Slide 1 of 40

NAEP
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Considerations for Analysis of NAEP Data

Slide 2 of 40

Module Objectives

- Describe the use of NAEPEX Software
- Describe the use of AM Software
- Describe considerations for analyzing NAEP data

Considerations for Analysis of NAEP Data

Slide 3 of 40

Use Caution When Analyzing NAEP Data

- Many statistical packages assume simple random sampling
- NAEP data should **not** be analyzed using such procedures
  - May result in biased estimates
  - Can also result in underestimated sampling errors that will produce incorrect p values, which can incorrectly indicate differences that are statistically significant when they are not
- It is necessary to use special statistical software that accounts for the NAEP study design
- This module will focus on teaching you two particular tools
  - NAEPEX Software
  - AM Software

Slide 4 of 40

NAEPEX Software

- NAEPEX is a data extraction program for choosing variables, extracting data, and generating SPSS, SAS, or Stata syntax
- It is part of the NAEP Data Toolkit
  - Or it can be accessed directly from the NAEP restricted-use CD-ROM
- NAEPEX also functions as NAEP's Electronic Code Book (ECB) and can be used to
  - Select variables you need for analysis
  - Produce syntax/script to generate
- NAEPEX Tutorial and User Guide

Slide 5 of 40

AM Software

- Statistical software developed by the American Institutes for Research (AIR)
- Designed for use with large-scale assessments
- AM Software can be downloaded from AIR's AM website
- AM uses SPSS, SAS, or Stata syntax to perform statistical analyses with NAEP data files
- Generates code that appropriately uses Plausible Values and takes into account complex sampling design in computation of sampling variance
- Use AM when analyzing NCES data that contain plausible values

Slide 6 of 40

Data Selection Methods and Sources

- Create SPSS, SAS, or Stata dataset using NAEPEX
  - Extract variables of interest from flat file on NAEP CD by creating SPSS, SAS, or Stata syntax file
  - Edit data (e.g., cleaning, recoding)
- Import SPSS, SAS, or Stata dataset into AM Software
- Analyze data with AM Software

Considerations for Analysis of NAEP Data

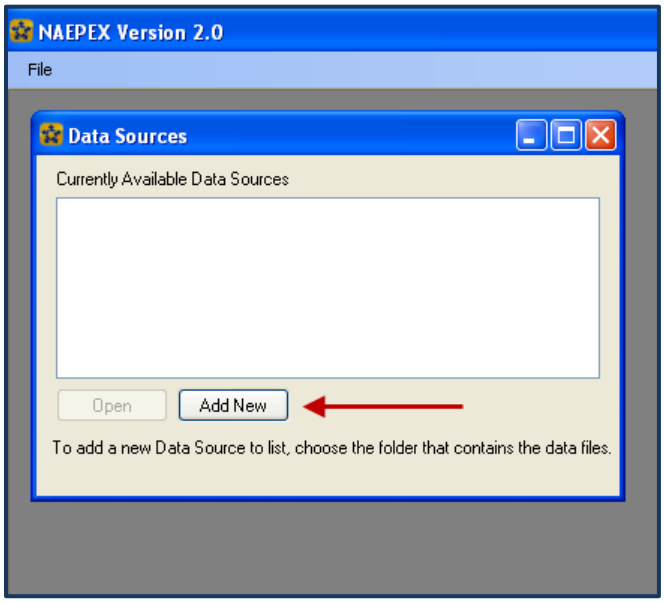
Slide 7 of 40

NIES 2007 Example Questions

- Is there a difference in 2007 reading achievement among public school 8th grade students based on the student gender and the amount of time the teacher spends on reading skills and strategies?
- What is the proportion distribution for days absent from school by gender for public school 8th grade students in 2007?
- Is there a relationship between 2007 reading achievement in the Midwest region among 8th grade public school students and absenteeism, and the number of books at home, while controlling for student's gender and the percentage of American Indian students in the school?

Slide 8 of 40

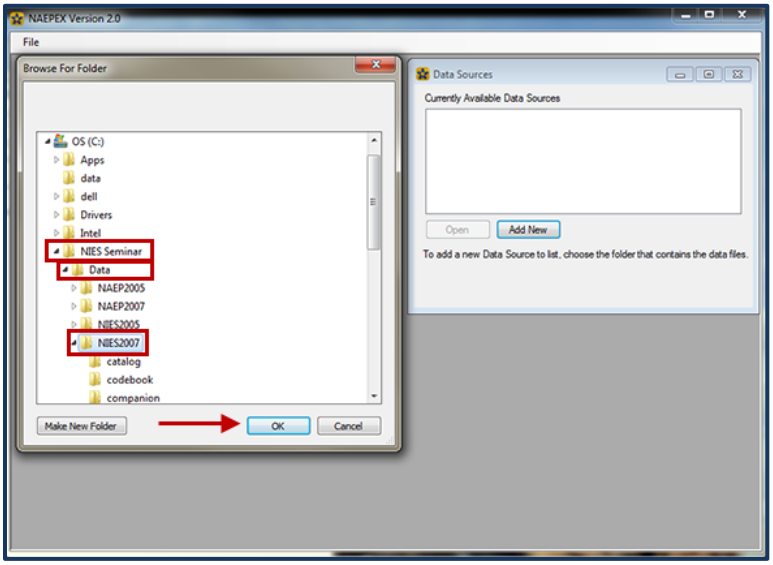
Load Data into NAEPEX



Considerations for Analysis of NAEP Data

Slide 9 of 40

Select Data Sources



For encrypted data source CD/DVD, first enter NCES-provided password to unzip CD/DVD to hard drive

Slide 10 of 40

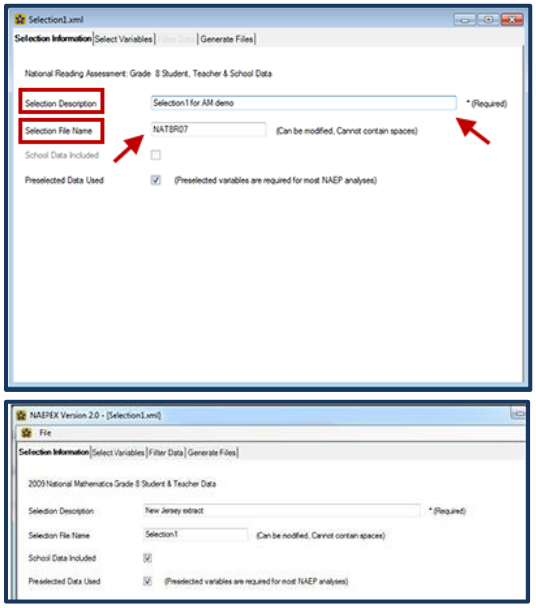
Open Data Set

- Select data source and click “Open”
  - Example: 2007 National Indian Education Study (NIES) parts I and II, grades 4 and 8
- Select data and click “Open Data File”
  - Example: National Reading Assessment: Grade 8 Student, Teacher & School data

Considerations for Analysis of NAEP Data

Slide 11 of 40

Define the Data File



Slide 12 of 40

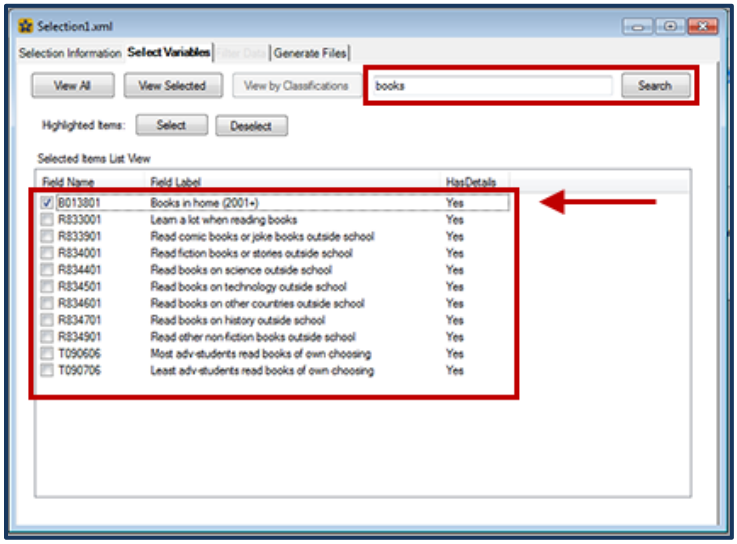
Pre-selected Variables

- RRPCM1-RRPCM5: Plausible NAEP reading values
- ORIGWT: Student weight
- JKUNIT: Jackknife variance unit
- REPGRP1: Jackknife variance stratum
- SRWT01-SRWT62: Replicate weights
- CENSREG: Census region of the country
- DSEX: Gender
- FIPS: FIPS state code

Considerations for Analysis of NAEP Data

Slide 13 of 40

Select Variables



Slide 14 of 40

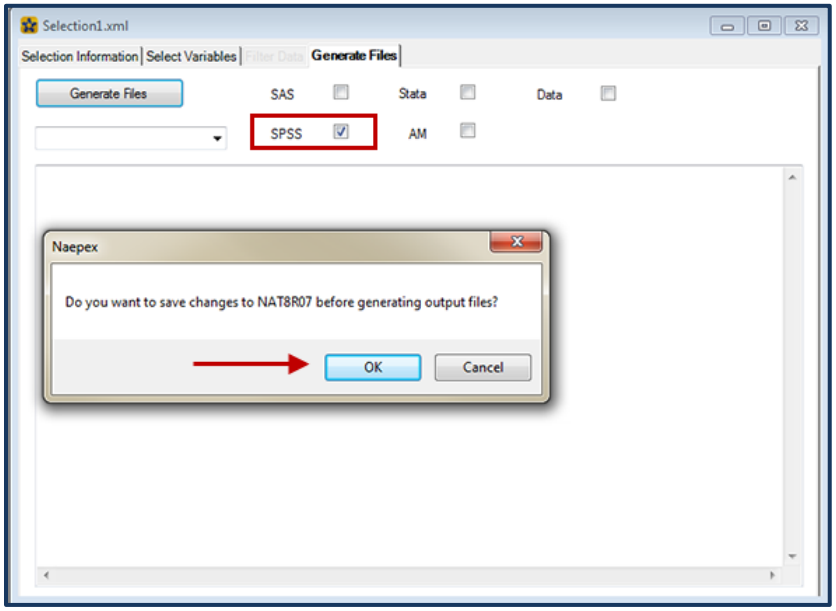
Select Variables (Continued)

- Variables for school type and percent American Indian selected (for example analysis)
- DSEX and CENSREG pre-selected
- **Must** select RPTSAMP (reporting sample flag)
- Select **all** needed variables before creating SPSS or SAS syntax

Considerations for Analysis of NAEP Data

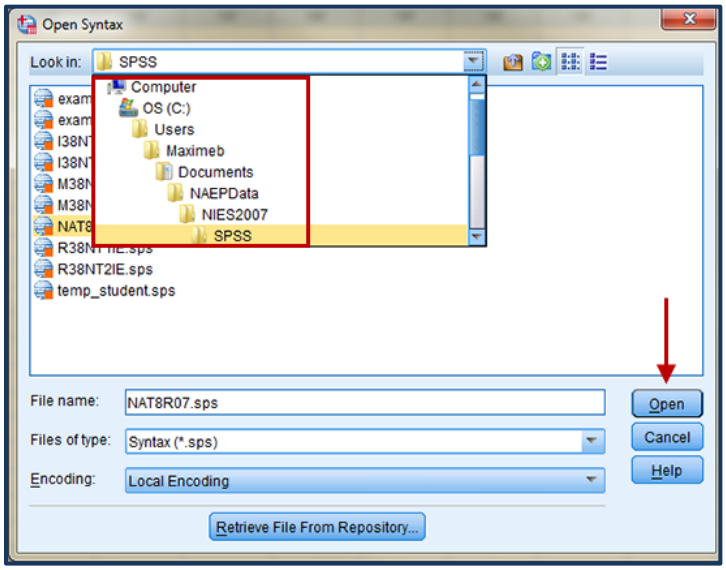
Slide 15 of 40

Generate Syntax File



Slide 16 of 40

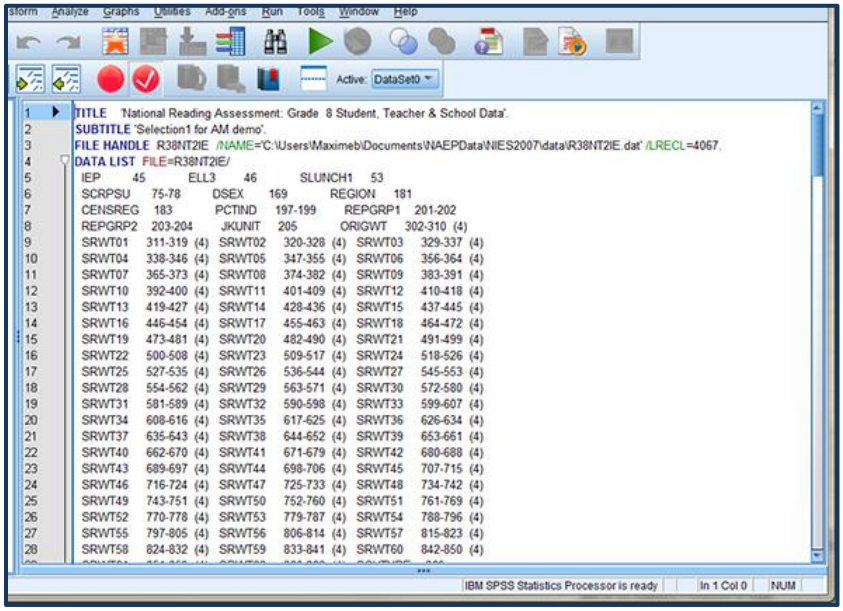
Open SPSS Syntax File



Considerations for Analysis of NAEP Data

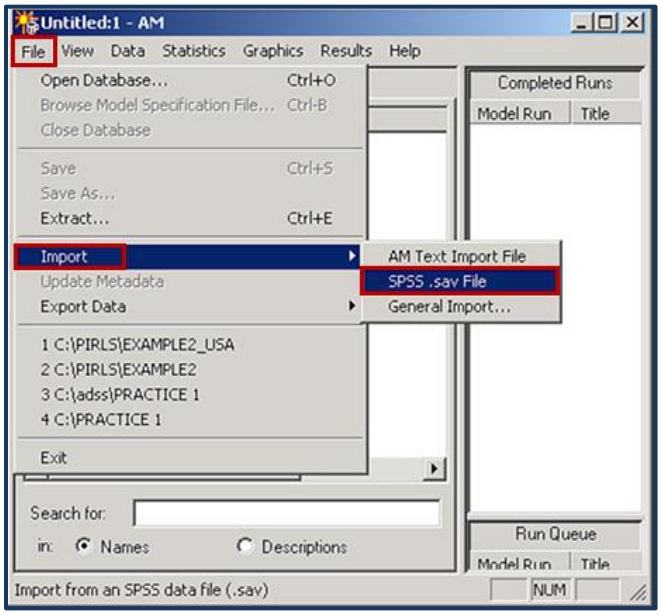
Slide 17 of 40

Open SPSS Syntax File (Continued)



Slide 18 of 40

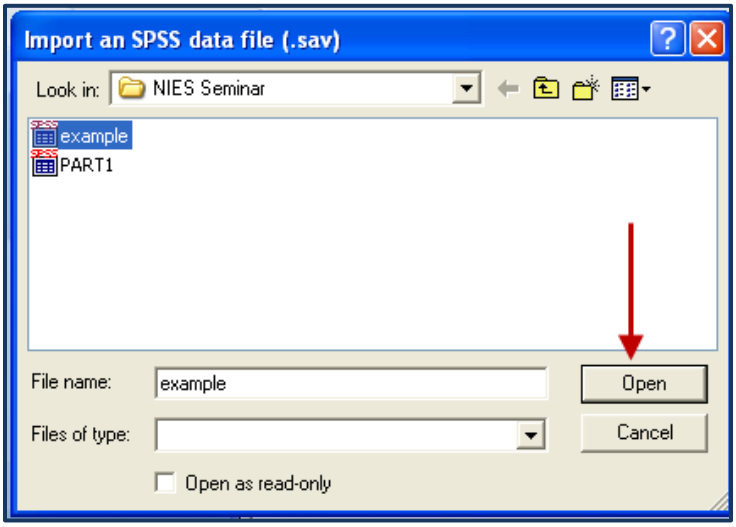
Starting AM Software



Considerations for Analysis of NAEP Data

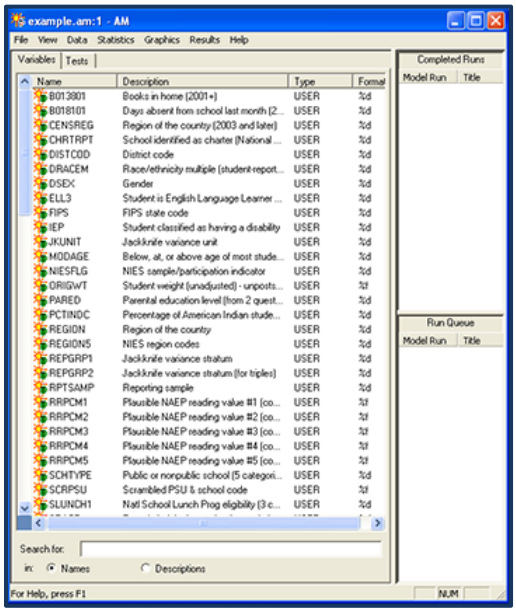
Slide 19 of 40

Import SPSS File into AM Software



Slide 20 of 40

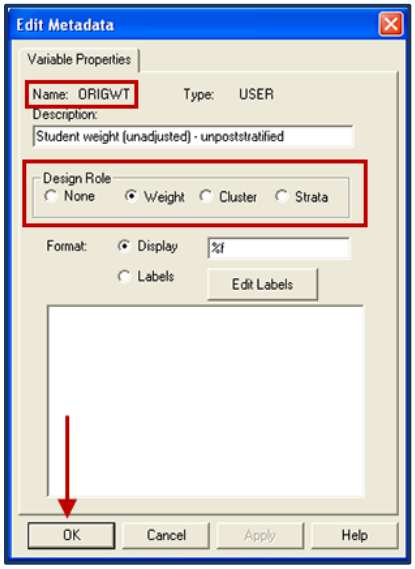
Save SPSS File in AM



Considerations for Analysis of NAEP Data

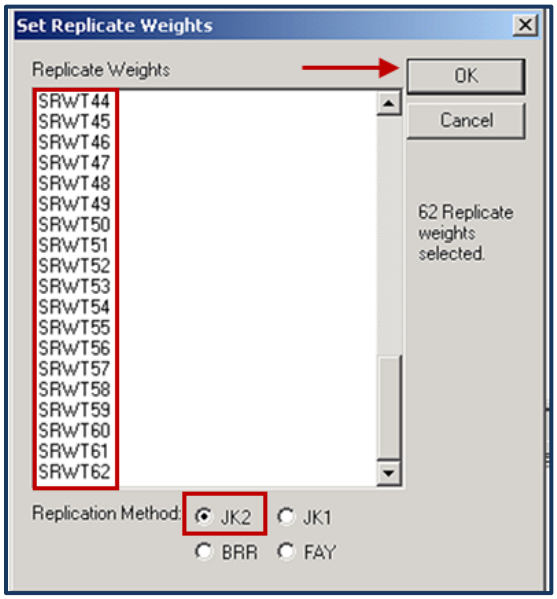
Slide 21 of 40

Identify Design Role Variables



Slide 22 of 40

Identify Replicate Weights



Considerations for Analysis of NAEP Data

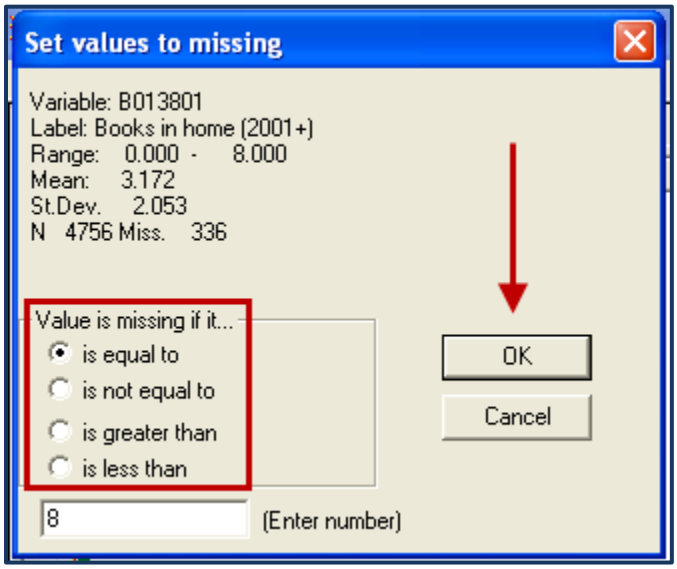
Slide 23 of 40

Define Missing Values

Value	Value Label	B013801	B018101	T083601
0	Multiple Response	1	2	8
8	Omitted	601	605	21

Slide 24 of 40

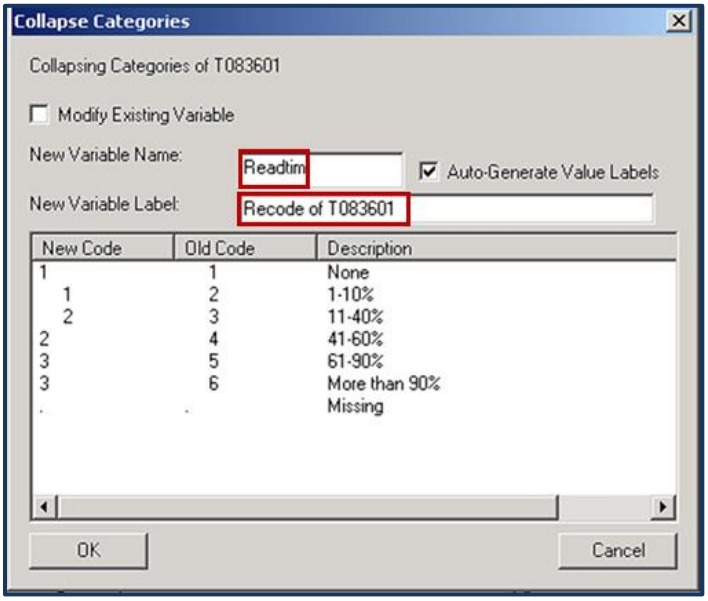
Define Missing Values (Continued)



Considerations for Analysis of NAEP Data

Slide 25 of 40

Collapse Categories



Slide 26 of 40

Dummy Code Variables

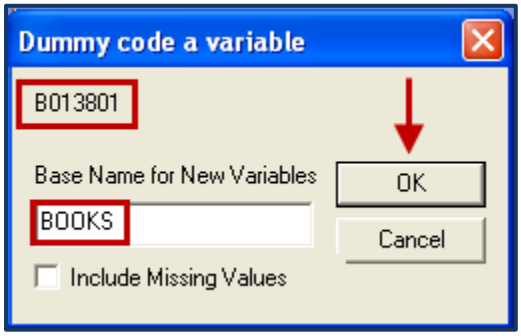
DSEX

1 → DSEX0 (Male)
2 → DSEX1 (Female)

Considerations for Analysis of NAEP Data

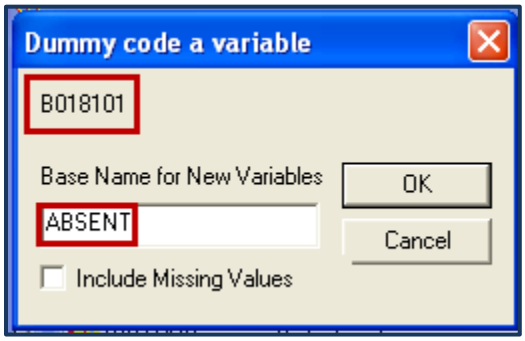
Slide 27 of 40

Dummy Code Variables (Continued)



Slide 28 of 40

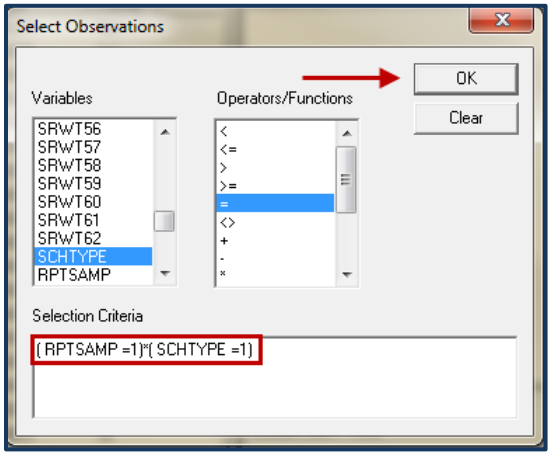
Create Dummy Variables



Considerations for Analysis of NAEP Data

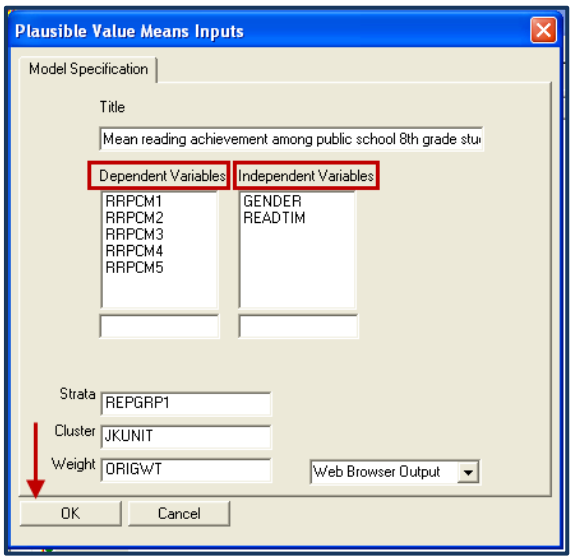
Slide 29 of 40

Set a Filter



Slide 30 of 40

Run a Crosstab



Is there a difference in 2007 reading achievement among public school 8th grade students based on the student gender and the amount of time the teacher spends on reading skills and strategies?

Considerations for Analysis of NAEP Data

Slide 31 of 40

Crosstab Results

Plausible Value Means: Mean reading achievement among public school 8th grade students by gender and time spent on reading skills and strategies

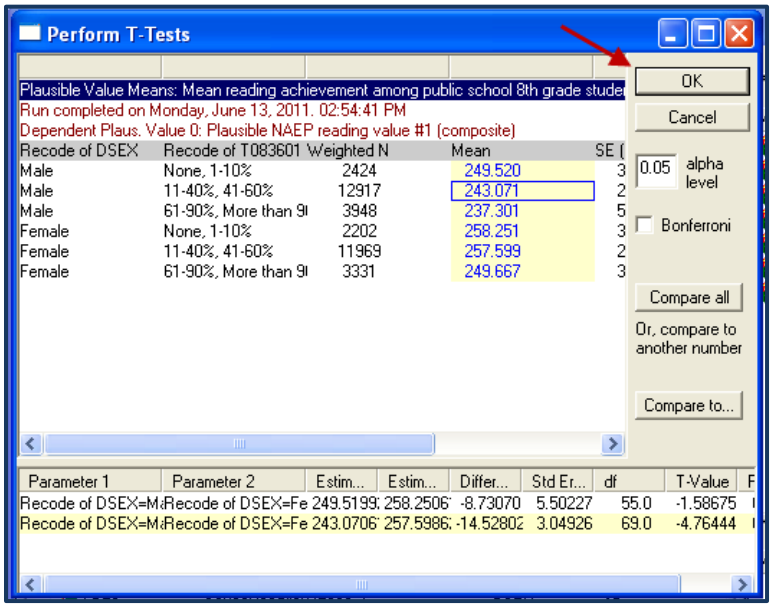
```
Run completed on Monday, June 17, 2013. 02:54:41 PM

Selection:            ALL
Observations:         3176
Strata Variable:      Jackknife variance stratum
Cluster Variable:     Jackknife variance unit
Weight Variable:      Student weight (unadjusted) - unpoststratified
Dependent Plaus. Value 0: Plausible NAEP reading value #1 (composite)
Dependent Plaus. Value 1: Plausible NAEP reading value #2 (composite)
Dependent Plaus. Value 2: Plausible NAEP reading value #3 (composite)
Dependent Plaus. Value 3: Plausible NAEP reading value #4 (composite)
Dependent Plaus. Value 4: Plausible NAEP reading value #5 (composite)
```

Recode of DSEX	Recode of T083601	Weighted N	Mean	SE (Mean)	Std. Dev
Male	None, 1-10%	2424	249.520	3.985	32.139
Male	11-40%, 41-60%	12917	243.071	2.189	38.416
Male	61-90%, More than 90%	3948	237.301	5.019	42.121
Female	None, 1-10%	2202	258.251	3.658	38.153
Female	11-40%, 41-60%	11969	257.599	2.113	37.972
Female	61-90%, More than 90%	3331	249.667	3.452	35.404

Slide 32 of 40

Conduct a T-Test



Considerations for Analysis of NAEP Data

Slide 33 of 40

T-Test Results

Plausible Value Means: Mean reading achievement among public school 8th grade students by gender and time spent on reading skills and strategies

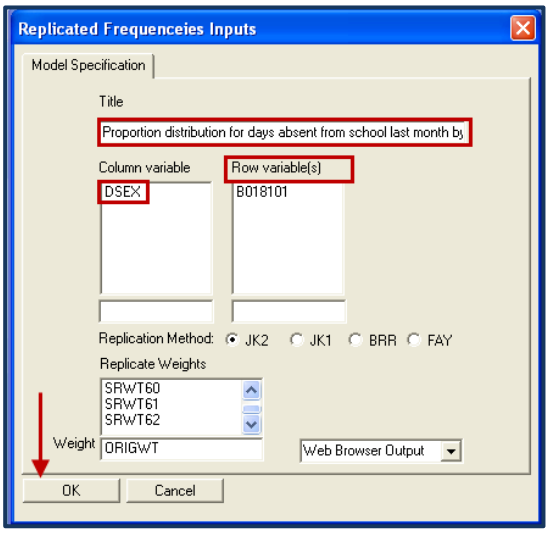
TTest Results

For run completed on Monday, June 17, 2013. 02:54:41 PM
Alpha = 0.05000

Parameter1	Parameter2	Mean 1	Mean 2	Difference	SE Difference	Deg. of freedom	T-statistic	p > t
Recode of DSEX=Male, Recode of T083601=None, 1-10%, Mean of Plausible NAEP reading value #1 (composite)	Recode of DSEX=Female, Recode of T083601=None, 1-10%, Mean of Plausible NAEP reading value #1 (composite)	249.520	258.251	-8.731	5.502	55.000	-1.587	0.118
Recode of DSEX=Male, Recode of T083601=11-40%, 41-60%, Mean of Plausible NAEP reading value #1 (composite)	Recode of DSEX=Female, Recode of T083601=11-40%, 41-60%, Mean of Plausible NAEP reading value #1 (composite)	243.071	257.599	-14.528	3.049	69.000	-4.764	0.000

Slide 34 of 40

Examine Proportion Distribution



- What is the proportion distribution for days absent from school by gender for public school 8th grade students in 2007?

Considerations for Analysis of NAEP Data

Slide 35 of 40

Frequencies Results

Replicated Frequencies: Proportion distribution for days absent from school last month by gender for public school 8th grade students

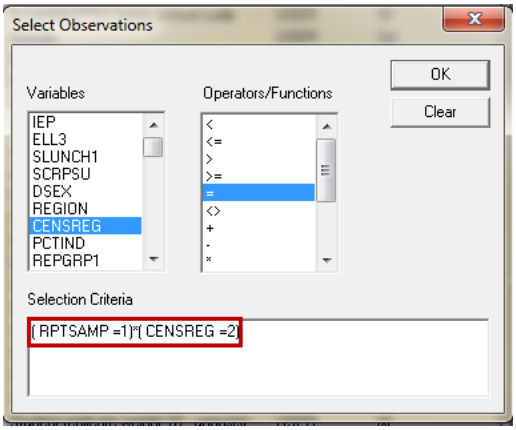
```
Run completed on Monday, June 17, 2013. 03:30:15 PM

Selection:            (  RPTSAMP=1 * SCHTYPE =1 )
Observations:         3102
Using 62 replicate weights
Estimates centered on overall weight:Student weight (unadjusted) - unpoststratified
```

Days absent from school last month (2001+)	Weighted N	Gender: Male	(se)	Female	(se)
None	13167	0.557	0.0183	0.443	0.0183
1-2 days	13809	0.516	0.0194	0.484	0.0194
3-4 days	6908	0.513	0.0327	0.487	0.0327
5-10 days	3729	0.452	0.0377	0.548	0.0377
More than 10 days	1306	0.636	0.0592	0.364	0.0592

Slide 36 of 40

Regression Analysis

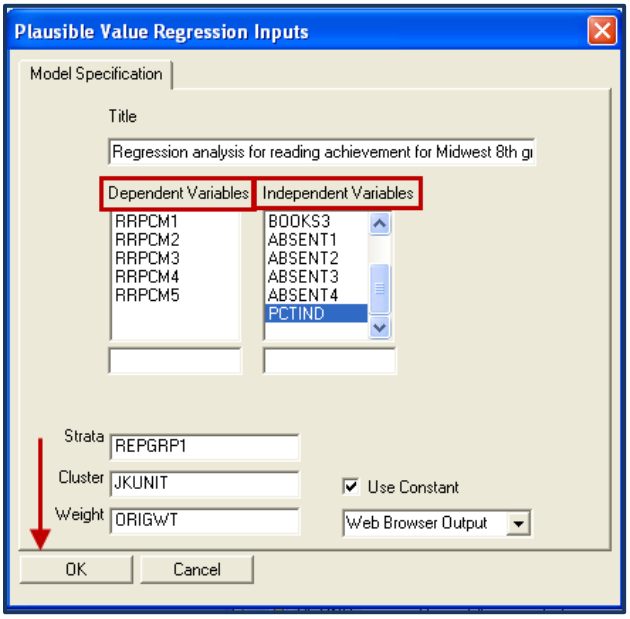


Is there a relationship between 2007 reading achievement in the Midwest region among 8th grade public school students and absenteeism, and the number of books at home, while controlling for student's gender and the percentage of American Indian students in the school?

Considerations for Analysis of NAEP Data

Slide 37 of 40

Regression Analysis (Continued)



Slide 38 of 40

Regression Result

Plausible Value Regression: Regression analysis for reading achievement for Midwest 8th grade students

```
 Run completed on Monday, June 17, 2013. 11:14:12 AM

WARNINGS were issued. Right click on completed run and selected View Warnings

 Selection:                ALL
 Observations:             1011
 Strata Variable:          Jackknife variance stratum
 Cluster Variable:         Jackknife variance unit
 Weight Variable:          Student weight (unadjusted) - unpoststratified

Adjusted Wald Test
F(9,70) = 10.3974
p(F > f) = 5.1769e-010
 Dependent Variable: Plausible NAEP reading value #1 (composite)
R-Square  =   0.204
```

Parameter Name	Estimate	Standard Error	z Score	p > |z|
Constant	242.261	6.469	37.452	0.000
Recode of DSEX	8.765	3.536	2.479	0.013
B013801=11-25 books	4.053	6.355	0.638	0.524
B013801=26-100 books	21.996	6.524	3.371	0.001
B013801=>100	24.589	7.055	3.485	0.000
B018101=1-2 days	-3.235	5.306	-0.610	0.542
B018101=3-4 days	-11.205	5.954	-1.882	0.060
B018101=5-10 days	-8.838	6.897	-1.281	0.200
B018101=More than 10 days	-23.108	8.641	-2.674	0.008
Percent American Indian	-0.164	0.042	-3.925	0.000
Root Mean Square Error	33.372	--	--	--

Slide 39 of 40

AM Software Resource

AM Online Manual/Help System

Slide 40 of 40

Module Summary and Resources

Summary

- Described the use of NAEPEX Software
- Described the use of AM Software
- Described considerations for analyzing NAEP data

Resources

- NAEP Data Toolkit
- NAEPEX Tutorial and User Guide
- AM Software Website
- NIES 2007 Example Questions
- AM Online Manual/Help System