

Assignment - Chapter 5

Complete the following problems related to Chapter 5. Upload to Canvas a single pdf containing your work before the deadline.

1. Assume that a study was conducted to analyze the relationship between the average weekly time spent on different hobbies (in hours) and the percentage of people reporting low stress levels. The data collected is as follows:

| Hobby | Average Weekly Time Spent (hours) | Percentage Reporting Low Stress Levels |
|---------------------|-----------------------------------|--|
| Gardening | 4 | 75 |
| Painting | 3 | 70 |
| Reading | 5 | 80 |
| Cooking | 6 | 65 |
| Playing Music | 2 | 85 |
| Watching TV | 10 | 40 |
| Exercising | 4 | 78 |
| Socializing | 7 | 60 |
| Meditation | 3 | 90 |
| Playing Video Games | 8 | 50 |

Make a scatterplot using the data above. Comment on the overall pattern you observe in the scatterplot. Describe whether you observe a positive or negative association between the time spent on hobbies and the percentage reporting low stress levels.

2. Think of any two scenarios where you expect to see a positive association such as temperature and ice-cream sales. Similarly, think of any two scenarios where you expect to see a negative association such as elevation and temperature.
3. For each of the following correlation coefficients, create a scatterplot that roughly illustrates the linear relationship:

$$r = -0.98, -0.5, 0, 0.5, 0.98$$

4. The following dataset was collected in an introductory statistics class to investigate the relationship between the hours studied for an exam and the exam score. The results are summarized in the table below.

| Student | Hours Studied (X) | Exam Score (Y) |
|---------|-------------------|----------------|
| 1 | 1 | 50 |
| 2 | 2 | 55 |
| 3 | 3 | 60 |
| 4 | 4 | 62 |
| 5 | 5 | 65 |
| 6 | 6 | 68 |
| 7 | 7 | 72 |
| 8 | 8 | 75 |
| 9 | 9 | 78 |
| 10 | 10 | 80 |

Using technology such as Wolfram Alpha, calculate the correlation coefficient between hours studied and exam score.

Let's introduce our 11th data point, which is an outlier, $(0, 98)$, i.e. 0 hours studied and an exam score of 98. Now, calculate the new correlation coefficient with the inclusion of this outlier. How did it change the correlation coefficient? Comment on your observations.