SPSS for Windows: Assignment 2

1. Saving standard scores as new variables
2. Selecting cases for analysis
3. Creating new variables using a compute statement
4. Creating new variables using an if-then statement

This assignment uses the bpdata dataset.

Please save your data files and output files every time you do something to change them. The only way I have of knowing that you did something is if your work has been saved!

Converting raw scores to standard scores and saving the standard scores as new variables

- Click Analyze
- Click Descriptive Statistics
- Click Descriptives
  - Highlight the name of a variable to be converted and then click the right arrow.
  - For example, select two variables – diabp1 and diabp2
  - Click the open box beside “Save standardized values as variables”
  - Click OK
  - Minimize the output window to bring back the spreadsheet.
  - Move to the last two columns of the spreadsheet. Two new variables have been created (adding two new columns) – zdiabp1 and zdiabp2

- Now let’s look at the means and standard deviations for these two new variables.
  - Click Analyze
  - Click Descriptive Statistics
  - Click Descriptives
  - Add the names for the two new variables to the selected variables box (there should now be four variable names in there).
  - Click off the “Save standardized values” option (because you don’t need to create the same standard scores a second time).
  - Under options make sure the mean and standard deviation are selected. Click continue to leave the options window.
  - Click OK under the descriptives window.
  - Check out the means and standard deviations of the four variables.
  - Notice that the mean for the variable zdiabp1 is 2.34E-15. This translates to 2.34 X 10^-15 or .00000000000000234 or, in other words, ZERO. The standard deviation is 1.00.

To select a subset of cases for analysis

For sure, you’re going to run into a situation where you want to conduct an analysis on only some of the subjects in your dataset. For example, let’s say that you
only want to get descriptive statistics for the variable diabp1 for subjects over the age of 40. To tell SPSS which subjects to include in the analysis…

- Click Data
- Click Select Cases
- In the Select Cases window you see the option to use all of the cases or to select cases “if condition is satisfied”.
- Click If condition is satisfied
- Click If
- Now in the Select Cases: If window, type age >= 40 in the open box at the top and middle of the window.
- Click continue
- Click OK
- In the spreadsheet you should now see a diagonal line in the case numbers for the subjects not included in the new data set.
- Now get the mean and standard deviation for diabp1 for subjects greater than or equal to 40 years of age.
- In the output window you notice that there are 44 people included in this analysis.
- To go back to include all of the subjects in later analyses, you have to go back to the select cases window and click on all cases. Go ahead and do this.

Creating a new variable using a compute statement

A compute statement is a tool for creating a new variable from mathematical manipulations and combinations of variables that already exist in the data set. For example, let’s say that we wanted to create a new variable for blood pressure that takes both systolic and diastolic blood pressure into account. So you decide to simply add the two measures of blood pressure together: Systolic BP + Diastolic BP = BP total. To do this in SPSS…

- Click Transform
- Click Compute
- In the Compute Variable window, type in the name for the new variable in the Target Variable box. Try typing bptot1 in this box. The “1” is suggested because we’re only going to use the blood pressure data collected at time 1.
- In the Numeric Expression box, type diabp1 + sysbp1
- Click OK
- The new variable bptot1 should now appear as the last variable in the spreadsheet.

If you want to get the mean of a set of two variables, the text in the Numeric Expression box would go (var1 + var2)/2. The format for the text that appears in this box basically follows common algebraic expression (i.e., the stuff in the parentheses gets computed first).
Creating a new variable using an if-then statement

An if-then statement is just what it sounds like: it says if this happens, then do this. A place where this type of statement comes in handy is in creating a discontinuous variable from a continuous variable.

For example, let’s say that you want to compare the younger subjects in the study to the older subjects in the study. You know the ages of every subject in the study, but there is no such variable as Age Group in the set. Age is a continuous variable, but you want to create a new variable that has only two levels: younger and older. Verbally you could say that “If a subject is less than or equal to 55 years of age, then give them a score of 1 for a new variable called Age Group”. “If a subject is greater than 55 years of age, then give them a score of 2 for a new variable called Age Group”. This let’s you convert the continuous variable of Age to the discontinuous variable of Age Group (with scores of only 1 or 2).

To create this new variable in SPSS…

- Click **Transform**
- Click **Compute**
- Let’s say that the name for the target variable is `agegrp` – short for Age Group – so that’s what you type in the Target Variable box.
- Now type a “1” in the Numeric Expression box. The combination of the target variable `agegrp` and the numeric expression “1” is read as “`agegrp` equals 1”. But we want to assign a subject to level 1 of `agegrp` only if they fit a certain criterion – if they are less than or equal to 55 years of age.
- Now click the **If** button.
- Click **Include if case satisfies condition**
- Type `age <= 55`
- Type **Continue**
- Click **OK**. A new variable, `agegrp`, is created in the spreadsheet. But notice that the value for many of the subjects is missing. That’s because we’ve only told SPSS how to label cases for subjects that are less than or equal to 55 years of age. To fill in the cells for the other subjects we need to repeat the procedure above, but just change some of the text in the if-then statement.
- Click **transform**
- Click **Compute**
- Remember that what we want to say now is that if the subjects are over 55, then the numerical label for those subjects will be a 2. Make sure that the target variable is still `agegrp`. Now put in a “2” for the Numeric Expression.
- Click **If** and then type `age > 55` in the box contained in the Compute Variables: If cases window.
- Click **Continue**
- Click **OK** – then click Yes, you want to change the existing variable.
• Now in the spreadsheet, every subject has been assigned a value of 1 or 2. Every subject has been placed in the younger (age <= 55) or older (age > 55) group.

Exercises

1. Use the **compute** statement to create a new variable called **tranxmn** by calculating the mean Trait Anxiety score at times 1 and 2.

2. Create a new variable called **dSYMGR1** that (a) assigns a score of 1 if subjects have scores that are less than or equal to 65 on the digit symbol test at time 1 and (b) assigns a score of 2 if subjects have scores on the digit symbol test at time 1 that are greater than 65. Remember, it’s the identical procedure to creating the discontinuous variable **AGEGRP** in the example above.

3. Use the **Select Cases** procedure to get descriptive statistics (i.e., mean and standard deviation) on the new variable **tranxmn** for subjects with a value of 1 for the new variable **dSYMGR1**

4. Turn the **Select Cases** selection criterion off so that you can use the data from all of the participant’s again. Create graph that displays the histogram for the scores for the variable **tranxmn**. Provide an overlay of the normal curve in this graph. Use the Chart Editor to change the title of the graph to say something like “[MyName]’s histogram”. In other words, in my chart I’d make the title say “Tom’s histogram”.

5. Use the Chart Editor to generate three additional graph’s that display the frequency distribution using 5, 10, and 15 intervals. Does the distribution appear to be normal? Why do you think so? Please provide your answer in a Text Box located at the bottom of your output window. To create a text box…

   • Go to the branching tree list of the various sections of your output window (located on the left-hand side of the output window). Click on the name (row) for the section of output above where you want your textbox to be located.
   • Click **Insert**
   • Click **New Text**
   • Now double-click inside the new box. Type whatever you want to add to the output window. When you click anywhere outside of the textbox, your text will still appear in the output window, but the box itself will not be visible. If you want to add anything else to the box, double-click anywhere on the text and the outline of the box will reappear so that you can change your text.

Please make sure that the output window is displaying the “notes” section. Please hand in your diskette with the revised bpdata datafile and the output window for the assignment.