SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA ANALYSIS IN EDUCATION AGENCIES

www.gse.harvard.edu/sdp/tools
SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA ANALYSIS IN EDUCATION AGENCIES
MISSION

Transform the use of data in education to improve student achievement.
STRATEGIC DATA PROJECT

Theory of Action

Right People

Right Decisions

Right Data

Right Analysis
I. Fellows
Create and support a national network of high quality data analysts who will influence policy at the local, state, and national levels.

Place Fellows in partner agencies.
I. Fellows
Create and support a national network of high quality data analysts who will influence policy at the local, state, and national levels.

2. Diagnostics
Create policy- and management-relevant standardized analyses for districts and states.

Perform Diagnostics in partner agencies.

- Human Capital
  Teacher Effectiveness
- College-Going Success
I. Fellows
Create and support a national network of high quality data analysts who will influence policy at the local, state, and national levels.

2. Diagnostics
Create policy- and management-relevant standardized analyses for districts and states.

3. Scale
Improve the way data is used in the education sector.

Achieve broad impact through wide dissemination of analytic tools, methods, and best practices.
SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA ANALYSIS IN EDUCATION AGENCIES

Human Capital
Teacher Effectiveness

College-Going Success
What is the relationship between 8th grade test scores and college enrollment rates?

Distribution of College Enrollment Rates by Prior Student Achievement

Seamless Enrollers

<table>
<thead>
<tr>
<th>8th Grade CRCT Composite Score Quartile</th>
<th>Percent of High School Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom Quartile</td>
<td>47</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>73</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>85</td>
</tr>
<tr>
<td>Top Quartile</td>
<td>90</td>
</tr>
</tbody>
</table>

Average for Quartile
What is the relationship between 8th grade test scores and college enrollment rates?

Distribution of College Enrollment Rates by Prior Student Achievement

Seamless Enrollers

<table>
<thead>
<tr>
<th>8th Grade CRCT Composite Score Quartile</th>
<th>Percent of High School Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom Quartile</td>
<td>32</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>47</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>70</td>
</tr>
<tr>
<td>Top Quartile</td>
<td>91</td>
</tr>
</tbody>
</table>

Orange: Average for Quartile

Blue: Individual High School
1. Identify essential data elements

- Test Scores
- College Enrollment
- High School Graduation
- Student School Enrollment

Distribution of College Enrollment Rates by Prior Student Achievement

Seamless Enrollers

8th Grade CRCT Composite Score Quartile

<table>
<thead>
<tr>
<th>Quartile</th>
<th>Average for Quartile</th>
<th>Individual High School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom Quartile</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Top Quartile</td>
<td>64</td>
<td></td>
</tr>
</tbody>
</table>
1. Identify essential data elements
2. Clean check, and build variables for your datasets
3. Connect relevant datasets from different sources
4. Analyze your datasets
5. Adopt best practices to facilitate shared and replicable data analysis
SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA ANALYSIS IN EDUCATION AGENCIES

Identify: Data Specification Guide
Clean: Data Building Tasks
Connect: Data Linking Guide
Analyze: Diagnostic Analyses Guide
Adopt: Coding Style Guide
SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA
ANALYSIS IN EDUCATION AGENCIES

Identify: Data Specification Guide
Clean: Data Building Tasks
Connect: Data Linking Guide
Analyze: Diagnostic Analyses Guide
Adopt: Coding Style Guide
SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA ANALYSIS IN EDUCATION AGENCIES

Identify: Data Specification Guide
Clean: Data Building Tasks
Connect: Data Linking Guide
Analyze: Diagnostic Analyses Guide
Adopt: Coding Style Guide
Successful data analysis begins with proper identification of data elements necessary to answer key questions of interest.
5 W’s of Data Collection

**WHY**
are we collecting data?

Research questions?
## 5 W’s of Data Collection

<table>
<thead>
<tr>
<th>WHY</th>
<th>WHAT</th>
<th>WHERE</th>
<th>WHO</th>
<th>WHEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why are we collecting data?</td>
<td>Data is needed?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research questions?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>Student demographics</strong> (ethnicity, gender, birth date)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>Teacher demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>Student learning classifications</strong> (ELL, SpEd, FRPL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>Teacher experience, pay, certifications</strong> (HR data)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>School enrollment, attendance, and graduation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>Class enrollment, class grades, student/teacher links</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>Test scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• <strong>College Enrollment</strong> (NSC data)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 5 W’s of Data Collection

<table>
<thead>
<tr>
<th>WHY</th>
<th>WHAT</th>
<th>WHERE</th>
<th>WHO</th>
<th>WHEN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WHY</strong></td>
<td><strong>WHAT</strong></td>
<td><strong>WHERE</strong></td>
<td><strong>WHO</strong></td>
<td><strong>WHEN</strong></td>
</tr>
<tr>
<td>are we collecting data?</td>
<td>data is needed?</td>
<td>does the data live?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research questions?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Student demographics** (ethnicity, gender, birth date)
- **Teacher demographics**
- **Student learning classifications** (ELL, SpEd, FRPL)
- **Teacher experience, pay, certifications** (HR data)
- **School enrollment, attendance, and graduation**
- **Class enrollment, class grades, student/teacher links**
- **Test scores**
- **College Enrollment** (NSC data)

- **Student Information System** (SIS)
- **Longitudinal Data Store** (LDS) or **Data Warehouse** (DW)
- **HR Systems**
- **Excel Spreadsheets** or **MS Access**
- **On paper!**
5 W’s of Data Collection

**WHY**
- Research questions?

**WHAT**
- Student demographics (ethnicity, gender, birth date)
- Teacher demographics
- Student learning classifications (ELL, SpEd, FRPL)
- Teacher experience, pay, certifications (HR data)
- School enrollment, attendance, and graduation
- Class enrollment, class grades, student/teacher links
- Test scores
- College Enrollment (NSC data)

**WHERE**
- Does the data live?
  - Owns these systems and is responsible for delivering the data?
  - Student Information System (SIS)
  - Longitudinal Data Store (LDS) or Data Warehouse (DW)
  - HR Systems
  - Excel Spreadsheets or MS Access
  - On paper!
5 W’s of Data Collection

**WHY**
are we collecting data?
Research questions?

- **Student demographics** (ethnicity, gender, birth date)
- **Teacher demographics**
- **Student learning classifications** (ELL, SpEd, FRPL)
- **Teacher experience, pay, certifications** (HR data)
- **School enrollment, attendance, and graduation**
- **Class enrollment, class grades, student/teacher links**
- **Test scores**
- **College Enrollment** (NSC data)

**WHAT**
data is needed?

**WHERE**
does the data live?

**WHO**
Owns these systems and is responsible for delivering the data?

**WHEN**
(over what date range/school years) do we need data for, in case of a longitudinal analysis?

- **Student Information System** (SIS)
- **Longitudinal Data Store** (LDS) or **Data Warehouse** (DW)
- **HR Systems**
- **Excel Spreadsheets or MS Access**
- **On paper!**

Reliability of historical data elements?
5 W's of Data Collection

- **WHY**
  - are we collecting data?
  - Research questions?

- **WHAT**
  - data is needed?

- **WHERE**
  - does the data live?

- **WHO**
  - Owns these systems and is responsible for delivering the data?

- **WHEN**
  - (over what date range/school years) do we need data for, in case of a longitudinal analysis?

  - Reliability of historical data elements?

- **HOW**
  - should the data appear?

- **Student demographics** (ethnicity, gender, birth date)
- **Teacher demographics**
- **Student learning classifications** (ELL, SpEd, FRPL)
- **Teacher experience, pay, certifications** (HR data)
- **School enrollment, attendance, and graduation**
- **Class enrollment, class grades, student/teacher links**
- **Test scores**
- **College Enrollment** (NSC data)

- **Student Information System (SIS)**
- **Longitudinal Data Store (LDS) or Data Warehouse (DW)**
- **HR Systems**
- **Excel Spreadsheets or MS Access**
- **On paper!**
### Student Attributes

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Values or Data Type</th>
<th>Definition</th>
<th>Importance</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sid</strong></td>
<td>numeric</td>
<td>Student identifier unique to each student. This identification number is typically assigned to a student upon enrollment in your agency. State agencies may have different identification numbers than district agencies for the same student.</td>
<td>5</td>
<td>Cannot Be Missing</td>
</tr>
<tr>
<td><strong>male</strong></td>
<td>0 = female 1 = male</td>
<td>Student gender.</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>race_ethnicity</strong></td>
<td>1 = African American 2 = Asian American 3 = Hispanic 4 = American Indian 5 = White, not Hispanic 6 = Other 7 = Multiple</td>
<td>For systems or school years within systems where race and ethnicity are treated as a combined variable. If the system allows the indication of multiple categories simultaneously (e.g., African American and white) report “multiple.”</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>race</strong></td>
<td>1 = African American 2 = Asian American 3 = American Indian 4 = White 5 = Other 6 = Multiple</td>
<td>For systems or school years within systems where race and ethnicity are treated as separate variables. If the system allows for the indication of multiple categories simultaneously (e.g., African American and white) report “multiple.”</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>ethnicity</strong></td>
<td>0 = Hispanic 1 = not Hispanic</td>
<td>For systems or school years within systems where race and ethnicity are treated as separate variables and Hispanic or Latino origin is asked as a separate question.</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>birth_date</strong></td>
<td>date format (yyyy-mm-dd)</td>
<td>Student birth_date.</td>
<td>2</td>
<td>Good to Have</td>
</tr>
<tr>
<td><strong>first_9th_school_year_reported</strong></td>
<td>spring calendar year</td>
<td>The school year during which the student was a 9th grader for the first time. For this variable, report what the system explicitly recorded for first 9th grade school year. Not all systems will record this information.</td>
<td>1</td>
<td>Not Essential</td>
</tr>
<tr>
<td><strong>hs_diploma</strong></td>
<td>0 = no high school diploma 1 = has high school diploma</td>
<td>Indicator variable equal to 1 if the student has received a high school diploma from the system.</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>hs_diploma_type</strong></td>
<td>use local values</td>
<td>Any locally defined description of the type of diploma the student received. Include instances in which more than one type of diploma is observed, for example, Honors diploma, College Prep diploma, or General Education Diploma (GED) diploma.</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>hs_diploma_date</strong></td>
<td>date format (yyyy-mm-dd)</td>
<td>The date on which the student received a high school diploma. If only a month and year, or only a school year is known report that partial information.</td>
<td>4</td>
<td>Absolutely Necessary</td>
</tr>
<tr>
<td><strong>zip_code</strong></td>
<td>xxxx or xxxxx-yyyy</td>
<td>The zip code of the student’s home address.</td>
<td>1</td>
<td>Not Essential</td>
</tr>
</tbody>
</table>

- **Identifies uniqueness of observations** in each file to avoid data duplication (double counting)
- **Standardizes encoded values** for certain data points for consistency (i.e. ethnicity, subject matter)
- **Importance listed** for each data element
### TABLE OF CONTENTS

#### STUDENT DATA FILES

<table>
<thead>
<tr>
<th>Data File</th>
<th>Description</th>
<th>CG</th>
<th>HK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Attributes</td>
<td>Time invariant demographic, cohort, and graduation data for students.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student School Year</td>
<td>Yearly classification and attendance data for students.</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Student School Enrollment</td>
<td>School enrollment/withdrawal data for students.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Class Enrollment</td>
<td>Class enrollment, grades, and credits earned data for students.</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Student Test Scores</td>
<td>Standardized test data for students (state standardized tests, advanced placement, SAT, ACT, etc). Every attempt at a test by a student should be recorded.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student NSC Enrollment</td>
<td>The National Student Clearinghouse Student Tracker student-level data report providing information on postsecondary outcomes.</td>
<td>16</td>
<td>18</td>
</tr>
</tbody>
</table>

#### SCHOOL DATA FILES

<table>
<thead>
<tr>
<th>Data File</th>
<th>Description</th>
<th>CG</th>
<th>HK</th>
</tr>
</thead>
<tbody>
<tr>
<td>School</td>
<td>Yearly location and classification information for schools.</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Class</td>
<td>Class level scheduling data.</td>
<td></td>
<td>21</td>
</tr>
</tbody>
</table>

#### STAFF DATA FILES

<table>
<thead>
<tr>
<th>Data File</th>
<th>Description</th>
<th>CG</th>
<th>HK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff Attributes</td>
<td>Time invariant demographic and recruitment data related to staff.</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>Staff School Year</td>
<td>Yearly pay, experience, school placement, and job codes for staff.</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Staff Degrees</td>
<td>Educational achievement for staff. Each degree a staff member has received should be recorded once.</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>Staff Certifications</td>
<td>Teaching certifications received by staff.</td>
<td></td>
<td>28</td>
</tr>
</tbody>
</table>

- **Broad range of data points** – covers many research questions, many data points common to different analyses
- **Structured in a way that facilitates analyses** and similar to many existing data systems
• Specification meant to be a format to *extract* research data files to, not as a stand alone database structure

• May need to *merge* data across data stores to reach final layout

• **Software**
  - *ETL*: MS SSIS, Oracle WB, Informatica, DataStage
  - *Reporting*: Cognos, Crystal Reports, Oracle BI
  - *Statistical*: Stata, SAS, SPSS
5. Adopt
Coding Style Guide

To ensure that statistical code is easily shared across a team and is replicable by future users, SDP and the Center for Education Policy Research (CEPR) recommends that you follow best coding, programming, and data management practices.
Coding Style

Why yes, I do have nice handwriting... sort of...
if $teacher == 1 {
    local numyrs = 4
    mat out = J(`numyrs',2,.)
    local row = 1
    local col = 1

    foreach subj in math read {
        use "$data/student_teacher_`subj'_vam.dta", clear
        forval yr = 2(1)`numyrs' {gen late_exp_`yr' = ever_late_hire*t_exp`yr'}
    }
}

if $teacher == 1 {

    local numyrs = 4

    // define empty matrix of Yr x Subj
    mat out = J(`numyrs',2,.)
    local row = 1
    local col = 1

    foreach subj in math read {
        use "$data/student_teacher_`subj'_vam.dta", clear

        forval yr = 2(1)`numyrs' {
            gen late_exp_`yr' = ever_late_hire*t_exp`yr'
        }
    } // end of loop on subject
} // end of teacher processing
if $teacher == 1 {
local numyrs = 4
mat out = J(`numyrs',2,.)
local row = 1
local col = 1

drop if _n_ > `numyrs'

drop if _n_ < `numyrs'

drop if _n_ == `numyrs'

drop if _n_ == 0

drop if _n_ == `numyrs'+1

drop if _n_ == `numyrs'+2

drop if _n_ == `numyrs'+3

foreach subj in math read {
use "$data/student_teacher_`subj’_vam.dta", clear
forval yr = 2(1)`numyrs’ {
gen late_exp_`yr' = ever_late_hire*t_exp`yr’
}
}
}

if $teacher == 1 {
local numyrs = 4

// define empty matrix of Yr x Subj
mat out = J(`numyrs',2,.)
local row = 1
local col = 1

drop if _n_ > `numyrs'

drop if _n_ < `numyrs'

drop if _n_ == `numyrs'

drop if _n_ == 0

drop if _n_ == `numyrs'+1

drop if _n_ == `numyrs'+2

drop if _n_ == `numyrs'+3

foreach subj in math read {
use "$data/student_teacher_`subj’_vam.dta", clear
forval yr = 2(1)`numyrs’ {
gen late_exp_`yr' = ever_late_hire*t_exp`yr’
}
}
}

// end of loop on subject
// end of teacher processing
FILE: crosswalk_masked_ids.do

Author(s): JSilver

Date: 5/27/11

Description: This program creates the crosswalk of student ids to random research ids by:

1. Inputting the universe of student ids
2. Filtering the distinct set of student ids
3. Generating random ids and associating to student ids

Inputs: ..raw/students/studentyearsch.dta
..raw/students/englang.dta

Outputs: ..data/bps_student_school_year.dta

Update 1: TKawakita, 6/1/11 - Added check to ensure random ids are unique

clear
set more off
capture log close
set mem 8000mb

global raw "/cepr-files/projects/DCPS/Raw"
global data "/cepr-files/projects/DCPS/Data"
global log "/cepr-files/projects/DCPS/Log Files"

//***** Step 1: Input universe of student ids *****
...
//***** Step 2: Filter distinct set of student ids *****
...
//***** Step 3: Generate random ids and associate to student ids *****
...
//***** Update 1: Add check to ensure ids unique *****
...
```plaintext
gen t_late_hire = 0

replace t_late_hire = 0 if t_hiredate <= td(1sep2006) & t_hiredate != & t_year==2007
replace t_late_hire = 1 if t_hiredate > td(1sep2006) & t_hiredate <= td(lapr2007) ///
   & t_hiredate!= & t_year==2007
replace t_late_hire = 0 if t_hiredate > td(lapr2007) & t_hiredate!= & t_year==2009

replace t_late_hire = 0 if t_hiredate <= td(lsep2007) & t_hiredate != & t_year==2008
replace t_late_hire = 1 if t_hiredate > td(lsep2007) & t_hiredate <= td(lapr2008) ///
   & t_hiredate!= & t_year==2008
replace t_late_hire = 0 if t_hiredate > td(lapr2008) & t_hiredate!= & t_year==2008

replace t_late_hire = 0 if t_hiredate <= td(lsep2008) & t_hiredate != & t_year==2009
replace t_late_hire = 1 if t_hiredate > td(lsep2008) & t_hiredate <= td(lapr2009) ///
   & t_hiredate!= & t_year==2009
replace t_late_hire = 0 if t_hiredate > td(lapr2009) & t_hiredate!= & t_year==2009

replace t_late_hire = 0 if t_hiredate <= td(lsep2009) & t_hiredate != & t_year==2010
replace t_late_hire = 1 if t_hiredate > td(lsep2009) & t_hiredate <= td(lapr2010) ///
   & t_hiredate!= & t_year==2010
replace t_late_hire = 0 if t_hiredate > td(lapr2010) & t_hiredate!= & t_year==2010

local num_yrs "4"
local first_yr "2007"
local cutoff1 "1sep"
local cutoff2 "1apr"

gen t_late_hire = 0

forval yr =`firstyr’(1)(`first_yr’+`numyrs’-1) {
   replace t_late_hire = 0 if t_hiredate <= td(`cutoff1’`yr’) & t_hiredate != & t_year==`yr’
   & t_year==`yr’
   replace t_late_hire = 1 if t_hiredate > td(`cutoff2’`yr’) ///
   & t_hiredate<= td(`cutoff2’`yr’) & t_hiredate!= & t_year==`yr’
   replace t_late_hire = 0 if t_hiredate > td(`cutoff2’`yr’) ///
   & t_hiredate!= & t_year==`yr’
}
```

# TABLE OF CONTENTS

## 5. Adopt: CEPR Coding Style Guide

To ensure that statistical code is easily shared across a team and is replicable by future users, SDP and the Center for Education Policy Research (CEPR) recommends that you follow best coding, programming, and data management practices.

### INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>4</td>
</tr>
<tr>
<td>Scope</td>
<td>4</td>
</tr>
<tr>
<td>Intended Audience</td>
<td>4</td>
</tr>
<tr>
<td>Document Structure</td>
<td>4</td>
</tr>
<tr>
<td>Terminology</td>
<td>4</td>
</tr>
</tbody>
</table>

### NAMING CONVENTIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Naming Conventions</td>
<td>5</td>
</tr>
<tr>
<td>Abbreviations and Acronyms</td>
<td>5</td>
</tr>
<tr>
<td>Folder Naming and Structure</td>
<td>6</td>
</tr>
<tr>
<td>File Naming</td>
<td>7</td>
</tr>
<tr>
<td>Variable Naming</td>
<td>7</td>
</tr>
</tbody>
</table>

### COMMENTING AND READABILITY

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td>8</td>
</tr>
<tr>
<td>General Commenting Guidelines</td>
<td>8</td>
</tr>
<tr>
<td>File Headers</td>
<td>11</td>
</tr>
<tr>
<td>White Space and Readability</td>
<td>11</td>
</tr>
</tbody>
</table>

### CODING GUIDELINES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initializing Your Environment (Stata)</td>
<td>15</td>
</tr>
<tr>
<td>Logging Output (Stata)</td>
<td>15</td>
</tr>
<tr>
<td>Global Macros as Switches</td>
<td>16</td>
</tr>
<tr>
<td>Conditions</td>
<td>17</td>
</tr>
<tr>
<td>Hard Coding vs Macros</td>
<td>18</td>
</tr>
<tr>
<td>Macros as File Paths</td>
<td>19</td>
</tr>
<tr>
<td>Closing</td>
<td>19</td>
</tr>
</tbody>
</table>
Q&A
SDP TOOLKIT
FOR EFFECTIVE DATA USE
A GUIDE FOR CONDUCTING DATA ANALYSIS IN EDUCATION AGENCIES

Will Be Released Prior to Webinar On:

Thursday, February 2

Thursday, February 9

Thursday, February 16

Identify: Data Specification Guide
Clean: Data Building Tasks
Connect: Data Linking Guide
Analyze: Diagnostic Analyses Guide
Adopt: Coding Style Guide
Thank You

The toolkit is currently in **BETA**.

Please send us your feedback at goo.gl/AAvdF.

Check [www.gse.harvard.edu/sdp/tools](http://www.gse.harvard.edu/sdp/tools) for the most recent toolkit version.

Please contact us at [sdp@gse.harvard.edu](mailto:sdp@gse.harvard.edu) if you have any questions about the toolkit.