PowerStats

LEARN BY DOING:

Running a linear regression and interpreting results

Need help?
nces.info@rti.org

http://nces.ed.gov/datalab

THE REGRESSION YOU'LL CREATE

This tutorial will guide you through the steps taken to create the linear regression shown below.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

HYPOTHESIS TESTING RESULTS

BASIC STEPS REQUIRED TO CREATE A LINEAR REGRESSION

1. Choose a group of students (dataset).
2. Choose a type of regression.
3. Select a dependent variable.
4. Select one or more independent variables.
5. Run your regression.

1. CHOOSE A GROUP OF STUDENTS (DATASET)

INSTRUCTIONS

1. Drag *All undergraduates* to the Group box.

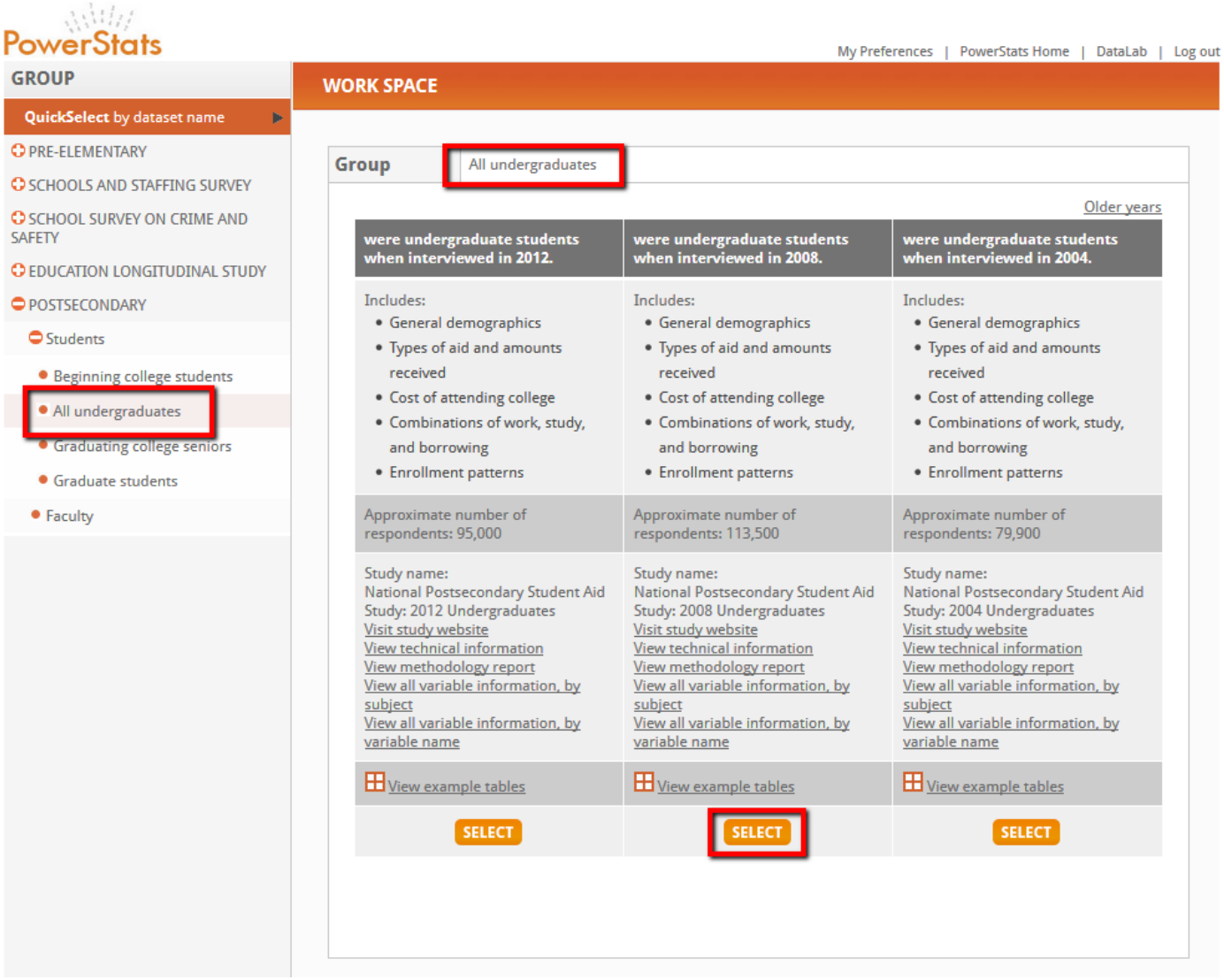
   The Group box is updated with your selection.

   Information about the dataset and the option to select it appear in the Work Space.

2. Click *Select* in the box labeled "were undergraduate students when interviewed in 2008."

TIP
Use the *QuickSelect by dataset name* option if you are familiar with NCES postsecondary studies.

NOTE
If a dataset contains more than one weight variable, you will be prompted to select one.

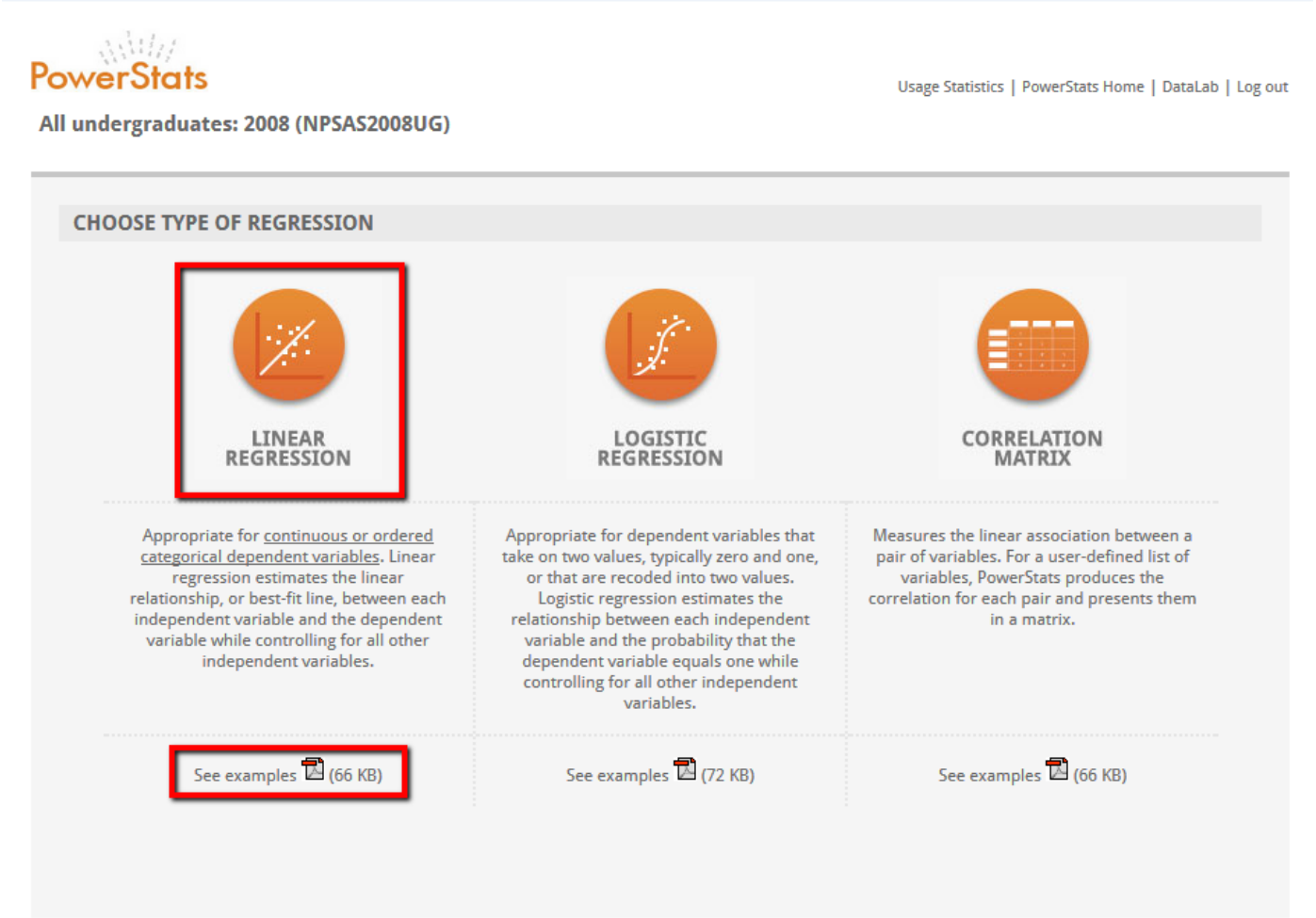


2. CHOOSE A TYPE OF REGRESSION

INSTRUCTIONS

1. Click the *Linear Regression* icon.

TIP
Click *See examples* under each regression type to learn more.



3. SELECT A DEPENDENT VARIABLE

INSTRUCTIONS

1. Click next to *Financial aid* to view available variables.

2. Click *Aid total amount*.

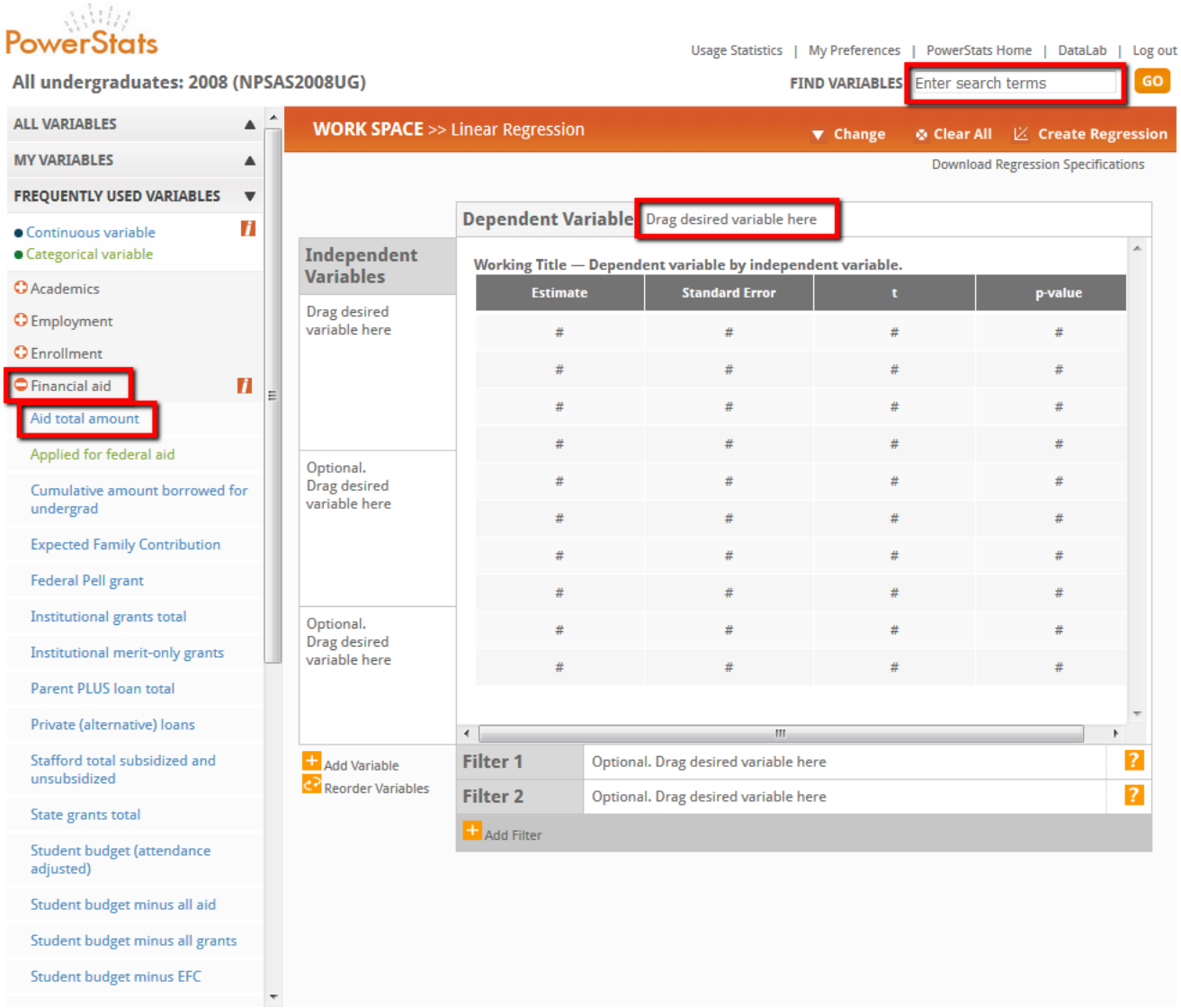
   The Variable Description screen appears.

3. Click *View descriptive statistics.* Note that *Aid total amount* is a continuous variable with a minimum value of 0 and a maximum value of 59,020.

4. Drag *Aid total amount* to the Dependent Variable box.

   The Usage Option layer appears.

TIP
Use the *Find Variables* option to search for variables by name and keyword.



3. SELECT A DEPENDENT VARIABLE, cont'd.

INSTRUCTIONS

5. Select *As a continuous variable*.

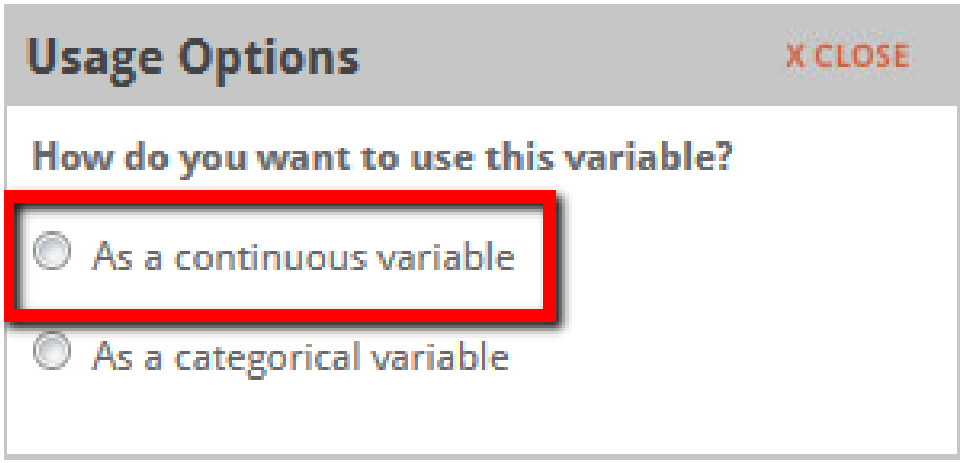
   The Continuous variable screen appears.

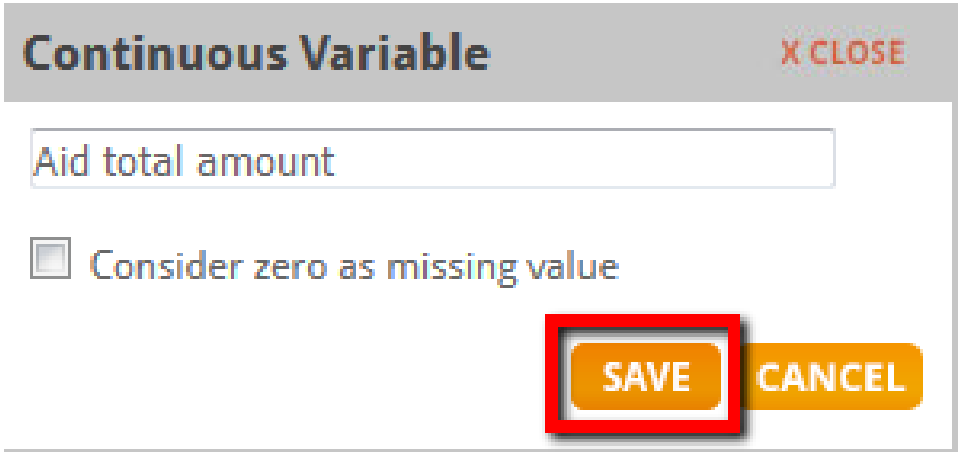
   Note: You're selecting this option because *Total aid amount* is the dependent variable in a linear regression.

6. Click *Save*.

   Do not check *Consider zero as a missing value* because it is possible for undergraduate students to receive no aid, and therefore, have zero as the aid amount.

   Note: This variable is continuous.





4. SELECT INDEPENDENT VARIABLE 1

INSTRUCTIONS

1. Click the ⊕ next to *Student's characteristics* to view available variables.

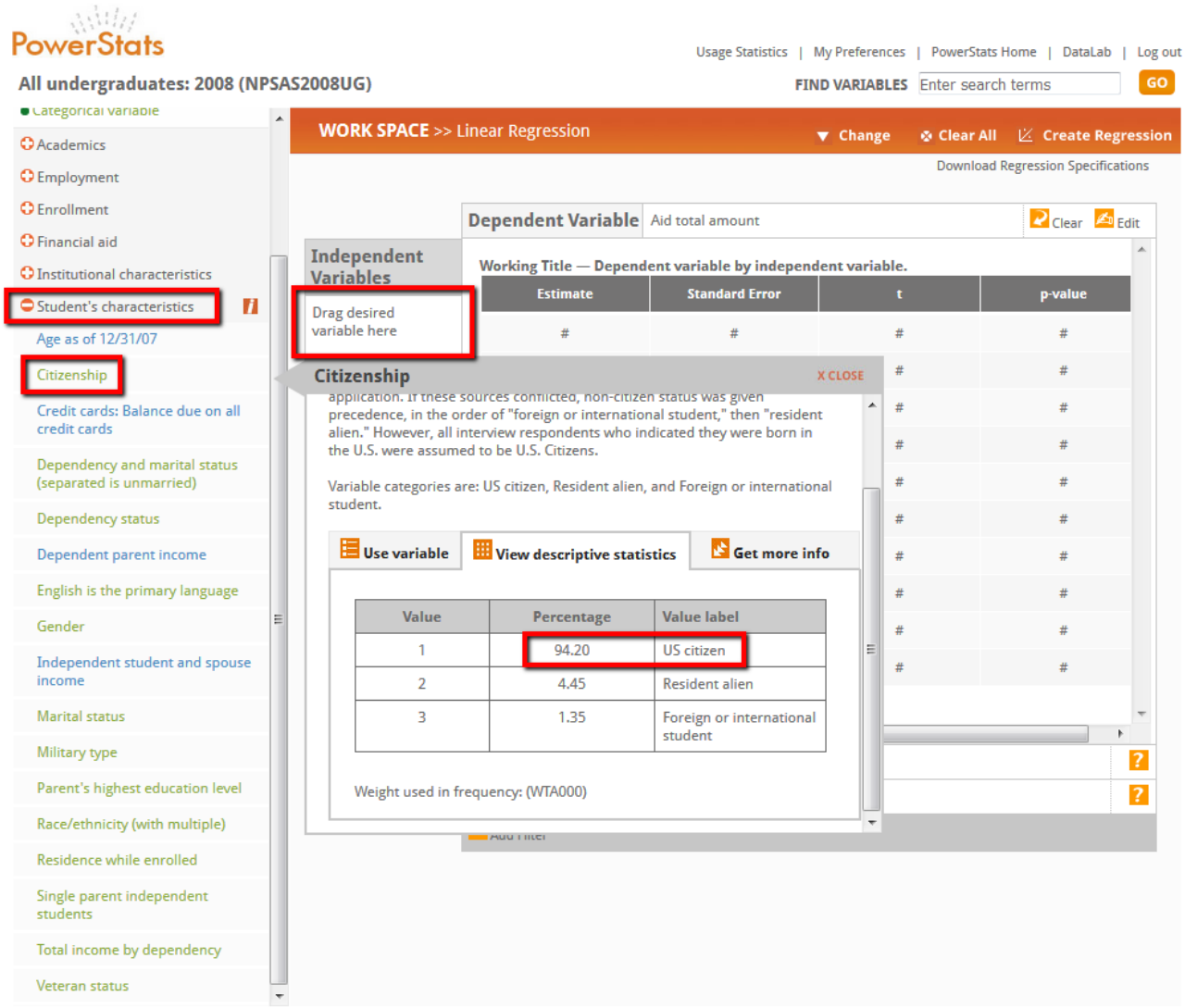
2. Click *Citizenship*.

   The Variable Information screen appears.

3. Click the *View descriptive statistics* tab. Note that 94% of the survey respondents are US citizens.

4. Drag the variable *Citizenship* to the first Independent Variables box.

   The Usage Option layer appears.



4. SELECT INDEPENDENT VARIABLE 1, cont'd.

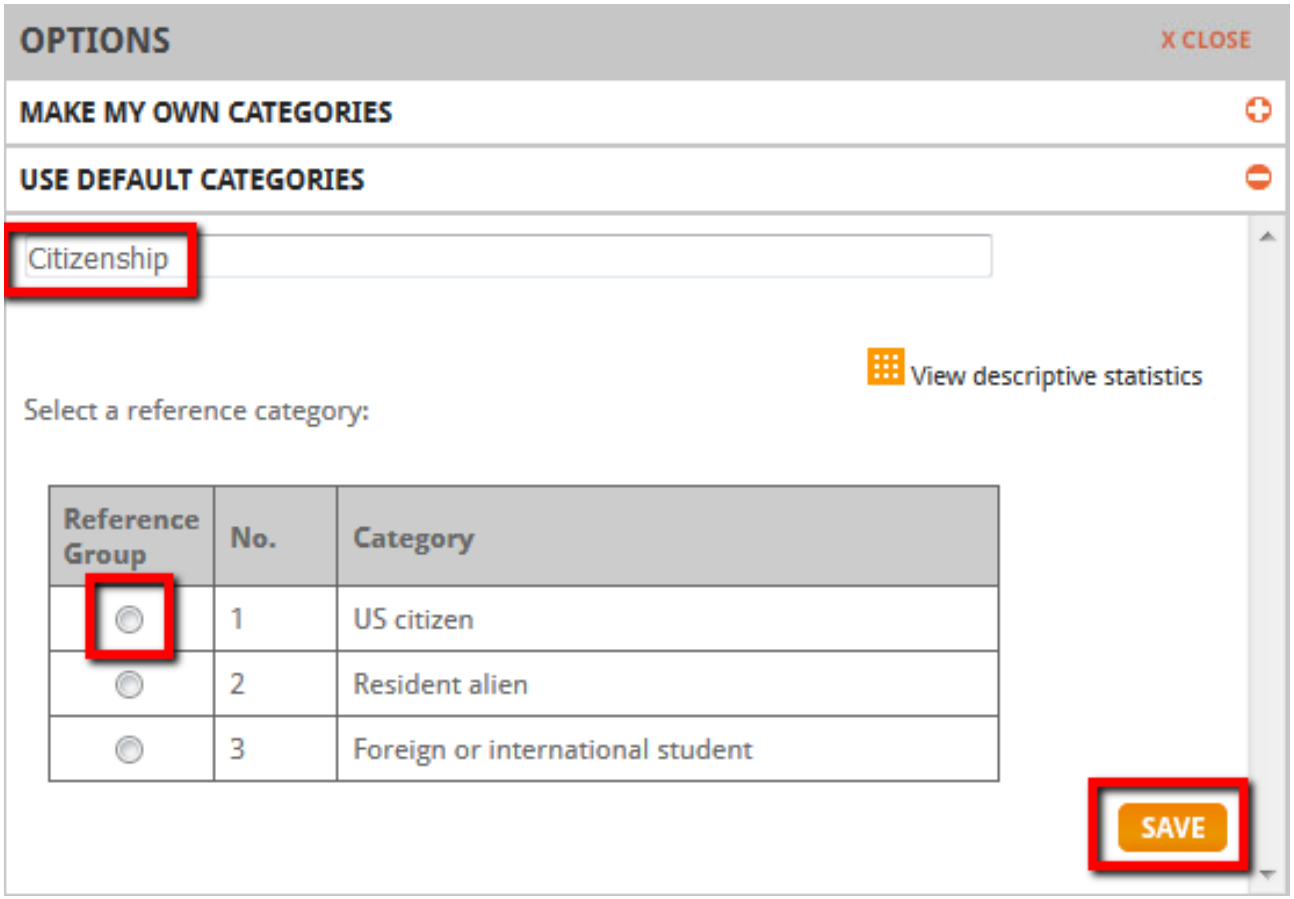
5. Click *US citizen* to use it as the reference group.

   Generally, you want to select the category with the largest number of responses as the reference group. Recall from the last page that 94% of survey respondents are US citizens.

   PowerStats will assign a value of zero to all US citizens and a value of one to all non-US citizens.

6. To make this distinction clear in the regression output, change the default variable label from *Citizenship* to *Non-citizen*.

7. Click *Save*.



4. SELECT INDEPENDENT VARIABLE 2

INSTRUCTIONS

1. In the Find Variables box, enter “income.”

   The matching variables appear in Search Results.

2. Click *Total income: Parents and independent.*

   The Variable Information screen appears.

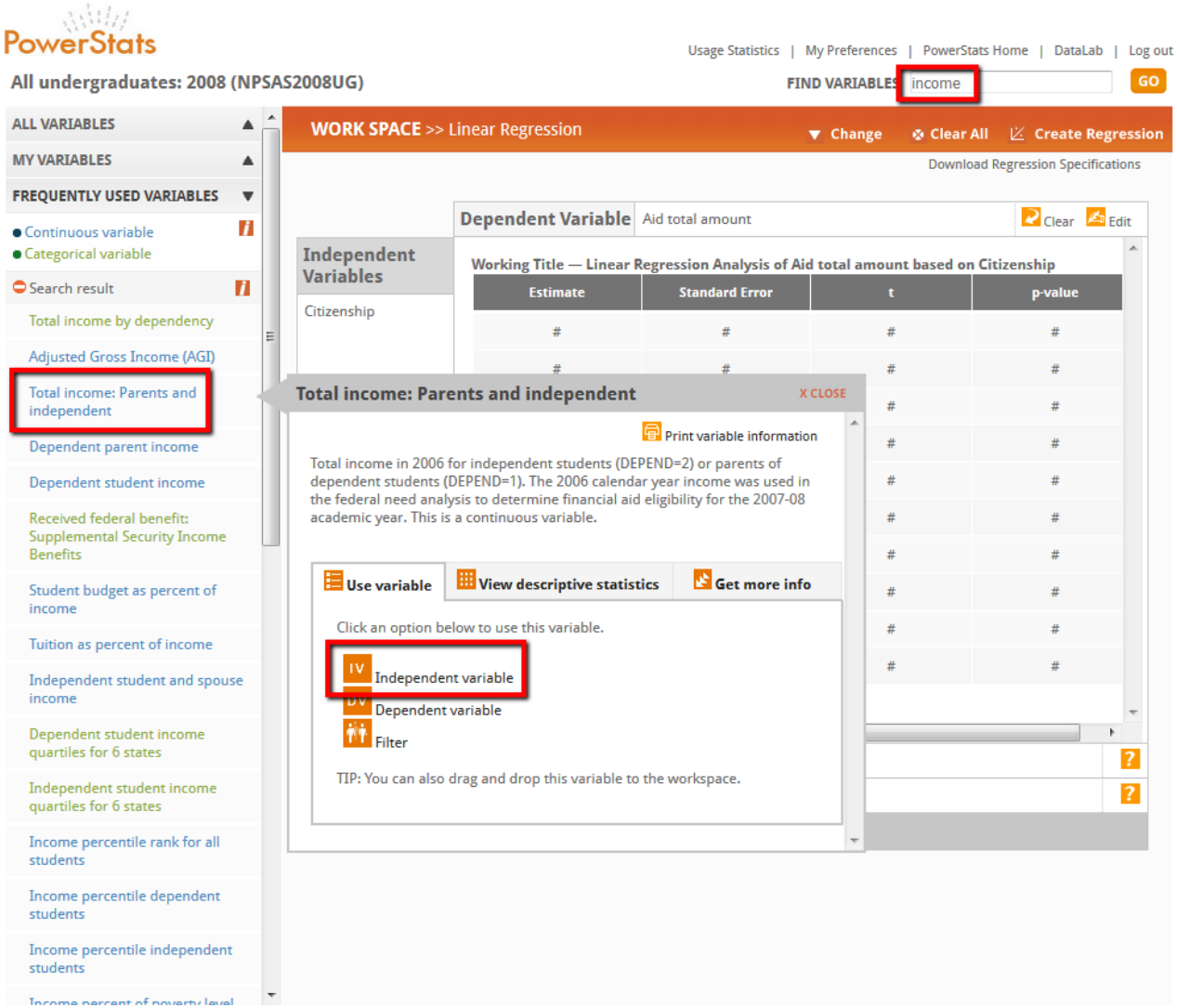
3. Click *Independent variable*.

   The Continuous Variable screen appears.

4. Select *As a continuous variable*.

   The Variable Options screen appears.

5. Leave *Consider zero as missing value* unchecked and click *Save*.



5. RUN THE REGRESSION

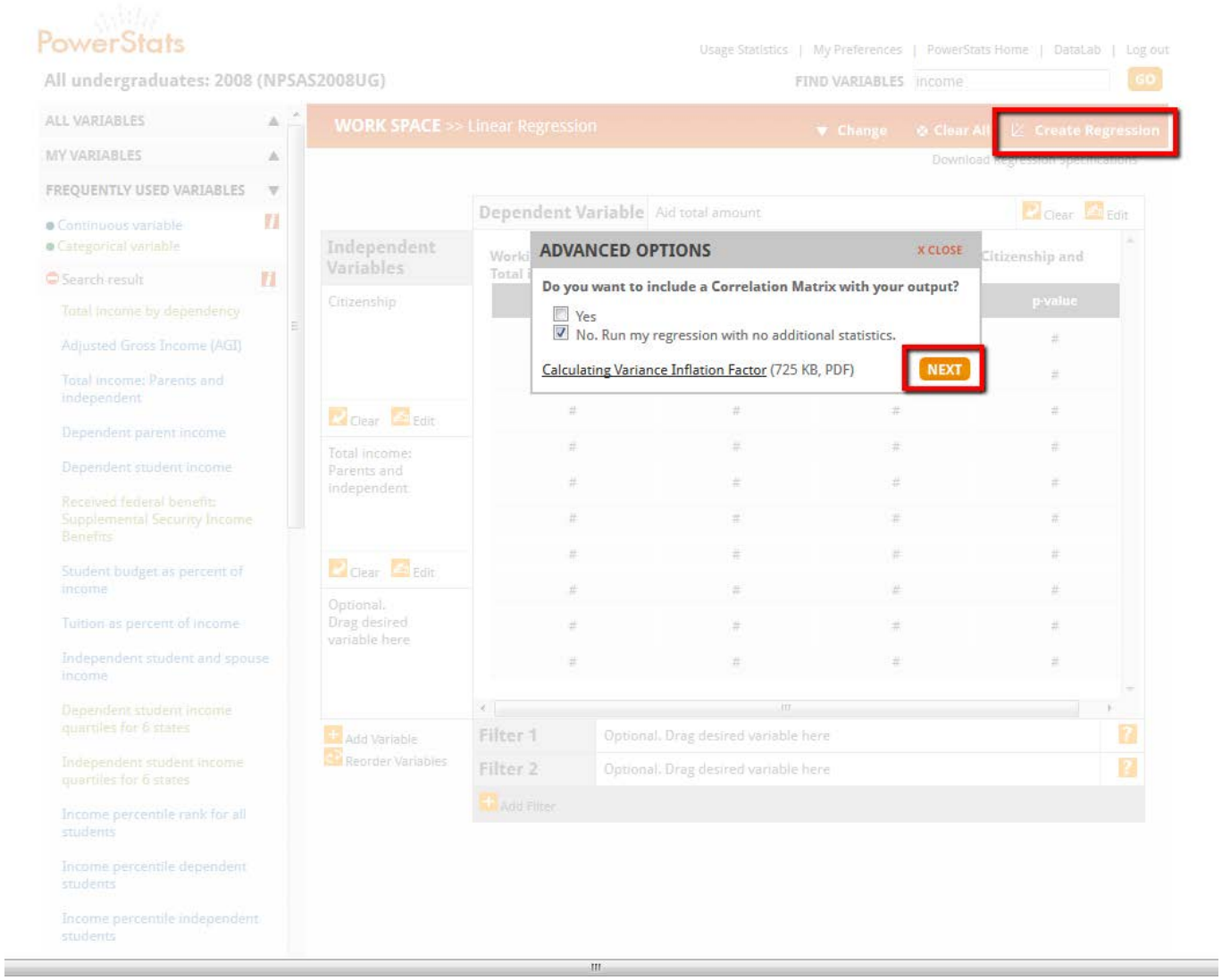
INSTRUCTIONS

1. Click *Create Regression*.

   The Advanced Options screen appears.

2. Without selecting additional statistics, click *Next*.

   PowerStats will begin to process the regression.



INTERPRETING LINEAR REGRESSION RESULTS

INTERPRETING COEFFICIENTS

The coefficient for the intercept or constant is the expected value of the dependent variable if all independent variables were set to zero. Here it implies that a student with a zero value for both *Non-citizen* (that is, a US citizen) and *Total income* would be expected to have an *Aid total* amount of $6,182.

TIP
Often it does not make sense to interpret the intercept at all. If age in years were one of the independent variables, for example, you should not calculate the expected value of *Aid total* for a person zero years old.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

HYPOTHESIS TESTING RESULTS

INTERPRETING LINEAR REGRESSION RESULTS, cont'd.

INTERPRETING COEFFICIENTS

Resident alien and *Foreign or international student* are treated as dummy variables, where 1 represents the presence of the category reported and zero represents its absence (or reference category). The -1286.157 coefficient for *Resident alien* implies that a resident alien would receive $1,286 less total aid than a US citizen, all else equal. Similarly, the -3230.233 value implies that a foreign or international student would receive $3,230 less total aid than a US citizen, all else equal.

TIP

"All else equal" means controlling for all other independent variables in the regression.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

Notice of Revision: In August 2013, NPSAS:08 weights were revised to match weighting procedures used in NPSAS:12.

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

INTERPRETING LINEAR REGRESSION RESULTS, cont'd.

INTERPRETING COEFFICIENTS

The -0.003 coefficient for *Total income* implies that for every additional dollar of total income a student has (as defined by NPSAS:08), they would be expected to receive $0.003 less in total aid, all else equal. In other words, for every additional $10,000 of income, a student would be expected to receive $30 less in total aid.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

Notice of Revision: In August 2013, NPSAS:08 weights were revised to match weighting procedures used in NPSAS:12.

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

INTERPRETING LINEAR REGRESSION RESULTS, cont'd.

INTERPRETING STANDARD ERRORS

A standard error (S.E.) represents the likely deviation of a sample statistic from the true population parameter due to sampling error. S.E.s are used to create the 95% confidence interval.

For illustration, the S.E. for *Resident alien* is 160.139, and the 95% confidence interval (the coefficient plus or minus 1.96 S.E.s) runs from -1601.952 to -970.363.

NOTE
PowerStats computes S.E.s using formulae that take the complex sampling design of the survey into account.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

Notice of Revision: In August 2013, NPSAS:08 weights were revised to match weighting procedures used in NPSAS:12.

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

INTERPRETING LINEAR REGRESSION RESULTS, cont'd.

INTERPRETING T-STATISTICS

Student's *t* is the ratio of the estimated coefficient divided by the standard error. If the absolute value of *t* is greater than or equal to 1.96, the coefficient is said to be statistically significant at the 95% confidence level. This means that even if the true coefficient were as much as two standard errors higher or lower than that estimated by the regression, there is 95% probability that the true coefficient would still be different from zero.

In the case of *Resident alien*, the t-statistic is -8.031. Its absolute value, 8.031, is greater than 1.96, so the coefficient for *Resident alien* is statistically significant at the 95% level.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

Notice of Revision: In August 2013, NPSAS:08 weights were revised to match weighting procedures used in NPSAS:12.

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

INTERPRETING LINEAR REGRESSION RESULTS, cont'd.

INTERPRETING P-VALUES

The p-value measures the probability that a sample would have yielded a coefficient of this magnitude due to random sampling error (also called sampling variation) if the true value of the coefficient were zero. Typically, a result is considered statistically significant if the p-value is less than .05.

NOTE

The p-values of .000 in this regression do not imply a zero likelihood that the coefficients were due to sampling error, but instead represent very small positive values less than .0005, and rounded to .000.

Linear Regression Analysis of Aid total amount based on Citizenship and Total income: Parents and independent

EXPAND ALL

REGRESSION MODEL INFORMATION

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

	b	S.E.	t	p-value	Lower 95%	Upper 95%
Intercept	6181.912	47.273	130.769	0.000	6088.688	6275.135
Citizenship						
Resident alien	-1286.157	160.139	-8.031	0.000	-1601.952	-970.363
Foreign or international student	-3230.233	226.517	-14.260	0.000	-3676.925	-2783.541
Total income: Parents and independent	-0.003	0.001	-3.833	0.000	-0.004	-0.001

Dependent variable: Aid total amount, .
For Citizenship, reference category includes: US citizen.

The names of the variables used in this regression are: CITIZEN2, TOTAID and CINCOME. The variable names are unique identifiers. To locate these variables, enter the variable name in the search box.

Source: U.S. Department of Education, National Center for Education Statistics, 2007-08 National Postsecondary Student Aid Study (NPSAS:08).

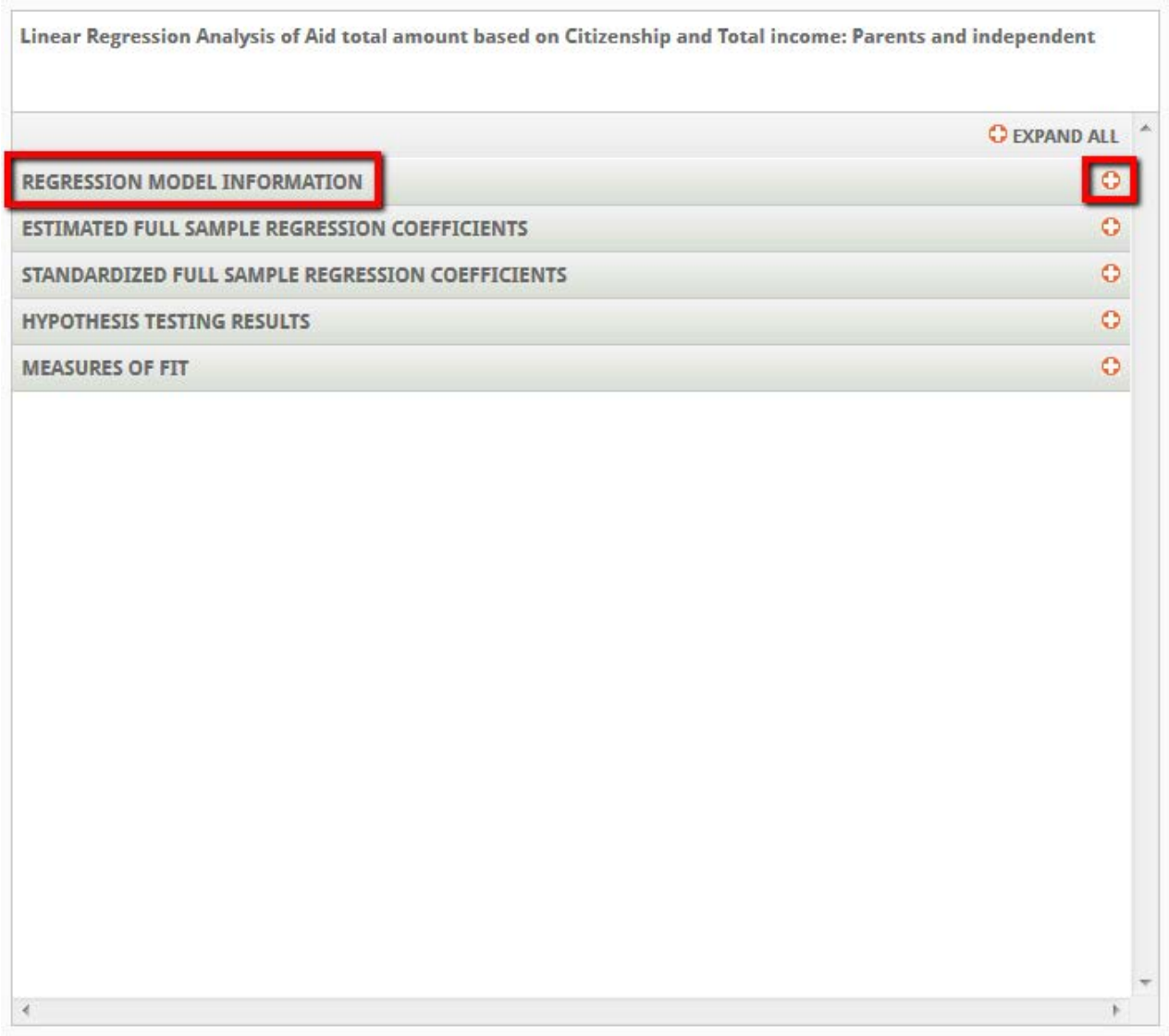
Notice of Revision: In August 2013, NPSAS:08 weights were revised to match weighting procedures used in NPSAS:12.

STANDARDIZED FULL SAMPLE REGRESSION COEFFICIENTS

OTHER FEATURES IN LINEAR REGRESSION

OTHER FEATURES

PowerStats reports other information from a linear regression, organized in the following areas:

1. Model information (sample size, variable names, estimation method)
2. Standardized coefficients
3. Statistics for hypothesis testing
4. Measures of goodness of fit (including R^2).

Click on the corresponding to view this information.