Instructions for Creating a New Categorical Variable

1. Select a dataset from the pull-down menu.
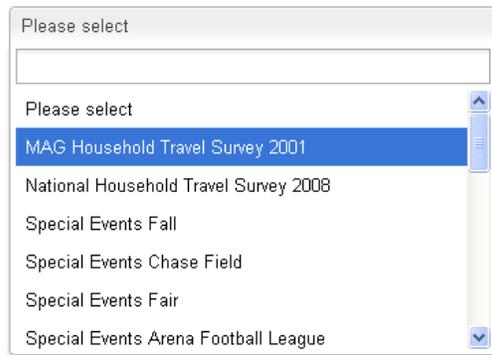
    Please select

    Please select
    MAG Household Travel Survey 2001
    National Household Travel Survey 2008
    Special Events Fall
    Special Events Chase Field
    Special Events Fair
    Special Events Arena Football League

2. Select a table from the pull-down menu.

    Please Select

    Please Select
    Household table for NHTS 2008
    Person table for NHTS 2008
    Vehicle table for NHTS 2008
    Trip table for NHTS 2008

3. Enter the name of a new *categorical* variable that you would like to create. Suppose you would like to create a new categorical variable called "hi_income" which is representative of households that have a high income, say a household family income that is greater than $100,000. Review the codebook to find the income variable in the data set. Note that the variable "hhfaminc" corresponding to a value of 18 implies that the household has an income over $100,000. Suppose the new variable we wish to create, "hi_income" takes a value of 1 if hhfaminc=18 and takes a value of 0 otherwise. Your entry would appear as follows:

    Name:

    hi_income

4. Next click on the button, **Add a Category**. A series of operators and fields will appear. You can use these operations to define the condition for each categorical value of the new variable.

5. Start by entering the first category value, say 1.

    Code: 1

6. Next, select a variable for defining the condition. In this example, we would like to define a condition that the new variable "hi_income" is equal to 1 if the variable "hhfaminc" is equal to 18. Choose "hhfaminc" from the dropdown menu and click the **Add** button to include it in the expression.

7. After the variable is selected and added, it will appear in the Expression box.
    At any point, you may click the **Clear** button to start over.

8. Choose the "=" operator among the numerical and algebraic operators available to you.   Also, enter a value of 18 (the value of the hhfaminc variable corresponding to household income greater than $100,000) in the Numeric Value field and click the **Add** button to include the numeric value in the expression box. Use the period to denote a decimal point. Use the comma to denote situations where multiple inputs are provided for an operation. For example, to compute x^y, one would have to enter **power (x,y)**.
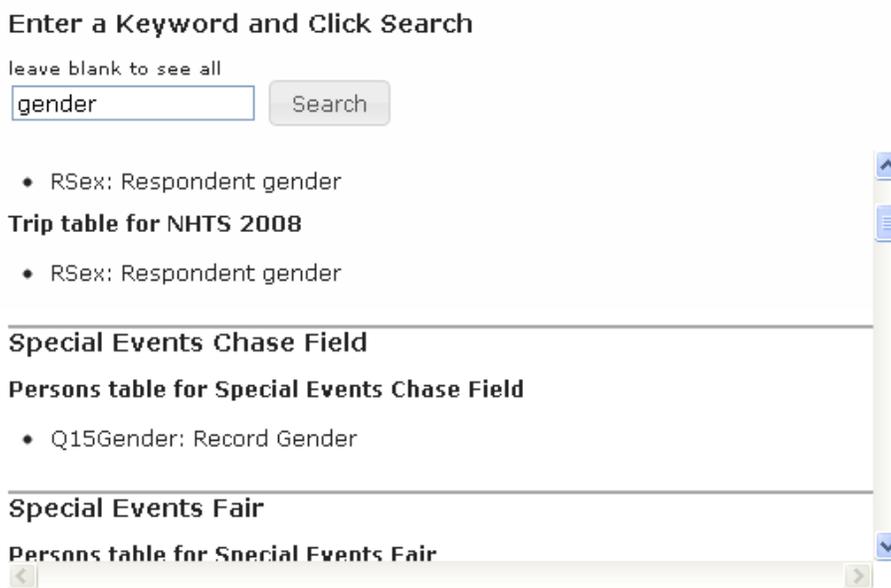

9. The Expression box should now be populated as follows.

The new variable takes on a code of 1 if hhfaminc = 18. Clicking the **Clear** button will result in the Expression being cleared so that you may re-enter the expression from scratch if desired. Clicking the **Delete** button will result in the elimination of the entire category/code.

10. Click on the button, **Add a Category**, to add another category for the new categorical variable. This process can be continued for as many categories as desired. After all desired categories are added, click on the **Save Variable** button to have the new variable appear in the right hand side list of Variables Added (tab).

   - hi_income: 1:Hhfaminc=18

11. Use the link to the <u>codebook</u> to find a variable of interest that exists in a data set table and its values. Alternatively, it is possible to search for a variable using the Variable Search tab on the right hand side of the screen.

Enter a Keyword and Click Search

leave blank to see all

```
gender
```
   Search

   - RSex: Respondent gender

Trip table for NHTS 2008

   - RSex: Respondent gender

___

Special Events Chase Field

Persons table for Special Events Chase Field

   - Q15Gender: Record Gender

___

Special Events Fair

Persons table for Special Events Fair

Using the "Variable Search" feature will allow you to identify variables that you may wish to use in creating a new categorical variable through the Expression generation process.

Note that new categorical variables will appear and be available for analysis only in the Cross Tabulation feature of the website. Categorical variables will not be available for univariate analysis or correlation analysis.