

Research Methods

Correlation analysis

This worksheet will allow you to practice the analysis of 'correlational' data.

There are a number of different data-sets available at the end of this worksheet to practice with. Each set can be described and assessed for the amount of correlation (or association) between the two variables. We will use the Pearson's Product Moment Correlation (or Pearson's correlation for short) to assess whether each correlation is significant, and the strength and direction of the correlation.

For each data set you should input the data in the format shown below, with a column for each of the variables. The first example depicts two measures taken for 'level of impulsivity' and 'frequency of gambling behaviour' for each participant.

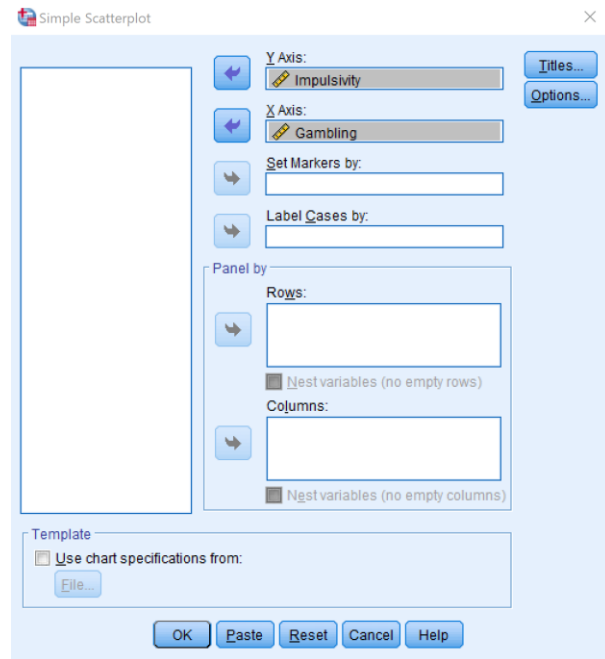
On the 'variable view', create your two variables. Toggle back to 'data view' to input the data as shown below.

	 Impulsivity	 Gambling
1	32.00	6.00
2	17.00	3.00
3	28.00	4.00
4	11.00	2.00
5	24.00	7.00
6	20.00	6.00
7	36.00	10.00
8	15.00	2.00
9	23.00	3.00
10	31.00	9.00
11		

Once you have input the first set of data above, produce a scatterplot:

- 'Graphs' on the top menu
- 'Legacy dialogues'
- 'Scatter/dot'
- Select the 'simple scatter' option and click 'define'

Move the 'impulsivity' variable into the box for the y-axis and the 'gambling' variable into the box for the x-axis (as shown), then press 'OK'. You can also change the scale of each axis if necessary by clicking on the x or y axis once it is in the output page.



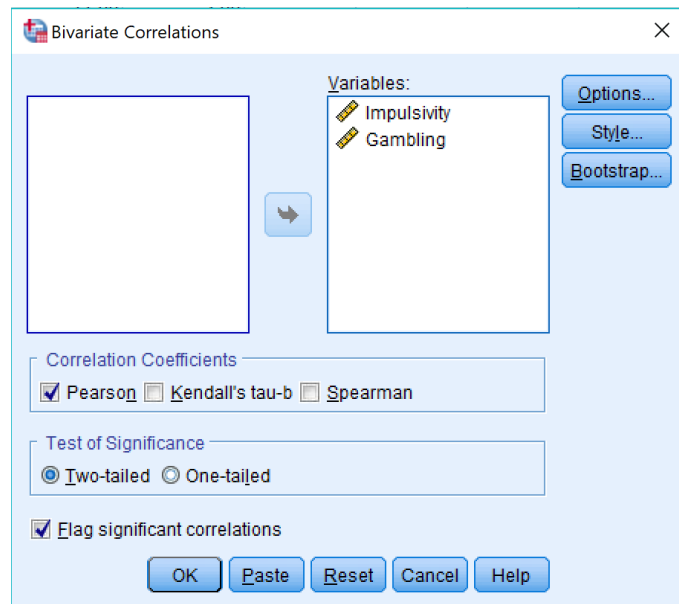
By looking at the scatterplot produced write down whether it appears a correlation exists between the two measures, and if so what direction is this correlation? You can then add a fit-line to the scatterplot to show the regression line for the variables. To do this, double-click on the scatterplot and select the small 'fit-line' icon  from the pop-up graph menu. Once the fit-line is added you can close the pop-up graph.

By looking at both the scatterplot and the fit-line produced you can determine whether it appears there is a correlation between the two measures, and the direction of the correlation. BUT! Remember that we do not know yet if the correlation is significant, so we need to conduct the statistical test to make this judgement.

Now, we can carry out the statistical test to make a judgement about how 'significant' any relation is between the variables.

1.
 - Select 'Analyze' on the top menu bar
2.
 - Hover over the 'Correlate' option
3.
 - Select the 'Bivariate' option

Move both variables across to the 'Variables' box for analysis, as shown below. Select the 'Pearson' coefficient, and the Two-tailed test. Then click 'OK'.



The output for this analysis will provide you with two tables of data.

The information in the **Descriptive Statistics** table will give you the mean and SD scores for each of your variables.

The **Correlations** table will provide you with the information you require to determine whether the relationship between the variables is significant or not.

Correlations

		Impulsivity	Gambling
Impulsivity	Pearson Correlation	1	.820**
	Sig. (2-tailed)		.004
	N	10	10
Gambling	Pearson Correlation	.820**	1
	Sig. (2-tailed)	.004	
	N	10	10

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

The correlation (r) value comes from here

The significance (p) value comes from here

You can work out df from the value of N

Note: A minus sign would tell you that there is a **negative** correlation – NOT that there is NO correlation

The Pearson Correlation value (r) indicates the strength of the correlation.
The Sig. value (p) will tell you if the two variables are significantly related.

Write the values for each of the main elements required, and then put the statistics into the correlation equation as demonstrated below:

r (df) = _____, p < / > .05 OR p = _____

r =
 p =
df =

 r (df) = _____, p

Finally, write out the **results** of this study (including the statistical equation above) and explain what the findings of the correlation analysis mean.

Now we will look at some additional datasets, to describe and analyse any possible correlations between the variables. For each dataset, create and input the data for both variables, produce a scatterplot of the data, and carry out the Pearson correlation analysis as before. Write out the statistical analysis, giving the r , N and p -values for each correlation, and interpret each set of result (state exactly what the findings mean)

Data-set 1		Data-set 2		Data-set 3	
Age	IQ	Height	Shoe size	Happiness	Spending
21	110	153	5	21	89
32	98	154	6	32	73
55	132	160	6	55	24
77	120	168	9.5	77	21
23	100	159	7	23	63
45	96	170	10.5	45	35
23	123	167	9.5	23	81
55	130	166	10	55	32
64	86	154	4	64	29
35	111	183	12.5	35	36

Dataset 1:

$r =$

$p =$

df =

$r(df) =$, p

Dataset 2:

$r =$

$p =$

df =

$r(df) =$, p

Dataset 3:

$r =$

$p =$

df =

$r(df) =$, p