0-9 Measures of Center, Spread, and Position

Find the mean, median, and mode for each set of data.

1. number of pages in each novel assigned for summer reading:
224, 272, 374, 478, 960, 394, 404, 308, 480, 624

SOLUTION:

To find the mean divide the sum of all the pages divided by the total of novels.

mean
$= \frac{224+272+374+478+960+394+404+308+480+624}{10}$
$= \frac{4518}{10}$
= 451.8 pages

To find the median. First arrange the page numbers in order.
224, 272, 308, 374, 394, 404, 478, 480, 624, 960.
Because there is an even number of pages, find the mean of the middle two.

median
$= \frac{394+404}{2}$
$= \frac{792}{2}$
=399 pages

The mode is the value that occurs the most often in the data set. There is not a repetitive number of pages, so there is no mode.

2. height in centimeters of bean plants at the end of an experiment:
14.5, 12, 16, 11, 14, 11, 10.5, 14, 11.5, 15, 13.5

SOLUTION:

To find the mean divide the sum of all the heights divided by the total of plants.

mean
$= \frac{14.5+12+16+11+14+11+10.5+14+11.5+15+13.5}{11}$
$= \frac{143}{11}$
= 13 cm

To find the median, arrange the heights in order.
10.5, 11, 11, 11.5, 12, 13.5, 14, 14, 14.5, 15, 16.

Because there is an odd number of plants, the median is the middle number. So, the median is 13.5 cm
The mode is the value that occurs the most often in the data set. In this case, it is 11 cm and 14 cm.

3. number of text messages sent each day during the last two weeks:
18, 35, 53, 44, 26, 57, 23, 27, 47, 33, 4, 35, 39, 41

SOLUTION:

To find the mean, divide the sum of number of all text messages divided by the 14 days.

mean
$= \frac{18+35+53+44+26+57+23+27+47+33+4+35+385+41}{14}$
$= \frac{482}{14}$
$\approx$ 34.4 text messages

To find the median, arrange the values in order.
4, 18, 23, 26, 27, 33, 35, 35, 38, 47, 41, 44, 53, 57.

Because there is an even number of values, find the mean of the middle two values.

median
$= \frac{35+35}{2}$
$= \frac{70}{2}$
= 35 text messages

The mode is the value that occurs the most often in the data set. In this case, it is 35 text messages.

4. **State whether the data in sets A and B represent *sample* or *population* data. Then find the range, variance, and standard deviation of each set. Use the standard deviations to compare the variability between the data sets.**

Wait Times (min)					
Ride A			Ride B		
45	22	40	35	50	32
48	11	51	31	35	45
36	55	60	45	49	40
32	24	37	43	37	45

SOLUTION:

12 wait times for Ride A and Ride B are given. For an amusement park ride, there must be more than 12 rides a day, so the data must be a sample.

The range is the difference between the greatest and least values in the set.
Ride A: $60 - 11 = 49$ min
Ride B: $50 - 31 = 19$ min

In order to find the variance and standard deviation, the mean is needed.

mean Ride A
$= \frac{45+22+40+48+11+51+36+55+60+32+24+37}{12}$
$= \frac{461}{12}$
$\approx$ 38.4 min

x	$x - \bar{x}$	$(x - \bar{x})^2$
45	45 – 38.4 = 6.6	$6.6^2 = 43.56$
22	22 – 38.4 = –16.4	$(-16.4)^2 = 268.96$
40	40 – 38.4 = 1.6	$1.6^2 = 2.56$
48	48 – 38.4 = 9.6	$9.6^2 = 92.16$
11	11 – 38.4 = –27.4	$(-27.4)^2 = 750.76$
51	51 – 38.4 = 12.6	$12.6^2 = 158.76$
36	36 – 38.4 = –2.4	$(-2.4)^2 = 5.76$
55	55 – 38.4 = 16.6	$16.6^2 = 275.56$
60	60 – 38.4 = 21.6	$21.6^2 = 466.56$
32	32 – 38.4 = –6.4	$(-6.4)^2 = 40.96$
24	24 – 38.4 = –14.4	$(-14.4)^2 = 207.36$
37	37 – 38.4 = –1.4	$(-1.4)^2 = 1.96$
		Sum = 2314.92

The variance of Ride A is $\frac{2314.92}{11}$ = 210.45 min

The standard deviation of Ride A is $\sqrt{210.45}$ = 14.5 min

mean Ride B
$= \frac{35+50+32+31+35+45+45+49+40+43+37+45}{12}$
$= \frac{487}{12}$
$\approx$ 40. 6 min

x	$x - \bar{x}$	$(x - \bar{x})^2$
35	35 – 40.6 = –5.6	$(-5.6)^2 = 31.36$
50	50 – 40.6 = 9.4	$9.4^2 = 88.36$
32	32 – 40.6 = –8.6	$(-8.6)^2 = 73.96$
31	31 – 40.6 = –9.6	$(-9.6)^2 = 92.16$
35	35 – 40.6 = –5.6	$(-5.6)^2 = 31.36$
45	45 – 40.6 = 4.4	$4.4^2 = 19.36$
45	45 – 40.6 = 4.4	$4.4^2 = 19.36$
49	49 – 40.6 = 8.4	$8.4^2 = 70.56$
40	40 – 40.6 = –0.6	$(-0.6)^2 = 0.36$
43	43 – 40.6 = 2.4	$2.4^2 = 5.76$
37	37 – 40.6 = –3.6	$(-3.6)^2 = 12.96$
45	45 – 40.6 = 4.4	$4.4^2 = 19.36$
		Sum = 464.92

The variance of Ride B is $\frac{464.92}{11}$ = 42.3 min

The standard deviation of Ride B is $\sqrt{42.3}$ = 6.5 min

Since the sample standard deviation of Ride A is greater than that of Ride B, there is more variability in the sample wait times for Ride A than Ride B.

5.

Number of Sponsors Obtained by Participants					
Charity Walk A			Charity Walk B		
44	14	61	8	28	15
22	27	25	100	42	19
38	50	49	25	75	82

SOLUTION:

9 sponsors for Walk A and Walk B are given. A

charity walk will have more than 9 walkers; the data given is for a sample.

Walk A: 47, ≈242.0, ≈15.6;
Walk B: 92, ≈1115.4, ≈33.4;

The range is the difference between the greatest and least values in the set.
Walk A: 61 – 14 = 47 sponsors
Walk B: 100 – 8 = 92 sponsors

In order to find the variance and standard deviation, the mean is needed.
mean Walk A
$= \frac{44+14+61+22+27+25+38+50+49}{9}$
$= \frac{330}{9}$
≈ 36.7 sponsors

x	$x-\bar{x}$	$(x-\bar{x})^2$
44	44 – 36.7 = 7.3	7.3^2 = 53.29
14	14 – 36.7 = –22.7	$(-22.7)^2$ = 515.29
61	61 – 36.7 = 24.3	24.3^2 = 590.49
22	22 – 36.7 = –14.7	$(-14.7)^2$ = 216.09
27	27 – 36.7 = –9.7	$(-9.7)^2$ = 94.09
25	25 – 36.7 = –11.7	$(-11.7)^2$ = 136.89
38	38 – 36.7 = 1.3	1.3^2 = 1.69
50	50 – 36.7 = 13.3	13.3^2 = 176.89
49	49 – 36.7 = 12.3	12.3^2 = 151.29
		Sum = 1936.01

The variance of Walk A is $\frac{1936.01}{8}$ = 242.0 sponsors

The standard deviation of Walk A is $\sqrt{242.0}$ = 15.6 sponsors

mean Walk B
$= \frac{8+28+15+100+42+19+25+75+82}{9}$
$= \frac{394}{9}$
≈ 43.8 sponsors

x	$x-\bar{x}$	$(x-\bar{x})^2$
8	35 – 43.8 = –35.8	$(-35.8)^2$ = 1281.64
28	50 – 43.8 = –15.8	$(-15.8)^2$ = 249.64
15	32 – 43.8 = –28.8	$(-28.8)^2$ = 829.44
100	31 – 43.8 = 56.2	56.2^2 = 3158.44
42	35 – 43.8 = –1.8	$(-1.8)^2$ = 3.24
19	45 – 43.8 = –24.8	$(-24.8)^2$ = 615.04
25	45 – 43.8 = –18.8	$(-18.8)^2$ = 353.44
75	49 – 43.8 = 31.2	31.2^2 = 973.44
82	45 – 43.8 = 38.2	38.2^2 = 1459.24
		Sum = 8923.56

The variance of Walk B is $\frac{8923.56}{8}$ = 1115.4 sponsors

The standard deviation of Walk B is $\sqrt{1115.4}$ = 33.4 sponsors

Since the sample standard deviation of Walk B is greater than that of Walk A, there is more variability in the number of sponsors obtained by participants in Walk B than in Walk A.

6.

Number of Days Each Student Missed This Year							
Class A							
10	8	5	9	7	3	6	8
5	13	0	15	9	7	9	10
14	11	8	4	7	8	2	
9	11	14	8	12	10	1	
Class B							
5	8	13	7	9	4	10	2
12	6	7	8	11	12	8	9
12	9	6	11	3	8	5	
3	10	5	13	9	1	8	

SOLUTION:
The data given is the of days absent for each student in Class A and Class B. Since there is data for each student, we are given data for the population.

Class A: 15, ≈13.7, 3.7;
Class B: 12, ≈10.5 , ≈3.2;

The range is the difference between the greatest and least values in the set.
Class A: 15 – 0 = 15 days
Class B: 13 – 1 = 12 days

In order to find the variance and standard deviation, the mean is needed.

mean Class A
$= \frac{10+8+5+9+7+3+6+8+14+11+8+4+7+8+2+5+13+0+15+9+7+9+10+9+11+14+8+12+10+1}{30}$

$= \frac{243}{30}$
$\approx$ 8.1 days

x	$x - \bar{x}$	$(x - \bar{x})^2$
10	10 – 8.1 = 1.9	$(1.9)^2 = 3.61$
8	8 – 8.1 = –0.1	$(–0.1)^2 = 0.01$
5	5 – 8.1 = –3.1	$(–3.1)^2 = 9.61$
9	9 – 8.1 = 0.9	$(0.9)^2 = 0.81$
7	7 – 8.1 = –1.1	$(–1.1)^2 = 1.21$
3	3 – 8.1 = –5.1	$(–5.1)^2 = 26.01$
6	6 – 8.1 = –2.1	$(–2.1)^2 = 4.41$
8	8 – 8.1 = –0.1	$(–0.1)^2 = 0.01$
14	14 – 8.1 = 5.9	$(5.9)^2 = 34.81$
11	11 – 8.1 = 2.9	$(2.9)^2 = 8.41$
8	8 – 8.1 = –0.1	$(–0.1)^2 = 0.01$
4	4 – 8.1 = –4.1	$(–4.1)^2 = 16.81$
7	7 – 8.1 = –1.1	$(–1.1)^2 = 1.21$
8	8 – 8.1 = –0.1	$(–0.1)^2 = 0.01$
2	2 – 8.1 = –6.1	$(–6.1)^2 = 37.21$
5	5 – 8.1 = –3.1	$(–3.1)^2 = 9.61$
13	13 – 8.1 = 4.9	$(4.9)^2 = 24.01$
0	0 – 8.1 = –8.1	$(–8.1)^2 = 65.61$
15	15 – 8.1 = 6.9	$(6.9)^2 = 47.61$
9	9 – 8.1 = 0.9	$(0.9)^2 = 0.81$
7	7 – 8.1 = –1.1	$(–1.1)^2 = 1.21$
9	9 – 8.1 = 0.9	$(0.9)^2 = 0.81$
10	10 – 8.1 = 1.9	$(1.9)^2 = 3.61$
9	9 – 8.1 = 0.9	$(0.9)^2 = 0.81$
11	11 – 8.1 = 2.9	$(2.9)^2 = 8.41$
14	14 – 8.1 = 5.9	$(5.9)^2 = 34.81$
8	8 – 8.1 = –0.1	$(–0.1)^2 = 0.01$
12	12 – 8.1 = 3.9	$(3.9)^2 = 15.21$
10	10 – 8.1 = 1.9	$(1.9)^2 = 3.61$
1	1 – 8.1 = –7.1	$(–7.1)^2 = 50.41$
		Sum = 410.7

The variance of Class A is $\frac{410.7}{29}$ = 14.2 days

The standard deviation of Class A is $\sqrt{14.2}$ = 3.8 days

mean Class B
$= \frac{5+8+13+7+9+4+10+2+12+9+6+11+3+8+5+12+6+7+8+11+12+8+9+3+10+5+13+9+1+8}{30}$
$= \frac{234}{30}$
$\approx$ 7.8 days

x	$x - \bar{x}$	$(x - \bar{x})^2$
5	5 – 7.8 = –2.8	$(–2.8)^2 = 7.84$
8	8 – 7.8 = 0.2	$(0.2)^2 = 0.04$
13	13 – 7.8 = 5.2	$(5.2)^2 = 27.04$
7	7 – 7.8 = –0.8	$(–0.8)^2 = 0.64$
9	9 – 7.8 = 1.2	$(1.2)^2 = 1.44$
4	4 – 7.8 = –3.8	$(–3.8)^2 = 14.44$
10	10 – 7.8 = 2.2	$(2.2)^2 = 4.84$
2	2 – 7.8 = –5.8	$(–5.8)^2 = 33.64$
12	12 – 7.8 = 4.2	$(4.2)^2 = 17.64$
9	9 – 7.8 = 1.2	$(1.2)^2 = 1.44$
6	6 – 7.8 = –1.8	$(–1.8)^2 = 3.24$
11	11 – 7.8 = 3.2	$(3.2)^2 = 10.24$
3	3 – 7.8 = –4.8	$(–4.8)^2 = 23.04$
8	8 – 7.8 = 0.2	$(0.2)^2 = 0.04$
5	5 – 7.8 = –2.8	$(–2.8)^2 = 7.84$
12	12 – 7.8 = 4.2	$(4.2)^2 = 17.64$
6	6 – 7.8 = –1.8	$(–1.8)^2 = 3.24$
7	7 – 7.8 = –0.8	$(–0.8)^2 = 0.64$
8	8 – 7.8 = 0.2	$(0.2)^2 = 0.04$
11	11 – 7.8 = 3.2	$(3.2)^2 = 10.24$
12	12 – 7.8 = 4.2	$(4.2)^2 = 17.64$
8	8 – 7.8 = 0.2	$(0.2)^2 = 0.04$
9	9 – 7.8 = 1.2	$(1.2)^2 = 1.44$
3	3 – 7.8 = –4.8	$(–4.8)^2 = 23.04$
10	10 – 7.8 = 2.2	$(2.2)^2 = 4.84$
5	5 – 7.8 = –2.8	$(–2.8)^2 = 7.84$
13	13 – 7.8 = 5.2	$(5.2)^2 = 27.04$
9	9 – 7.8 = 1.2	$(1.2)^2 = 1.44$
1	1 – 7.8 = –6.8	$(–6.8)^2 = 46.24$
8	8 – 7.8 = 0.2	$(0.2)^2 = 0.04$
		Sum = 314.8

The variance of Class B is $\frac{314.8}{29}$ = 10.9 days

The standard deviation of Class B is $\sqrt{10.9}$ = 3.3 days

Since the sample standard deviation of Class A is greater than that of Class B, there is more variability in the number of days that students missed during the school year for Class A than for Class B.

Find the minimum, lower quartile, median, upper quartile, and maximum of each data set. Then interpret this five-number summary.

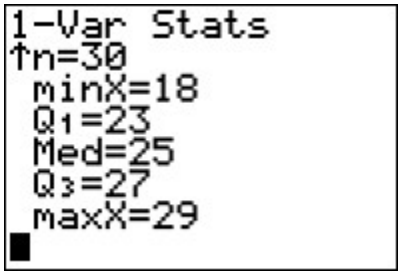
7.

Number of Students In Each Math Class at Central High							
25	27	26	26	19	27	24	23
24	26	18	28	29	29	26	24
19	28	25	24	20	22	22	
24	23	23	25	25	29	28	

SOLUTION:

Enter the data into **L1**.

Press STAT ▸ ENTER to display the 1-Var statistics.



The minimum is 18.
The lower quartile is 23.
The median is 25.
The upper quartile is 27.
The maximum is 29.

There are 18 students in the smallest math class at Central High and 29 students in the largest class. 25% of the classes have less than 23 students, 50% of the classes have less than 25 students, and 75% of the classes have less than 27 students.

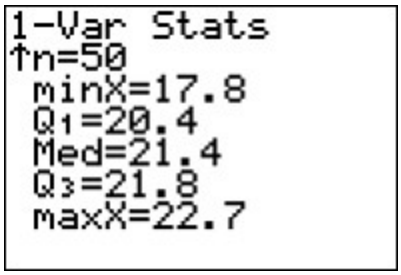
8.

State Mean ACT Scores				
20.2	21.3	21.5	20.4	21.6
20.0	21.7	21.3	20.2	21.6
20.8	22.4	21.4	22.2	18.8
20.1	22.3	20.3	21.2	21.4
21.5	20.5	20.3	21.5	22.7
20.3	22.5	21.5	17.8	20.5
22.0	21.6	20.3	19.8	22.6
21.5	21.7	21.2	22.5	21.2
20.6	22.5	21.8	21.9	19.3
20.9	22.5	22.2	21.4	20.7

SOLUTION:

Enter the data into **L1**.

Press STAT ▸ ENTER to display the 1-Var statistics.



The minimum is 17.8.
The lower quartile is 20.4.
The median is 21.4.
The upper quartile is 21.8.
The maximum is 22.7.

The lowest mean score for a state is 17.8 and the highest mean score is 22.7. 25% of the states have a mean score that is less than 20.4, 50% of the states have a mean score that is less than 21.4, and 75% of the states have a mean score that is less than 22.7.

Identify any outliers in each data set, and explain your reasoning. Then find the mean, median, mode, range, and standard deviation of the data set with and without the outlier. Describe the effect on each measure.

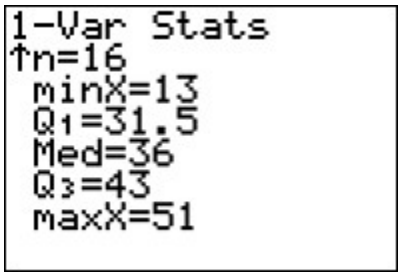
9. fuel efficiency in miles per gallon of 15 randomly selected automobiles:
40, 36, 29, 45, 51, 36, 48, 34, 36, 22, 13, 42, 31, 44, 32, 34

SOLUTION:

Enter the data into **L1**.

Keystrokes: STAT ▸ ENTER

Use the 1-Var statistics to identify Q_1 and Q_3.



$Q_1 = 31.5$ and $Q_3 = 43$. IQR = $Q_3 - Q_1$ or $43 - 31.5 = 11.5$.

Find and use the IQR to find the values beyond which any outlier would lie.

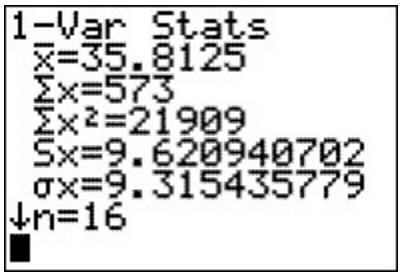
$Q_1 - 1.5(\text{IQR})$ and $Q_3 + 1.5(\text{IQR})$

$31.5 - 1.5(11.5)$ $\quad$ $43 + 1.5(11.5)$

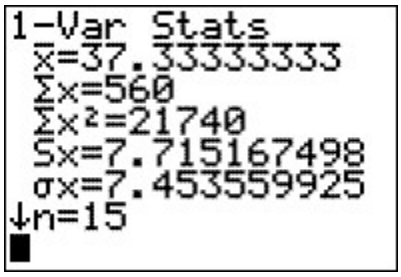
14.25 $\quad$ 60.25

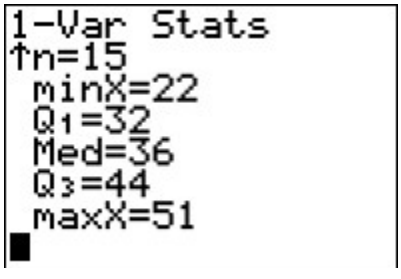
The only outlier on the low end is 13. There are no outliers on the upper end.

With outlier:



Without outlier:





Data Set	Mean	Median	Mode	Range	Standard Deviation
with outlier	≈35.8	36	36	38	≈9.3
without outlier	≈37.3	36	36	29	≈7.5

Removing the outlier did not affect the median or mode. However, the removal did affect the mean, standard deviation, and range. The mean and standard deviation increased, and the range decreased.

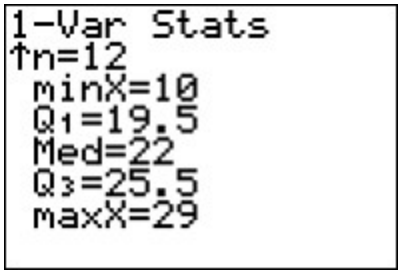
10. number of posts to a certain blog each month during a particular year:
25, 23, 21, 27, 29, 19, 10, 21, 20, 18, 26, 23

SOLUTION:

Enter the data into **L1**.

Keystrokes: STAT ▸ ENTER

Use the 1-Var statistics to identify Q_1 and Q_3.



$Q_1 = 19.5$ and $Q_3 = 25.5$. IQR = $Q_3 - Q_1$ or $25.5 - 19.5 = 6$.

Find and use the IQR to find the values beyond which any outlier would lie.

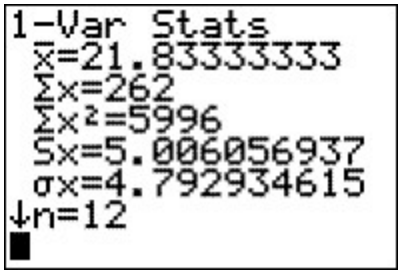
$Q_1 - 1.5(\text{IQR})$ and $Q_3 + 1.5(\text{IQR})$

$19.5 - 1.5(6)$ $\quad$ $25.5 + 1.5(6)$

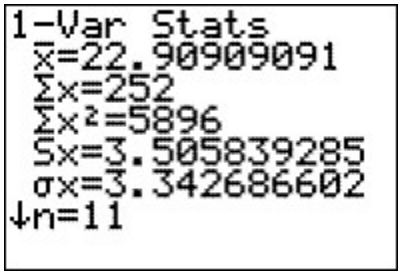
10.5 $\quad$ 34.5

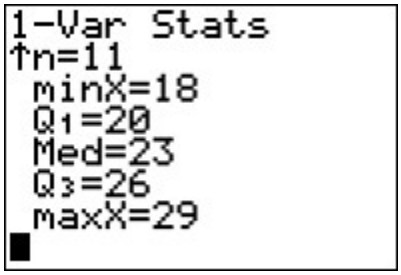
The interval beyond which any outliers would lie is $10.5 < x < 34.5$. Since $10 < 10.5$, it is an outlier. There are no outliers on the upper end.

With outlier:



Without outlier:





Data Set	Mean	Median	Mode	Range	Standard Deviation
with outlier	≈21.8	22	21	19	≈4.8
without outlier	≈22.9	23	21	11	≈3.3

Removing the outlier did not affect the mode. However, the removal did affect the mean, median, standard deviation, and range. The mean and median increased, and the standard deviation and range decreased.

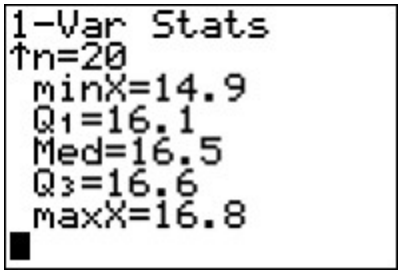
11. **CEREAL** The weights, in ounces, of 20 randomly selected boxes of a certain brand of cereal are shown.
16.7, 16.8, 15.9, 16.1, 16.5, 16.6, 16.5, 15.9, 16.7, 16.5,
16.6, 14.9, 16.5, 16.1, 15.8, 16.7, 16.2, 16.5, 16.4, 16.6
a. Identify any outliers in the data set, and explain your reasoning.
b. If the outlier was removed and an additional cereal box that was 17.35 ounces was added, would this value be an outlier of the new data set? Explain.
c. What are some possible causes of outliers in this situation?

SOLUTION:

a. Enter the data into **L1**.

Keystrokes: STAT ▸ ENTER

Use the 1-Var statistics to identify Q_1 and Q_3.



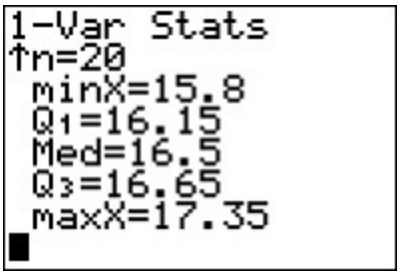
$Q_1 = 16.1$ and $Q_3 = 16.6$. IQR = $Q_3 - Q_1$ or $16.6 - 16.1 = 0.5$.
Find and use the IQR to find the values beyond which any outlier would lie.

$Q_1 - 1.5(\text{IQR})$ and $Q_3 + 1.5(\text{IQR})$

$16.1 - 1.5(0.5)$ $\quad$ $16.6 + 1.5(0.5)$
15.35 $\quad$ 17.35

The interval beyond which any outliers would lie is $15.35 < x < 17.35$. Since $14.9 < 15.35$, it is an outlier.

b.



$Q_1 = 16.15$ and $Q_3 = 16.65$. IQR = $Q_3 - Q_1$ or $16.65 - 16.15 = 0.5$.
Find and use the IQR to find the values beyond which any outlier would lie.

$Q_1 - 1.5(\text{IQR})$ and $Q_3 + 1.5(\text{IQR})$

$16.15 - 1.5(0.5)$ $\quad$ $16.65 + 1.5(0.5)$
15.4 $\quad$ 17.4

The new interval beyond which any outliers would lie would be $15.4 < x < 17.4$. Since $17.4 > 17.35$, it would not be an outlier.
c. Outliers can be caused by data recording errors or manufacturing errors.