

Module 6

Creating New Variables

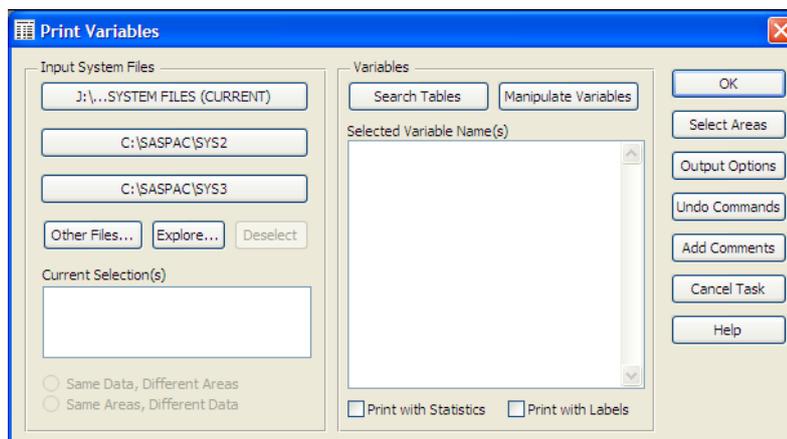
Even though the Census Area Statistics and Standard Tables between them contain some 75,000 variables, it is often the case that the information required is not directly available. The user may require a variable that is the sum of two or more other variables, or more commonly, may need to know a percentage or proportion. SASPAC provides facilities for these to be calculated and subsequently used in the same way as variables which already exist in the dataset.

Calculations within SASPAC are undertaken using the standard mathematical operators such as '+', '-', '*', and '/'. The use of these operators is facilitated by a set of on-screen buttons.

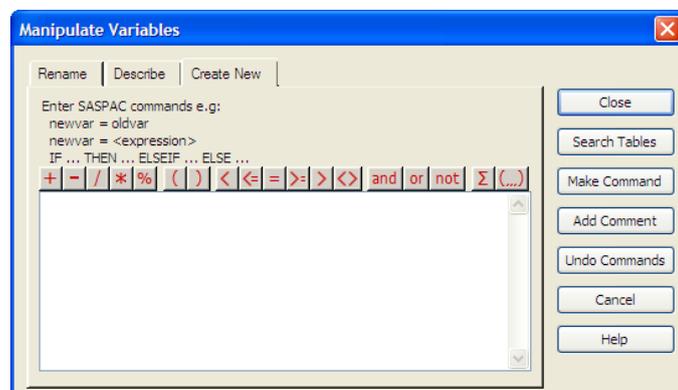
Example

For all Output Areas in the London Borough of Enfield, calculate the percentage of the population that is aged 11 to 13, and output, along with the total population, to a print file. Improve the column headings and change the report heading.

1. Again, the task is defined as producing variables as printed output, even though the variables required do not exist as such in the database. There is a need to create the variable which represents the required percentage. This is done through use of the 'Manipulate Variables' Command Button which appears on the main task window.



2. Use of this command button presents a window with three Tabs, and for this task, where we need to create a new variable, we need to use the third of these which bears the title 'Create New' as shown in the screen image.

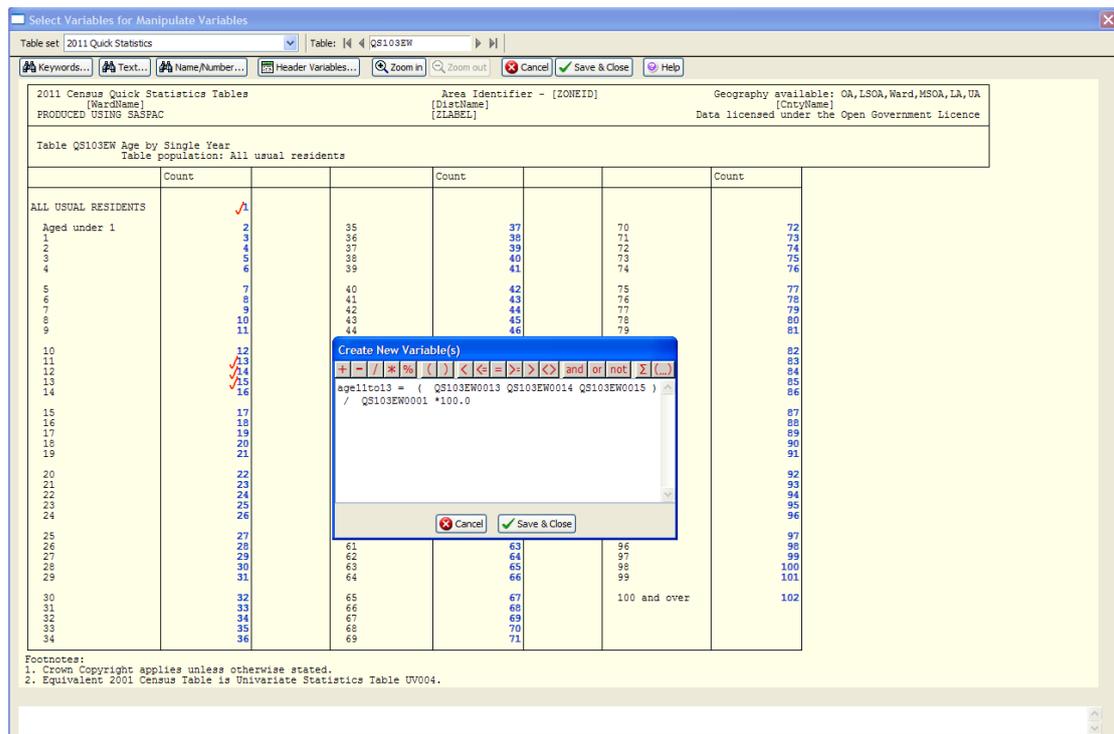


3. In this window, the new variable is created by first allocating a descriptive name to it. This name can have a maximum of 12 characters which can only be numeric or

alphabetic. No blanks or special characters may be used. We shall call it 'age15to24', and this name is typed into the text box. We need to create a command of the form:

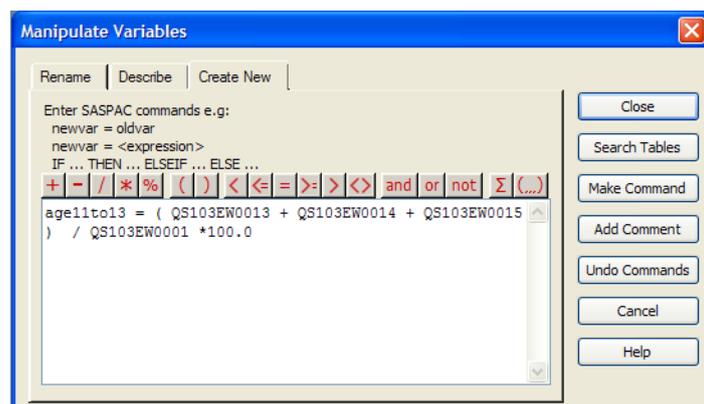
$$\text{age11to13} = \text{<mathematical calculation>}$$

Assistance in creating this mathematical calculation can be provided through use of the 'Search Tables' Command Button on the right of the window, and the symbol buttons above the text box. The Command Button performs the same functions as that to be found elsewhere, such as on the 'Print Variables' main task window. Using it allows the user to select the required variables on-screen as shown below.

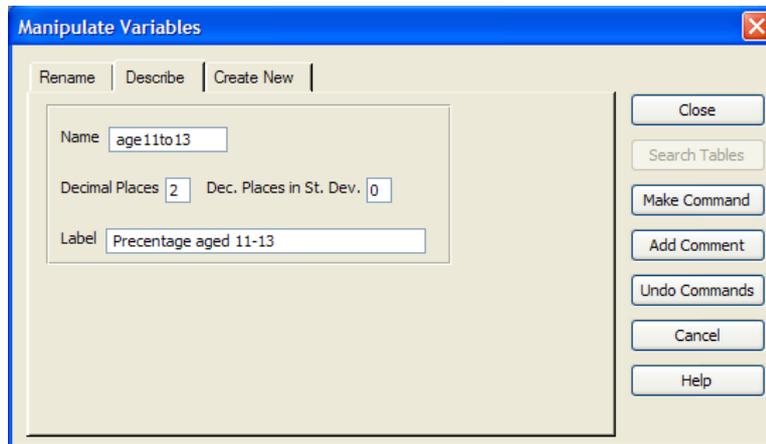


Note that to create the total aged 11 to 13, we need the single year of age data contained in tabel QS103EW. To calculate the percentage, we then need the total population from the same table. By selecting the variables in this order, they are made available in the correct order for creation of the new variable.

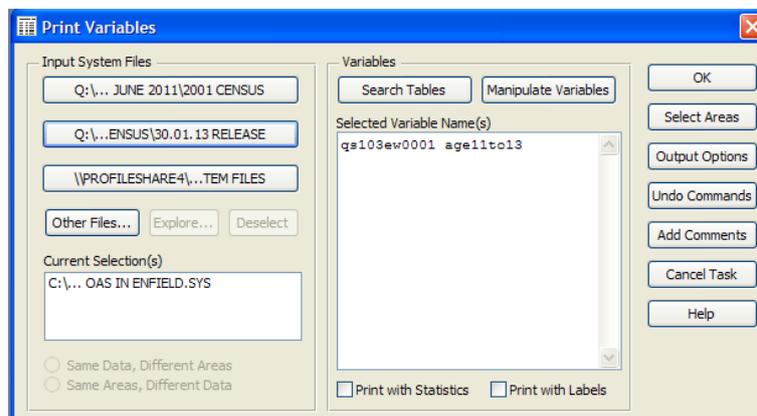
- When the 'Save & Close' Command Button is used, the selected variables are copied into the 'Manipulate Variables' text box, where the appropriate mathematical symbols can be added as shown here.



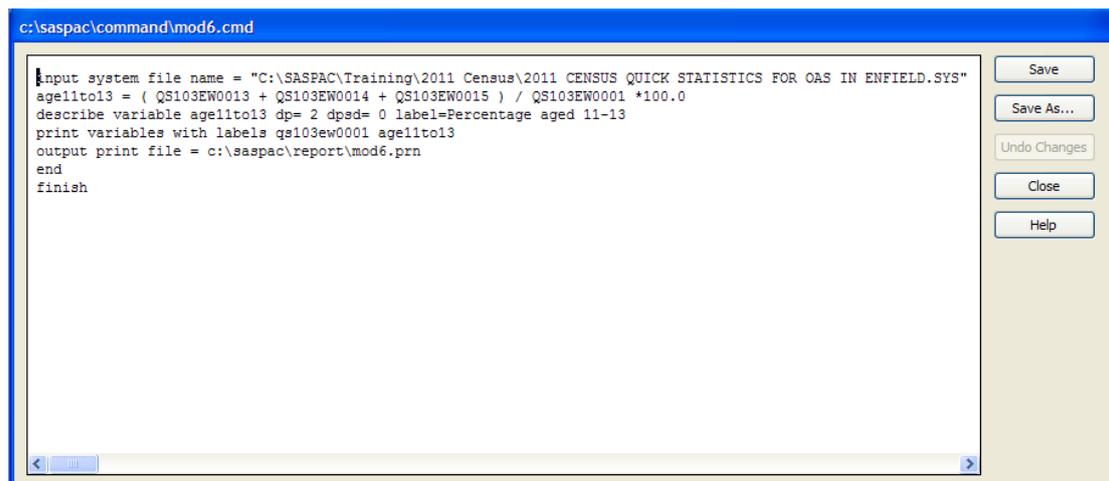
Once the command is complete, the 'Make Command' Command Button must be used to create the command, otherwise the command will be lost. Once this button is pressed, the text box is cleared, and the user has the option to create another new variable. In this case we have created the only new variable that we need, and so we could now use the 'Close' Command Button to return to the 'Print Variables' main task window. However, one of the other tabs in this window allows us to describe a variable by adding a label. This is done as shown in the following window.



4. After making this command and closing the window, we must ensure that the appropriate variables are listed for output, and that a suitable input system file has been selected as shown in the image below. As we have described one of the variables, we can use the 'Print with Labels' checkbox to make use of that label. The 'OK' Command Button may be used to proceed to the 'Task' window.



6. In the 'Task' window, a user-defined name is allocated, but instead of using the 'Save and Run' Radio Button, we will use 'Save Only', as there is a need to edit the generated command file to allow for the non-division by zero clause in the task. This is achieved by using the 'File / Open / Open Command File' menu option to open a dialog box in which the required file is identified and opened. (Alternatively, the file may be opened by accessing the 'Latest Files' list under the 'File' menu.). Whichever method is employed, the window shown on the next page will be presented to the user.



```

c:\saspac\command\mod6.cmd
input system file name = "C:\SASPAC\Training\2011 Census\2011 CENSUS QUICK STATISTICS FOR OAS IN ENFIELD.SYS"
age11to13 = ( QS103EW0013 + QS103EW0014 + QS103EW0015 ) / QS103EW0001 *100.0
describe variable age11to13 dp= 2 dpsd= 0 label=Percentage aged 11-13
print variables with labels qs103ew0001 age11to13
output print file = c:\saspac\report\mod6.prn
end
finish

```

This is a text box into which any editing may be typed directly. Note the presence of two consecutive forward slashes '/' on the first and second lines. SASPAC has inserted these automatically as line continuation symbols. When SASPAC encounters a '/' symbol at the end of a line it will automatically add the text on the next line to the first line.

In this case we have the need to divide by CS0010001 only if the value of CS0010001 is not zero. This means that the line which reads:

$$\text{age11to13} = (\text{QS103EW0013} + \text{QS103EW0014} + \text{QS103EW0015}) / \text{QS103EW0001} * 100.0$$

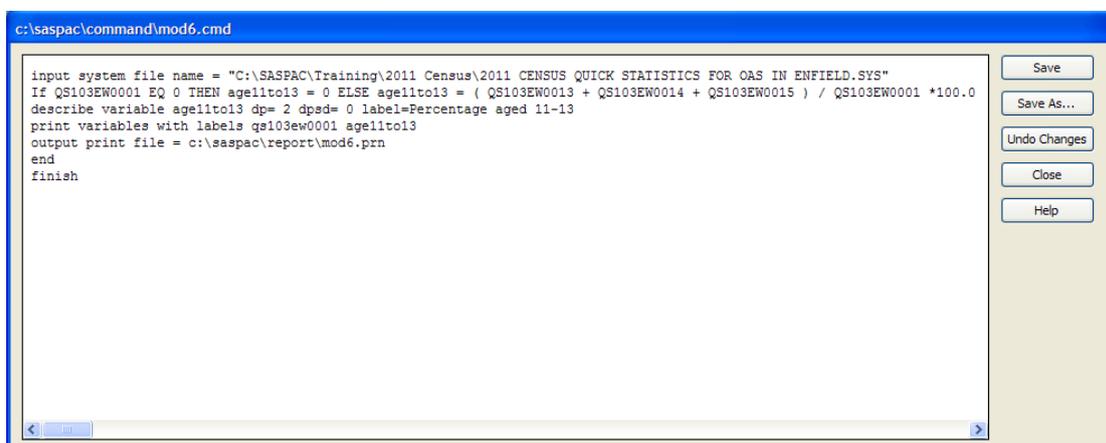
must be edited to read:

```

If QS103EW0001 EQ 0 THEN age11to13 = 0 ELSE
age11to13 = ( QS103EW0013 + QS103EW0014 + QS103EW0015 ) /
QS103EW0001 * 100.0

```

As a single line as shown in the next image. The user need not have any concerns as to the length of the resulting line. SASPAC will split it with a continuation symbol if necessary.



```

c:\saspac\command\mod6.cmd
input system file name = "C:\SASPAC\Training\2011 Census\2011 CENSUS QUICK STATISTICS FOR OAS IN ENFIELD.SYS"
If QS103EW0001 EQ 0 THEN age11to13 = 0 ELSE age11to13 = ( QS103EW0013 + QS103EW0014 + QS103EW0015 ) / QS103EW0001 *100.0
describe variable age11to13 dp= 2 dpsd= 0 label=Percentage aged 11-13
print variables with labels qs103ew0001 age11to13
output print file = c:\saspac\report\mod6.prn
end
finish

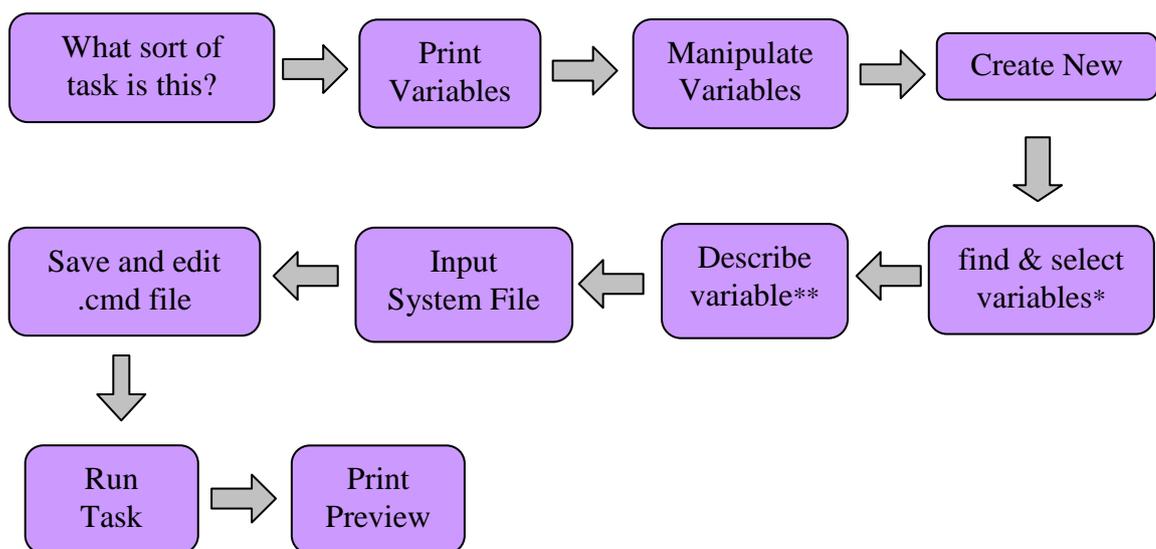
```

- When this command file is run by using the 'Save' and 'Close File and Run Task' menu option, the output that is obtained is as shown below. If the value of CS0010001 is zero for an area, the value of the new variable is set to 0.00, as defined, rather than the missing value symbol of '-'. Also note that the variable that was not described has the default label at the head of the column.

ZONE ID	QS103EW0001	Percentage aged 11-13
E00006910	310	1.61
E00006911	235	4.68
E00006912	336	2.68
E00006913	368	2.17
E00006914	284	1.41
E00006915	333	3.60
E00006916	298	1.34
E00006917	370	1.62
E00006918	362	2.21
E00006919	546	4.76
E00006920	396	1.77
E00006921	393	4.58
E00006922	338	2.96
E00006923	408	4.41
E00006924	311	3.22
E00006925	278	1.80
E00006926	327	2.14
E00006927	299	4.35
E00006928	372	6.99
E00006929	454	2.64
E00006930	516	3.29
E00006931	314	2.23
E00006932	344	0.87
E00006933	311	2.25

User Note: There may be a need to edit the label by adding in blanks at suitable places to ensure that the label splits appropriately over two lines as 12 by 12 characters.

Summary of task sequence:



*remember to 'make command' after creating your new variable

**remember to 'make command' after describing your new variable