UNIQ_KEY on the school files and in this file.
Digit 7 indicates the interview number and respondent type.
  1 = 1st regular classroom/teacher sampled
  2 = 2nd regular classroom/teacher sampled
  .
  .
  .
  8 = 8th regular classroom/teacher sampled
  S = 1st specialist sampled
  T = 2nd specialist sampled

Digit 8 indicates the component.
  8 = Physical education
  9 = Health education

ST_CODE   This is a 2-digit variable (e.g., 14).
This is equivalent to ST_CODE contained in the state status file.

DSDS_ID   This variable is mis-named/labeled. It actually has nothing to do with districts. It is a 4-digit variable (e.g., 1067) generated internally, and is equivalent to the last four digits of UNIQ_KEY contained in this file.

DISTRICT   This is a 6-digit variable produced by QED. It should be used in conjunction with ST_CODE and SCH_ID when merging classroom data sets to a state, district, or school data file.

SCH_ID   This is a 6-digit variable and is equivalent to UNIQ_KEY in this file.


There is one identification variable on each of the two classroom-level questionnaire files. Each of the classroom-level questionnaire files contains a similar variable; Health Education (HED) will be used in the example below.

HECLID   This is an 8-digit variable (e.g., 01100103) equivalent to CLASSID contained in the classroom status file.

To merge classroom-level questionnaire and status files, link these two variables.