JAMES COOK UNIVERSITY SINGAPORE

Correlation

Learning Centre

JAMES COOK UNIVERSITY AUSTRALIA

Cairns
Singapore
Townsville

www.jcu.edu.sg

Contents

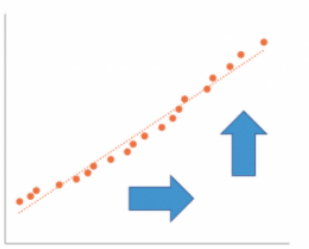
1. Correlation and its purpose
2. Describing correlations
3. An illustrated example
4. Reporting

What is Correlation?

JAMES COOK UNIVERSITY SINGAPORE

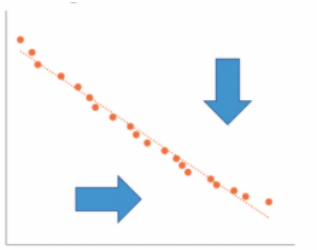
- A correlation analysis is used to assess the strength and direction of a relationship between two variables
- A correlation coefficient is most commonly annotated using Pearson's *r*
- Pearson's *r* can range from -1 to 1. A negative coefficient indicates a negative association, while a positive coefficient indicates a positive association
- A common misconception is that correlation equals causation. This is not the case.

How do we describe correlations?



Positive Correlation

An increase in one variable that is accompanied with a increase in another variable



Negative Correlation

An increase in one variable that is accompanied with a decrease in another variable

No Correlation

Two variables do not covary

Pictures taken from https://astutesolutions.com/blog/articles/causation-vs-correlation

Correlation ≠ Causation

JAMES COOK UNIVERSITY SINGAPORE

- A correlation simply looks at the strength and direction of a relationship
  - For example, a study found that ice cream sales was strongly positively correlated with shark attacks. The more ice creams were sold, the more shark attacks occurred. But, selling ice creams *do not* and *cannot* cause shark attacks.
- Causation implies a cause and effect relationship: a change in Variable B is caused by a change in Variable A, or vice versa
  - For example, the more I exercise, the more I feel the fatigue after the exercise. How tired I feel is directly affected by how much I have worked out.

Correlation ≠ Causation

JAMES COOK UNIVERSITY SINGAPORE

If we look at the ice cream example, do ice cream sales cause shark attacks? Or do shark attacks cause more ice cream sales?

Neither is true! In fact, an increase in ice cream sales is actually caused by hot weather during summer, and during summer, more people go to the beach to enjoy water sports. This then leads to higher probability of shark attacks.

There is a third or hidden variable (i.e., hot summer weather), that affects our 2 variables, so we can only say that ice cream sales are positively correlated with shark attacks, but not that one causes the other.

Location of SPSS Data Files

JAMES COOK UNIVERSITY SINGAPORE

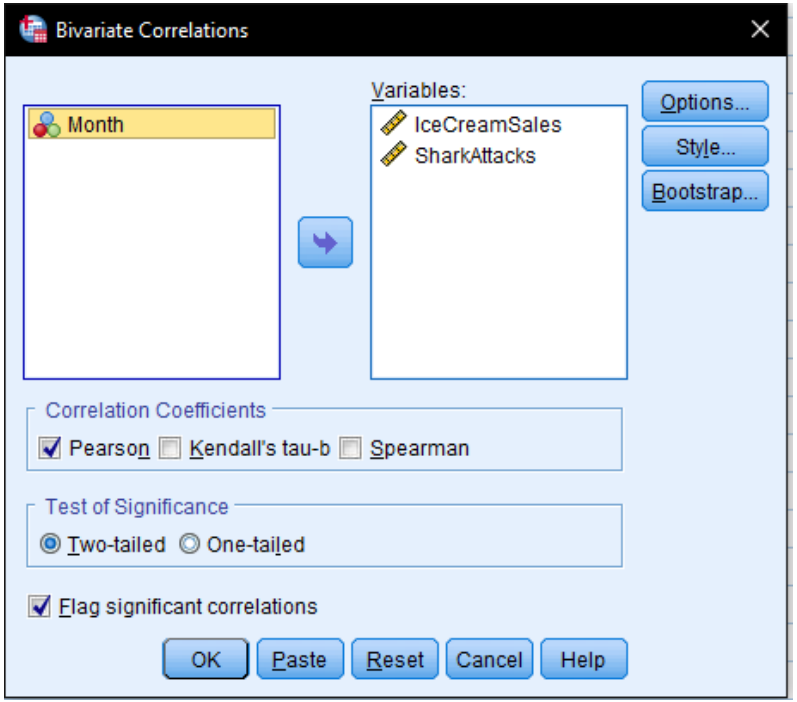
Example SPSS data for practice are available on LearnJCU:

Log in to LearnJCU -> Organisations -> Learning Centre JCU Singapore -> Statistics Support -> Statistics Resources -> SPSS Data for Practice

SPSS time!

Using the ice cream and shark example, let us now conduct a correlation analysis

- Go to Analyze -> Correlate -> Bivariate
- Shift the 2 variables of interest to the right column
- Click OK!



SPSS time!

- Looking at the output, we have $r = .941$, and the p value of this correlation less than $< .001$
- We may classify r in a few ways:
  - If r is below |.3|, it is a weak correlation
  - If r is between |.4| to |.6|, it is a moderate correlation
  - If r is above |.7|, it is a strong correlation

Correlations

Correlations

		IceCreamSales	SharkAttacks
IceCreamSales	Pearson Correlation	1	.941**
	Sig. (2-tailed)		.000
	N	12	12
SharkAttacks	Pearson Correlation	.941**	1
	Sig. (2-tailed)	.000	
	N	12	12

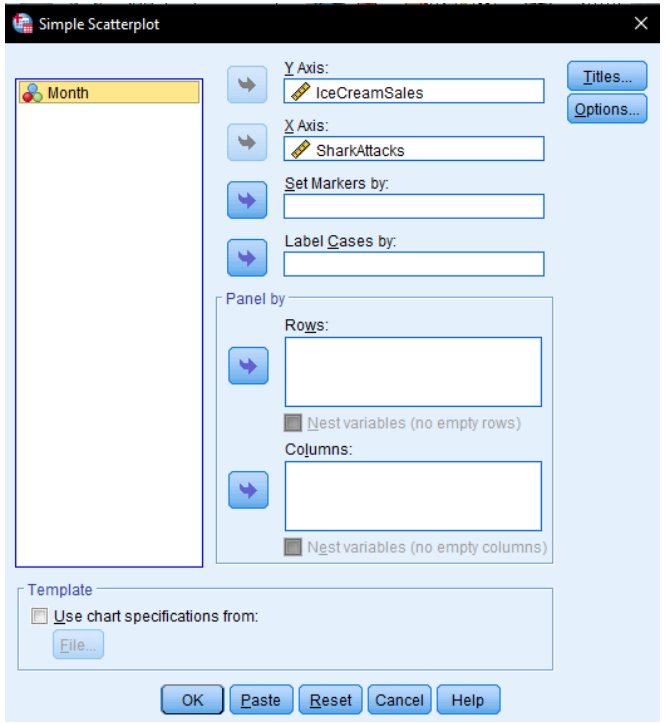
**. Correlation is significant at the 0.01 level (2-tailed).

**Note that different sources will provide different benchmarks on what is a weak, moderate, or strong correlation*

A Visual Representation

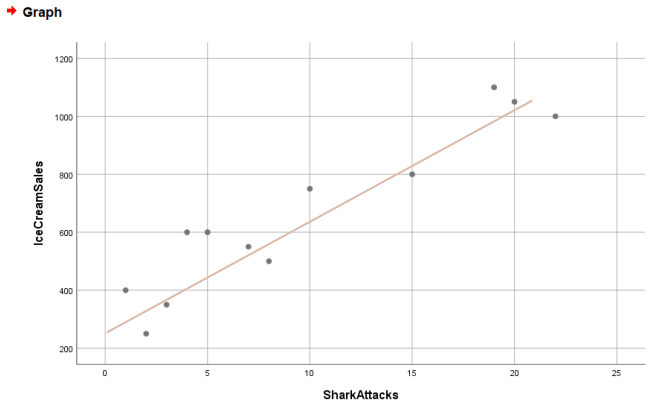
JAMES COOK UNIVERSITY SINGAPORE

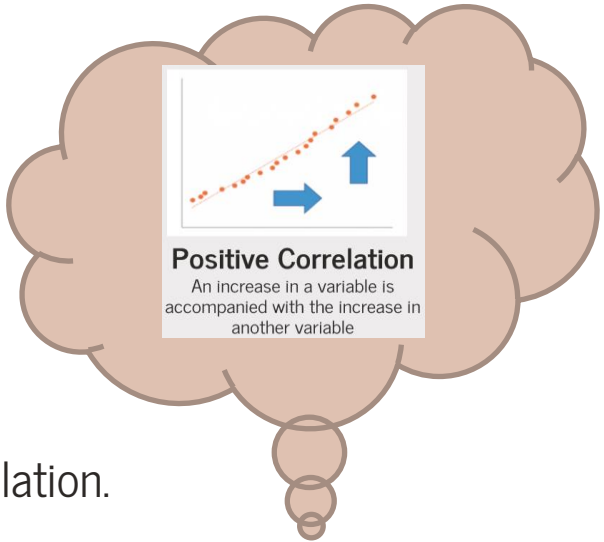
- We can also plot a **graph** to visually inspect the data
  - Go Graphs -> Legacy Dialogs -> Scatter/Dot ->
  - Select the simple/scatter option -> Define
  - Move IceCreamSales under Y Axis
  - Move SharkAttacks under X axis.
  - You can swap the axes if you wish
  - Click OK!



A Visual Representation

Here our graph! If we draw a line of best fit, we would get something like this:





- Recall that this line shows a positive correlation.
- In our example, $r = .941$, which is very close to 1 (look at Slide 4!).

Reporting

JAMES COOK UNIVERSITY SINGAPORE

An example write-up can be found on:

JCUS Learning Centre website -> Statistics and Mathematics Support

JAMES COOK UNIVERSITY SINGAPORE

Questions?

learningcentre-singapore@jcu.edu.au

JAMES COOK UNIVERSITY AUSTRALIA

Cairns
Singapore
Townsville

www.jcu.edu.sg