Slide 1 of 20

Getting Started with the HSLS:09 Data

Slide 2 of 20

Module Objectives

- Introduce users to the data collected across the data collection rounds of the High School Longitudinal Study of 2009 (HSLS:09)
- Describe the process of downloading the available public-use HSLS:09 data
- Describe the resources available within the restricted-use DVD
- Describe the contents of the data files
- Describe variable information and naming conventions
- Provide resources that are available to learn more about the study, the data, and the data files

Slide 3 of 20

HSLS:09 Data Collection Components

HSLS:09 data are obtained directly from one or more respondents across multiple points in time in the study

Respondents:	Data collections across the study:	Collection methods:
• Students • Parents • Math Teachers* • Science Teachers* • School Counselors • School Administrators • High School and Postsecondary Education Transcripts	• 2009 (9th grade) • 2012 • 2013 • 2016 • 2025	• Direct Assessment • Web Survey • Telephone Survey • School Records

*Only for the 2009 round

Consider: Most of the data collection instruments that were used to directly obtain data from one or more respondents in the study are available online via the HSLS Questionnaires page

Slide 4 of 20

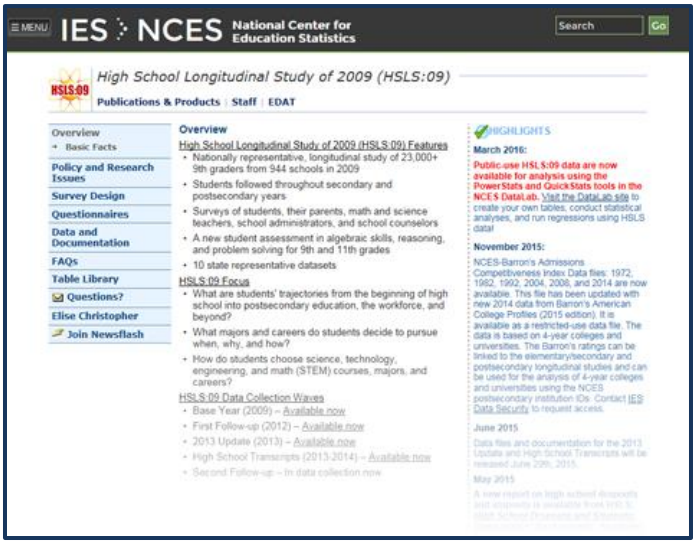
Data File Basics

- Data are available in both public- and restricted-use formats
- Public-use data provide data that most users will need for most analyses; they are available in three software-based formats (SPSS, SAS, and Stata)
- Restricted-use files for a given round of data collection contain more data and a wider range of data values than are included in the public-use files
- Researchers who are uncertain of which data file to use should first examine the corresponding public-use data file to ascertain whether it can be used to meet their specific analytical objectives

Slide 5 of 20

Downloading Public-use HSLS:09 Data

- Go to the Education Data Analysis Tool (eDAT)
  - Download the syntax file for the statistical software package of choice
- Go to the Data and Documentation page on the HSLS:09 homepage
  - Download the Data File Documentation



Slide 6 of 20

Using QuickStats and PowerStats to Analyze HSLS:09 Data

- Analyses of HSLS:09 data can be conducted using tools in the NCES DataLab
  - QuickStats
    - Generate percentages, averages tables, and simple graphs
  - PowerStats
    - May be used to calculate complex averages, percentages, medians, and centiles tables
    - May also be used to conduct linear and logistic regressions
- For more information about using the NCES DataLab tools, review the DLDT module DataLab Tools: QuickStats, PowerStats, and TrendStats

Slide 7 of 20

Accessing the Restricted-use HSLS:09 Data

Restricted-use HSLS:09 data contains additional resources within the DVD

- README file
    - Most important resource on the DVD; always start with the README file
    - Information about DVD contents and structure
    - Installation information for the Electronic Codebook (ECB) software

Slide 8 of 20

Accessing the Restricted-use HSLS:09 Data

- Electronic Codebook (ECB)
- Reference materials
    - Source documents, etc.
- Reports
    - Methodological and/or analytical

Slide 9 of 20

Contents of the Base Year Data File

Student-level data file (all merged to the student)	School-level data file (all merged to the school)
• Derived variables • Student-level analytic weights • Student assessment • Student questionnaire • Parent questionnaire • Math teacher questionnaire • Science teacher questionnaire • School counselor questionnaire • School administrator questionnaire	• Derived variables • School-level analytic weights • School counselor questionnaire • School administrator questionnaire

Slide 10 of 20

Contents of the First Follow-up Data File

Student-level data file (all merged to the student)

- Derived variables
- Student-level analytic weights
- Student assessment
- Student questionnaire
- Parent questionnaire
- School counselor questionnaire
- School administrator questionnaire

Slide 11 of 20

Contents of the 2013 Update Data File

- Derived variables
- Student-level analytic weights
- Student and parent questionnaires
- High school transcript school files
- High school transcript course files

Slide 12 of 20

Contents of the Second Follow-up Data File

The HSLS:09 Second Follow-up Data File includes

- Derived variables
- Student-level analytic weights
- Student questionnaires

Slide 13 of 20

Variable Naming

Character 1	Character 2	Characters 3-12
Component identifiers Composite variables = X Student = S Parent = P Math teacher = M Science teacher = N Administrator = A Counselor = C Weights = W Transcripts = T	Round identifier (i.e., 1, 2, 3), in which all base year variables are "1" and subsequent rounds will follow sequentially (e.g., first follow-up is "2")	Indicates a descriptive name for the variable

Slide 14 of 20

Variable Labels

- Provide more description than the variable name
- For convenience, retain the first two characters of the variable name
- Examples
  - Variable name: P1JOBNOW1
  - Variable label: P1 C05 Parent 1 currently holds a job

Slide 15 of 20

Variable Naming Conventions Examples

Name	Label
X1SEX	X1 Student's sex
X1RACE	X1 Student's race/ethnicity composite
S1SUREHSGRAD	S1 F10 How sure the 9th grader is that he/she will graduate from high school
P1HOMELANG	P1 B21 Language other than English is regularly spoken in the home
M1TEACHING	M1 B01A Math teachers in this school set high standards for teaching
N1TEACHING	N1 C01A Science teachers in this school set high standards for teaching
A1ADA	A1 A19 Average daily attendance percentage for high school students
C1FTCNSL	C1 A01A Number of full-time high school counselors
W1STUDENT	W1 Base year student analytic weight
T3SGPAO	T3 Grade point average overall

Slide 16 of 20

Composite/Derived Variables

- What are composite/derived variables?
  - Variables not directly collected as part of instruments
- Why use them?
  - Users may not have all data necessary to create similar variables
  - Convenience (time/effort)
  - Fewer missing values (imputation)
- How were they created?
  - Check data file documentation

Slide 17 of 20

Examples of Composite/Derived Variables

Algebra Assessment

Name	Label
X1TXMTH	X1 Mathematics theta score
X1TXMSEM	X1 Mathematics standard error of measurement for raw theta score
X1TXMSCR	X1 Mathematics IRT-estimated number right score (of 72 base year items)
X1TXMTSCOR	X1 Mathematics standardized theta score
X1TXMQUINT	X1 Mathematics quintile score
X1TXMPROF1	X1 Mathematics proficiency probability score: level 1
X1TXMPROF2	X1 Mathematics proficiency probability score: level 2
X1TXMPROF3	X1 Mathematics proficiency probability score: level 3
X1TXMPROF4	X1 Mathematics proficiency probability score: level 4
X1TXMPROF5	X1 Mathematics proficiency probability score: level 5

Slide 18 of 20

Examples of Composite/Derived Variables (Continued)

Family Characteristics

Name	Label
X1P1RELATION	X1 Parent 1: relationship to 9th grader
X1PAR1EDU	X1 Parent 1: highest level of education
X1PAR1EMP	X1 Parent 1: employment status
X1PAR1RACE	X1 Parent 1: race/ethnicity
X1P2RELATION	X1 Parent 2: spouse's relationship to 9th grader
X1PAR2EDU	X1 Parent 2: highest level of education
X1PAR2EMP	X1 Parent 2: employment status
X1PAR2RACE	X1 Parent 2: race/ethnicity
X1PAREDU	X1 Parents'/guardians' highest level of education
X1PARPATTERN	X1 P1-P2 relationship pattern
X1MOMREL	X1 Mother/female guardian's relationship to 9th grader
X1MOMEDU	X1 Mother/female guardian's highest level of education
X1MOMEMP	X1 Mother/female guardian's employment status
X1MOMRACE	X1 Mother/female guardian's race/ethnicity
X1DADREL	X1 Father/male guardian's relationship to 9th grader
X1DADEDU	X1 Father/male guardian's highest level of education
X1DADEMP	X1 Father/male guardian's employment status
X1DADRACE	X1 Father/male guardian's race/ethnicity

Slide 19 of 20

Resources Available From Other Sources

- To learn more about the study, the data, and the data file
  - Read Base Year Data File Documentation
  - Read First Follow-up Data File Documentation
  - Read 2013 Update and High School Transcripts Data File Documentation
  - Review the Questionnaires
  - Read HSLS:09: A First Look at Fall 2009 9th-Graders
  - Read HSLS:09: A First Look at Fall 2009 Ninth-Graders in 2012
  - Read HSLS:09: A First Look at Fall 2009 Ninth-Graders in 2013
- Visit HSLS:09 main study page for study overview, publications, and staff contact information

Slide 20 of 20

Module Summary and Resources

Summary

- Introduced users to the data collected across the data collection rounds of the High School Longitudinal Study of 2009 (HSLS:09)
- Described the process of downloading the available public-use HSLS:09 data
- Described the resources available within the restricted-use DVD
- Described the contents of the data files
- Described variable information and naming conventions
- Provided resources that are available to learn more about the study, the data, and the data files

Resources

- HSLS:09 Homepage
- Public-use Data
- eDAT Homepage
- Base Year Data File Documentation
- DataLab Tools: QuickStats, PowerStats, and TrendStats
- First Follow-up Data File Documentation
- 2013 Update and High School Transcripts Data File Documentation
- HSLS:09: A First Look at Fall 2009 9th-Graders
- HSLS:09: A First Look at Fall 2009 Ninth-Graders in 2012
- HSLS:09: A First Look at Fall 2009 Ninth-Graders in 2013