- mean: 68.92 bpm
- standard deviation: 8.9 bpm
- minimum: 55 bpm
- Q1: 61 bpm

- median: 70 bpm
- Q3: 75.5 bpm
- maximum: 85 bpm

The mean decreased by 0.86 bpm and the median remained the same. The standard deviation decreased by 1.81 bpm which is about 17% of its previous value. Based on the standard deviation, the data set with the outlier removed shows much less variability than the original data set containing the outlier. Since the mean and standard deviation use all of the numerical values, removing one very large data point can affect these statistics in important ways.

The median remained the same after the removal of the outlier and the IQR increased slightly. These measures of center and variability are much more resistant to change than the median and standard deviation. The median and IQR measure the middle of the data based on the number of values rather than the actual numerical values themselves, so the loss of a single value will not often have a great effect on these statistics.

The source of any possible errors should always be investigated. If the measurement of 112 beats per minute was found to be taken under the right conditions and merely included an athlete whose heart rate did not slow as much as the other athletes, it should not be deleted so that the data reflect the actual measurements. If the situation cannot be revisited to determine the source of the outlier, it should not be removed. To avoid tampering with the data and to report accurate results, data values should not be deleted unless they can be confirmed to be an error in the data collection or data entry process.

Glossary

- outlier

Lesson 14 Practice Problems
Problem 1
Statement

The number of letters received in the mail over the past week is recorded.

| 2 | 3 | 5 | 5 | 5 | 15 |

Which value appears to be an outlier?

A. 2

B. 3

C. 5

D. 15

Solution

D

Problem 2

Statement

Elena collects 112 specimens of beetle and records their lengths for an ecology research project. When she returns to the laboratory, Elena finds that she incorrectly recorded one of lengths of the beetles as 122 centimeters (about 4 feet). What should she do with the outlier, 122 centimeters, when she analyzes her data?

Solution

Sample response: Elena should go back to the 112 specimens and measure them again to find the error. If she cannot do this, then she should eliminate the outlier from her analysis, because this is an error in recording.

Problem 3

Statement

Mai took a survey of students in her class to find out how many hours they spend reading each week. Here are some summary statistics for the data that Mai gathered:

mean: 8.5 hours   standard deviation: 5.3 hours

median: 7 hours   Q1: 5 hours   Q3: 11 hours

a. Give an example of a number of hours larger than the median which would be an outlier. Explain your reasoning.

b. Are there any outliers below the median? Explain your reasoning.

Solution

a. Sample response: 22 hours, since Q3 together with one and a half times the interquartile range is 20.

b. No, because outliers below the median must be less than -4 ($5 - 1.5 \cdot (11 - 5) = $ -4), and it is not possible to read for a negative number of hours.

Problem 4

Statement

The box plot shows the statistics for the weight, in pounds, of some dogs.

weight of dog (pounds)
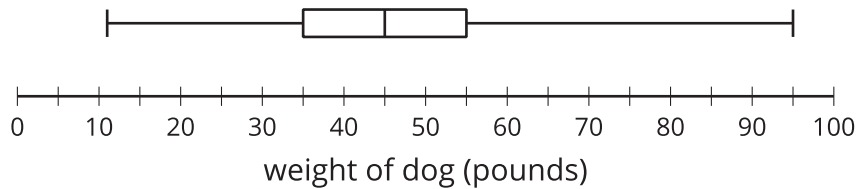
Are there any outliers? Explain how you know.

Solution

Yes, there is at least one outlier that is too high. The dog that weighs about 95 pounds is an outlier. The IQR is about 20 and $1.5 \cdot 20 = 30$. The dog that weighs about 95 pounds is about 40 pounds above the third quartile.

Problem 5

Statement

The mean exam score for the first group of twenty examinees applying for a security job is 35.3 with a standard deviation of 3.6.

The mean exam score for the second group of twenty examinees is 34.1 with a standard deviation of 0.5. Both distributions are close to symmetric in shape.

a. Use the mean and standard deviation to compare the scores of the two groups.

b. The minimum score required to get an in-person interview is 33. Which group do you think has more people get in-person interviews?

Solution

a. Sample response: The first group scored higher on average based on the mean, but showed much greater variability than the second group, because the first group had a higher standard deviation.

b. Sample response: I think more people in the second group get in-person interviews. The minimum score is 33 and they have a mean of 34.1 with a standard deviation of 0.5, so the bulk of their scores had to be above 33. If not, they would have had a standard deviation greater than 1.1.

(From Unit 1, Lesson 13.)

Problem 6

Statement

A group of pennies made in 2018 are weighed. The mean is approximately 2.5 grams with a standard deviation of 0.02 grams.

Interpret the mean and standard deviation in terms of the context.

Solution

The average penny weighs 2.5 grams. There is a little bit of variation in the weights, but most of the pennies weigh very close to 2.5, within 0.02 grams on average as measured by the standard deviation.

(From Unit 1, Lesson 13.)

Problem 7

Statement

These values represent the expected number of paintings a person will produce over the next 10 days.

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

    a. What are the mean and standard deviation of the data?

    b. The artist is not pleased with these statistics. If the 5 is increased to a larger value, how does this impact the median, mean, and standard deviation?
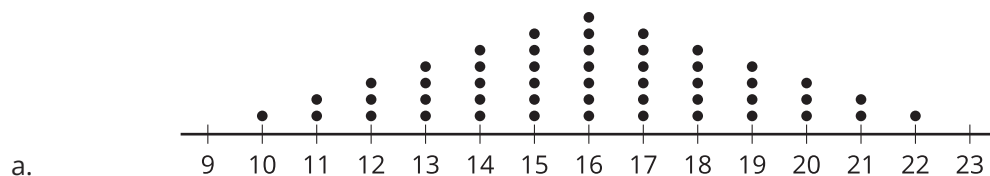
Solution

    a. Both the mean and the standard deviation are 1.5 paintings.

    b. The median will still be 1 painting. The mean will increase, and so will the standard deviation, because the data is more spread out.
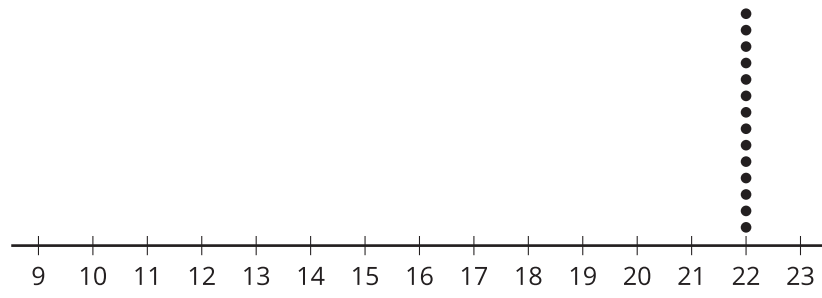
(From Unit 1, Lesson 12.)

Problem 8

Statement

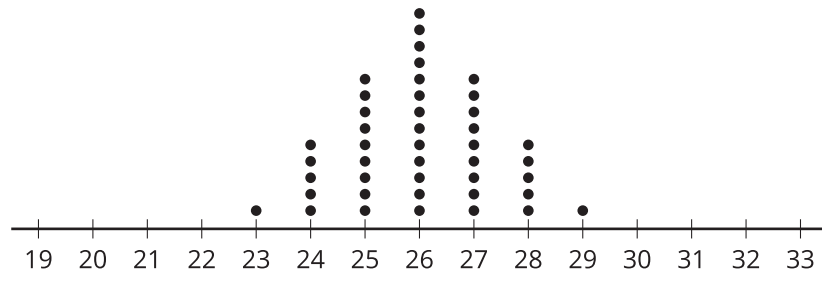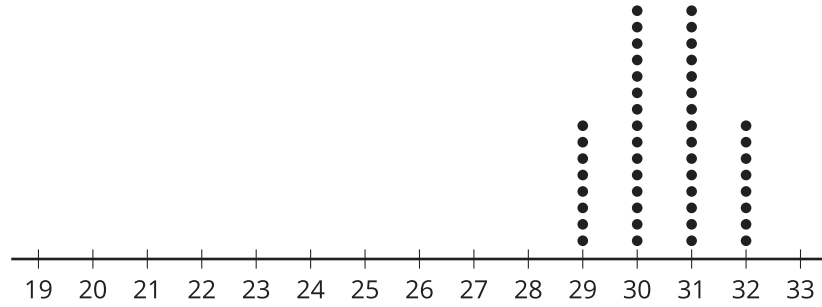List the four dot plots in order of variability from least to greatest.

a.

b.

c.

d.

Solution

b, d, c, a

(From Unit 1, Lesson 11.)