End of Topic Test Form A

1a. The mode is the number or numbers that occur the most in a data set.

b. The median of a data set is the middle number of a data set when the data are arranged in order from least to greatest.

c. The mean of a data set is the arithmetic average of the data set. It also indicates the balance point on the graph of the data set.

2a. There are 2 modes: 25 minutes and 40 minutes.

b. The modes tell Justin that he spent 25 minutes or 40 minutes exercising on more days than he spent any other amounts of time.

c. There are 14 numbers in the data set. Arranged in increasing order, they are 0, 0, 20, 25, 25, 25, 30, 40, 40, 40, 45, 50, 60, 75. Because there is an even number of data values, the median is found by adding the two middle values and dividing by 2: $\frac{30 + 40}{2} = \frac{70}{2} = 35$, so the median time Justin spent exercising each day was 35 minutes.

d. The median tells Justin that half of the days he spent less than 35 minutes exercising and half of the days he spent more than 35 minutes exercising.

e. $\frac{0 + 0 + 20 + 25 + 25 + 25 + 30 + 40 + 40 + 40 + 45 + 50 + 60 + 75}{14}$
$= \frac{475}{14} \approx 33.9$

The mean is approximately 33.9 minutes. Rounded to the nearest minute, the mean is 34 minutes.

f. On average, Justin spent 34 minutes exercising each day.

g. The mean is less than the median. This indicates the data distribution is skewed left.

3a. The distribution of the data is symmetric.

b. Because the data distribution is symmetric, the most appropriate measure of center is the mean.

The mean = 5.

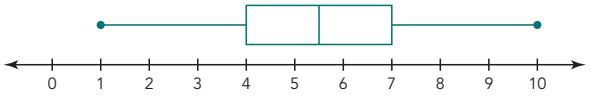
c. Because the data distribution is symmetric, the most appropriate measure of variation is the mean absolute deviation. The MAD = 1.45.

4a. The distribution of the data is skewed left.

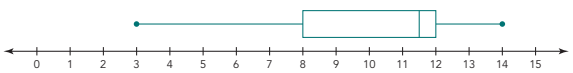
b. No. Because the data distribution is skewed left, the most appropriate measure of center is the median. The median = 97.

c. Yes. Because the data distribution is skewed left, the most appropriate measure of variation is the IQR. The IQR = 15.

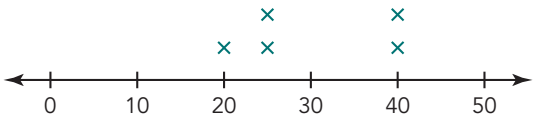
5.



6.

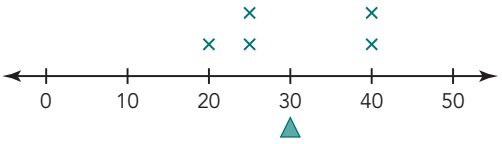


7a.



b. The median is 25.

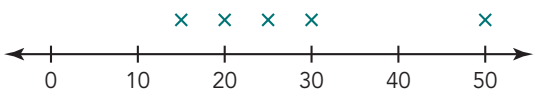
c.



The mean is 30.

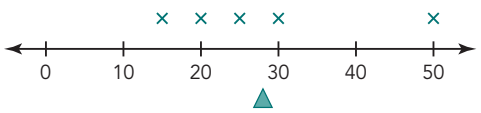
d. The modes are 25 and 40.

8a.



b. The median is 25.

c.



The mean is 28.

d. There is no mode for this data set.

9a. There are 19 players represented in the stem-and-leaf plot. I know this because there are 19 leaves in the plot.

b. The mode is 55 inches. This height appears three times. No other height appears more than twice.

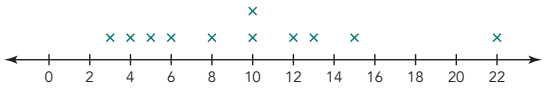
c. The median player height is 56 inches. I arranged the heights in order from least to greatest. The median is the height directly in the middle of these heights.

d. The distribution is symmetric.

e. Because the distribution is symmetric, the mean will be about the same as the median.

f. The mean height is 55.7 inches. I calculated it by adding all of the player heights and dividing by 19, the number of players.

10a.



b. Mode = 10 minutes

Median = 10 minutes

Mean:

$$\frac{3 + 4 + 5 + 6 + 8 + 10 + 10 + 12 + 13 + 15 + 22}{11}$$

$$= \frac{108}{11} \approx 9.8 \text{ minutes}$$

c. MAD = 4.2 minutes

IQR = 7 minutes

d. Either the MAD or the IQR could be used, but the IQR is a better choice because of the outlier at 22.

11a. There are 31 students represented in the histogram. The first bar represents 10 students, the second bar 7 students, the third bar 7 students, the fourth bar 5 students, the fifth bar 1 student, and the last bar 1 student.

$10 + 7 + 7 + 5 + 1 + 1 = 31$

b. The distribution is skewed right.

c. The mean would be greater than the median because the graph is skewed right. That means that the median is closer to the middle of most of the data, but the mean is greater than the median because of the greater values on the right side of histogram.

d. No. I cannot determine the measures of center because I do not know the data values. I only know how many values fall in each interval.

e. The median should be used because the distribution is skewed.

12a. There are 24 students represented in the stem-and-leaf plot. I know this because there are 24 leaves in the plot.

b. There are two modes, 76 and 88. Each of these scores appears twice. No other scores appear more than once.

c. The median test score is 80. I arranged the scores in order from least to greatest. Because there are 24 scores, the median will be halfway between the 12th and 13th scores, which are 79 and 81, so the median is 80.

d. The distribution is skewed left.

e. Because the distribution is skewed left, the mean will be less than the median.

f. To the nearest tenth, the mean test score is 77.8. I calculated it by adding all of the test scores and dividing by 24, the number of scores.

13a. Lopez family: The mean age is $\frac{16 + 8 + 11 + 5 + 5}{5} = \frac{45}{5} = 9$ years.

Holland family: The mean age is $\frac{9 + 10 + 7 + 10 + 4}{5} = \frac{40}{5} = 8$ years.

b.

Lopez family

Name	Age	Deviation from mean	Absolute Deviation from Mean
Rosa	16	7	7
Jose	8	−1	1
Lucia	11	2	2
Anget	5	−4	4
Cariotta	5	−4	4

Holland family

Name	Age	Deviation from mean	Absolute Deviation from Mean
Danielle	9	1	1
Eric	10	2	2
Alexis	7	−1	1
Joshua	10	2	2
Cody	4	−4	4

c. Lopez family: The mean absolute deviation is $\frac{7 + 1 + 2 + 4 + 4}{5} = \frac{18}{5} =$ 3.6 years.

Holland family: The mean absolute deviation is $\frac{1 + 2 + 1 + 2 + 4}{5} = \frac{10}{5} = 2$ years.

d. The smaller mean absolute deviation for the Holland family shows that their children's ages are closer together than the Lopez family.

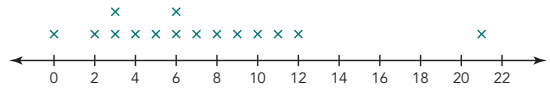
e. Lopez family: The median is 8 years. This is less than the mean of 9 years.

Holland family: The median is 9 years. This is greater than the mean of 8 years.

f. Lopez family: Because the median is less than the mean, the distribution for this family is skewed right.

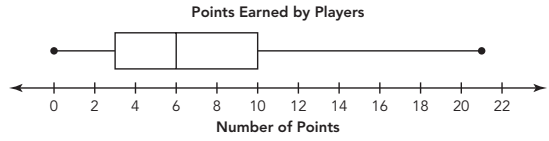
Holland family: Because the median is greater than the mean, the distribution for this family is skewed left.

14a.



b. Minimum = 0 Q1 = 3
Median = 6 Q3 = 10
Maximum = 21

c.



d. The right whisker is longer than the left whisker. This indicates that the data between Q3 and the maximum value are more spread out than the data between the minimum value and Q1.

e. The box and whisker plot show the data is skewed right. Therefore, the most appropriate measure of center to use is the median and the most appropriate measure of variation is the IQR.

15a. The mean for Week 1 is 86.9 minutes.

The mean for Week 2 is 90 minutes.

b. The range for Week 1 is 138 − 35 = 103 minutes.

The range for Week 2 is 160 − 25 = 135 minutes.

c. There are 7 days in a week, so for each week, the median will be the fourth largest (or fourth smallest) number.

For Week 1, the median is 90 minutes.

For Week 2, the median is 98 minutes.

d. For Week 1, Q1 = 57 and Q3 = 123. Twenty-five percent of the talk times in Week 1 are less than 57 minutes and 75% are greater than 57 minutes. Seventy-five percent of the talk times are less than 123 minutes and 25% are greater than 123 minutes.

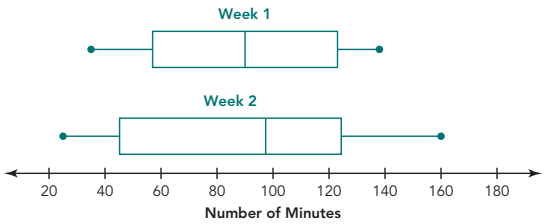
For Week 2, Q1 = 45 and Q3 = 124. Twenty-five percent of the talk times in Week 2 are less than 45 minutes and 75% are greater than 45 minutes. Seventy-five percent of the talk times are less than 124 minutes and 25% are greater than 124 minutes.

e. The IQR for Week 1 is 123 − 57 = 66 minutes.

The IQR for Week 2 is 124 − 45 = 79 minutes.

In Week 1, 50% of Allison's time on the phone was between 57 and 123 minutes per day. In Week 2, 50% of her time on the phone was between 45 and 124 minutes per day.

f.



g. Week 1. The data distribution for Week 1 is symmetric and the mean and median are almost equal. Therefore, you could use the mean absolute deviation as the measure of variation.