The Five Number Summary and Boxplots

Five-number summary

What is a Five-Number Summary?

Five Number summary: the five numbers that summarize the overall characteristics of a data set. These include the minimum, first quartile, median, third quartile, and maximum.

Minimum: the smallest value in the data set

First quartile: the value that separates the first 25% or first quarter of a set of data from the rest of the data. This is found by determining the median of the lower half of the data set.

Median: the middle value in a data set when ordered from least to greatest. If there are two middle values, the median is found by solving for the average of the two values

Third quartile: the value that separates the first 75% or three-quarters of a set of data from the rest of the data. This is found by determining the median of the upper half of the data set.

Maximum: the largest value in the data set

Example 1

Karen owns a bakery. She keeps track of how many items she sells each day. The number of items she sold over the last two weeks are: 25, 34, 28, 48, 52, 37, 29, 35, 54, 43, 64, 39, 58, 45. Summarize this data set with the five-number summary.

Step 1: Order the values from least to greatest.

Our list of values from least to greatest is: 25, 28, 29, 34, 35, 37, 39, 43, 45, 48, 52, 54, 58, 64

Step 2: Determine the minimum and maximum of the data set by identifying the lowest and highest values.

Based on our ordered list, the minimum value is 25 and the maximum value is 64.

Step 3: Find the median of the data set. Separate the lower half from the upper half.

There are two middle numbers in our list of values: 39 and 43. To find the median we calculate $39+432=822=41$

Our median is 41. We will use this to separate the data set into a lower half (25, 28, 29, 34, 35, 37, 39) and an upper half (43, 45, 48, 52, 54, 58, 64).

Step 4: Find the first and third quartiles by finding the median of the lower and upper half of the data.

The median, or middle number, of our lower half 25, 28, 29, 34, 35, 37, 39 is the value 34. Therefore our first quartile is 34.

The median, or middle number, of our upper half 43, 45, 48, 52, 54, 58, 64 is the value 52. Therefore our third quartile is 52.

Step 5: Summarize the data set by stating the minimum, first quartile, median, third quartile, and maximum.

Based on Steps 1 - 4, our five-number summary is:

- **Minimum: 25**
- **First Quartile: 34**
- **Median: 41**
- **Third Quartile: 52**
- **Maximum: 64**

Question:

Two dice were rolled 15 times, and their sum was recorded. The sums of the 15 throws are: 5, 7, 9, 7, 6, 12, 2, 4, 8, 11, 6, 7, 3, 10, 7.

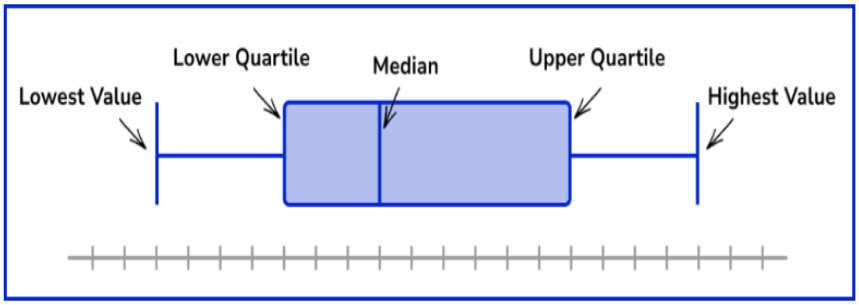
Summarize this data set with the five-number summary.

What is a box plot?

A box plot is a diagram showing the following information for a set of data.

1. Lowest value or smallest value (Minimum)
2. Lower quartile or first quartile (Q_1)
3. Median, middle number, or second quartile (Q_2)
4. Upper quartile or third quartile (Q_3)
5. Highest value or largest value (Maximum)

This set of descriptive statistics is called the five-number summary. The box plot must be featured on a scale to show these values clearly.

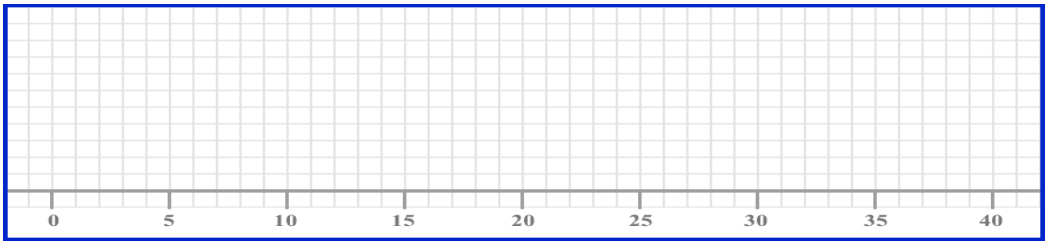


<u>Example 1:</u>

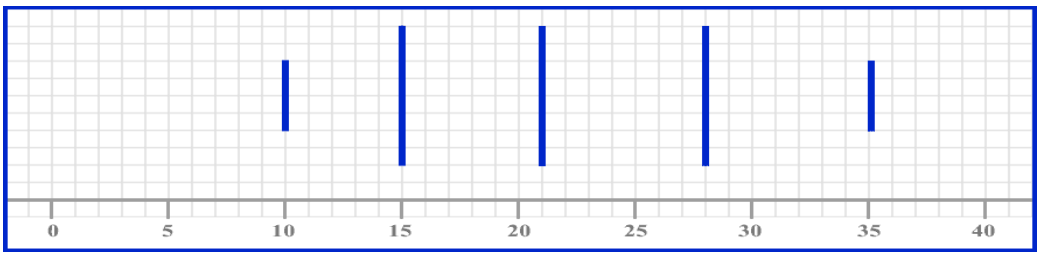
Draw a box plot using the following information.

Lowest Value	10
Lower Quartile	15
Median	21
Upper Quartile	28
Highest Value	35

The scale needs to be long enough to mark on the lowest and highest values, so in this example, we use 0 to 40.

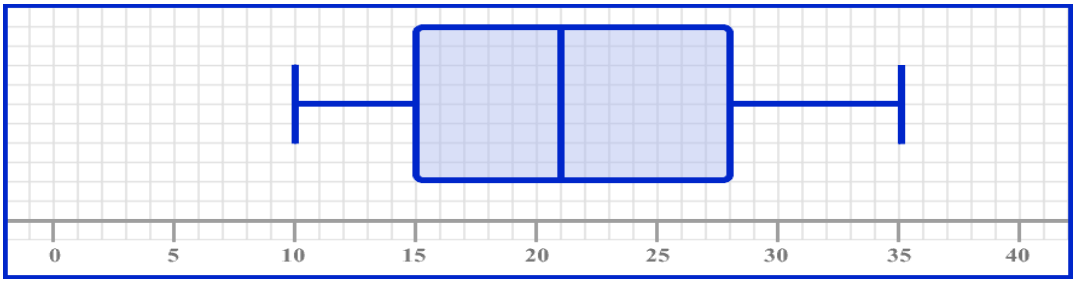


Mark the five key values with vertical lines.



Join the Q_1 and Q_3 to form the box and draw horizontal lines to the minimum and maximum values.

The box runs from the lower quartile (15) to the upper quartile (28). The whiskers join the box; the lower whisker ends at the minimum value (10) and the upper whisker ends at the maximum value (35).



Question

Draw a box plot for the following data points.

1,1,2,3,5,7,7,8,10,12,151,1,2,3,5,7,7,8,10,12,15