Correlations and Scatterplots in SPSS

Background

A correlation measures how strongly two variables are related or "go together." Correlations range from -1 to +1. A correlation of zero means two variables are not related.

Preparing Your Data

Before running correlations, you'll need to form composites for each variable you want to correlate with other variables. See the "Composites, Descriptive Statistics, and Histograms" handout for more information on how to do so.

Correlations

The syntax for running correlations is

```
correlations var1 [var2 …].
```

For example, if you have composites for each of the big five personality traits (e = extraversion, a = agreeableness, c = conscientiousness, s = emotional stability, o = openness), the syntax to look at how they all correlate with each other would be:

```
correlations e a c s o.
```

Reading the Output

SPSS's correlation output can be a bit confusing:

Correlations

		e	a	c	s	o
e	Pearson Correlation	1	.214	.134	.206	.071
	Sig. (2-tailed)		.001	.032	.001	.257
	N	254	254	254	254	254
a	Pearson Correlation	.214	1	.237	.165	.062
	Sig. (2-tailed)	.001		.000	.008	.328
	N	254	254	254	254	254
c	Pearson Correlation	.134	.237	1	.194	.064
	Sig. (2-tailed)	.032	.000		.002	.312
	N	254	254	254	254	254
s	Pearson Correlation	.206	.165	.194	1	-.007
	Sig. (2-tailed)	.001	.008	.002		.908
	N	254	254	254	254	254
o	Pearson Correlation	.071	.062	.064	-.007	1
	Sig. (2-tailed)	.257	.328	.312	.908	
	N	254	254	254	254	254

Each correlation is listed twice. Once above the diagonal (the cells that all contain "1" and run from the top-left cell to the bottom-right cell) and once below. So, we can safely mentally ignore the "upper matrix:"

Correlations

		e	a	c	s	o
e	Pearson Correlation	1				
	Sig. (2-tailed)					
	N	254				
a	Pearson Correlation	.214	1			
	Sig. (2-tailed)	.001				
	N	254	254			
c	Pearson Correlation	.134	.237	1		
	Sig. (2-tailed)	.032	.000			
	N	254	254	254		
s	Pearson Correlation	.206	.165	.194	1	
	Sig. (2-tailed)	.001	.008	.002		
	N	254	254	254	254	
o	Pearson Correlation	.071	.062	.064	-.007	1
	Sig. (2-tailed)	.257	.328	.312	.908	
	N	254	254	254	254	254

Looking only at the lower matrix, there are three rows for each correlation. The "Pearson Correlation" rows contain the correlation coefficients (r), the "Sig." rows contain the statistical significance (p), and the "N" rows contain the sample size for the correlation. Thus, the matrix below shows *only* the correlation coefficients (rs). When copying correlations, make sure to copy the correct number!

Correlations

		e	a	c	s	o
e	Pearson Correlation	1				
	Sig. (2-tailed)					
	N					
a	Pearson Correlation	.214	1			
	Sig. (2-tailed)					
	N					
c	Pearson Correlation	.134	.237	1		
	Sig. (2-tailed)					
	N					
s	Pearson Correlation	.206	.165	.194	1	
	Sig. (2-tailed)					
	N					
o	Pearson Correlation	.071	.062	.064	-.007	1
	Sig. (2-tailed)					
	N					

Reporting Correlations

Correlations are generally reported to two-decimal places—without a leading zero. For example, the correlation between extraversion and agreeableness should be reported as r = .21. Whenever you report a correlation in the main text of your manuscript, you should also report a 95% confidence interval (CI). Although SPSS does not output CIs for correlations, you can find online calculators (such as the one at http://vassarstats.net/rho.html) by Google searching for "correlation confidence interval calculator." To compute the CI, you'll need to enter the correlation coefficient and the sample size (N; in the above example, the N for all correlations was 254).

> Sample Manuscript Text:
> Extraversion and agreeableness were moderately correlated, r = .21, 95% CI [.09, .33].

It's common to include correlation tables in manuscripts. When you do so, be sure to bold all correlations that are statistically significant (i.e., $p < .05$).

Sample table:

	Correlations				
Variable	1	2	3	4	5
1. Extraversion	-				
2. Agreeableness	**.21**	-			
3. Conscientiousness	**.13**	**.24**	-		
4. Emotional Stability	**.21**	**.17**	**.19**	-	
5. Openness	.07	.06	.06	-.01	-

Note: The 95% confidence intervals for correlations in **boldface** do not include zero.

Creating Scatterplots

It's often an extremely good idea to create histograms to look at the scatterplots for important correlations in your manuscript.

1. Select "Graphs > Chart Builder" from the main menu
2. Click the "Scatter/Dot" category and double click the "Simple Scatter" option
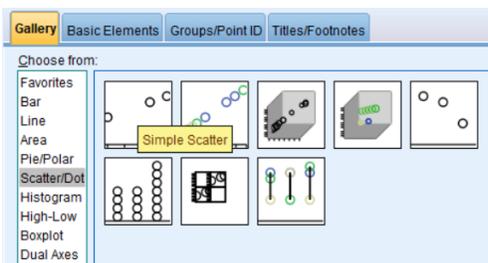
3. Drag the desired variable from the "Variables" list onto the X and Y axes:
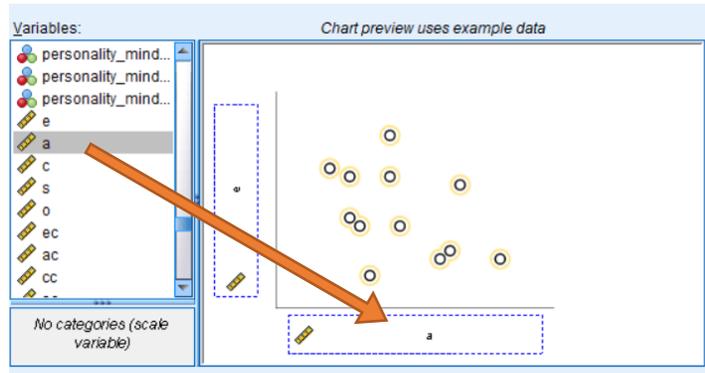
4. Click "OK"

You can now examine your scatterplots for any problems. Do you see evidence of restriction of range? Does the "cloud" of dots look roughly like a football? These are all important questions to ask when looking at histograms.