**Correlations How To in SPSS with HS&B Data**

High School & Beyond (HS&B) Features

* Nationally representative, longitudinal study of 10th and 12th graders in 1980
* Follow-up surveys conducted throughout their postsecondary years
* Surveys of students, teachers, and parents of sampled students
* High school and postsecondary transcripts to enhance analyse

HS&B Focus

* What are students' trajectories after leaving high school into postsecondary education, the workforce, and beyond?
* What factors influence the students' educational and career outcomes after passing through the American educational system?

SPSS:

* In general, make sure your data is properly set up
  + What are your variables?
* What do you want to know? Is there a relationship between \_\_\_ and \_\_\_?
  + ***For our purposes today, stick to the scale variables!!!***
* Analyze 🡪 correlate 🡪 bivariate
  + Select the variables you are assessing
  + Make sure ‘Pearson’ is the correlation coefficient that is marked off
  + Make sure the test of significance is ‘two-tailed’
  + If you click on ‘options’, you can select to get the means and SDs
  + Click OK
* Read output!

**Example from HSB2 data (high school and beyond)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | writing score | reading score |
| writing score | Pearson Correlation | 1 | .597\*\* |
| Sig. (2-tailed) |  | .000 |
| N | 200 | 200 |
| reading score | Pearson Correlation | .597\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| N | 200 | 200 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

Here, you can see that the correlations are replicated on each side. This is because the output is given as a matrix, so the relationship of writing score to writing score is 1 (it is the same thing), but then you have the relationship of writing score to readings score. This is repeated again below since each of the variables are written across the top and across the side.

If you look at the Sig., you will see the *p* value. Is it less than 0.05? Is your correlation significant?

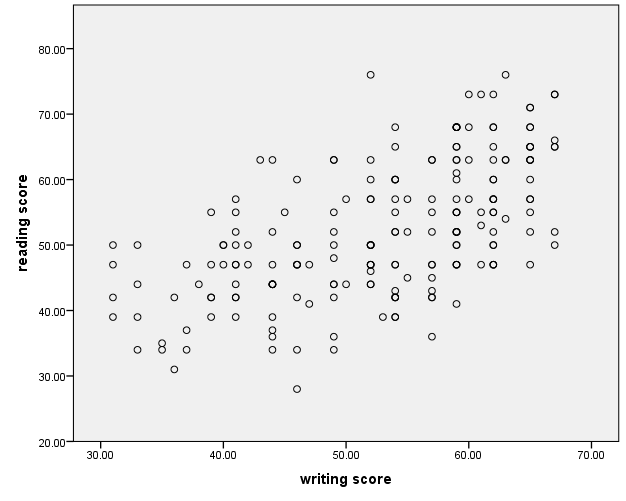
We could write up these results in the following way:

A Pearson correlation was run to examine the relationship between scores on a writing exam and scores on a reading exam. There was a moderate, positive correlation between writing scores and reading scores, which was statistically significant (*r* = .597, *n* = 200, *p* < .01).

*Remember that a correlation ranges from -1 to +1, where the closer the # is to 1, the stronger the correlation. Here, with a value of .597, we have a significant relationship with a moderate correlation.*

Correlation strength, as described by Evans (1996):

• .00-.19 “very weak”  
• .20-.39 “weak”  
• .40-.59 “moderate”  
• .60-.79 “strong”  
• .80-1.0 “very strong”



In order to create a scatterplot:

* Graphs 🡪 chart builder
* Select type of graph you want from the options along the bottom left
* For a correlation, you want a scatterplot
* Drag the scatterplot picture (the first one should be appropriate for your data) to the chart preview area above
* Then, select the variables you are analyzing from the variables list on the left and drag them to the X and Y axis (for this example, writing score on the X and reading score on the Y)
  + Generally, your independent variable, or the ‘cause’ variable would go on the x-axis and your dependent variable, or the ‘effect’ variable on the y. For example, if you were looking at the relationship between height and distance jumped on a long jump, you would expect that the height would influence the distance jumped – so the height is you independent variable and what you expect to impact the results, so it would go on the x-axis. The distance jumped would be your dependent variable and that would go on the y.