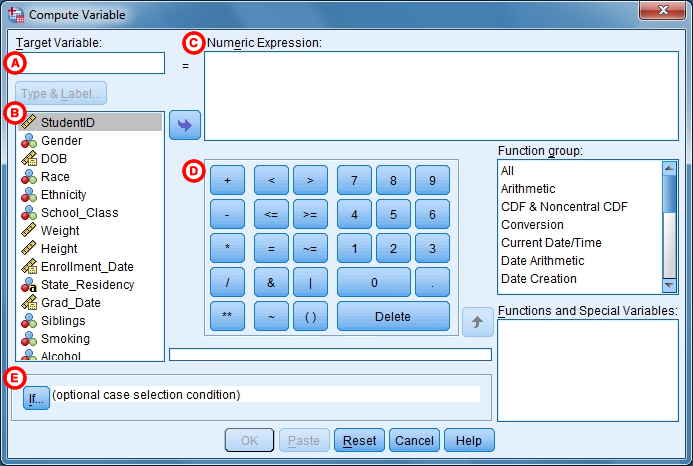
http://libguides.library.kent.edu/SPSS/ComputeVariables

Sometimes you may need to compute a new variable based on existing information (from other variables) in your data. For example, you may want to convert the units of a variable from feet to meters, or use a subject's height and weight to compute their BMI. In this tutorial, we'll discuss how to compute variables in SPSS using numeric expressions, built-in functions, and conditional logic.

To compute a new variable, click **Transform** **> Compute Variable**.

http://lgimages.s3.amazonaws.com/data/imagemanager/68594/spss_transform_compute_menu_opt.jpg

The Compute Variable window will open where you will specify how to calculate your new variable.



**A.** **Target Variable:** The name of the new variable that will be created during the computation. Simply type a name for the new variable in the text field. Once a variable is entered here, you can click on “Type & Label” to assign a variable type and give it a label. The default type will be numeric.

**B.** The left column lists all of the variables in your dataset. Select the variable(s) that will be used in your computation and click the arrow to move the selected variable(s) to the**Numeric Expression** text field, (3).

**C.** **Numeric Expression:** Specify how to compute the new variable by writing a numeric expression.

**D.** The center of the window includes a collection of arithmetic operators, Boolean operators, and numeric characters, which you can use to specify how your new variable will be calculated. There are many kinds of calculations you can specify by selecting a variable (or multiple variables) from the left column, moving them to the center text field, and using the blue buttons to specify values (e.g., “1”) and operations (e.g., +, \*, /).

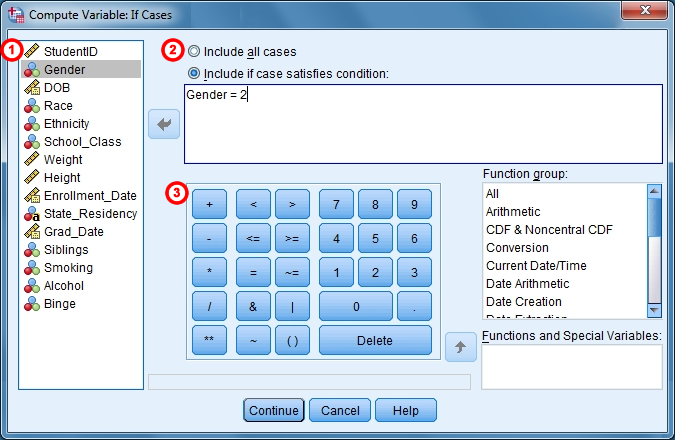
You can also use the built-in functions in the **Function group** list on the right-hand side of the window. The function group contains many useful, common functions that may be used for calculating values for new variables (e.g., mean, logarithm). To find a specific function, simply click one of the function groups in the **Function Group** list. You will now see a list of functions that belong to that function group in the**Functions and Special Variables** area. If you click on a specific function, a description of that function will appear in the text field to the left.

**E.** **If:**The **If** option allows you to specify the conditions under which your computation will be applied.

**Function group**and **Functions and Special Variables**: (Not labeled above) The **Function group** area shows different classes of SPSS's built-in functions. Clicking on one of the group names will populate the **Functions and Special Variables**box with a list of corresponding built-in functions.

**Using Conditional Logic to Compute New Variables**

Sometimes you may wish to compute values for a new variable that you are computing only when other conditions in your data are satisfied. To do this, click **If** (indicated by letter E in the above image) to open the Compute Variable: If Cases window.



**1.** The left column displays all of the variables in your dataset. You will use one or more variables to define the conditions under which your computation should be applied to the data.

**2.** The default specification is to **Include all cases**. To specify the conditions under which your computation should be applied, however, you will need to click **Include if case satisfies condition**. This will allow you to specify the conditions under which the computation will be applied to your data.

**3.** The center of the dialog box includes a collection of arithmetic operators, Boolean operators, and numeric characters, which you can use to specify the conditions under which your recode will be applied to the data. There are many kinds of conditions you can specify by selecting a variable (or multiple variables) from the left column, moving them to the center text field, and using the blue buttons to specify values (e.g., “1”) and operations (e.g., +, \*, /). You can also use the built-in functions in the **Function Group** list under the right column.

After you are finished defining the conditions under which your computation will be applied to the data, click **Continue**.

When you specify a condition in the Compute Variable: If Cases window, the formula you specify in the Compute Variable window will be applied to only those cases meeting the specified condition. If a case does not meet that condition, it will be assigned a missing value.

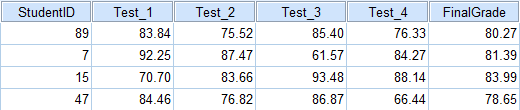
**Example: Computing a New Variable Using an Equation**

Now we will use what we have learned throughout this tutorial to demonstrate how to compute a new variable. In this example, we wish to compute a new variable called *FinalGrade* that is the average of four test scores—variables *Test\_1*, *Test\_2*, *Test\_3*, and *Test\_4*.

1. Click **Transform** **>** **Compute Variable**.
2. In the **Target Variable** field, type a name for the new variable that will be computed. Let's call our new variable *FinalGrade.*
3. Highlight each variable—*Test\_1*, *Test\_2*, *Test\_3*, and *Test\_4*—from the list on the left and click the arrow to move each variable to the **Numeric Expression** field. (Alternatively, you can double-click on the variable name to move it to the **Numeric Expression** field.) Make sure you click the spacebar to create a space between each variable.
4. Now your four variables will appear in the **Numeric Expression** field. Move your cursor between each set of variables and click the blue “+” sign to add the symbol for addition to the numeric expression. Now your expression should appear as “Test\_1 + Test\_2 + Test\_3 + Test\_4.”
5. Now insert parentheses around the expression so that it appears as “(Test\_1 + Test\_2 + Test\_3 + Test\_4).”
6. At the end of the expression, add the “/” sign and the number “4.” Now your expression should appear as “(Test\_1 + Test\_2 + Test\_3 + Test\_4) / 4.”
7. The final expression indicates that the new variable, *FinalGrade* will be calculated as the average of the four test scores.
8. Click **OK** to complete the computation and apply the changes to the data.
9. Finally, let’s make sure that a new variable called *FinalGrade* was successfully created.  
   * We can find the new variable in the last column in Data View or in the last row of Variable View. If you do not see the new variable, the computation was unsuccessful.
   * We can check the syntax that was executed by looking at the log in the Output Viewer window. After running Compute Variable, the syntax that should have appeared in the output window is:
   * **COMPUTE FinalGrade1=(Test\_1 + Test\_2 + Test\_3 + Test\_4) / 4.**
   * **EXECUTE.**

If there was an error in how the computation was specified, the log in the Output Viewer will often show an error message.

* + It is also useful to explore whether the computation you specified was applied correctly to the data. You can spot-check the computation by viewing your data in the Data View tab. To check that the new variable computed correctly, you can manually calculate the averages for a few cases in your dataset just to spot-check that the computation worked correctly.



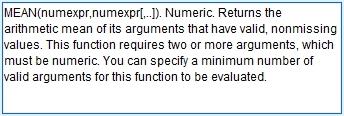
Let's check the case associated with StudentID 89:

http://lgimages.s3.amazonaws.com/data/imagemanager/68594/spss_transform_computevars_checkaverages_opt.gif

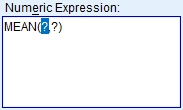
**Example: Computing a New Variable Using a Built-In Function**

Let's instead try computing the average test score using the built-in mean function.

1. Click **Transform > Compute Variable**.
2. In the **Target Variable** area, type a name for the new variable that will be computed; let's call the new variable *FinalGrade2.*
3. In the **Function group** list, click **All**.
4. In the **Functions and Special Variables**list, scroll down until you find “Mean”, then click on it. A description of this function will appear in the text box to the left. In this example, the description reads:



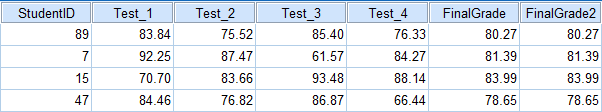
1. Double-click “Mean” under in the **Functions and Special Values** list. The basic setup for using this function will now appear in the**Numeric Expression**field.



1. Now add each of the test variables (i.e., *Test\_1* to *Test\_4*) to the numeric expression by double-clicking on the variable name in the left list. The variable names should be separated by commas, and all of the variable names should remain inside the parentheses.
2. Your final numeric expression should appear as MEAN(Test\_1,Test\_2,Test\_3,Test\_4). The final expression indicates that the new variable, *FinalGrade2*, will be calculated as the average of the four test scores.
3. Click **OK** to complete the computation and apply the changes to the data.
4. Finally, let’s make sure that a new variable called *FinalGrade2* was successfully created.  
   * We can find the new variable in the last column in Data View or in the last row of Variable View. If you do not see the new variable, the computation was unsuccessful.
   * We can check the syntax that was executed by looking at the log in the Output Viewer window. After running Compute Variable, the syntax that should have appeared in the output window is:
   * **COMPUTE FinalGrade2=MEAN(Test\_1,Test\_2,Test\_3,Test\_4).**
   * **EXECUTE.**

If there was an error in how the computation was specified, the log in the Output Viewer will often show an error message.

* + It is also useful to explore whether the computation you specified was applied correctly to the data. You can spot-check the computation by viewing your data in the Data View tab. To check that the new variable computed correctly, you can manually calculate the averages for a few cases in your dataset just to spot-check that the computation worked correctly.



If you've already verified the computation for *FinalGrade*, then all you need to do here is verify that *FinalGrade* and *FinalGrade2* are equal.

**Example: Computing a New Variable from a Range of Variables**

Notice that in the sample dataset, the test score variables in the sample dataset are all next to each other, and their naming scheme is sequential: *Test\_1*, *Test\_2*, ..., *Test\_4*.

In the previous example, we explicitly specified all four test score variables in the **MEAN** function. But what if there had been ten or twenty test score variables? It would take much longer to manually enter all twenty variable names.

What if we wanted to refer to the entire range of test score variables, from *Test\_1* through *Test\_4*, without having to type out each variable's name?

When using SPSS's special built-in functions, you can refer to a range of variables by using the statement **TO**. Let's repeat the previous example and show how the **TO** statement is used to refer to a range of variables inside a function.

1. Click **Transform > Compute Variable**.
2. In the **Target Variable** area, type a name for the new variable that will be computed; let's call the new variable *FinalGrade3.*
3. In the **Function group** list, click **All**.
4. In the **Functions and Special Variables**list, scroll down until you find “Mean”, then click on it.
5. Double-click “Mean” under in the **Functions and Special Values** list. The basic setup for using this function will now appear in the**Numeric Expression**field.
6. Inside the MEAN function, change the arguments to **Test\_1 TO Test\_4**.
7. Your final numeric expression should appear as **MEAN(Test\_1 TO Test\_4)**. The final expression indicates that the new variable,*FinalGrade3*, will be calculated as the average of all the variables between *Test\_1* and *Test\_4* in the dataset.
8. Click **OK** to complete the computation and apply the changes to the data.
9. Finally, let’s make sure that a new variable called *FinalGrade3* was successfully created.  
   * We can find the new variable in the last column in Data View or in the last row of Variable View. If you do not see the new variable, the computation was unsuccessful.
   * We can check the syntax that was executed by looking at the log in the Output Viewer window. After running Compute Variable, the syntax that should have appeared in the output window is:
   * **COMPUTE FinalGrade3=MEAN(Test\_1 TO Test\_4).**
   * **EXECUTE.**

If there was an error in how the computation was specified, the log in the Output Viewer will often show an error message.

* + It is also useful to explore whether the computation you specified was applied correctly to the data. You can spot-check the computation by viewing your data in the Data View tab. To check that the new variable computed correctly, you can manually calculate the averages for a few cases in your dataset just to spot-check that the computation worked correctly.

If you've already verified the computation for *FinalGrade* or *FinalGrade2*, then you should be able to verify that *FinalGrade*, *FinalGrade2*, and *FinalGrade3* are all equal.

*WARNING*: This method is dependent on the positions of the variables in the dataset. If the variables are not in sequential order, this method may not work correctly.