

SOLUTIONS TO HOMEWORK: CORRELATION

(Problem 1) The data below shows the reading and math scores of 12 students. Calculate the correlation coefficient.

Reading Scores	Math Scores
1	4
1	7
2	3
3	8
3	5
4	7
5	9
6	4
6	8
7	10
8	10
8	9

$$\sum X = 54; \sum Y = 84; \sum XY = 423; \sum X^2 = 314; \sum Y^2 = 654$$

$$r = \frac{540}{\sqrt{[852][792]}} = .657$$

If you test for significance, the calculated t with 10 degrees of freedom = 2.75. It is significant and we reject Ho that there is no correlation in the population between reading scores and math scores. There is a statistically significant positive correlation.

(Problem2)

The data below shows the hourly earnings (tips included) of 10 employees at a bar and their attractiveness scores (0 = not at all attractive ... 10= extremely attractive). What is the correlation between attractiveness and hourly earnings?

Attractiveness Score	Hourly Earnings
0	\$20
1	24
2	25
3	26
4	20
5	30
6	32
7	38
8	34
9	40

$$\sum X = 45; \sum Y = 289; \sum XY = 1472; \sum X^2 = 285; \sum Y^2 = 8801$$

$$r = \frac{1715}{\sqrt{[825][4489]}} = .891$$

If you test for significance, the calculated t with 8 degrees of freedom = 5.55. It is significant and we reject H_0 that there is no correlation in the population between attractiveness and hourly earnings. There is a statistically significant positive correlation.

(Problem 3)

The data below shows the score on a promotion test given to police officers and the number of hours studied. Calculate the correlation coefficient.

Hours Studied	Score on promotion test
0	0
1	0
2	1
3	4
4	5
6	6
8	8
16	8

$$\sum X = 40; \sum Y = 32; \sum XY = 262; \sum X^2 = 386; \sum Y^2 = 206$$

$$r = \frac{816}{\sqrt{[1488][624]}} = .847$$

(Problem 4)

The data below shows the height of students and the overall high school average. What is the correlation coefficient?

Height in Inches	High School Average
73	100
79	95
62	90
69	80
74	70
77	65
81	60
63	40
68	30
74	20

$$\sum X = 720; \quad \sum Y = 650; \quad \sum XY = 46,990; \quad \sum X^2 = 52,210; \quad \sum Y^2 = 49,150$$

$$r = \frac{1900}{\sqrt{[3700][69000]}} = 1900/15,978 = .119$$

This correlation coefficient is very low (not quite 0 but extremely low). If you do a statistical test on it, you will find that it is not significant, i.e., there is no significant correlation between height and high school average. The calculated t-value with 8 degrees of freedom is 0.34.

Statistically, speaking it is virtually impossible to get a sample correlation coefficient of exactly 0. This is why it is important to test the correlation coefficient for significance to make sure that you are not simply looking at a chance relationship.

(Problem 5)

The data below shows the number of pounds overweight and hourly wage of 10 employees working as secretaries in a law firm. Calculate the correlation coefficient.

Pounds Overweight	Hourly Wage
50	\$12
30	14
20	15
20	13
18	15
13	14
10	20
4	19
0	22
0	25

$$\sum X = 165; \quad \sum Y = 169; \quad \sum XY = 2308; \quad \sum X^2 = 4809; \quad \sum Y^2 = 3025$$

$$r = \frac{-4805}{\sqrt{[20865][1689]}} = -4805/5936.4 = -.809$$

$r = -.809$ There is a negative (inverse) correlation between pounds overweight and hourly wage.