**Measures of Central Tendency**

The mean, median and mode, as  **measures of central tendency,** provide us with a point of comparison. As an example, consider Company ABC where the average (mean) salary is $55,000/year. An employee earning $38,000/year might feel unjustly treated or at the very least the employee might explore the reasons for the substantial difference. If in the process the employee learns that the median salary at his workplace  is $26,000/year the employee would learn that relative to everyone else this employee’s  salary is in the upper half of the employee group.

To provide additional comparison the employee could consider other measures of position or location. Two such measures are percentiles and quartiles.

**Percentiles**

Percentile:

- A percentile is a value below which a certain percentage of observations lie.

- 95 percentile means that the person has got better marks then 95% of the entire students

Determining Percentiles

To determine the kth percentile that is represented by a particular data item ***x*,** the following formula can be used.



Step 1: If necessary order the data values from smallest to largest.

Step 2: Determine the total number of data values, n. This will be the denominator in the formula.

Step 3: Count the number of data values that are less than the value ***x.*** This will be the value in the numerator of the formula.

Step 4: Calculate the percentile, *k*, that is associated with a score of  ***x*** using the formula.

EXAMPLE 1

A class set of exam scores for 48 students are ranked from lowest to highest. Determine the percentiles associated with the scores of  a) 39%  b) 60% c) 94%.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 39 | 54 | 59 | 65 | 75 | 79 | 84 | 92 |
| 42 | 54 | 60 | 67 | 76 | 80 | 86 | 92 |
| 43 | 55 | 60 | 69 | 76 | 80 | 88 | 94 |
| 48 | 57 | 60 | 69 | 77 | 82 | 88 | 95 |
| 51 | 57 | 63 | 72 | 77 | 83 | 89 | 96 |
| 51 | 59 | 65 | 72 | 78 | 83 | 91 | 97 |

**Solution**

a) For a score of 39%:

Step 1: The data values are already ordered from smallest to largest.

Step 2: Determine the number of data values. Since there are 48 students n = 48.

Step 3: We count  0  data values that are less than 39

Step 4: Calculate the percentile, *k*, that is associated with a score of  x using the formula

image

k = (0/48)\*100% = 0%.

This means that the student who scored 39% is in the 0 percentile. A score of 39% is not higher than any other score.

b) For a score of 60%:

There are 13 scores lower than 60%  so    k = (13/48)\*100% = 27%. A score of  60% is in the 27th percentile which means that 27% (or just over one-fourth)  of the test scores are less than 60%.

c) For a score of 94%:

There are 44 scores less than 94%  so     k = (44/48)\*100% = 92%. A score of  94% is in the 92nd percentile which means that 92%  of the test scores are less than 94%.

**Quartiles**

Quartiles divide ordered data into quarters. Quartiles are special percentiles. The first quartile, *Q*1, is the same as the 25th percentile, and the third quartile, *Q*3, is the same as the 75th percentile. The median is a number that separates ordered data into halves. Half the values are the same as or smaller than the median, and half the values are the same as or larger than the median. The median can be called both the second quartile Q2 and the 50th percentile.

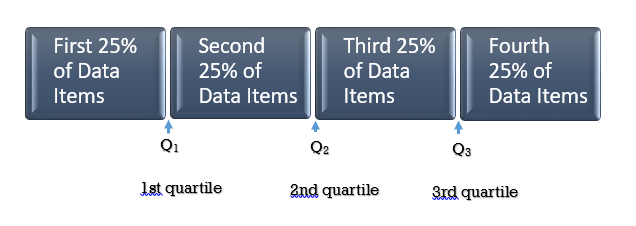
Quartiles

Quartiles divide the data set into **four** equal parts.

The first quartile, *Q*1, is the same as the 25th percentile, and the third quartile, *Q*3, is the same as the 75th percentile. The median can be called both the second quartile, Q2, and the 50th percentile.

As with the median, the quartiles may or may not be part of the data set.

As indicated in [Figure 1](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure1) each quartile divides a data set into four equal parts so that one-fourth of the data set is located in each part.

Fig. 1

**Determining Quartiles**

We will consider two methods for determining quartiles. As with percentiles, the data values must first be ordered from smallest to largest. The first method involves dividing the data set into four equal parts. The second method involves the use of formulas.

Determining Quartiles: Method 1

Step 1: Order the data from smallest to largest.

Step 2: Determine the number of data values **n**.

Step 3: Determine the median (Q2) of the data set. This will divide the data set into two equal parts.

Step 4: Determine Q1. This will divide the first half of the data set into two equal parts.

Step 5: Determine Q3. This will divide the second half of the data set into two equal parts.

**Note:** The median and the quartiles may not be actual observations from the data set.

**Method 1**

Consider the following data set:

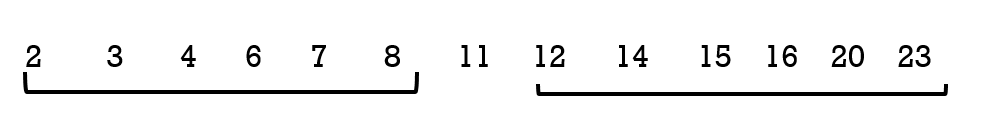
15       4       20       8      3     12      14      11      7     2     6     23     16

Step 1: To determine the  quartiles, order the data values from smallest to largest:

2      3       4     6     7       8      11     12      14      15    16     20     23

Step 2: The number of data values is 13.

Step 3: Determine the median, which measures the “centre” of the data. It is the number that separates ordered data into halves. Half the observations are the same number or smaller than the median, and half the observations are the same number or larger.



Since there are 13 observations, the median will be in the seventhh position. The median, and therefore the 2nd quartile Q2 , is eleven. The median is often referred to as  the “middle observation,” but it is important to note that it does not actually have to be one of the observed values.

Step 4: The first quartile, *Q*1, is the **middle value of the lower half** of the data.

To determine the**first quartile**, Q1, consider the lower half of the data observations:

2      3       4      6      7       8

Since there are six observations, the middle observation will be the average of the third and fourth data values  or  (4 + 6)/2 = 5  therefore   Q1  is 5

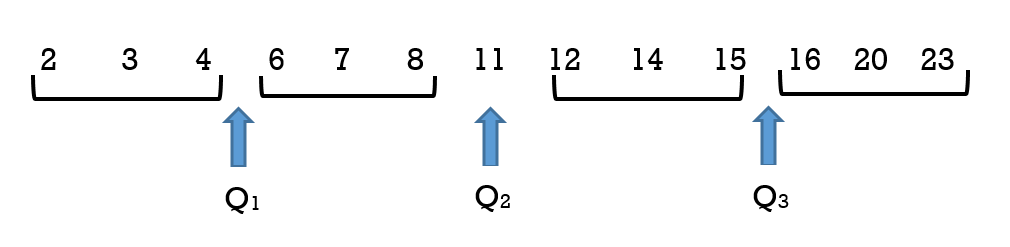
Step 5: The third quartile, *Q*3, is the **middle value of the upper half** of the data.

To determine the**third quartile**, Q3, consider the upper half of the data observations:

12      14      15     16    20    23

Since there are six observations, the middle observation will be the average of 15 and 16 , or 15.5 therefore    Q3  is 15.5.

[Figure 2](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure2) illustrates the three quartiles, which divide the data set into four equal parts.

Fig. 2

The number 4.5 is the first quartile, Q1. One-fourth of the entire set of observations lie below 4.5 and  three-fourths of the data observations lie above 4.5.

The third quartile, *Q*3, is 15.5. Three-fourths (75%) of the ordered data set lie below 15.5. One-fourth (25%) of the ordered data set lie above 15.5.

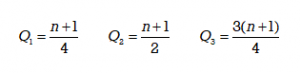
It is important to note that a quartile may not be a data observation. Sometimes there may be a need to average or weight the data values when determining the quartiles.

A second method for determining quartiles is to use a formula to determine the position of each quartile. This is especially useful when there is a large number of data items.

Determining Quartiles: Method 2

**Quartile Formula**

The following formulas, where **n** is the **number of data values**,  can be used to determine the **position** of the three quartiles.



It is important to note that these results indicate the **positions** of the quartiles, not the actual data obervations. If, for example, the calculation gives Q1=3, this indicates that the first quartile will be the data obervation in the 3rd **position.** If  Q3 = 32, this indicates that the third quartile will be the data observation in the 32nd **position.**

Step 1: Order the data from smallest to largest.

Step 2:  Determine**n.**

Step 3: Use the formula to determine the **position** for the median (Q2) of the data set. Count from left to right to determine the corresponding data value. If the position is a fraction then two data values will need to be weighted to determine the median value.

Step 4: Use the formula to determine the **position** for the first quartile Q1 of the data set. Count from left to right to determine the corresponding data value. If the position is a fraction then two data values will need to be weighted to determine the value of Q1.

Step 5: Use the formula to determine the **position** for the third quartile  Q3 of the data set. Count from left to right to determine the corresponding data value. If the position is a fraction then two data values will need to be weighted to determine the value of Q3.

**Method 2:**

Consider the following data set:

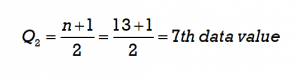
15       4       20       8      3     12      14      11      7     2     6     23     16

Step 1: To determine the  quartiles, order the data values from smallest to largest:

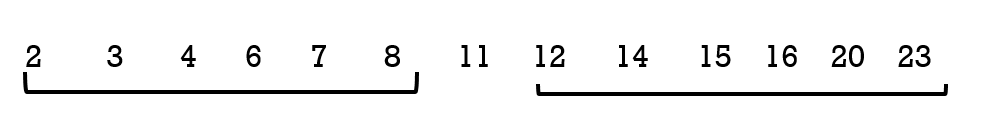
2      3       4     6     7       8      11     12      14      15    16     20     23

Step 2: The number of data values is 13.

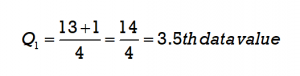
Step 3: Use the formula to determine the **position** for the median (Q2) of the data set.



Count from left to right to determine the corresponding data value in the 7th position. The corresponding value is 11.

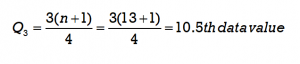


Step 4: Use the formula to determine the **position** for the first quartile (Q1) of the data set.



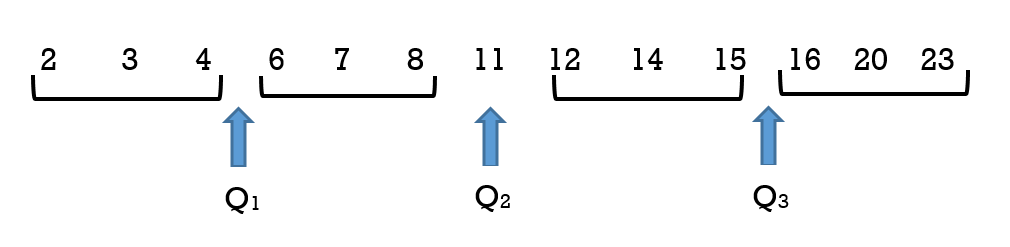
Since 3.5 is a fraction, the first quartile will be the average of the two data values that are in the 3rd and 4th positions. Count from left to right to determine the corresponding data values. The data value 4 is in the 3rd position and the data value 6 is in the 4th position so these will be averaged (4 + 6)/2 = 5. The first quartile will be 5.

Step 5: Use the formula to determine the **position** for the third quartile (Q3) of the data set.



Since 10.5 is a fraction, the third quartile will be the average of the two data values that are in the 10th and 11th positions. Count from left to right to determine the corresponding data values. The data value 15 is in the 10th position and the data value 16  is in the 11th position so these will be averaged (15 + 16)/2 = 15.5. The third quartile will be 15.5.

[Figure 3](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure3) illustrates the three quartiles, which divide the data set into four equal parts.

Fig. 3

EXAMPLE 2

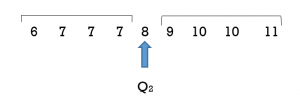
A shoe store wanted to determine the popularity of different shoe sizes for women’s tennis shoes. It planned to place its next order using this information. In  a five day period it sold nine pairs of women’s tennis shoes in the following sizes:    7,  8, 11,  10,  7,   6,   9, 10,  7

**Solution**

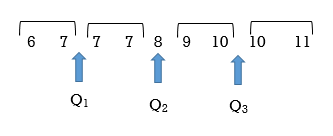
**Method 1:**

To determine the quartiles:

1. Order the shoe sizes from smallest to largest:  6,  7,   7,   7,  8,   9,  10,   10,   11
2. Count the number of values: n = 9
3. Determine Q2, the median, which is the middle observation. Since there are nine data observations (shoe sizes) the median, or second quartile, will be the 5th data value. The 5th data value is 8.



4. Determine the first quartile Q1.It will be the middle observation of the **lower half** of data values. This will be the average of the 2nd and 3rd data values  so (7 +7)/2 = 7.

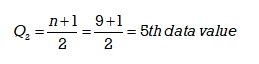


5. Determine the third quartile Q3. This will be the middle observation of the **upper half**. This will be the average of the 7th and 8th data values  so (10+10)/2 = 10

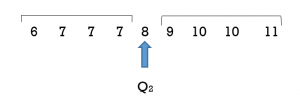
**Method 2:**

The formulas can be used to determine the quartiles.

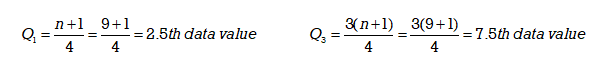
1. Order the shoe sizes from smallest to largest:  6,  7,   7,   7,  8,   9,  10,   10,   11 .
2. Determine the number of data values, n.     n = 9
3. Use the formula to determine the median. The median , or second quartile,  can be determined as follows:



Counting from left to right, the 5th data value is 8. The median, or 2nd quartile Q2, is 8.



4 & 5.  The first and third quartiles can be determined as follows:



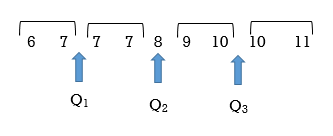
The first quartile is the 2.5th data value. To determine the 2.5th data value we must take the average of the 2nd and 3rd data values. The 2nd data value is 7 and the 3rd data value is 7 so  (7+7)/2 = 7.

The first quartile, Q1 = 7

The third quartile is the 7.5th data value. This will be the average of the 7th and 8th data values. The 7th data value is 10 and the 8th data value is also 10  so   (10+10)/2 = 10.

The third quartile, Q3 = 10

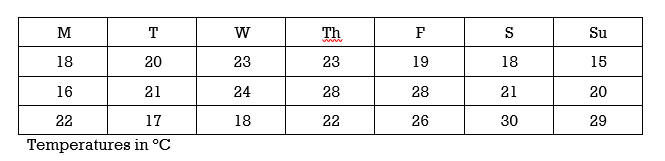
We can see that  Q2= 8  splits the data set into two halves. Q1= 7  is the middle value of the lower half of the data set and Q3 = 10 is the middle value of the upper half of the data set.



In Example 2 the number of data items was **odd.** When *n*  is odd the median or Q2 will be one of the data observations. When *n*is odd the formula for finding quartiles is straight forward.

TRY IT 2

Determine the quartiles for the following temperature data that was recorded over a 3-week period in May:



Show answer

It is important to note that a quartile may **not** be a data observation. When the number of data values *n*  is **even** the median or Q2 will **not** be one of the actual data observations. As a result, when *n*is **even** an adjustment must be made to the value of **n** that is to be used in the formula to determine the **first** and **third** quartiles.

**Method 1**:

Consider the following data set:

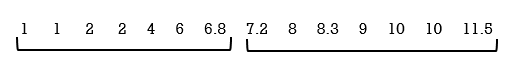
1;  11.5;  6;  7.2;  4;  8;  9;  10;  6.8;   8.3;   2;   2;   10;  1

Step 1: To determine the  quartiles, order the data values from smallest to largest:

1   1   2   2   4   6   6.8   7.2   8    8.3   9    10   10   11.5

Step 2:  The number of data values is 14

Step 3: Determine the median, which measures the “centre” of the data. It is the number that separates ordered data into halves. Half the observations are the same number or smaller than the median, and half the observations are the same number or larger.



Since there are 14 observations, the median lies between the seventh observation, 6.8, and the eighth observation, 7.2. To find the median, add the two values together and divide by two.   Median = (6.8 + 7.2)/2 = 7

The median, and therefore the 2nd quartile Q2 , is seven. It is important to note that the median is not actually one of the observed data values.

Step 4: The first quartile, *Q*1, is the **middle value of the lower half** of the data.

To determine the**first quartile**, Q1, consider the lower half of the data observations:

1     1     2     2    4    6    6.8.

Since there are seven observations, the middle observation will be the 4th item. The middle or 4th item of these data observations  is 2.

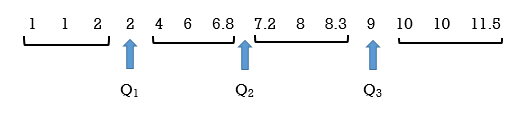
Step 5: The third quartile, *Q*3, is the **middle value of the upper half** of the data.

To determine the**third quartile**, Q3, consider the upper half of the data observations:

7.2     8     8.3    9    10    10     11.5.

Since there are seven observations, the middle observation will be the 4th item in the upper half. The middle item of these data observations is 9.

[Figure 4](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure4) illustrates the three quartiles, which divide the data set into four equal parts.

Fig. 4

The number 2 is the first quartile, Q1. One-fourth of the entire set of observations lie below 2 and  three-fourths of the data observations lie above 2.

The third quartile, *Q*3, is 9. Three-fourths (75%) of the ordered data set lie below 9. One-fourth (25%) of the ordered data set lie above 9.

**Method 2**:

Consider the following data set:

1;  11.5;  6;  7.2;  4;  8;  9;  10;  6.8;   8.3;   2;   2;   10;  1

Step 1: To determine the  quartiles, order the data values from smallest to largest:

1   1   2   2   4   6   6.8   7.2   8    8.3   9    10   10   11.5

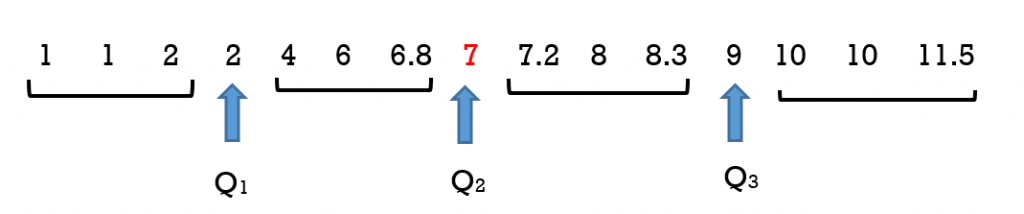
Step 2:  The number of data values is 14 so **n is an even number.**

Step 3: Use the formula to determine the position of  Q2, the median. The position will be (14 + 1)/2 = 7.5. This means that the median,  or Q2, will be in the 7.5th observation or halfway between the 7th and 8th position. The observation 6.8 is in the 7th position and the observation 7.2 is in the 8th position therefore the average of these (6.8 + 7.2)/ 2 is the median or Q2.

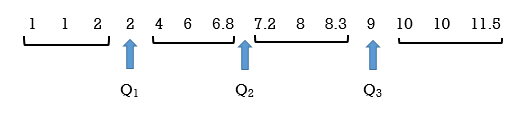
Note that the median is **not** an actual observation in the data set. If we use the formula to find Q1 and Q3 then we must adjust “n” to include this additional item so in effect “n” will be 15. This is done **only** when determining the positions of Q1 and Q3(and not for determining the position of Q2)

Step 4: Use the formula to determine the position of  Q1, the first quartile. Remember than**n** will now be 15, not 14. The position will be (15 + 1)/4 = 4 th. This means that  Q1 will be in the 4th position. Counting from the left, the data value 2 is in the 4th position so Q1= 2.

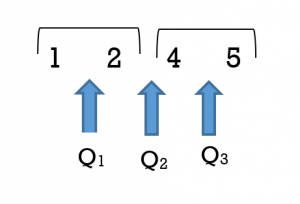
Step 5: Use the formula to determine the position of  Q3, the third quartile. Remember than**n** will now be 15, not 14. The position will be 3(15 + 1)/4 = 12th. This means that  Q3 will be in the 12th position. Refer to [Figure 5](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure5). Counting from the left, we include the median value of  7,  to determine that the data value in the 12th position. This value is 9  so  Q3 will be 9.

Fig. 5

It is also important to recognize that the median of 7 is not an actual data value in this set. It was included in [Figure 5](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure5) to illustrate that its **position** must be counted when determing the position of the third quartile. It is not actually part of the data set. The actual data set is illustrated in [Figure 6](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure6)  (and [Figure 4](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure4)).

Fig. 6

Consider [Figure 7](https://opentextbc.ca/businesstechnicalmath/chapter/8-1-percentiles-and-quartiles/#figure7) where the data set that has an even number of data values:   1   2    4    5

Fig. 7

In this data set  Q1 = 1.5,  Q2 = 3,  and Q3 = 4.5   This illustrates that quartile values need not be actual values in the data set. The second quartile Q2 is 3 which  is the average of the data values 2 and 4. Similarily the first quartile of 1.5 is the average of two data values 1 and 2 and the third quartile of 4.5 is the average of the two data values 4 and 5. Determining the quartile values can become complex as it may require different weightings of the data values but this is beyond the scope of this textbook.

Example 3 illustrates two techniques for determining quartiles when the number of data observations is **even**.

EXAMPLE 3

Consider again the  shoe store and a different week. Over a five day period it sold ten pairs of tennis shoes in the following sizes:

6,  8, 11,  10,  7,   6,   9, 10,  8,   9

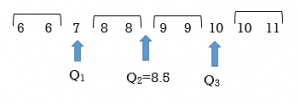
Note that there is an **even** number of data values  n = 10

**Solution**

**Method 1:**

To determine the quartiles:

1. Rank the sizes from smallest to largest:   6,  6,   7,   8,  8,   9,   9,   10,   10,  11 and divide the data set into four equal quarters.
2. n = 10
3. Start with the median which is the middle observation. The median, or second quartile, will lie between the 5th and 6th data values. The 5th data value is 8  and the 6th data value is 9  so the average of 8 and 9, or 8.5, is the median.
4. Determine the first quartile Q1.It will be the middle observation of the **lower half** of data values. This is the 3rd data value or the observation of  7.
5. Determine the third quartile Q3. This will be the middle observation of the **upper half**. This will be the data observation of 10.

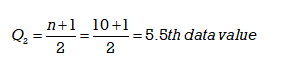


Note that each quartile divides the data values such that there are an equal number of data values in each of the four sections.

**Method 2:**

**An alternative is to use the formulas** to determine the quartiles.

To determine the quartiles:

1. Rank the sizes from smallest to largest:   6,  6,   7,   8,  8,   9,   9,   10,   10,  11
2. n = 10
3. Determine the position of the median  using the formula.

The 5.5th data value will be the average of the 5th and 6th data values. The 5th data value is 8  and the 6th data value is 9  so

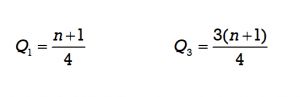
(8 + 9) /2 = 8.5  The median or Q2  is 8.5.

**Note:** Q2 is **not** one of the **actual** data values.  In this example Q2 is the 5.5th data value or 8.5. It is the data value that lies between the 5th and 6th data values but it is not one of the original data values.

4 & 5.   Determine the first quartile Q1and the third quartile Q3.

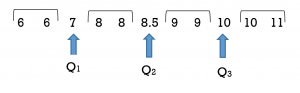
Since the number of data values ***n*** is **even** the median or  Q2 is **not** one of the **actual** data values so when we use the formula to determine Q1  and Q3 we must **increase the value of n by 1**. In effect the number of data values has increased by one and therefore the value of **n** in the formula must be increased by 1. This is done only when determining the positions of Q1 and Q3(and not for determining the position of Q2)

In this example, when determining Q1 and Q3 the original value of  **n = 10** will now be increased by 1. The new number for **n** to be used in the formula will be**n = 11.** Using the formula, the first and third quartile positions can be determined as follows:





Using the results from the formula we count to get the 3rd and 9th data values. When determining these values  be sure to include and count the **position occupied by the new median value** of  8.5. The 3rd data value is 7  and the 9th data value is 10.

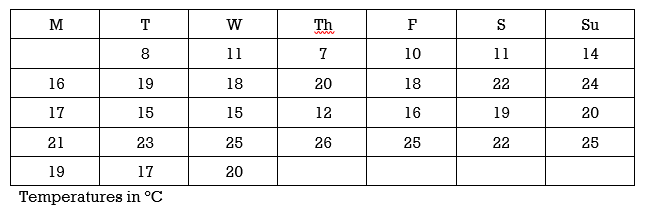


Note that Method 1 and Method 2 yield the same results.

We have seen that either of Method 1 or Method 2 will produce the same quartile values although the formula method can be less intuitive when **n** is **even.**

TRY IT 3

Use either technique to determine the quartiles for the following temperature data that was recorded over the month of April:



Show answer

EXAMPLE 4

Consider the data set:  3, 4, 5, 6, 7, 8, 8, 9, 10, 10, 11, 12, 13, 13, 14, 15.   Determine the three quartiles using either technique.

**Method 1:**

Step 1: Order the data values     3, 4, 5, 6, 7, 8, 8, 9, 10, 10, 11, 12, 13, 13, 14, 15

Step 2:  n = 16

Step 3:  The median will be the average of 9 and 10, so 9.5. This is not one of the observed values.

Step 4:  Q1 is the value that splits the lower half, which will be the average of 6 and 7, so 6.5.

Step 5:  Q3 is the value that splits the upper half, which will be the average of 12 and 13, so 12.5.

**Method 2:**

Step 1: Order the data values     3, 4, 5, 6, 7, 8, 8, 9, 10, 10, 11, 12, 13, 13, 14, 15

Step 2:  n = 16

Step 3: Use the formula  (16 + 1) /2 = 8.5. The median will be in the 8.5th position. This is the average of the 8th value of  9 and the 9th value of 10 so the median is 9.5

Step 4 and 5: Since n is **even,** we will use a value of 17, not 16, in the formulas to determine Q1 and Q3.

Q1 will be (17 + 1)/4 = 4.5 th. This means that  Q1 will be in the 4.5th position or the average of the 4th and 5th data values. The 4th value is 6 and the 5th value is 7 so  Q1= 6.5.

Q3 will be 3(17 + 1)/4 = 13.5 th. This means that  Q3 will be in the 13.5th position or the average of the 13th and 14th data values. Including the median’s position when we count, the 13th value is 12 and the 14th value is 13 so  Q3 = (12 + 13)/2 = 12.5.

**Key Concepts**

* **A data set can be divided into one hundred equal parts by ninety-nine percentiles P1 , P2 , P3 , … P99 . Percentiles are best used with large sets of data.**
* **Quartiles divide the data set into four equal parts. The first quartile, *Q*1, is the same as the 25th percentile, and the third quartile, *Q*3, is the same as the 75th percentile. The median can be called both the second quartile, Q2, and the 50th percentile.**
* **Quartiles may or may not be actual observations within a set of data.**