

# Composites, Descriptive Statistics, and Histograms in SPSS

## Background

In psychology studies, we often measure constructs using scales—a series of several questions that all ask about aspects of the same construct. For example, to measure extraversion, you might ask people to rate the following scale items from 1 (strongly disagree) to 5 (strongly agree):

1. I see myself as talkative
2. I see myself as quiet
3. I see myself as outgoing

In order to analyze a variable (including compute its descriptive statistics), you'll need to form a composite from all of the scale items by averaging them all together.

## Reversing Items

Before forming a composite for a scale, you'll need to make sure your items are all properly reversed, if necessary. "Reversing" items means that you need to make sure that all items are scored such that higher numbers represent higher levels of the construct. For example, in the sample extraversion scale above, higher responses to item #2 would indicate that people saw themselves as LESS extraverted (because they're agreeing that they're quiet). So, before computing reliability for the extraversion scale, you will need to reverse item #2. (If you've already reversed the scoring for item #2 in your web survey when you collected the data, you do not need to reverse it again!)

### How do I know if I need to reverse an item or not?

You have to look at your survey where you originally collected the data. You have to look at how people's responses were scored on each item. If higher numbers on an item mean LOWER scores on the variable, you need to reverse the item before including it in the composite. For example, consider the item "I see myself as quiet," rated on a scale from disagree (1) to agree (5). Higher numbers mean the person *agrees* that they're quiet—and if the person is quiet, it means that they are not extraverted. LOWER numbers mean that person thinks that they ARE extraverted. So, we have to reverse the item before including it in an extraversion composite.

If the item "I see myself as quiet" was already scored properly in your survey (e.g., if agree = 1 and disagree = 5), higher numbers ALREADY indicate greater levels of extraversion (because the person thinks that they're NOT quiet). So, you should NOT reverse the item in your syntax. It was already properly scored in the survey itself.

To reverse an item, the formula is:

$$(\text{Reversed Score}) = (\text{Scale Max}) + (\text{Scale Min}) - (\text{Original Score})$$

So, imagine that you have the three extraversion items in your dataset:

	e1	e2	e3
1	4.00	2.00	3.00
2	5.00	3.00	4.00
3	2.00	5.00	3.00
4	4.00	2.00	4.00

We need to reverse item e2. It was rated on a scale from one to five. So, using the above formula, we'll use the following syntax:

```
compute e2r = 5+1-e2.  
execute.
```

Whenever you reverse a variable, *always* create a new variable (e.g., e2r). Don't simply overwrite the old variable. You ***never*** want to overwrite your original data. ***Always*** add new variables instead.

## Forming Composites

To form a composite variable, simply average all of the variables together. The basic syntax is:

```
compute composite = mean(var1[, var2, ...]).  
execute.
```

So, for our extraversion example, the syntax would be:

```
compute extraversion = mean(e1, e2r, e3).  
execute.
```

Notice that we **ONLY** include the properly scored items (e.g., e2r). We do **NOT** include the original "negatively scored" item for items we reversed (e.g., we do **NOT** include the original e2 item in our composite).

Also notice that "compute" commands must be followed by an "execute" command!

## Computing Descriptive Statistics

The syntax for computing descriptive statistics is:

```
descriptives var1 [var2 ...].
```

For our extraversion example, the syntax would be:

```
descriptives extraversion.
```

## Reading the Output

The means and standard deviations for your variables will appear in the output table:

	N	Minimum	Maximum	Mean	Std. Deviation
extraversion	6	2.67	4.33	3.3889	.80046
Valid N (listwise)	6				

Remember that we *always* round means and standard deviations to two decimal places when reporting them. “The mean level of extraversion in the sample was  $M = 3.39$ , ( $SD = 0.80$ ).”

## Standardizing Variables

If you’d like to standardize variables, you can use the descriptives command with the /save option:

```
descriptives extraversion /save.
```

Running this command will show the descriptive statistics table in the output. But it will also create a new variable in the dataset called “Zextraversion.” This new variable is people’s standardized levels of extraversion.

Please note that if you were to try to standardize a variable more than once, SPSS will create new variables beginning with “Zsco” (e.g., “Zsco01,” “Zsco02”). For example:

```
descriptives extraversion /save.  
descriptives extraversion /save.
```

If you ran the above syntax, SPSS would create two new (identical) variables called “Zextraversion” and “Zsco01.” If you start to see random “Zsco” variables in your dataset, check your syntax to see if you’re trying to standardize the same variable more than once!

## What if I have more than one scale?

What happens if you have more than one scale in your dataset? For example, what if you have items measuring both extraversion and agreeableness?

	e1	e2	e3	a1	a2	a3
1	4.00	5.00	3.00	2.00	3.00	4.00
2	5.00	5.00	4.00	5.00	4.00	3.00
3	2.00	3.00	2.00	2.00	3.00	4.00
4	4.00	3.00	4.00	4.00	4.00	4.00

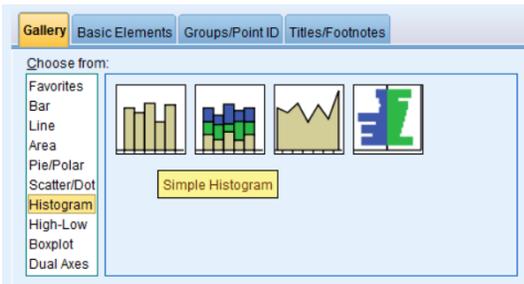
In this case, you will need to form composites for extraversion and agreeableness separately:

```
compute extraversion=mean(e1, e2, e3).  
compute agreeableness=mean(a1, a2, a3).  
execute.
```

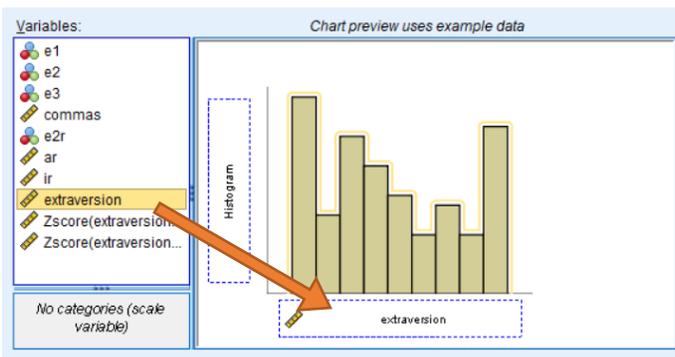
## Creating Histograms

It's often an extremely good idea to create histograms to look at the distribution of scores in your data. To do so, we'll need to use the drop-down menus (SPSS's graphing syntax is extremely complex).

1. Select "Graphs > Chart Builder" from the main menu
2. Click the "Histogram Category" and double click the "Simple Histogram" option:



3. Drag the desired variable from the "Variables" list onto the X axis:



4. Click "OK"

You can now examine your histograms for any problems. Do the scores look relatively normal? Or do they look skewed? Do you see evidence of ceiling or floor effects? These are all important questions to ask when looking at histograms.