You can find the mean, median, mode, and range to describe a set of data.

<table>
<thead>
<tr>
<th>Terry’s Test Scores</th>
<th>76</th>
<th>81</th>
<th>94</th>
<th>81</th>
<th>78</th>
</tr>
</thead>
</table>

The mean or average is the sum of the items divided by the number of items.

\[
76 + 81 + 94 + 81 + 78 = 410 \\
410 \div 5 = 82
\]

First, find the sum of the values. Then divide the sum by the number of values in the set of data.

The mean is 82 points.

The median is the middle value of an ordered set of data. If there are two middle values, the median is the mean of those two values.

76, 78, 81, 81, 94

Put the values in order first.

The median is 81 points.

The mode is the value that occurs most often in a set of data.

The mode is 81 points.

The range is the difference between the greatest and least values in the set of data.

\[
94 - 76 = 18
\]

Use subtraction to find the range.

The range is 18 points.

Find the mean, median, mode, and range of each set of values.

1. 23, 78, 45, 22
   - mean: ____________
   - median: ____________
   - mode: ____________
   - range: ____________

2. 102, 79, 82, 103, 79
   - mean: ____________
   - median: ____________
   - mode: ____________
   - range: ____________

3. 56, 99, 112, 112, 56
   - mean: ____________
   - median: ____________
   - mode: ____________
   - range: ____________
Data Collection and Analysis

Review for Mastery: Additional Data and Outliers

An outlier is a value in a set of data that is much greater or much less than the other values.

Number of Minutes Spent on Homework

<table>
<thead>
<tr>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thurs</th>
<th>Fri</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>42</td>
<td>45</td>
<td>46</td>
<td>10</td>
</tr>
</tbody>
</table>

The outlier is 10 minutes because it is much less than the other values in the set.

An outlier may affect the mean, median, or mode.

Data without Friday’s value: mean = 45  median = 45.5  no mode
Data with Friday’s value: mean = 38  median = 45  no mode

When Friday’s value is included, the mean decreases by 7 minutes, the median decreases by 0.5 minutes, and the mode stays the same. The mean is most affected by the outlier because it is less than every value except for the outlier itself.

Find the mean, median, and mode for the set of data with and without the outlier.

1. 22, 25, 48, 26, 21, 27, 26, 29

With outlier: ____________________________________________

Without outlier: ____________________________________________

When an outlier affects the mean, median, or mode, choose a value that best describes the data.

In the example above, the median best describes the data because 45 minutes is closer to most of the data values in the set.

Find the mean, median, and mode. Then decide which best describes the set of data.

2. 16, 12, 14, 17, 81, 18, 13, 19, 14, 19

__________________________________________________

__________________________________________________
6.3 Mastery: Measures of Variation

A box-and-whisker plot is a visual way to show the spread of data. It divides the data into fourths using quartiles.

This is the box-and-whisker plot for the data 0, 1, 3, 5, 7, 8, and 11.

1/4 of the data are between 0 and 1. 1/4 of the data are between 1 and 5.
1/4 of the data are between 5 and 8. 1/4 of the data are between 8 and 11.

The plot shows the least value (0), the greatest value (11), the median (5), the first quartile (1), and the third quartile (8). It does not show every data point from the set. The quartiles are the medians of each half of the data.

The interquartile range (IQR) describes the range of the middle 50% of the data. To find the IQR, subtract the first quartile from the third quartile. Use the values of the right and left edges of the box.

8 – 1 = 7
The interquartile range is 7.

The mean absolute deviation measures the mean distance of the data from the mean.

Use the data to draw a box-and-whisker plot. Then find the IQR and the absolute mean deviation for the data.

1. Theo measured 8 samples for an experiment. The weights of the samples (in grams) are 3, 2, 4, 1, 4, 5, 1, and 4.

   Box-and-Whisker Plot:

   IQR: ______________
   Absolute Mean Deviation: ____________________________
Julie picked the following cards from a deck.

You can make a frequency table to organize the data. Make a row for the cards. Then for each card, make a tally mark in the appropriate column.

<table>
<thead>
<tr>
<th>Julie’s Cards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Card</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
</tbody>
</table>

1. Make a frequency table to organize the data.

<table>
<thead>
<tr>
<th>Rolls of a Number Cube</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 3 6 5 1 4 1 3 5 1 6 1 4</td>
</tr>
</tbody>
</table>

A line plot gives a visual picture of data. To make a line plot of Julie’s data, draw a number line. Then use an X to represent each tally mark in the frequency table.

2. Make a line plot for the frequency table you made in Exercise 1.
Sometimes when organizing data, you may want to make a frequency table with intervals, or you may want to make a histogram.

A frequency table can organize the data with intervals.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–10</td>
<td>0</td>
</tr>
<tr>
<td>11–20</td>
<td>4</td>
</tr>
<tr>
<td>21–30</td>
<td>3</td>
</tr>
<tr>
<td>31–40</td>
<td>7</td>
</tr>
<tr>
<td>41–50</td>
<td>4</td>
</tr>
<tr>
<td>51–60</td>
<td>2</td>
</tr>
</tbody>
</table>

A histogram is a bar graph that shows the number of values that occur within each interval.

You make a histogram the same way you make any other bar graph, except that the bars touch. They do not overlap.

Here is a histogram for the frequency table above.

3. Use the data to make a histogram.

<table>
<thead>
<tr>
<th>Total Books Read by Participants in Summer Reading Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 3 8 7 6</td>
</tr>
<tr>
<td>2 9 10 1 2</td>
</tr>
<tr>
<td>4 5 7 3 5</td>
</tr>
<tr>
<td>3 1 0 10 4</td>
</tr>
<tr>
<td>3 5 8 2 1</td>
</tr>
<tr>
<td>1 7 0 4 11</td>
</tr>
</tbody>
</table>
A dot plot is a visual way to show the spread of data. A number line is used to show every data point in a set. When the data are symmetric about the center, and the median has the greatest number of data, then the shape is described as a normal distribution. Recall that symmetric means that the two halves are mirror images. In a data set with normal distribution, the mean, median, and mode are equal.

This dot plot shows a normal distribution.
• The data are symmetric about the center, 5.
• The median has the greatest number of data.
• The mean, median, and mode are all 5.

Data sets do not always have normal distribution. The data may cluster more to the left or right of center. This is called a skewed distribution. The measures of center for a skewed data set with skewed distribution are not all equal.

This dot plot shows a skewed distribution.
• The data are not symmetric.
• The mean, median, and mode vary.
• The data are skewed to the left.

Describe the shape of the data distribution for the dot plot.

1.