

The Five Number Summary

The five number summary of a set of observations on a single variable consists of the following statistics:

Maximum (max) – the largest observation

Upper Quartile (Q3) – a value that separates the largest 25% of the observations from the smallest 75%

Median (M) – a value that separates the largest 50% of the observations from the smallest 50%

Lower Quartile (Q1) – a value that separates the largest 75% of the observations from the smallest 25%.

Minimum (min) – the smallest observation

We typically will use a computer to calculate the five number summary, but it is good to know how to compute the summary by hand. To do so...

1. Arrange the observations in order from smallest on the left to largest on the right.
2. Record the minimum and maximum.
3. The median is the middle number in the list if there are an odd number of observations and the average of the two middle numbers if there are an even number of observations.
4. If the number of observations minus one is evenly divisible by four, the upper quartile is the median of the observations starting with the median and including all observations to the right. If the number of observations minus one is not evenly divisible by four, the upper quartile is the median of the observations to the right of the location of the overall median.
5. If the number of observations minus one is evenly divisible by four, the lower quartile is the median of the observations starting with the median and including all observations to the left. If the number of observations minus one is not evenly divisible by four, the lower quartile is the median of the observations to the left of the location of the overall median.

For example, suppose we wish to compute the five number summary for the following observations:

19 11 7 24 13 15 10 3 10 20

1. We order the observations.

3 7 10 10 11 13 15 19 20 24

2. The minimum and maximum are 3 and 24, respectively.
3. The median is $(11 + 13)/2 = 12$ because 11 and 13 are the two observations in the middle of the list.
4. The number of observations minus one ($10 - 1 = 9$) is not evenly divisible by four ($9/4 = 2.25$), so the upper quartile is the median of the observations to the right of 12, i.e.,

13 15 19 20 24

The median of these observations is 19 because 19 is the middle number in this list. Thus, the upper quartile is 19.

5. Likewise the lower quartile is the median of the observations to the left of 12, i.e.,

3 7 10 10 11

The first 10 in the list is the median of these observations, so the lower quartile is 10.

Now suppose we wish to compute the five number summary for the following observations:

6 22 13 2 18 28 26 9 18

1. We order the observations.

2 6 9 13 18 18 22 26 28

2. The min and max are 2 and 28, respectively.

3. The first 18 to appear in the list is the median because there are four numbers in the list to either side of the first 18, i.e., the first 18 is the middle number.

4. There are 9 observations. The number of observations minus one ($9 - 1 = 8$) is evenly divisible by four ($8/4 = 2$). Thus, the upper quartile is the median of the observations

18 18 22 26 28.

because these are the observations that include the overall median and all numbers to the right. The middle number in this list of numbers is 22, so 22 is the upper quartile.

5. The lower quartile is the median of the observations

2 6 9 13 18

because these are the observations that include the overall median and all observations to the left. Thus the lower quartile is 9 – the middle number in the list.

Please verify that the five number summary for the data set

13 17 7 10 2 9 3 26 18 24 2 29

is $\min=2$, $Q1=5$, $M=11.5$, $Q3=21$, $\max=29$.