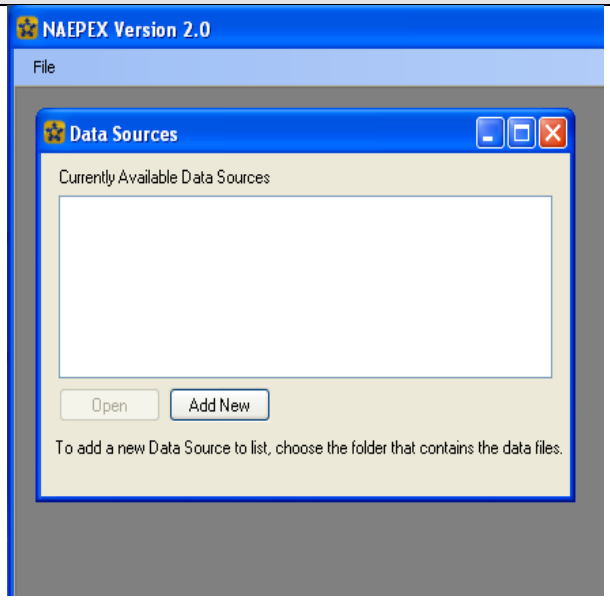
Resource Document to Facilitate Analysis of NAEP Data Using National Indian Education Study (NIES) Data and AM Software

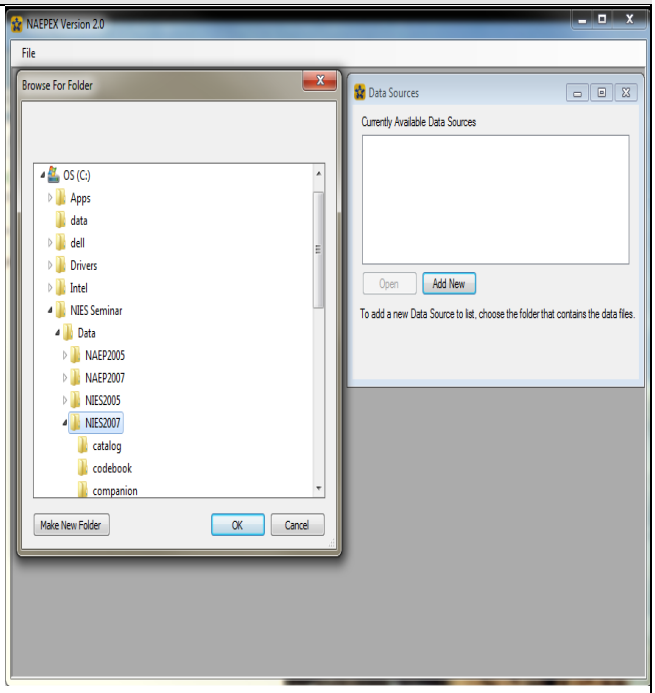
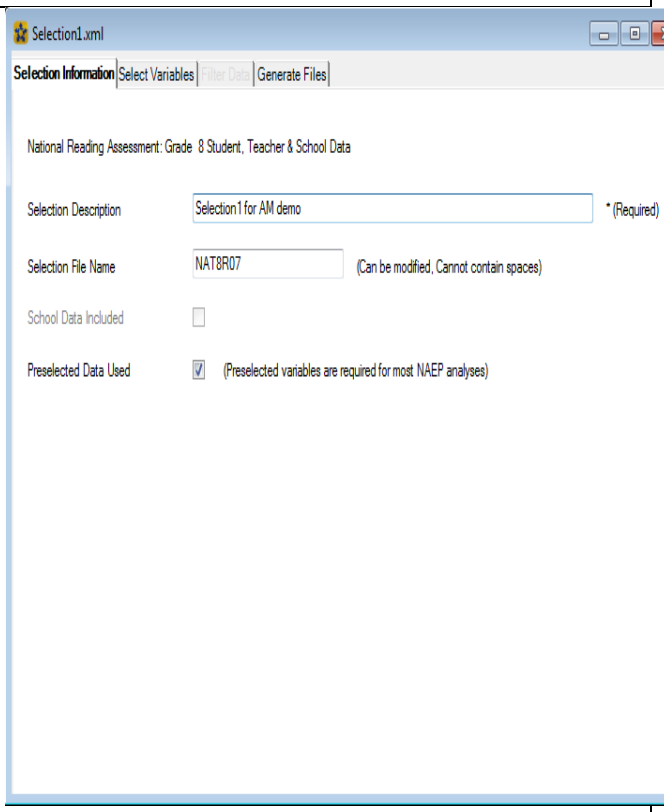
Purpose: To demonstrate how to conduct basic NIES dataset editing and analyses using AM software including how to:

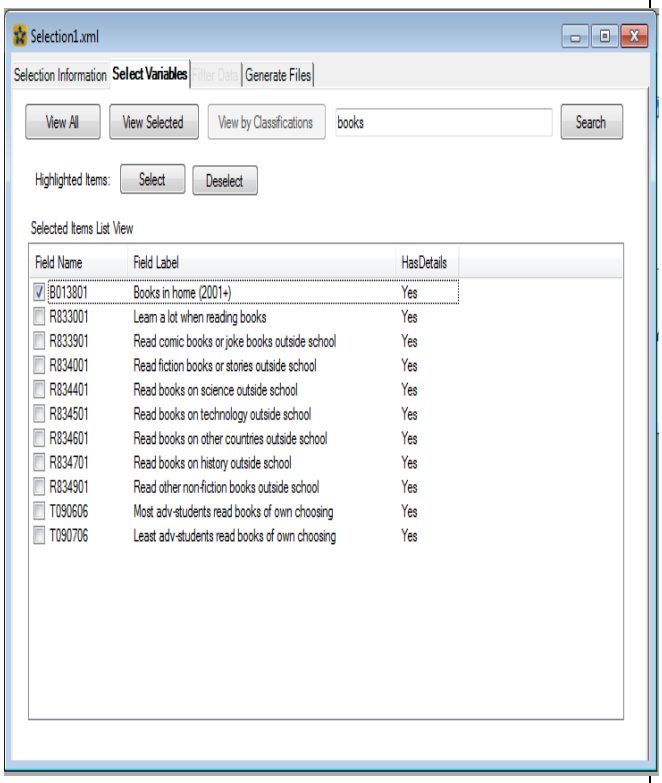
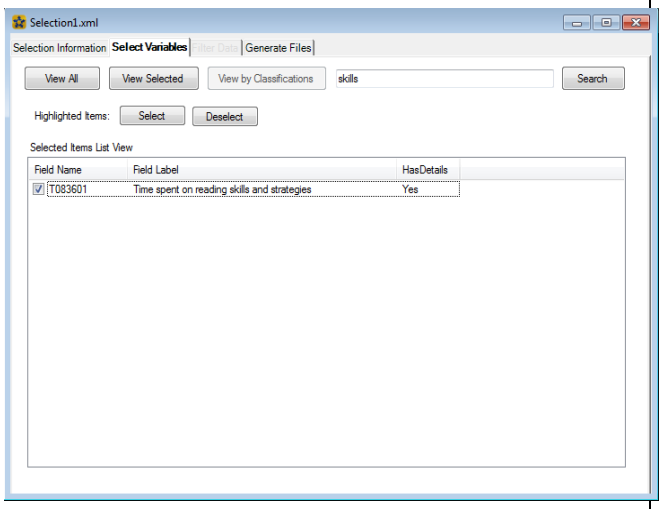
- Create an SPSS dataset using NAEPEX
- Import the SPSS dataset into AM
- Edit data with AM
- Develop descriptive statistics with AM
- Run a regression analysis with AM

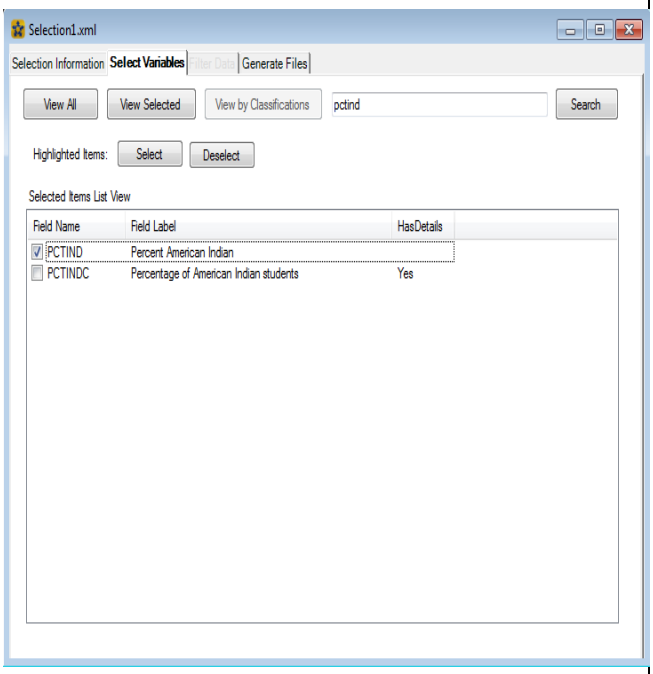
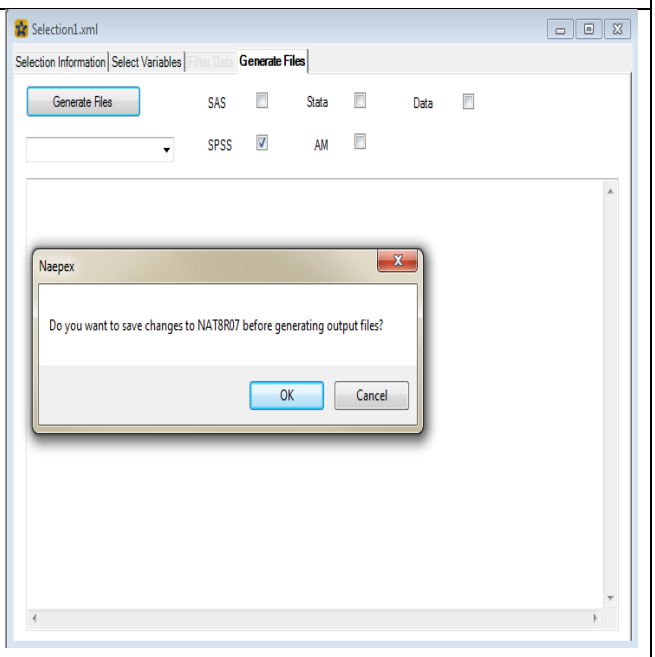
Analysis Questions to be addressed in this example:

1. 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?
2. What is the proportion distribution for days absent from school by gender for public school 8th grade students in 2007?
3. Is there a relationship between 2007 reading achievement in Midwest region among 8th grade public school students and absenteeism, the number of books at home, while controlling for student's gender and the percentage of American Indian students in the school?

General Instruction	Computer Screen
Step 1: Starting the NAEPEX Program On the desktop, click on **NAEPEX2.0.** Click on "**Add New**" in the Data Sources Window.	

General Instruction	Computer Screen
Step 1: Starting the NAEPEX (cont.) In the "**Browse For Folder**" window, navigate to the appropriate location, where the data is stored, that is C:\NIES Seminar\Data\NIES2007 Once you click OK, it will take some time for the data to be loaded. Once the loading is completed, click OK to have access to the Data Sources window where the newly registered data will appear, select **2007 National Indian Education Study (NIES) parts I and II grades 4 and 8** data source, and click on **Open**. Next, select the desired data which is: *National Reading Assessment: Grade 8 Student, Teacher & School data,* and then click **Open Data File**.	
Step 2: Defining Data File The NAEPEX window is displayed in the adjacent figure. 1) In the Selection Description box, type: *Selection1 for AM demo.* 2) In the Selection File Name box, type: *NAT8R07*.	

General Instruction	Computer Screen
Step 3: Selecting Variables *Select the following variables by clicking the check boxes:* B013801 → Number of books in home B018101 → Days absent from school last month T083601 → Percent of time spent on reading skills and strategies. Under the **Select Variables** tab, type **BOOKS** in the **Search String** box and click **Search**. The window will show that t**here were 11 records Found That Matched 'BOOKS'.** Click on the variable **B013801** to choose that variable. It will appear under the view selected tab. Similarly search for **ABSENT** and add the variable **B018101**. Click **OK.** Under the **Select Variables** tab, type **SKILLS** to search variables for "time spent on reading skills and strategies". Follow the same procedure above and search for **SKILLS,** and add the variable **T083601.** *Variables can be deselected by clicking a checked box, which will uncheck the box, deselecting the variable.*	 

General Instruction	Computer Screen
Step 3: Selecting Variables (cont.) Search and select the following 3 variables: SCHTYPE → School type PCTIND → Percent American Indian RPTSAMP → Reporting sample The following variables (necessary for our analysis) are already pre-selected: 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	
Step 4: Creating SPSS Syntax File To display all the selected variables, click on the **View Selected** tab. After examining the listing, click on "**Generate Files**" tab. Check the **SPSS** box and click on "**Generate Files**" button. Then click OK to save changes to NAT8R07. Minimize NAEPEX.	

General Instruction	Computer Screen

Step 5: Creating SPSS Data File

Open SPSS to run the syntax file and generate an SPSS dataset.

On the desktop, click on the **IBM SPSS** icon or go to **Start** button, and then point to **Programs**, then point to **IBM SPSS 19**.

On the Menu Bar, go to **File**, select **Open**, and then select **Syntax**.

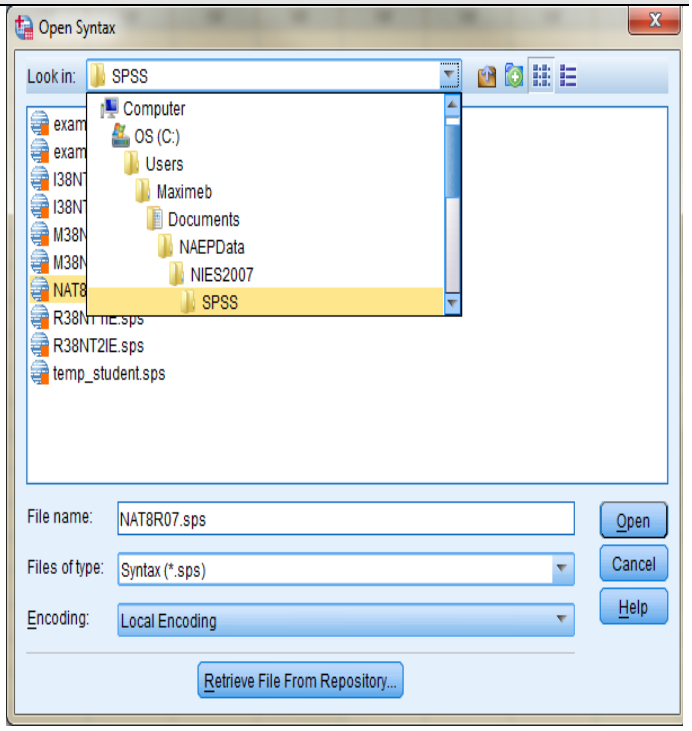
Navigate to the location, **'C:\users\user\Documents\NAEPData\NIES 2007\SPSS ',** where the syntax file generated by the NAEPEX was stored. Select **the syntax file, NAT8R07.sps**, and then click on **Open**.

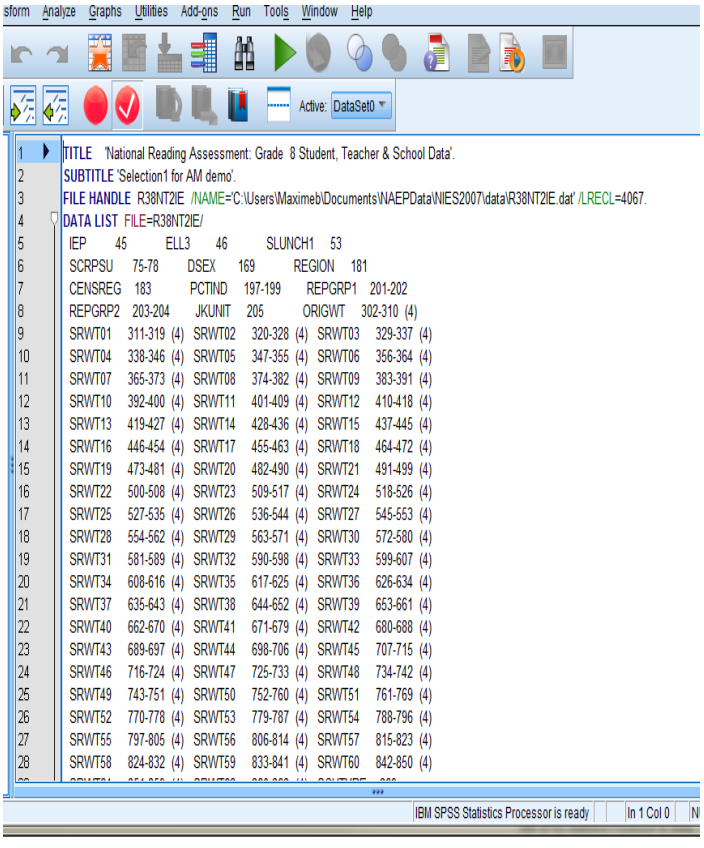
From the menu bar select **Run-ALL**.

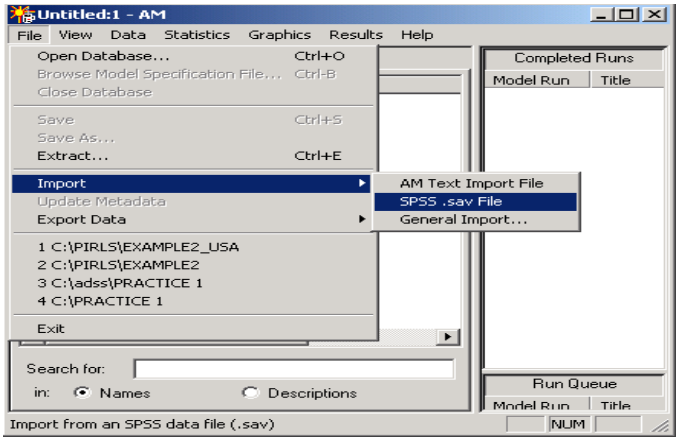
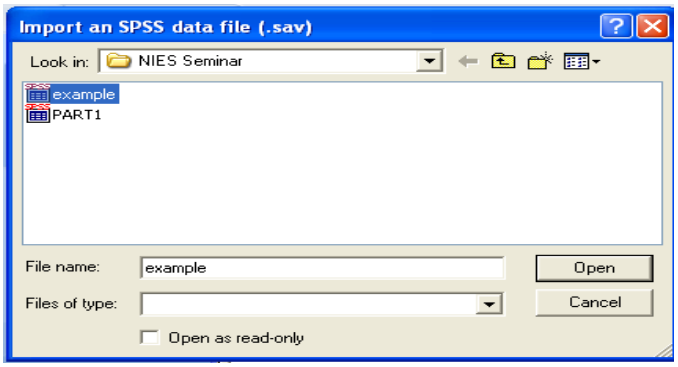
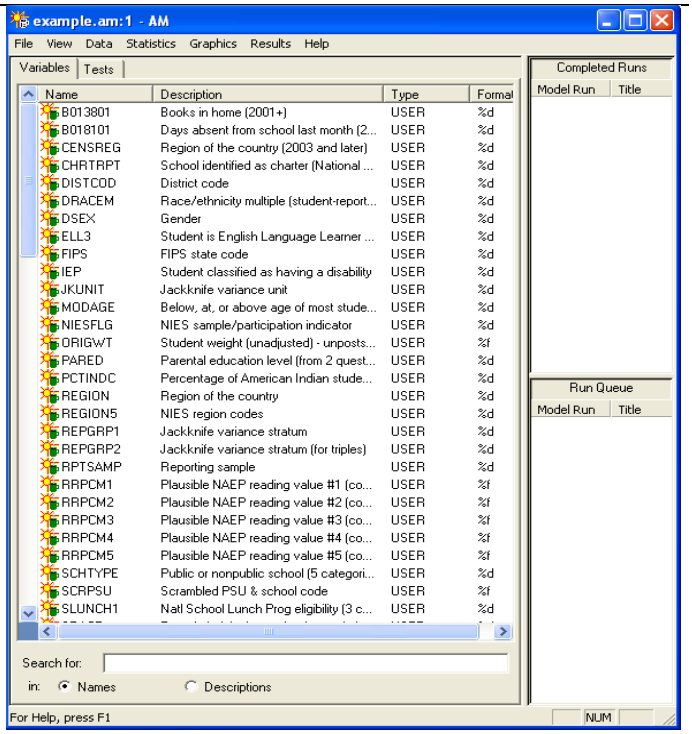
To save the file as SPSS data file, go to **File** on SPSS Data Editor, then **Save As…**

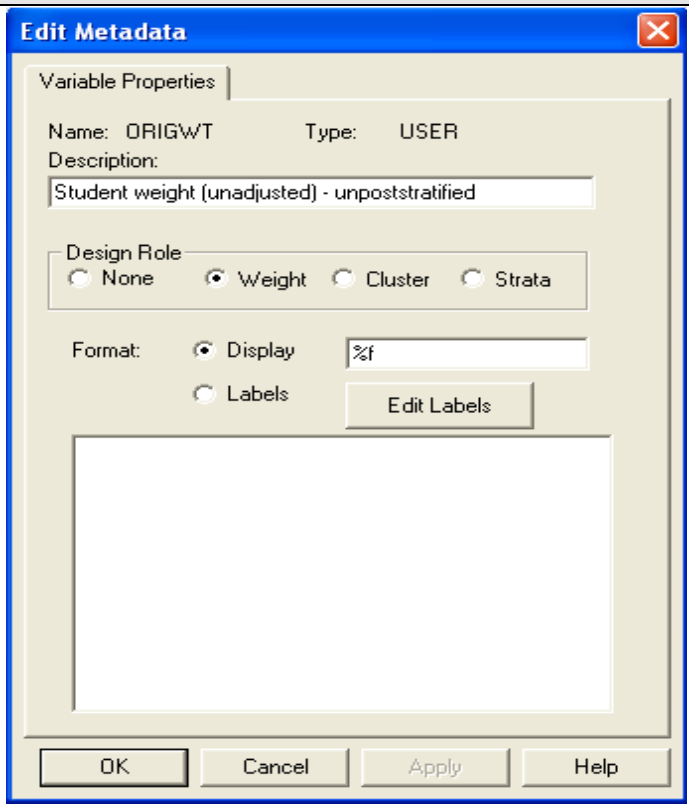
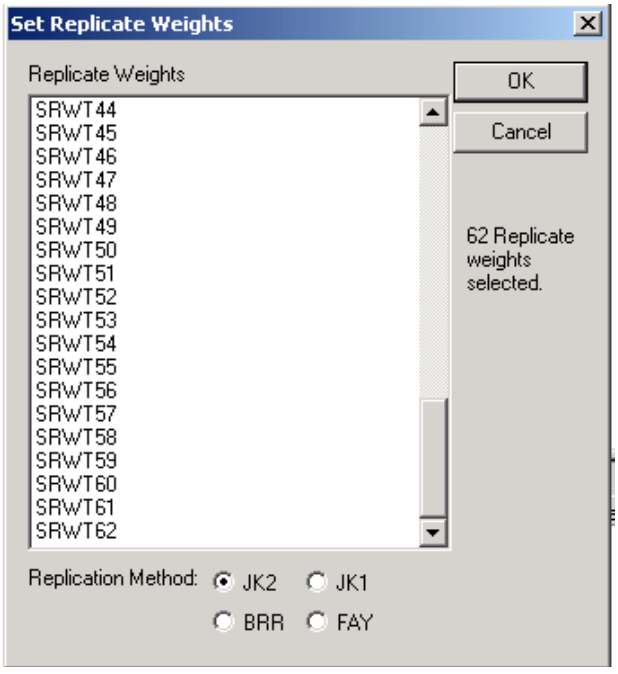
Save file as "**example.sav**" at the location "C:\NIES Seminar\".

Close SPSS.





General Instruction	Computer Screen
Step 6: Starting AM Software Click on the **Start** button, and then point to **Programs**, and click on **American Institutes for Research**, then click on **AM v.0.06.** The introductory screen will appear. Click on the screen to start. On the Menu Bar, go to **File**, select **Import**, and then select **SPSS.sav file.** A dialog box will appear. Give the path and file name of the SPSS data we saved in the last step; **C:\NIES Seminar\example.sav** Click on **Open**.	 
Step 7: Saving file in AM The window on the right will display. To save the file as an AM data file, go to **File**, then **Save As…** Save file as "**example.am**" at the location "C:\NIES Seminar\".	

General Instruction	Computer Screen
Step 8: Identifying the "Design Role" Variables In the window on the right, define **Weight**, **Cluster** and **Strata**. 1) Right click on the variable name **ORIGWT**, then click on "**Edit Metadata**", in the dialogue box, check the appropriate "**Design Role**" button -- **Weight**. Click **OK** 2) Right click on **REPGRP1** (Strata Variable) and **JKUNIT** (Cluster Variable) respectively, then click on "**Edit Metadata**", in the dialogue box, check "**Strata**" and "**Cluster**" button respectively. Click on the **OK**. *Notice that the little coffee cups next to the design role variables have changed from green to blue.*	
Step 9: Identifying replicate weights. Right click on any variable, and click on **Set Replicate Weights** from the menu. Highlight all replicate weight variables (SRWT01- SRWT62) and **drag** them into the Replicate Weights box of the Set Replicate Weights window. Set the replication method to **JK2.** Click on the **OK.** *Notice that the little coffee cups next to the replicate weight variables have changed from green to blue.*	

General Instruction	Computer Screen

Step 10: Data Editing

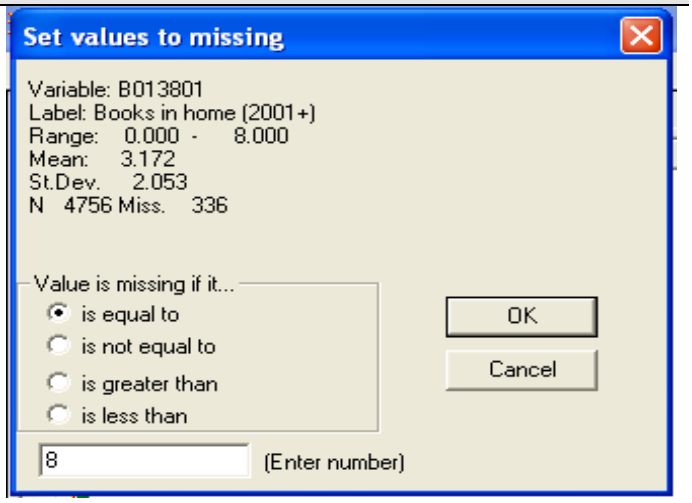
a) Define missing values:
Define missing values for the following variables:

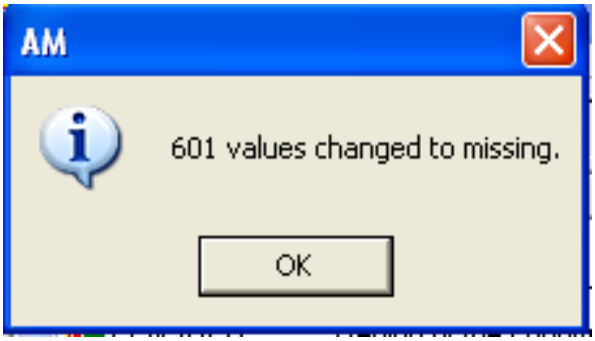
B013801 → Number of books at home *(**Set 0 & 8 to missing**);*
B018101 → Number of days absent from school last month (**Set 0 & 8 to missing**).
T083601 →Percent of time spent on reading skills and sstrategies (**Set 0 & 8 to missing**);

Right click on the variable name and choose "**Set values to missing**", the dialogue box on the right opens. Enter a number that represents missing values, then click on **OK**.

Repeat the procedure for other variables with missing values.

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





b) Recoding

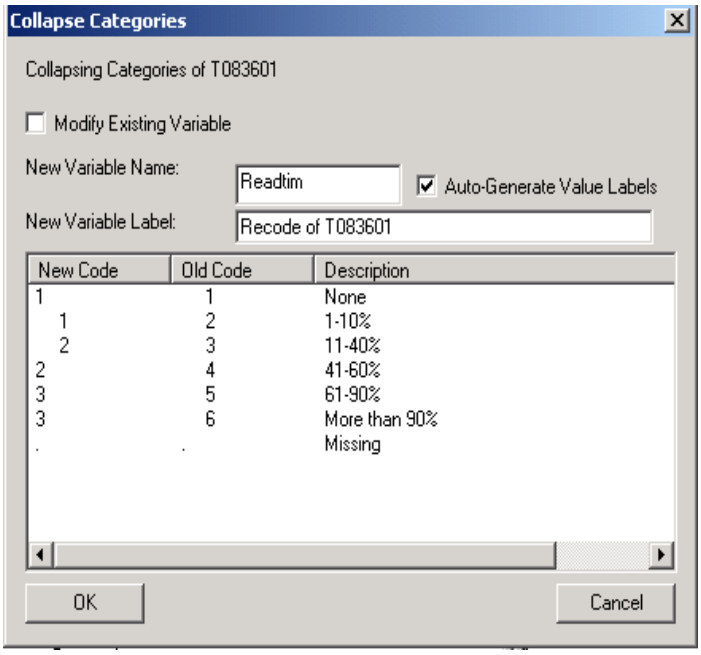
To recode time spent on reading skills and strategies (**T083601)**, **right click** the variable name, choose "**collapse categories**"; and then type a new name ***Readtim.***
Recode as follows:

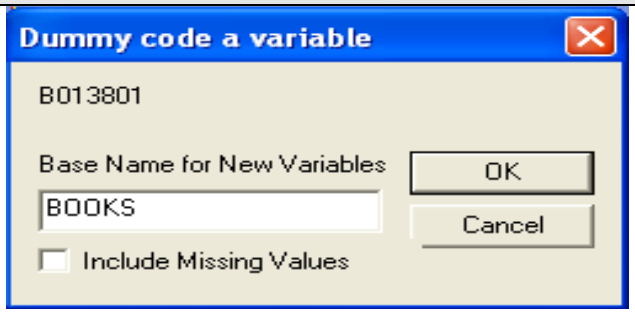
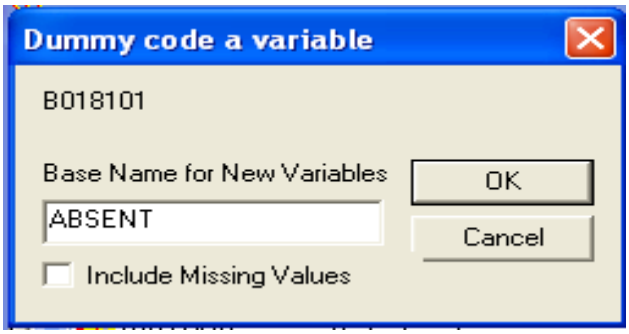
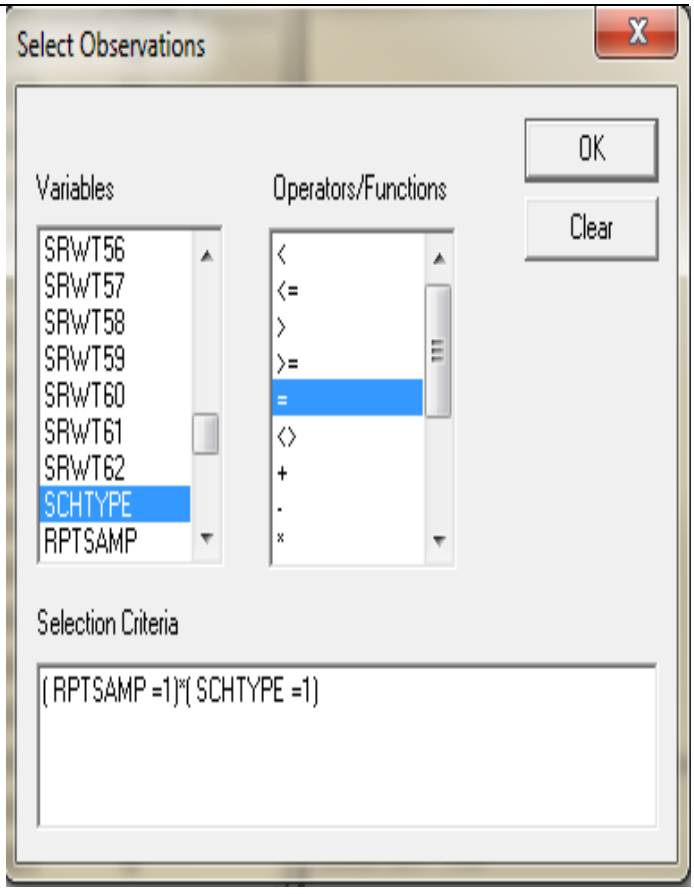
T083601		***Readtim***
1,2	→	1
3,4	→	2
5,6	→	3

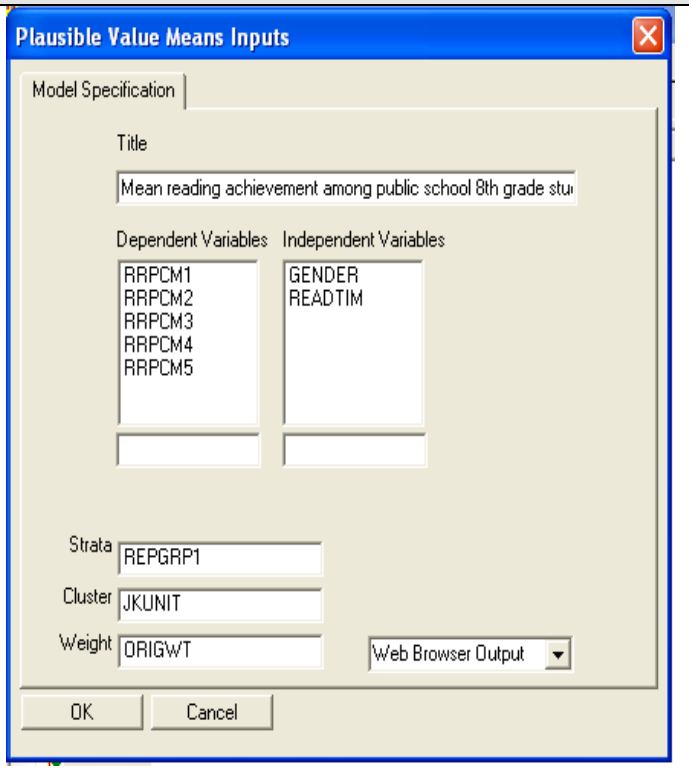
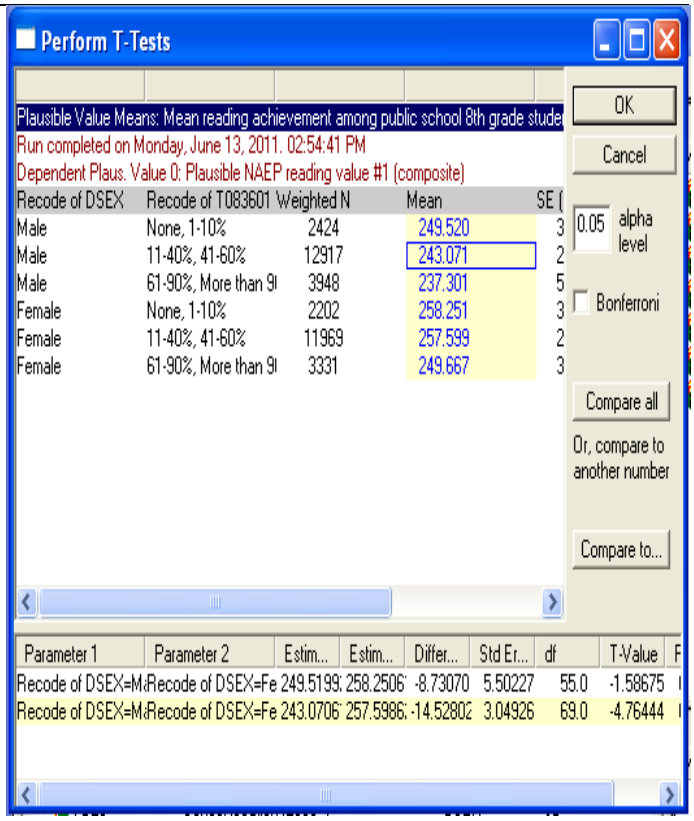
Click on **OK**.

Similary recode the varaible

DSEX		***GENDER***
1	→	0 (Male)
2	→	1 (Female)



General Instruction	Computer Screen
c) Creating Dummy variables Right click on the variable name **B013801** (Number of books at home) which has 4 categories, then click on **"dummy code this variable",** and then enter a new variable name **BOOKS** in the **"Base name for new variables"** window. AM then generates a series of dummy variables for each old variable by adding a number starting with 0 to the end of the base variable name. Scroll down to bottom of variable list to view new dummy variables. Click on **OK.** **Similarly, dummy code the variable, B018101;** and enter a new base variable **ABSENT**.	 
Step 11: Setting a Filter This analysis is based on **public school 8th grade (**SCHTYPE=1**)** students who were in the **reporting sample (**RPTSAMP=1**)** To select this dataset, go to **"DATA" → "FILTER OBSERVATIONS"** Type in the **SELECTION CRITERIA** window: (RPTSAMP=1)*(SCHTYPE =1) Click on the **OK.**	

General Instruction	Computer Screen
Step 12: Running a Crosstab (1st Analysis) In the main menu, go **to "Statistics"—>" Plausible Values Procedures"—>"Means"** Type in Title "***Mean reading achievement among public school 8th grade students by gender and time spent on reading skills and strategies***" Drag the dependent variables (**RRPCM1-5**) into the "**Dependent Variables**" box and the Independent Variables (**GENDER** and **READTIM**) into the "**Independent Variables**" box, and click on **OK**. ***View Output****: Results are opened in an internet explorer window once you click OK in the last step. Hard copy of the output is provided as an attachment. You can also choose to view an output in Spreadsheet or Text format by changing output options.*	
Step 13: Conducting a t-Test (1st Analysis) In the AM window—**Completed Runs** on the right hand side, Right click "PV Means Plausible Value Means", and click on "**t-Test**". The "**Perform T-Tests**" window opens, click on the mean of the given group—"**MALE - NONE, 1-10%"**, then click on the mean of a contrast group—"**FEMALE - NONE, 1-10%"** at the bottom of the screen, t-Test results appears. The t-test indicates that the difference is **not significant** at a 0.05 level of significance. Click on the mean of another contrast group—"**MALE - 11-40%, 41-60%**", then on "**FEMALE - 11-40%, 41-60%**" again a t-Test appears. It indicates that the difference is **significant** at a 0.05 level of significance; and click on **OK**. ***Notice that the yellow highlight indicates that there is a significant difference.***	

1st Analysis: Running a Crosstab

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

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

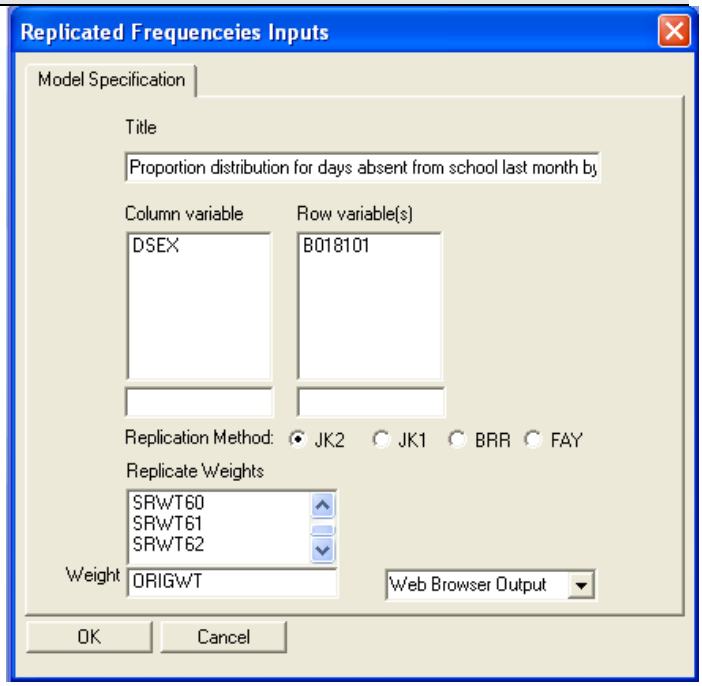
1st Analysis: Conducting a t-Test

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

TTest Results

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

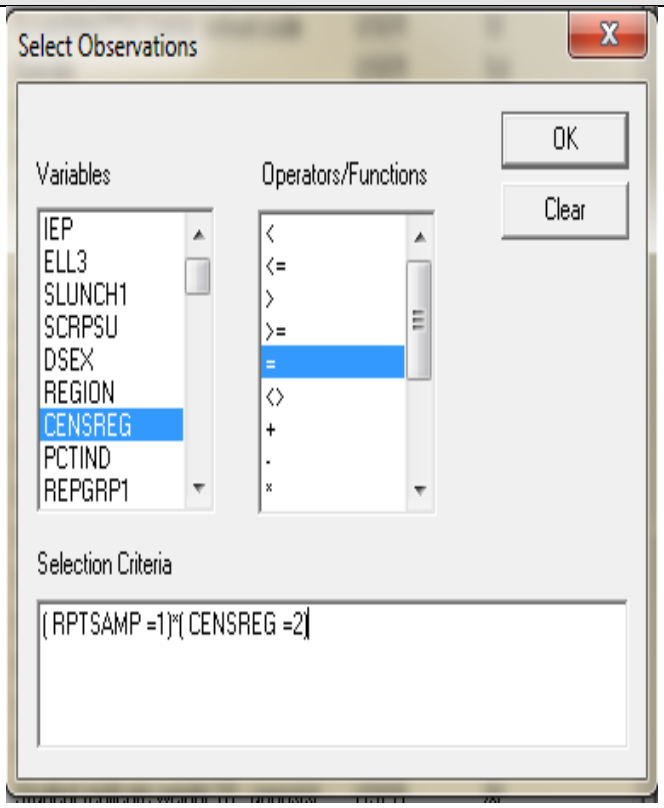
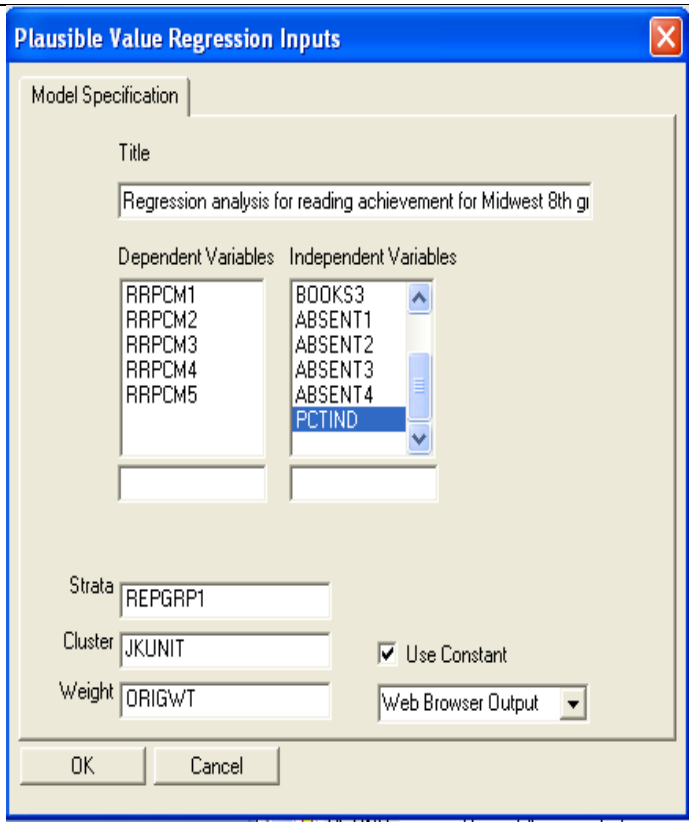
General Instruction	Computer Screen
Step 14: Proportion Distribution (2nd Analysis) In the main menu, go **to "Statistics"—>" Replication Procedures for Basic Statistic" —>"Frequencies"** Type in Title "***Proportion distribution for days absent from school by gender for public school 8th grade students*** " Drag the variable **B018101** (Days absent from school last month) into the "**Row Variable"** window, and **DSEX (**Gender**)** into the "**Column Variable**" window. Click the **OK**. ***Within the AM software package, it is important to take care when dragging dependent and independent variables into differnt model specification windows.***	

2nd Analysis: Proportion Distribution

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

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

		Gender			
Days absent from school last month (2001+)	**Weighted N**	**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

General Instruction	Computer Screen
Step 15: Regression Analysis (3rd Analysis) This analysis is based on 8th grade students in Midwest region who were in the reporting sample. Set the filter for: RPTSAMP=1 → Reporting sample CENSREG=2 → Midwest To select this dataset, go to **"DATA" → "FILTER OBSERVATIONS"** Type in the **SELECTION CRITERIA** window: (RPTSAMP=1)*(CENSREG=2) Click on the **OK.**	
Step 15: Regression Analysis (3rd Analysis) (cont.) In the main menu, go to "**Statistics**"—"**Plausible Values Procedure**"—"**Regression**" Type in title "***Regression analysis for reading achievement for Midwest 8th grade students"*** Drag the dependent variables (**RRPCM1-5**) into the "**Dependent Variable**" box; drag the independent variables (**GENDER, BOOKS1-3, ABSENT1-4** and **PCTIND**) into the "**Independent Variables**" box. Click on the **OK** to start the procedure.	

3rd Analysis: Regression Analysis

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

```
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	--	--	--

Notes about AM Software:

- Different analyses can be done using AM:
  - Selection of groups from the Data/Filter Observation menu
  - Using “Descriptive Statistics”
  - Frequency Tables/Cross Tabulations
  - Multiple regression analysis
- Variables can be manipulated using AM:
  - Recoding/Collapsing categories
  - Automatic dummy coding

More detail on these and other features are available in the AM online help, and within the AM User Guide. Both of these resources can be accessed via the AM website at http://am.air.org. Updated versions of AM will are made available (for free) from the website as they are released.