

AP Statistics - Scatterplots and Correlation - Worksheet

Common Core AP STAT 2.4, AP STAT 2.5, HSS-ID.B.6, HSS-ID.C.8, HSS-ID.C.9 · maths-pdf.fr/en/

Name: _____ Class: _____ Date: ____/____/____

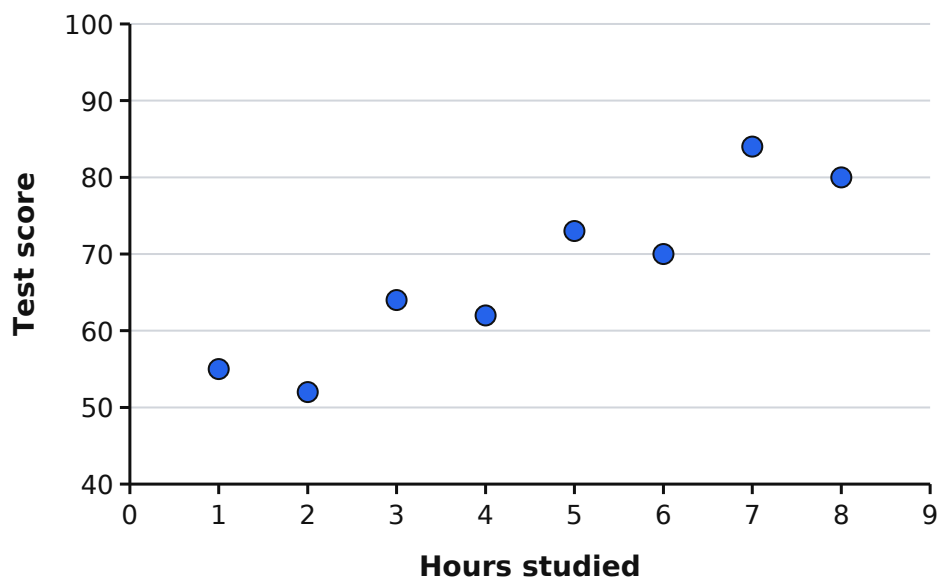
These 20 problems practice AP Statistics topics 2.4 and 2.5. You will describe scatterplots, match plots with correlations, compute r from z-scores, and explain how outliers, units, and lurking variables affect correlation.

AP Statistics · Common Core State Standards: **AP STAT 2.4, AP STAT 2.5, HSS-ID.B.6, HSS-ID.C.8, HSS-ID.C.9**

Exercise 1 - Explanatory and response

Name the explanatory and response variables: (a) predicting a student's final exam score from the number of absences; (b) predicting the price of a used car from its mileage.

Exercise 2 - Describe a scatterplot



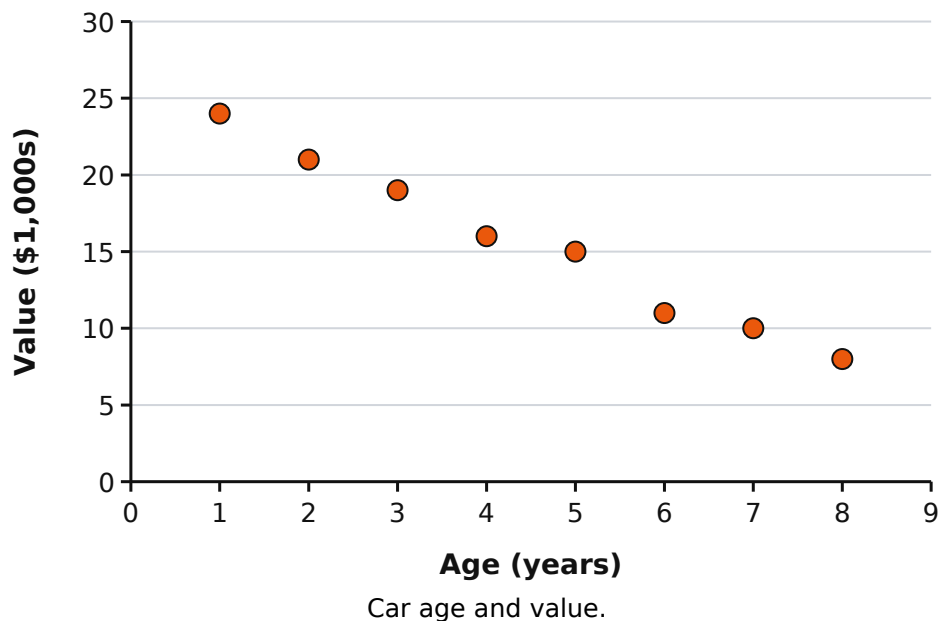
Positive, roughly linear, strong association between hours studied and test score ($r = 0.934$).

Describe the association shown in the scatterplot of hours studied and test score.

Exercise 3 - Direction and form

Describe the direction and form: (a) outdoor temperature and heating costs; (b) hours of sleep and number of mistakes on a test; (c) a person's age and height from age 1 to 60.

Exercise 4 - Match r to the plot



Which is the best estimate of r for the plot: -0.98, -0.5, 0, or 0.9? Explain.

Exercise 5 - Compute r by hand

Five students: (hours, score) = (1, 2), (2, 4), (3, 5), (4, 4), (5, 5). The means are 3 and 4; the standard deviations are 1.581 and 1.095. Find r from z-scores.

Exercise 6 - Interpret r

A study of 40 cities reports $r = -0.72$ between average winter temperature and heating costs. Interpret r in context.

Exercise 7 - Change of units

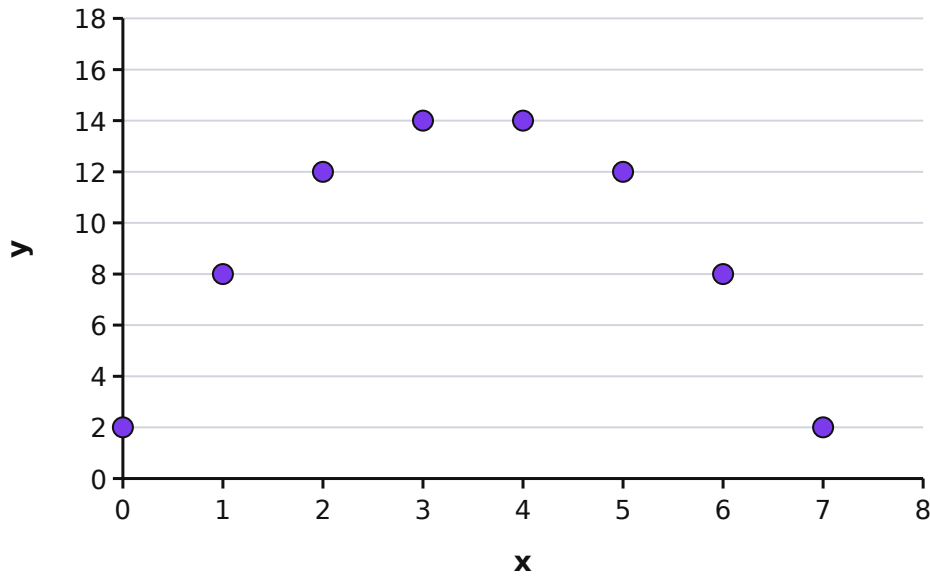
The correlation between height in inches and weight in pounds is $r = 0.62$. What is r if height is

measured in centimeters and weight in kilograms? Explain.

Exercise 8 - Swap the variables

The correlation between x and y is 0.45. What is the correlation between y and x ?

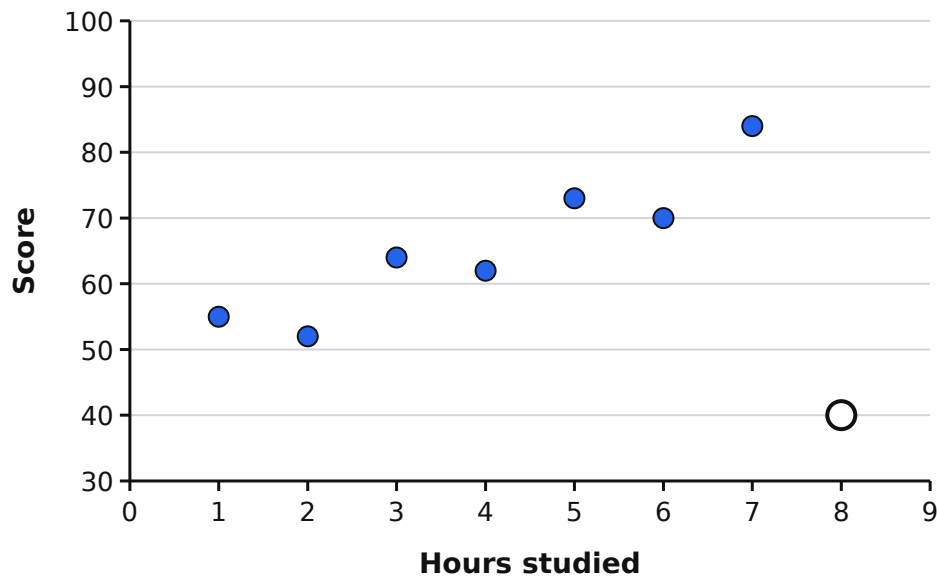
Exercise 9 - Zero but related



A curved pattern.

The scatterplot shows a clear arch. r is 0.000. Is there no relationship between the variables? Explain.

Exercise 10 - Effect of an outlier



Outlier at (8, 40).

The original data (hours, score) had $r = 0.934$. One score is changed to (8, 40). The new r is 0.179. Explain the change.

Exercise 11 - Correlation and causation

The monthly number of ice cream sales and the number of sunburns are strongly positively correlated. Does eating ice cream cause sunburn? Name a lurking variable.

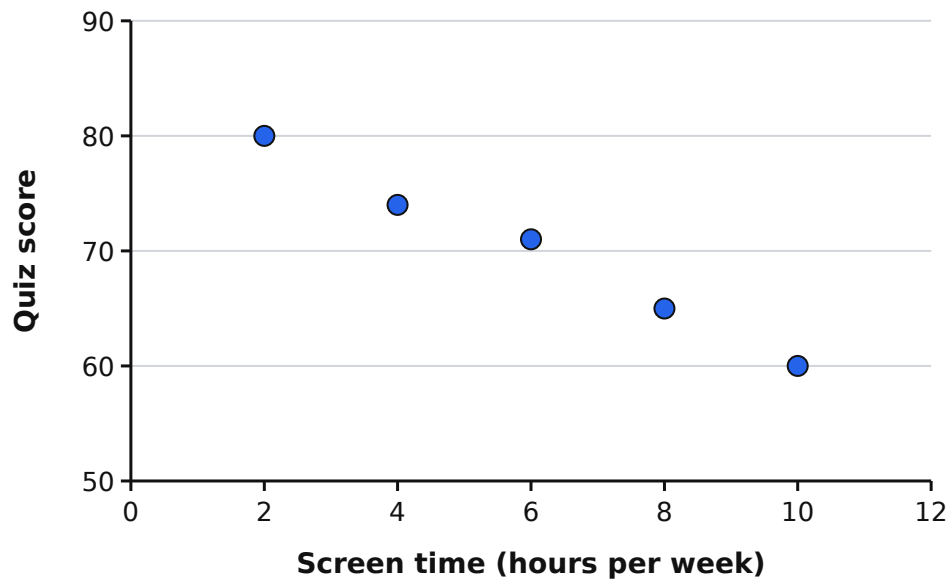
Exercise 12 - Steep is not strong

Maya says that because the points on a scatterplot rise steeply, r must be close to 1. Is she right?

Exercise 13 - Rank the strength

Order from weakest to strongest linear association: $r = 0.8$, $r = -0.3$, $r = -0.95$, $r = 0.1$.

Exercise 14 - Screen time



Five students' weekly screen time (hours) and quiz scores are (2, 80), (4, 74), (6, 71), (8, 65), (10, 60). Describe the association and give r from a calculator.

Exercise 15 - Which variable on the x-axis?

A researcher wants to see whether the number of hours of exercise per week helps explain resting heart rate. Which variable goes on the horizontal axis? Would a positive or negative association be expected?

Exercise 16 - Two clusters

A scatterplot of height and weight for all students in a school shows two clusters of points, one for younger and one for older students, each with a weak trend. Explain how this could change r .

Exercise 17 - Calculator output

A calculator gives $r = 0.87$ and $r^2 = 0.757$ for a data set. Describe the association, and state what the correlation tells you about causation.

Exercise 18 - Weaken or strengthen

A scatterplot of 10 points has $r = 0.82$. An 11th point is added far above the line at the right end. Will r likely increase or decrease? Explain.

Exercise 19 - Describe in context

A scatterplot of 50 cities shows average income (x) and average rent (y). Points rise in a straight pattern with moderate scatter. One city has very high income and rent. Write a description.

Exercise 20 - Predictor choice

A coach records each runner's weekly mileage and 5K time. Which variable is explanatory? Would you expect a positive or a negative association, and would r be closer to -1 or to 0 for runners of a similar level?