

AP Statistics - Scatterplots and Correlation - Answer Key

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/

Each solution states direction, form, strength, and unusual features where a description is asked, and shows the z-score products when r is computed by hand. Values are rounded to three decimals.

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

(a) Explanatory: absences; response: final exam score. (b) Explanatory: mileage; response: price.

Exercise 2 - Describe a scatterplot

As study hours increase, scores increase. The points follow a straight pattern closely and there are no outliers.

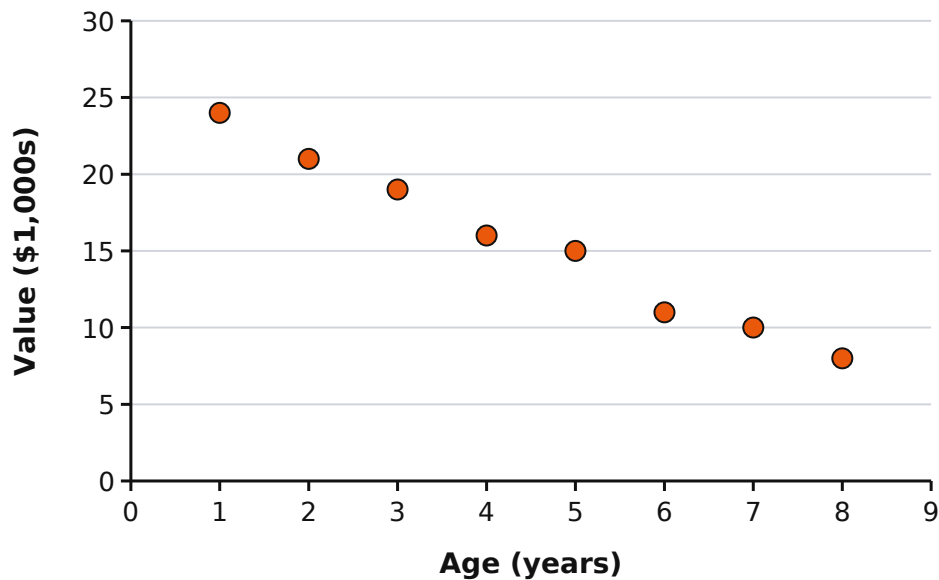
Positive, linear, strong association; $r = 0.934$ with no unusual points.

Exercise 3 - Direction and form

(a) As temperature rises, heating costs fall: negative, roughly linear. (b) More sleep, fewer mistakes: negative. (c) Height rises quickly then levels off: positive, curved.

(a) negative linear; (b) negative; (c) positive and curved.

Exercise 4 - Match r to the plot



Car age and value.

The points fall steadily with very little scatter.

r is about -0.994, closest to -0.98: strong and negative.

Exercise 5 - Compute r by hand

x	y	z of x	z of y	product
1	2	-1.265	-1.633	2.066
2	4	-0.632	0.000	-0.000
3	5	0.000	0.816	0.000
4	4	0.632	0.000	0.000
5	5	1.265	0.816	1.033
			sum	3.098

z-scores and products.

$$r = \frac{1}{4}(3.098) = 0.775.$$

$r = 0.775$, a fairly strong positive linear association.

Exercise 6 - Interpret r

The sign is negative and the size is moderately strong.

There is a moderately strong, negative, linear association: cities with warmer winters tend to have lower heating costs.

Exercise 7 - Change of units

Changing units is a linear transformation of each variable. z-scores do not change.

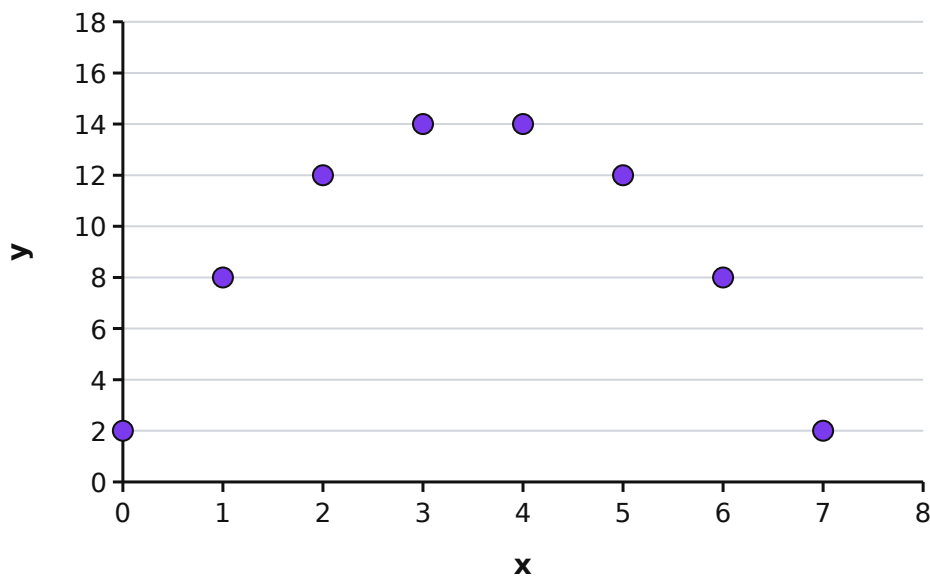
$r = 0.62$; correlation has no units and is unchanged by linear changes of units.

Exercise 8 - Swap the variables

The formula multiplies z-scores, and multiplication is commutative.

0.45.

Exercise 9 - Zero but related

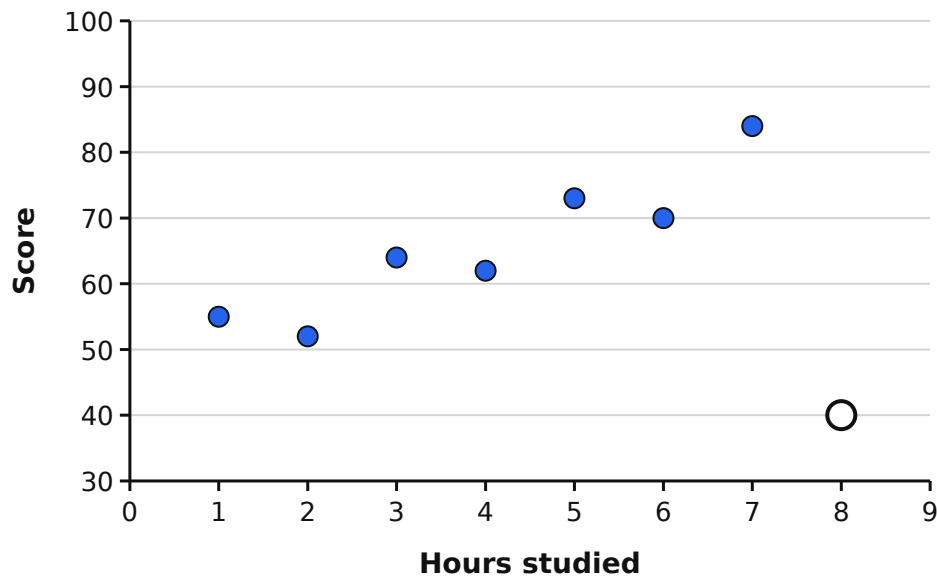


A curved pattern.

The correlation is close to zero because the pattern is curved, not linear.

No: there is a strong curved relationship. r measures only linear association.

Exercise 10 - Effect of an outlier



Outlier at (8, 40).

The point (8, 40) lies far from the trend, so it weakens the linear pattern.

r falls from 0.934 to 0.179. Correlation is not resistant to outliers.

Exercise 11 - Correlation and causation

A third variable could drive both.

No. Hot, sunny weather is a lurking variable: it increases both ice cream sales and sunburns.

Exercise 12 - Steep is not strong

Strength depends on how tightly points cluster around a line, not on the slope.

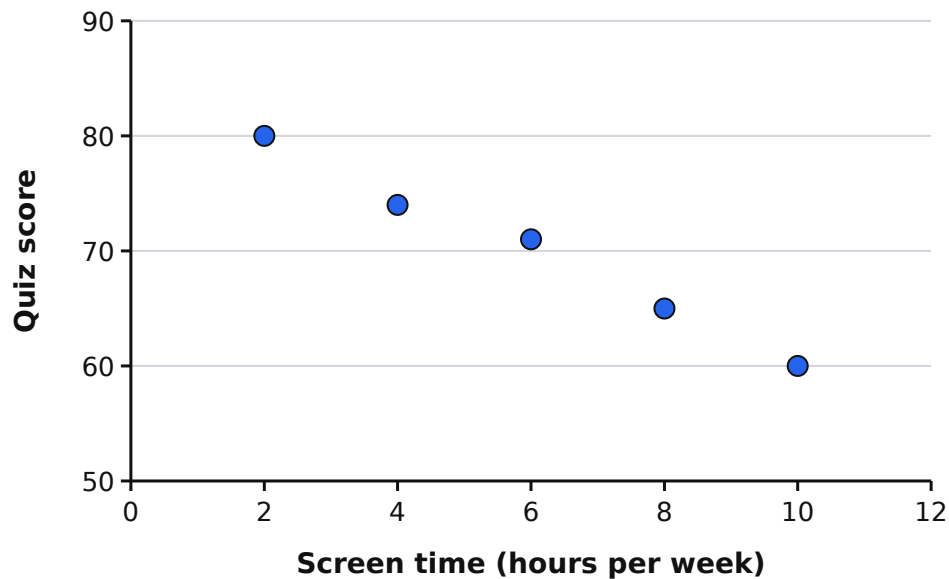
No. A steep cloud of widely scattered points has a weak r , and a nearly flat tight line has r near 1 or -1 in absolute value.

Exercise 13 - Rank the strength

Strength depends on $|r|$: 0.8, 0.3, 0.95, 0.1.

0.1, -0.3, 0.8, -0.95.

Exercise 14 - Screen time



Screen time and score.

Scores decrease steadily as screen time rises.

Negative, linear, very strong; $r = -0.996$.

Exercise 15 - Which variable on the x-axis?

The explanatory variable is on the x-axis.

Hours of exercise on the x-axis; a negative association is expected (more exercise, lower resting heart rate).

Exercise 16 - Two clusters

Together, the clusters line up from lower left to upper right even if each has a weak trend inside it.

The combined r is stronger than within each cluster; the lurking variable is age. Describe the clusters as an unusual feature.

Exercise 17 - Calculator output

$r = 0.87$ is strongly positive.

There is a strong, positive linear association. Correlation alone does not show that changes in x cause changes in y .

Exercise 18 - Weaken or strengthen

The new point is far from the existing pattern (far above the line).

r will likely decrease, because the point adds vertical scatter around the linear trend.

Exercise 19 - Describe in context

Use direction, form, strength, unusual features, and context.

There is a moderately strong, positive, linear association between average income and average rent in these cities; one city with very high income and rent is an unusual point.

Exercise 20 - Predictor choice

Mileage is used to explain 5K time. More weekly mileage usually lowers the time.

Explanatory: weekly mileage. Negative association; for runners at a similar level r would be weak, closer to 0, because other factors matter more.