

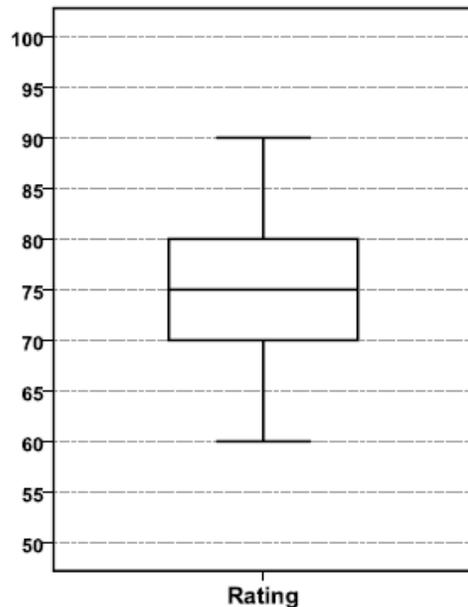


1. The head circumference (in centimeters) of 15 college-age males was obtained, resulting in the following measurements: 55, 56, 56, 56.5, 57, 57, 57, 57.5, 58, 58, 58, 58.5, 59, 59, 63. If the last measurement (63 cm's) were incorrectly recorded as 73, which one of the following statistics would change?

- a) Q1 (1st quartile)
- b) Standard deviation**
- c) Median
- d) Q3 (3rd quartile)

2. The following boxplot gives the distribution of the ratings of a new brand of peanut butter for 50 randomly selected consumers (100 points possible with higher points corresponding to a more favorable rating).

Identify the five-number summary and the description of each of them using the example.



- Min – 60 points**
- Q1 – 70 points**
- Median – 75 points**
- Q3 – 80 points**
- Max – 90 points**

This is a standard deviation contest. You must choose four numbers from the whole numbers 0 to 10, with repeats allowed.

1. Choose four numbers that have the smallest possible standard deviation.
- a) 7, 7, 7, 8
 - b) 3, 5, 7, 9
 - c) 1, 1, 1, 1**
 - d) 1, 2, 3, 4

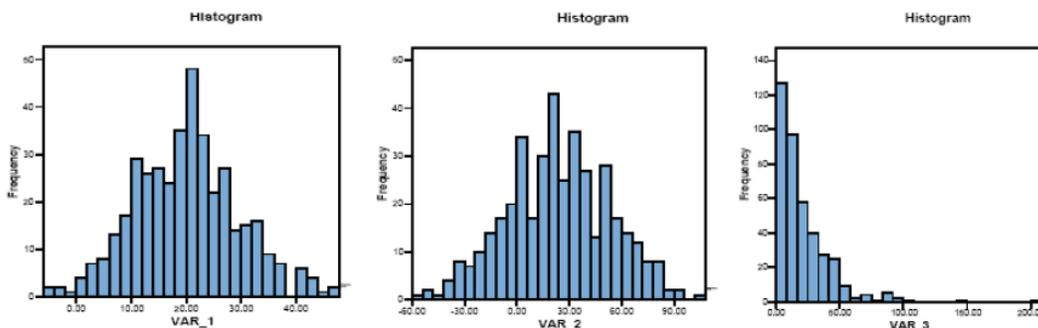
2. Choose four numbers that have the largest possible standard deviation.
- a) 0, 3, 6, 10
 - b) 9, 9, 10, 10
 - c) 1, 4, 7, 10
 - d) 0, 0, 10, 10**

Researchers are interested in how crime rates are different for southern states.

3. The standard deviation of unemployment for 35 to 39-year-old for southern states is 8.01 and the standard deviation for non-southern states is 8.76. What do these standard deviations tell you about the southern and non-southern states?
- a. The average unemployment rate for 35 to 39-year old in southern states is about the same as in non-southern states.
 - b. The average unemployment rate for 35 to 39-year old in southern states is slightly lower than in non-southern states.
 - c. The average spread from the mean rate for 35 to 39-year-old in southern states is about the same as in non-southern states.**

Below, you are given the data summary of three datasets: the mean, median and standard deviation of each one.

DATASET	A	B	C
Mean	23.61	20.49	20.03
Median	22.49	13.80	19.06
Standard Deviation	29.54	21.78	9.42



4. Match each data summary to its corresponding histogram (Plot 1, Plot 2, Plot 3)



- a) Data set A Plot 2
- b) Data set B Plot 3
- c) Data set C Plot 1

In each of the following situations, is it more reasonable to simply explore the relationship between the two variables or to view one of the variables as an explanatory variable and the other as a response variable? In the latter case, which is the explanatory variable and which is the response variable? Are they categorical or quantitative (quantitative means "numerical")?

- 5. The typical number of calories a person consumes per day and that person's percent of body fat.
 - a) Number of calories consumed per day: response, quantitative. Percent of body fat: explanatory, quantitative.
 - b) Number of calories consumed per day: explanatory, quantitative. Percent of body fat: response, quantitative.**
 - c) Number of calories consumed per day: response, quantitative. Percent of body fat: explanatory, categorical.
 - d) Number of calories consumed per day: explanatory, categorical. Percent of body fat: response, categorical.

- 6. Water temperature controlled at different levels and growth (measured by weight) of corals in aquariums.
 - a. Water temperature: response, quantitative. Growth: explanatory, categorical.
 - b. Water temperature: explanatory, categorical. Growth: response, categorical.
 - c. Water temperature: response, categorical. Growth: explanatory, quantitative.
 - d. Water temperature: explanatory, quantitative. Growth: response, quantitative**

Coffee is a leading export from several developing countries. When coffee prices are high, farmers often clear forest to plant more coffee trees. Here are data on prices paid to coffee growers in Indonesia and the rate of deforestation in a national park that lies in a coffee-producing region, for five years:



Price (cents per pound)	Deforestation (percent)
29	0.49
40	1.59
54	1.69
55	1.82
72	3.10

7. Coffee is currently priced in dollars. If it were priced in euros, and the dollar prices in the above table were translated into the equivalent prices in euros, would the correlation between coffee price and percent deforestation change?
- a) The correlation would remain zero, because the two variables are independent
 - b) Yes, units affect correlation
 - c) No, units do not affect correlation
 - d) It is impossible to calculate the correlation, because coffee price is categorical.

A study shows that there is a positive correlation between the size of a hospital (measured by its number of beds (x)) and the median number of days (y) that patients remain in the hospital.

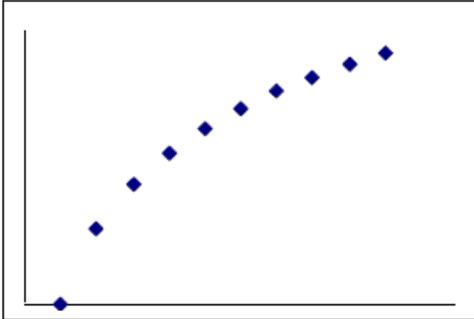
8. What lurking variable could be present in this study?
- a) cost: its more expensive to run larger hospitals.
 - b) severity of disease: since large hospitals have better facilities and more doctors to cope with severe illness
 - c) number of visitors: since larger hospitals receive more visitors.
 - d) facilities: since larger hospitals have better facilities, patients choose to stay longer

Over the past decade, there has been a strong positive correlation between teacher salaries and prescription drug costs.

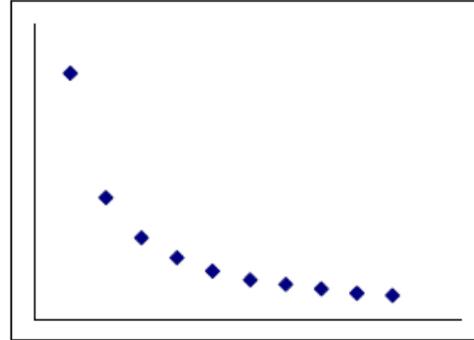
9. Do you think paying teachers more causes prescription drugs to cost more?
- a) Yes. A strong correlation always implies causation.
 - b) No. A strong correlation can never go along with causation.
 - c) Yes. In this case, a strong correlation likely implies causation.
 - d) No. In this case, a strong correlation does not imply causation.

10. Which of the following plots will have a correlation coefficient of .85? **D**

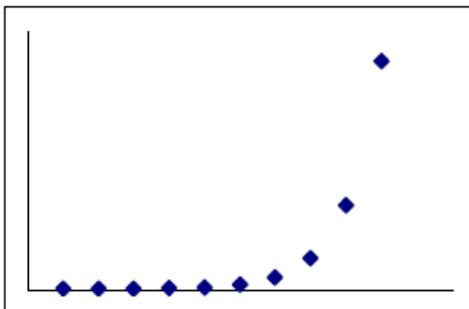
A.



B.



C.



D.

