

Dr. Peggy Kern's Capstone Statistics
Practice #3: Correlations - ANSWERS

Statistics are learned best by doing! Here is a set of exercises to let you practice what we have covered on the correlations. See if you can work through them before checking your responses.

1. For the following correlations (r): 0.29 -0.63 0.15 -0.34 0.04

a. Which is the **strongest** correlation?

With correlation, remember that we can think about both strength and direction of the effect, so look at the absolute value distance.

Strongest is -.63.

b. Which is the **weakest** correlation?

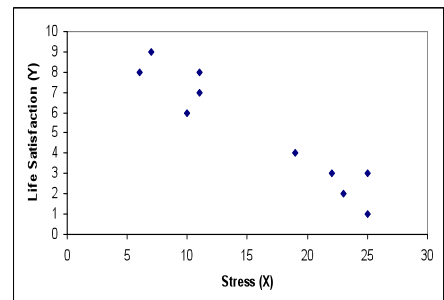
Weakest is .04

2. Janice and Paul did a study on feelings of stress and life satisfaction. Participants completed a measure on how stressed they were feeling (on a 1 to 30 scale) and a measure of how satisfied they felt with their lives (measured on a 1 to 10 scale). The table below indicates the participants' scores. **Using this data, answer the following questions:**

Participant #	Stress score (X)	Life Satisfaction (Y)
1	11	7
2	25	1
3	19	4
4	7	9
5	23	2
6	6	8
7	11	8
8	22	3
9	25	3
10	10	6
Σ	159	51
Mean	15.9	5.1
SD	7.23	2.70

a) On a scrap paper, try to draw a rough scatterplot of the data, just to get an idea of what these look like.

You didn't have to include this, but essentially a scatterplot would look something like this:



There are multiple ways to calculate a correlation coefficient r (that is, a standardized indicator of the relation between two variables). We can calculate the **covariation** between two variables (X and Y), and then adjust this by the standard deviation and sample size.

$$r = \frac{\text{COV}(x,y)}{s_x s_y} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{(N)s_x s_y}$$

Alternatively, we can calculate r from the Z scores:

$$r = \frac{\sum(Z_x Z_y)}{N}$$

b) Using either method, **calculate the correlation (r) between stress and life satisfaction.**

This was probably the most time intensive element of this assignment. We can use either formula – I’ve done both here as an example, with the covariance method in green and the Z score method in red.

#	X	X-M _x	Z _x	Y	Y-M _y	Z _y	(X-M _x)(Y-M _y)	Z _x Z _y
1	11	-4.9	-0.68	7	1.9	0.70	-9.31	-0.48
2	25	9.1	1.26	1	-4.1	-1.52	-37.31	-1.91
3	19	3.1	0.43	4	-1.1	-0.41	-3.41	-0.17
4	7	-8.9	-1.23	9	3.9	1.44	-34.71	-1.78
5	23	7.1	0.98	2	-3.1	-1.15	-22.01	-1.13
6	6	-9.9	-1.37	8	2.9	1.07	-28.71	-1.47
7	11	-4.9	-0.68	8	2.9	1.07	-14.21	-0.73
8	22	6.1	0.84	3	-2.1	-0.78	-12.81	-0.66
9	25	9.1	1.26	3	-2.1	-0.78	-19.11	-0.98
10	10	-5.9	-0.82	6	0.9	0.33	-5.31	-0.27
Σ	159	0	0.00	51	0.00	0.00	-186.90	-9.57
Mean	15.9			5.1				
ŝ	7.23			2.70				

Then we add these to the formulas.

Using the covariance: $r = \frac{(-186.90)}{(10 * 7.23 * 2.70)} = -0.957$

Using the Z scores: $r = \frac{-9.57}{10} = -0.957$

Round to 2 decimal points, so $r = .96$

c) Write a brief interpretation of this correlation, including the strength, direction, and an explanation of the effect.

There is a very strong, negative correlation between stress and life satisfaction, such that as stress increases, life satisfaction decreases.

d) Can you say that being more stressed causes a lower level of life satisfaction? Why or why not?

No, because correlation does not equal causation.

3. At Hogwarts School of Witchcraft and Wizardry, students often have a lot of homework. The table below indicates the number of hours students studied, and how they performed on an exam in two of their classes.

Student	Potions		Defense against the dark arts	
	study hours	exam score	study hours	exam score
1	3	75	4	70
2	15	95	12	98
3	6	65	9	85
4	8	70	6	80
5	4	85	2	65
6	2	80	3	75
7	10	65	10	92
Mean	6.86	76.43	6.57	80.71
SD	4.22	10.25	3.54	10.95

- a. Find the correlations between hours spent studying and how students performed in their potions and defense against the dark arts classes.

This requires computing two sets of correlations, 1 for each class. Again, you could use the covariance method of the Z score method. I've done both here as an example, with the covariance method in green and the Z score method in red. In the tables, X indicates hours studied and Y indicates exam scores

Potions Class

#	X	X-M _x	Z _x	Y	Y-M _y	Z _y	(X-M _x)(Y-M _y)	Z _x Z _y
1	3	-3.86	-0.91	75	-1.43	-0.14	5.51	0.13
2	15	8.14	1.93	95	18.57	1.81	151.22	3.49
3	6	-0.86	-0.20	65	-11.43	-1.11	9.80	0.23
4	8	1.14	0.27	70	-6.43	-0.63	-7.35	-0.17
5	4	-2.86	-0.68	85	8.57	0.84	-24.49	-0.57
6	2	-4.86	-1.15	80	3.57	0.35	-17.35	-0.40
7	10	3.14	0.74	65	-11.43	-1.11	-35.92	-0.83
Σ	48	0	0.00	535	0.00	0.00	81.43	1.88
Mean	6.86			76.43				
SD	4.22			10.25				

Then we add these to the formulas.

Using the covariance: $r = \frac{81.43}{(7 * 4.22 * 10.25)} = .27$

Using the Z scores: $r = \frac{1.88}{7} = .27$

Defense Against the Dark Arts Class

#	X	X-M _x	Z _x	Y	Y - M _y	Z _y	(X-M _x)(Y-M _y)	Z _x Z _y
1	4	-2.57	-0.73	70	-10.71	-0.98	27.55	0.71
2	12	5.43	1.53	98	17.29	1.58	93.84	2.42
3	9	2.43	0.69	85	4.29	0.39	10.41	0.27
4	6	-0.57	-0.16	80	-0.71	-0.07	0.41	0.01
5	2	-4.57	-1.29	65	-15.71	-1.43	71.84	1.85
6	3	-3.57	-1.01	75	-5.71	-0.52	20.41	0.53
7	10	3.43	0.97	92	11.29	1.03	38.69	1.00
Σ	46	0	0.00	565	0.00	0.00	263.14	6.79
Mean	6.57			80.71				
SD	3.54			10.95				

Then we add these to the formulas.

Using the covariance: $r = \frac{263.14}{(7 * 3.54 * 10.95)} = .97$

Using the Z scores: $r = \frac{6.79}{7} = .97$

- b. Which class was more strongly correlated with studying?

Studying was positively related to performing better on the exam in both classes, but there was a much stronger correlation between studying and exam scores in the Defense Against the Dark Arts class ($r = .97$) than in Potions ($r = .27$).