Correlation and Scatter Plots

Statistics Worksheet · Grade 9–12

Name: _____ Date: _____

Learning Objectives

- Identify the direction and strength of a linear relationship from a scatter plot
- Interpret the correlation coefficient r and its range from -1 to 1
- Distinguish between strong, moderate, and weak correlations using r values

Problems

1. The correlation coefficient r measures two things about a linear relationship between two quantitative variables. Name both things that r measures.

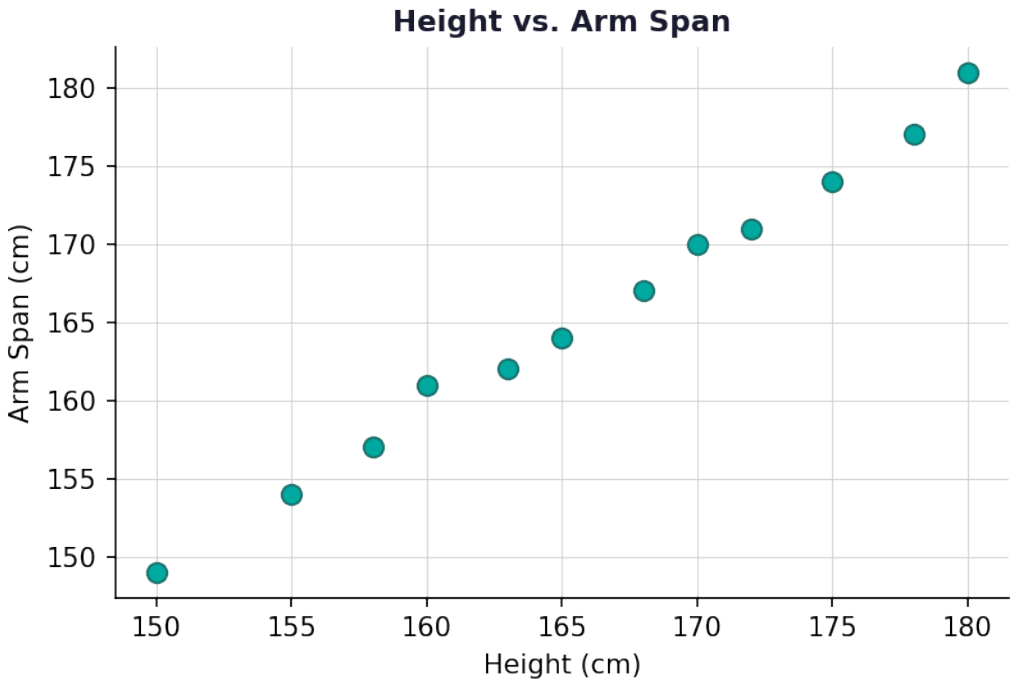
2. What is the full range of possible values for the correlation coefficient r? Choose the correct interval.

$$-1 \le r \le 1$$

3. Match each correlation value to the correct description: r = 0.97, r = 0.30, r = -0.85, r = 0.00. Descriptions: No relationship, Strong positive, Weak positive, Strong negative.

4. The scatter plot below shows the heights and arm spans of 12 students. Based on the scatter plot, describe the direction of the relationship and estimate whether r is closest to 0.85, -0.85, or 0.05.

Scan to watch



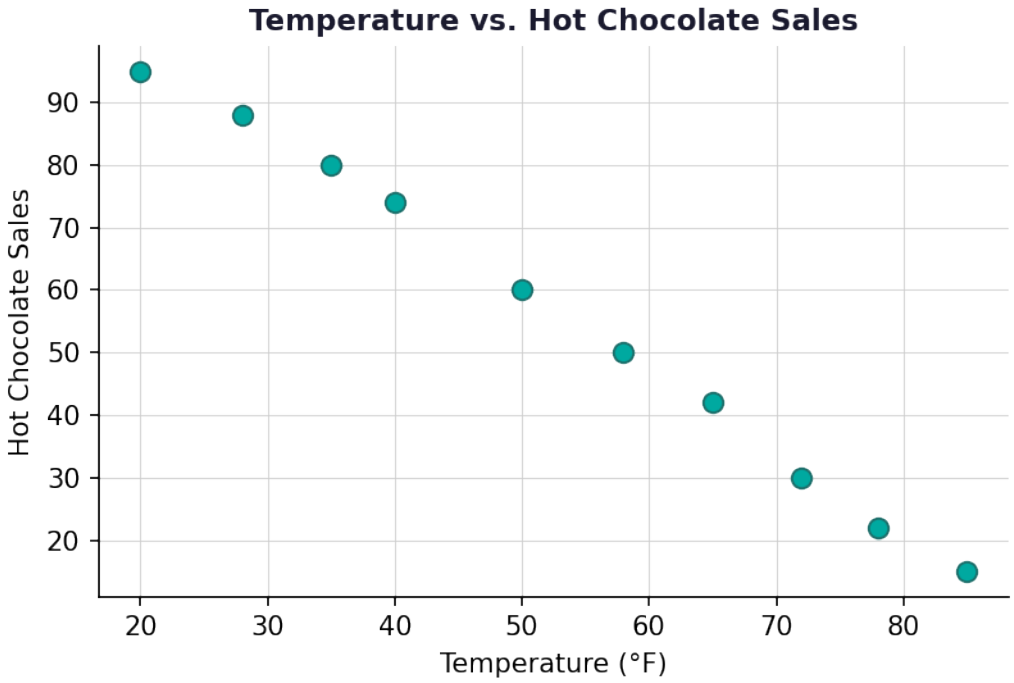
5. A researcher collects data on the number of hours students study per week and their exam scores. Use the data in the table to identify which pair of r values below is more reasonable for this data set: r = 0.78 or r = -0.78. Explain your reasoning.

Hours Studied	Exam Score
2	58
4	65
5	70
6	75
8	82
10	91

6. Two scatter plots (Graph A and Graph B) display the same data points but Graph A uses a much wider scale on the axes than Graph B. A student claims Graph B shows a stronger correlation because the points look tighter. Is the student correct? Explain using the definition of correlation.

Scan to watch

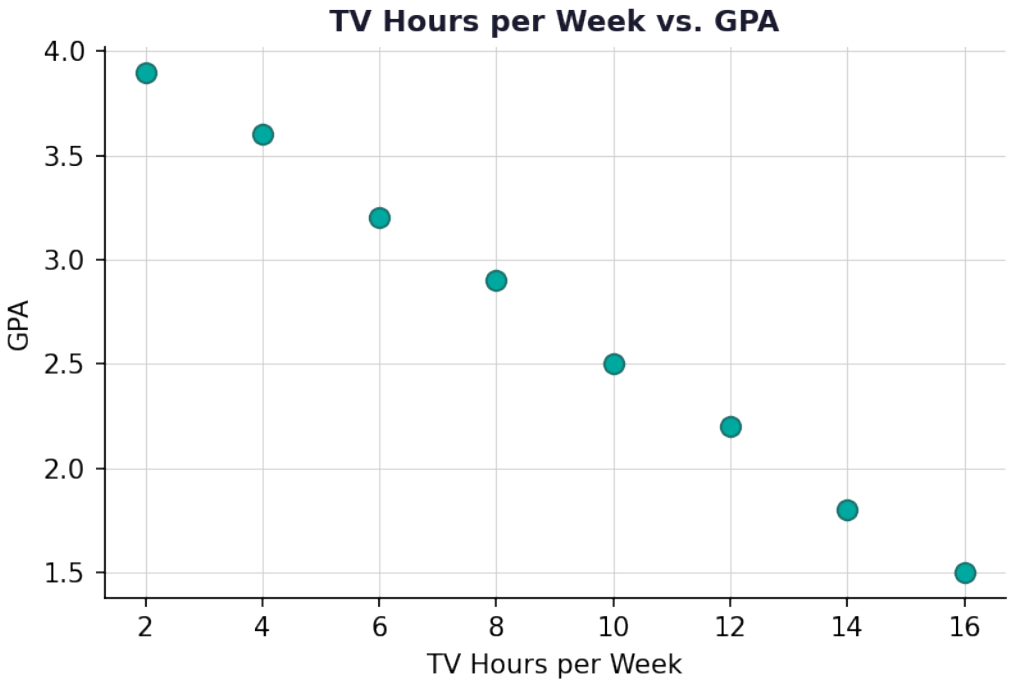
7. The scatter plot below shows daily temperature (in degrees Fahrenheit) versus hot chocolate sales at a cafe. Estimate the correlation coefficient r, state the direction, and describe the strength of the relationship.



8. A student computes correlation coefficients for four different data sets. Order the following r values from weakest to strongest relationship: r = -0.91, r = 0.45, r = -0.10, r = 0.76.

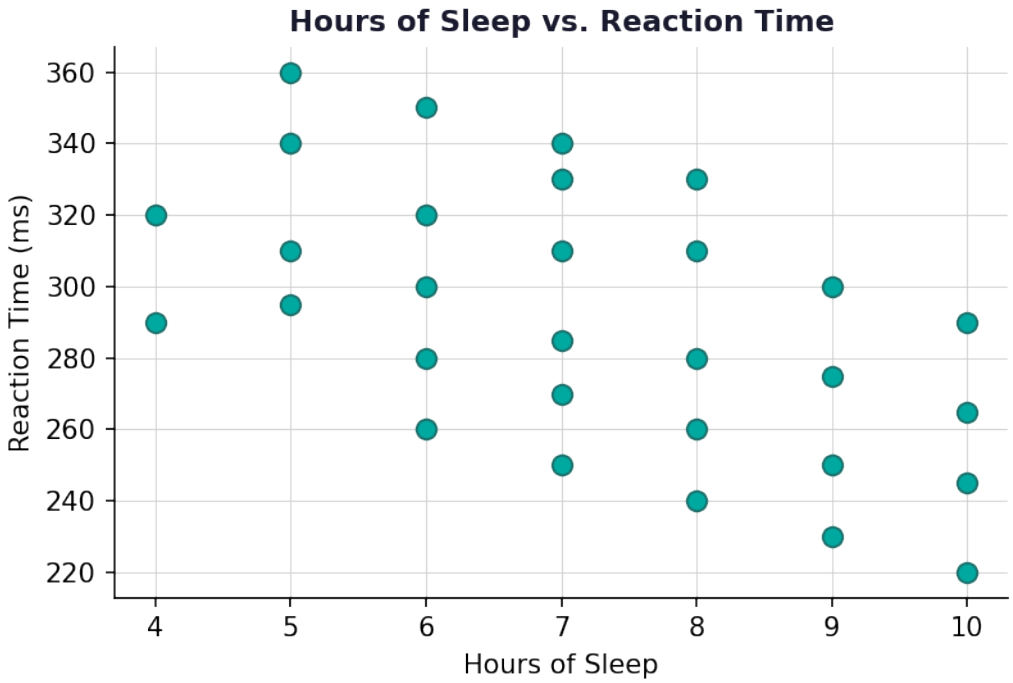
9. The scatter plot below shows hours of television watched per week versus GPA for 8 high school students. Estimate the best-fit line and the correlation coefficient r. Then describe what happens to r as the data points spread farther from the line.

Scan to watch



10. A researcher studying 30 students records hours of sleep and reaction time (in milliseconds). The computed correlation is r = 0.299. Another researcher claims that adding 20 more students whose data spreads widely across the scatter plot will increase r toward 1. Is this claim correct? Justify your answer using what you know about how scatter and sample size affect r. Also state what r = 0.299 tells us about the relationship.

Scan to watch



Scan to watch

Correlation and Scatter Plots — Answer Key

Statistics Worksheet · Grade 9–12

Answer Key

1. Answer: Direction and strength

- The correlation coefficient r measures the direction of the linear relationship (positive or negative).
- It also measures the strength of the linear relationship (how closely the data follows a straight line).

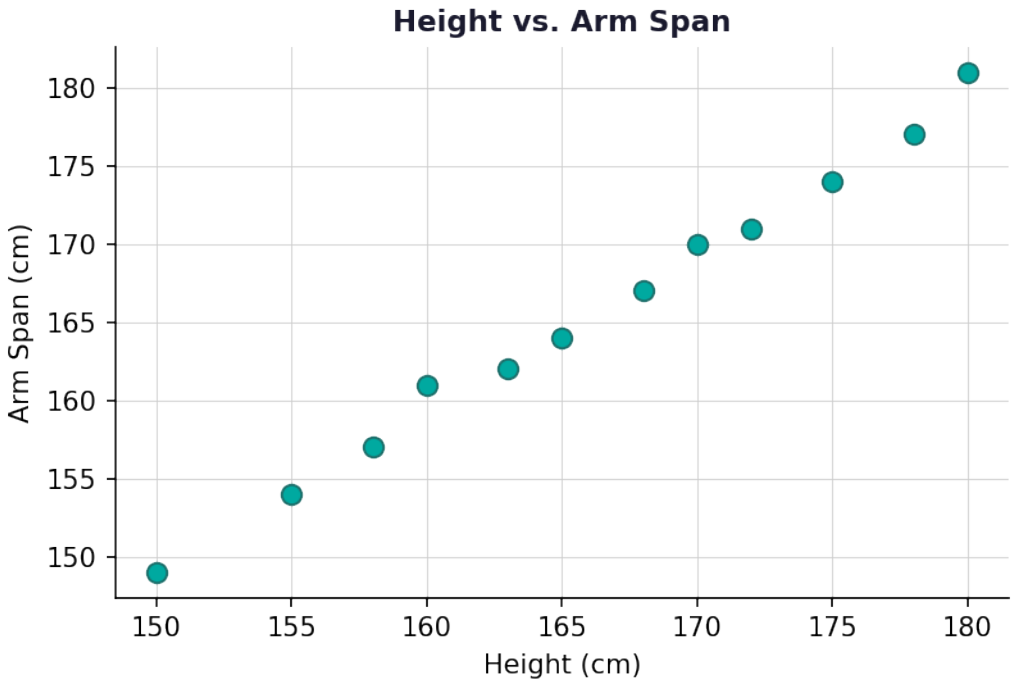
2. Answer: r ranges from -1 to 1, inclusive

- The correlation coefficient r can never be less than -1 or greater than 1.
- Therefore the complete range is $-1 \leq r \leq 1$.

3. Answer: r = 0.97 → Strong positive; r = 0.30 → Weak positive; r = -0.85 → Strong negative; r = 0.00 → No relationship

- Values close to 1 indicate a strong positive relationship: r = 0.97 is Strong positive.
- Values close to 0 but positive indicate a weak positive relationship: r = 0.30 is Weak positive.
- Values close to -1 indicate a strong negative relationship: r = -0.85 is Strong negative.
- r = 0.00 means no linear relationship at all.

4. Answer: Positive direction; r ≈ 0.85



- As height increases, arm span also increases, so the direction is positive.
- The points cluster fairly closely around a line, indicating a strong relationship.
- Therefore r is closest to 0.85 (strong positive).

Scan to watch

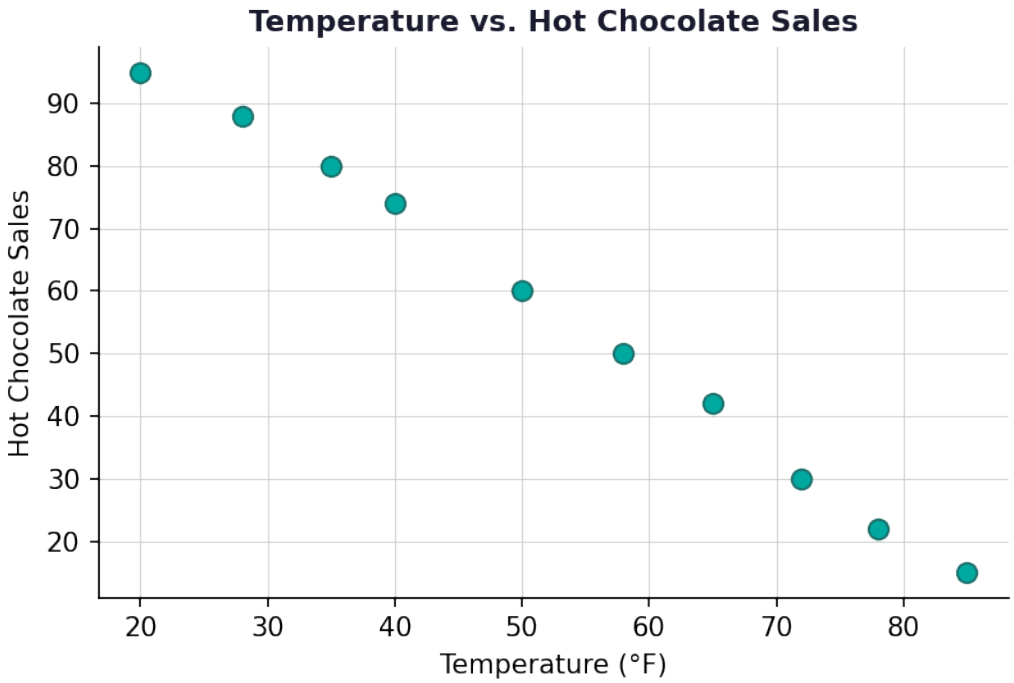
5. Answer: r = 0.78 (strong positive); as study hours increase, exam scores increase

- Examine the table: as hours studied increase (2 → 10), exam scores also increase (58 → 91).
- This is a positive relationship, so r must be positive.
- The points are fairly close to a line, suggesting a strong correlation, making r = 0.78 the reasonable choice.

6. Answer: No. The correlation r does not change with axis scaling. Both graphs have the same r value.

- Correlation r is a numerical measure computed from the actual data values, not the appearance of the graph.
- Changing the scale of the axes makes points look closer together or farther apart visually, but does not change r.
- Therefore both Graph A and Graph B have identical r values even though they look different.

7. Answer: r ≈ -0.99; strong negative linear relationship



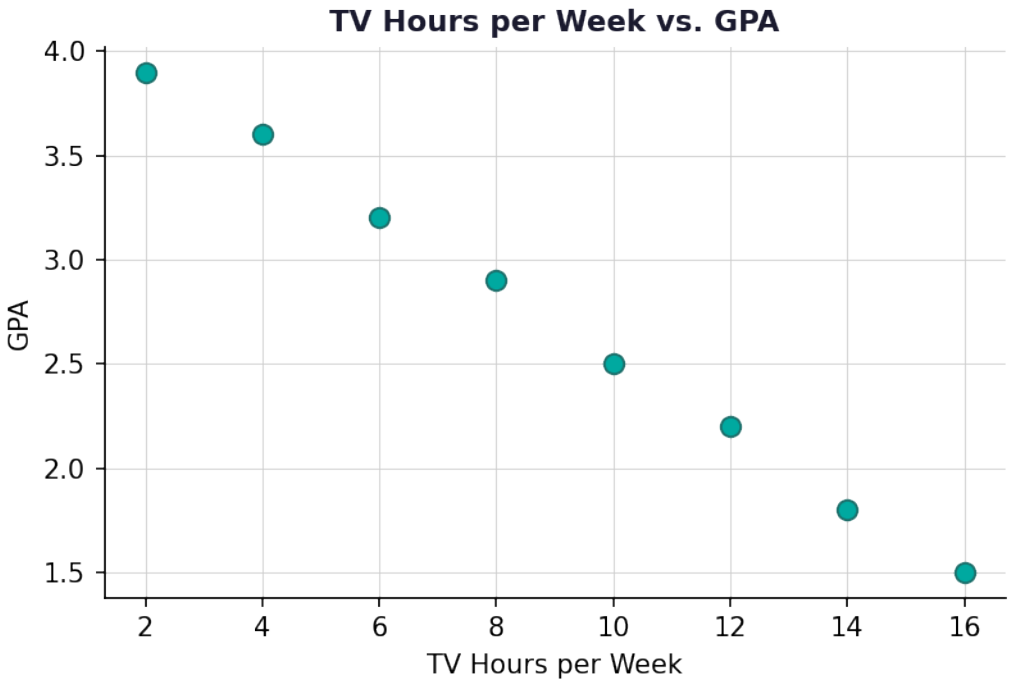
- As temperature increases, hot chocolate sales decrease — this is a negative direction.
- The points fall very close to a straight line, indicating a very strong relationship.
- Therefore r is close to -1, approximately -0.99.

8. Answer: Weakest to strongest: r = -0.10, r = 0.45, r = 0.76, r = -0.91

- Strength is determined by the absolute value |r|, regardless of sign.
- |–0.10| = 0.10 (weakest), |0.45| = 0.45, |0.76| = 0.76, |–0.91| = 0.91 (strongest).
- Ordered from weakest to strongest: r = -0.10, r = 0.45, r = 0.76, r = -0.91.

9. Answer: r ≈ -0.99; strong negative. As points spread from the line, |r| decreases toward 0.

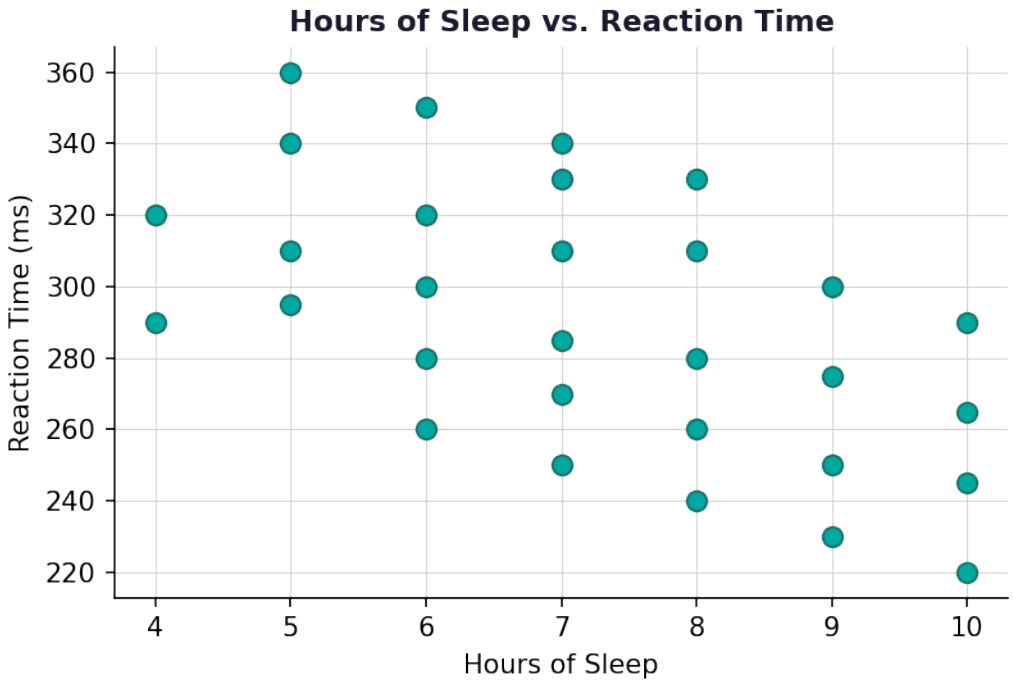
Scan to watch



- As TV hours increase, GPA decreases — a negative direction.
- The points lie very close to a straight line, so r ≈ -0.99 (strong negative).
- When data points spread farther from the line (more scatter), the r value moves closer to 0, indicating a weaker relationship.

10. Answer: Claim is incorrect. Widely spread points pull r closer to 0, weakening the relationship. r = 0.299 indicates a weak positive linear relationship.

Scan to watch



• r = 0.299 is close to 0, which means the current relationship between sleep hours and reaction time is weak and positive.
• Adding students whose data is spread widely across the scatter plot increases scatter around the line of best fit.
• Greater scatter causes the correlation r to move closer to 0, making the relationship weaker — not stronger.
• Therefore the researcher's claim is incorrect; adding widely scattered points would decrease |r|, not increase it toward 1.

Scan to watch