

Lab 8: Correlation and Simple Regression

Now that we have completed some inferential analyses (i.e., t-test and ANOVA), we are able to indicate whether group differences exist in a data set. However, sometimes our analyses are more exploratory. That is, we would like to know how one variable is related to a second variable. For this lab, we will be investigating variables' relationships with one another.

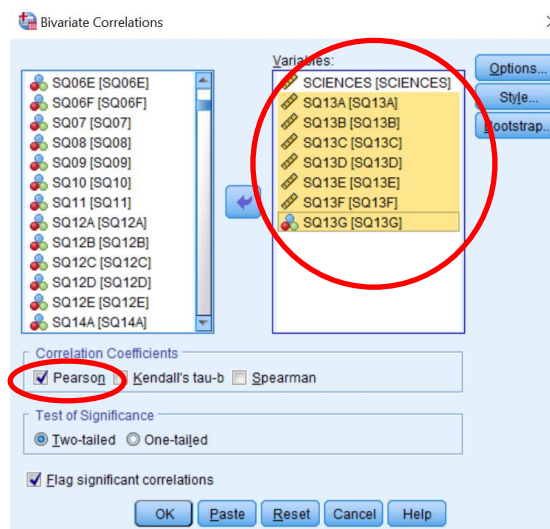
During this lab, we are interested in investigating the relationship between students' responses to the bullying item and their overall science achievement. Literature has suggested that students' academic achievement in school is hindered by negative emotions that are often related to being bullied. That is, students who are bullied often have lower grades in school.

1. Correlation

In this task, we are trying to answer the research question: What is the relationship between students' being bullied and their science achievement? Correlation works best when all the variables are on continuous scales. When answering this question, we do not know whether one variable is causing the second variable. As such, we are **not** able to **make** any **causal claims!**

Correlation \neq Causality

To run the correlation please select Analyze \rightarrow Correlate \rightarrow Bivariate \rightarrow Move all eight relevant variables into the Variables box \rightarrow check off Pearson in the Correlation Coefficients box \rightarrow click OK



Questions:

1. Please write the null hypothesis and alternative hypothesis for the research problem.

2. When you ran the correlation analysis, you selected Pearson, but what does each correlation coefficient (i.e., Pearson, Kendall’s tau-b, and Spearman) represent? That is, under what conditions should we select each type of correlation coefficient?

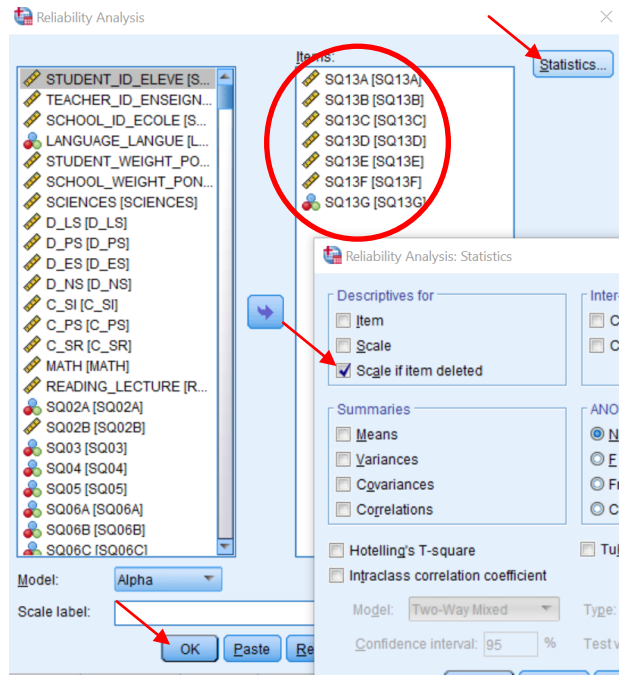
3. Please report your correlation in a table. The table below may be helpful as you develop your table.

Table 1

Title (please make a title for this table)

	Science	SQ13A	SQ13B	SQ13C	SQ13D	SQ13E	SQ13F	SQ13G
Science								
SQ13A								
SQ13B								
SQ13C								
SQ13D								
SQ13E								
SQ13F								
SQ13G								

Since the correlation of the bullying items is relatively high and, more importantly, statistically significant, we can group the bullying items into one variable we will call “Bullying” for the analyses. To ensure the items are indeed measuring the same concepts, we will also use Cronbach alpha to determine whether or not we should combine sets of items. To do this, please select Analyze → Scale → Reliability Analysis → move the bullying items into the Items box → click Statistics → check Scale if item deleted → click Continue → click OK.



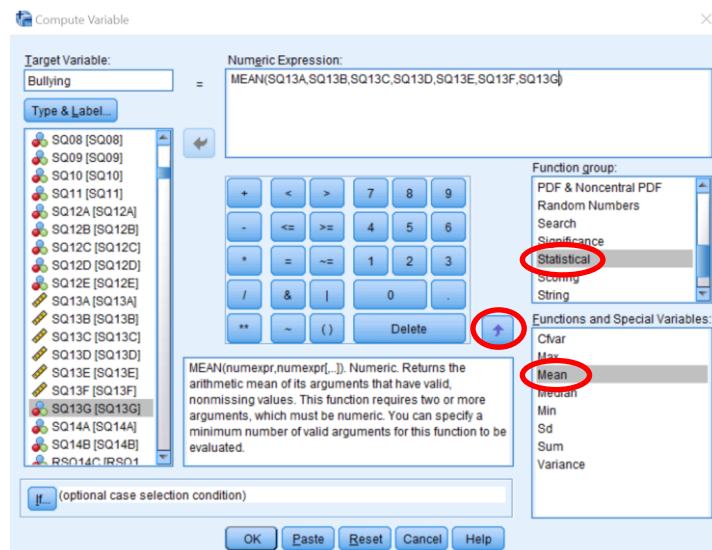
Questions:

4. Please define Cronbach alpha and indicate what it is used for.

5. What is the Cronbach alpha for the bullying items?

6. Looking at the Cronbach's Alpha if Item Deleted column of the Item-Total Statistics table for the bullying items, what is the best possible alpha value that could be achieved? Which item(s) would you recommend we combine in our mean to achieve this value?

To combine the bullying items you recommended from question 6, please select Transform → Compute Variable → under Target Variable please type *Bullying* → under Function group select Statistical, and then under Functions and Special Variables select Mean → please hit the Up arrow located beside the Delete button → please move each bullying item into the bracket area in the Numeric Expression box using a comma to separate each item.



Now, please rerun your correlation analysis using only science achievement and your new bullying variable.

Question:

7. Please answer the research question using the sentence “There was a statistically [significant/non-significant] and [positive/negative] relationship between students’ science achievement and being bullied, r =[insert your Pearson correlation], p [</>]0.05.”

2. Regression: simple

In this lab, we will only do a simple regression, while the multiple regression will be completed in the next lab. A simple regression is used to determine whether one variable predicts another variable. Again, when one variable can predict another variable at a high rate, it still **does not indicate causality!** Now that we know there is a relationship between being bullied and science achievement, we would like to know how well one variable predicts the other. Our new research question is now: Does being bullied predict science achievement scores?

Question:

8. Please indicate which variable is the dependent variable and independent variable.

To run a regression, click Analyze → Regression → Linear → move *the appropriate variable* into the Dependent box and *the other variable* into the Independent(s) box → click OK.

Question:

9. Looking at the **Model Summary** results, what is the difference between R, R Squared,

and Adjusted R square (i.e., what does each of these represent)? (You may need to use Google for this question.)

Often, readers prefer to *see* a relationship with a diagram, so it is good practice to include a scatterplot. To make a scatterplot in SPSS, click Graphs → Legacy Dialogs → Scatter/Dot → select Simple Scatter → click Define → move “dependent variable” into the Y-Axis box and “independent” into the X-Axis box → click OK.

Question:

10. Is your overall model statistically significant? Please answer this question using the sentence “Being bullied is [able/not able] to predict science achievement significantly using statistical analyses, r^2 =[insert your R Square value], F ([df_{regression}, df_{residual}])=[insert your F-value], p [</>]0.05.” (You may need to use Google to help you read the SPSS output for this question.)
11. Is the regression coefficient (i.e., Bullying) statistically significant? Please answer the question using the sentence “The regression coefficient of Bullying is statistically [significant/non-significant], b =[insert your b value], t =[insert your t value], p [</>]0.05” (You may need to use Google to help you read the SPSS output for this question.)
12. Using the results shown in *Coefficients*, please write out your regression equation in the format “ y =[slope] x + [intercept]” and “science achievement=[slope]Bully+ [intercept]”. In an article, you would only write out the second formula to help you readers better understand that the y-axis represents the science achievement and the x-axis represents the bullying. (You may need to use Google to help you read the SPSS output for this question.)
13. What are your F-, t-, and t^2 -values (you will need to manually calculate t^2)? What is the relationship between F and t?
14. Please write a *short* paragraph discussing your correlation and regression results. (There should be no technical jargon and your answer should be a maximum of *five* sentences.)

You have now completed the lab.