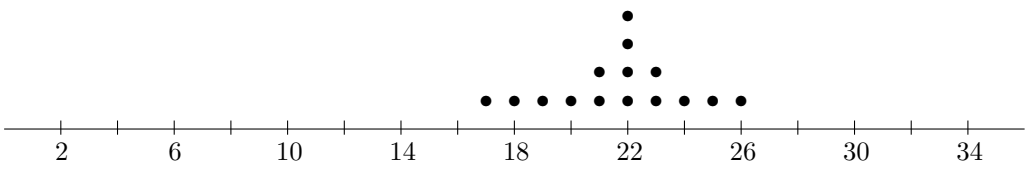
We consider here different ways of representing data sets.

1. Consider the following 15 numbers.

$$22, 17, 21, 23, 22, 25, 24, 18, 19, 20, 21, 22, 23, 26, 22$$

To make the *dotplot* of the 15 numbers, we can first put them in order (but you don't have to):

$$17, 18, 19, 20, 21, 21, 22, 22, 22, 22, 23, 23, 24, 25, 26$$



2. Consider the numbers

$$32, 45, 67, 42, 33, 23, 37, 44, 33, 36, 45, 41, 33, 35, 22, 15, 9, 8, 17, 33$$

To make the *stem-and-leaf plot* of the 20 numbers, we can first put them in order(but you don't have to):

$$08, 09, 15, 17, 22, 23, 32, 33, 33, 33, 33, 35, 36, 37, 41, 42, 44, 45, 45, 67$$

We make the first digit in the two digit number the stem and the second number in the leaf. The stem goes on the left of the line, and the leaves go on the right of the line.

0	89
1	57
2	23
3	23333567
4	12455
5	
6	7

Since the rows of 3 and 4 are so long, sometimes we can break those rows into two or more smaller rows to make it easier to read.

0	89
1	57
2	23
3	2333
3	3567
4	124
4	55
5	
6	7

3. Consider the following numbers (already in order)

4 9 13 15 17 20 24 25 26 29

We give what is called a five number summary.

$$\begin{aligned} \text{Max} &= 29 \\ Q_3 = \text{third quartile} &= 25 \\ Q_2 = \text{Median} &= \frac{17+20}{2} = 18.5 \\ Q_1 = \text{first quartile} &= 13 \\ \text{Min} &= 4 \end{aligned}$$

You are already familiar with the min, max, and median, but how do we obtain the first quartile (denoted Q_1) and the third quartile (denoted Q_3)?

To calculate the first quartile, look at all the number to the left[1] of the median 18.5:

4 9 13 15 17

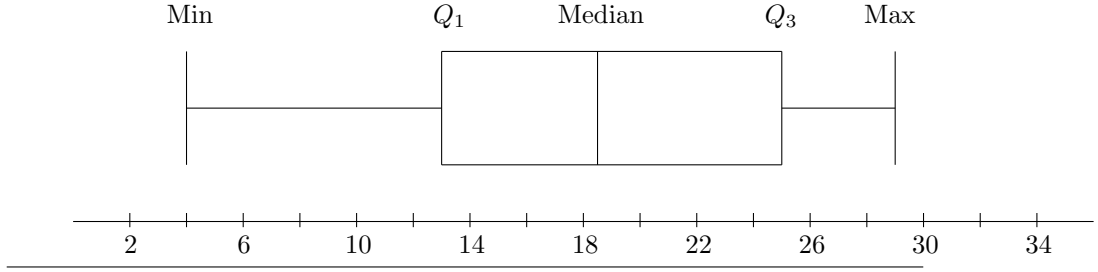
The median of this set of numbers, which is 13, is the first quartile.

To calculate the third quartile, look at all the number to the right[2] of the median 18.5:

20 24 25 26 29

The median of this set of numbers, which is 25, is the third quartile.

To make a *boxplot* (see below), we have vertical lines representing each of the five numbers. There is a box made between Q_1 and Q_3, which denote the interquartile range – the middle 50% of the data. The two tails on the ends represent the lower 25% and the upper 25%.



Mail to rstephens@colgate.edu

[1]Be careful if there are numbers equal to the median. For instance, if the four middle numbers are 8, 8, 8, 8, the median would be 8, and the left two 8's would be considered when finding the first quartile. Also, if you have an odd number of data values and your median is one of your data values, then you consider all the numbers strictly to the left of the median.

[2]See above footnote. The right two 8's would be considered when finding the third quartile. In the case of an odd number of data values, you would consider all the numbers strictly to the right of the median.