

1. E.

The mean is the sum of all the values divided by the number of values. It can be strongly affected by outliers.

F. The median is the middle value in a data set. It is not affected by outliers.

F. The mode is the most common value in a data set. It is not affected by outliers.

2. A.

An outlier in a data set is a value that is much higher or much lower than almost all other values. An outlier can change the mean of a data set, but does not affect the median or mode.

3. A. A single-peaked distribution is symmetric if its left half is a mirror image of its right half.

B. The heights of a sample of 100 women is a symmetric distribution.

C.

The speed of cars on a road where a visible patrol car is using radar to detect speeders is a left-skewed distribution.

C. The number of books read during the school year by fifth graders is a right-skewed distribution.

4. D.

The statement makes sense because the mean of a data set can be pulled far to the left or right of the middle of the range of values by outliers.

5. A.

This does not make sense because the mean and median should lie somewhere to the left of the mode if the distribution is left-skewed.

6. A. The mean is 55.4. (Round to one decimal place as needed.)

A. The median is 56.5. (Round to one decimal place as needed.)

B. There is no mode.

7. 0.188

0.165

A. The mode(s) is(are) 0.16 sec. (Use a comma to separate answers as needed.)

8. 58.6

57.5

A. The mode(s) is(are) 49 sec. (Use a comma to separate answers as needed.)

9. 0.81227

0.8158

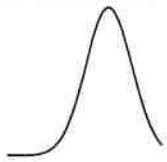
A. The outlier is 0.7903. (Type an integer or a decimal.)

A. The mean without the outlier is 0.81593. (Round to five decimal places as needed.)

A. The median without the outlier is 0.81605. (Type an integer or a decimal.)

10. C.

There would probably be one peak because there are no obvious reasons why the exam scores would form different groups.



D.

D.

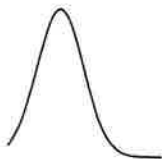
The distribution would probably be left-skewed because many of the students got an A, and very few got a C.

B.

The variation would probably be large because many students got an A, some got a B, and a small number got a C, and so the data are not clustered.

11. A.

There would probably be one peak because there are no obvious reasons why the values of the annual rainfall would form different groups.



B.

A.

The distribution would probably be right-skewed because there is a clear minimum amount of rainfall but no obvious maximum amount.

D.

The variation would probably be large because there would likely be several cities with very different amounts of rainfall.

12. B.
There would probably be one peak because there are no obvious reasons why the prices of dog food would form different groups.

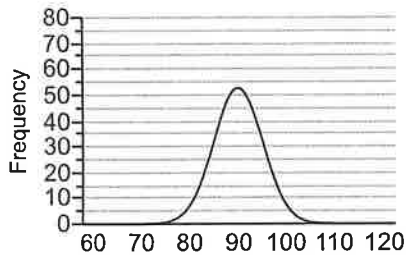


- C.
B.
The distribution would probably be symmetric because there are no obvious factors to indicate that there would be more high or low priced items.
C.
The variation would probably be moderate because there are no obvious reasons to expect an especially large or small amount of variation.
-

13. (1) right-skewed
(2) low
-

14. A. one peak, right-skewed, moderate variation
-

15.

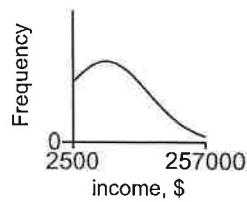


C. Weight (kilograms)

1

- (1) symmetric,
(2) fairly low
-

16.



D.

right-skewed

B. 150 families, because the median is the middle value in the sorted data set.

A. No, because the number of families that earned more than \$47,000 depends on the distribution.

17. 3.44

18. F.

At least one student drinks more than 19 sodas per week. An outlier value that is higher than most of the other values would explain why the mean is higher than the median.

19. B.

The driving time will have the highest variation at 7:00 on a Sunday morning. Sometimes the road might be clear, but other times there will be congestion because of holiday travel, among other reasons.

20. A.

The customers would have more complaints about the waiting time at the store that has more variation because some customers would have longer waits and might think they are being treated unequally.

21. A.

The statement makes sense because the range of data for the heights of a group of 5-year-old children is smaller than the range of data for the heights of a group of children who range in age from 3 to 15.

22. C. The heights of all elementary-school children, since it has the largest range.

23. A.

A normal distribution is a symmetric, bell-shaped distribution with a single peak. Its peak corresponds to the mean, median, and mode of the distribution. Its variation is characterized by the standard deviation of the distribution.

24. C. A standard score is the number of standard deviations a data value lies above or below the mean.

D. The standard score for a particular data value is given by $z = \frac{\text{data value} - \text{mean}}{\text{standard deviation}}$.

25. C.

This makes sense. The 50th percentile height means that 50% of all the heights in the data set are less than or equal to Jack's height. The median is the middle value, which means it splits the distribution in half. These two statements are the same.

26. (1) (a) and (b)

(2) (a)

27. 50.0

16.0

81.5

28. - 3

29. - 0.11

30. - 1.8

3.59

31. 99.7

32. 0.15

56

33. F.

About 16%, because 43 miles per gallon is 1 std. deviation above the mean. By the 68-95-99.7 rule, about 68% of the distribution lies within 1 std. deviation of the mean. So 32% lies outside of this range, 16% in each tail.

34. B.

Statistical inference is the process of making a conclusion about a population from results for a sample. It is important because the goal of most statistical studies is to learn something about an entire population.

35. A.

If a test is statistically significant, then the results are substantial enough to reject the original hypothesis about the population and accept an alternative hypothesis.
