## Data Display

## Box-and-Whisker Plots

**Introduction:** There are many ways to organize and display data. A box-and-whisker plot is a display used when the purpose is to show the data's distribution in quarters, using the median, quartiles, and extremes. The display visually represents quartiles, the median, and the extremes. Half of the data points fall in the box while the whiskers show the distance to the extremes.

**Instruction:** When data is displayed on a box-and-whisker plot, the plot will divide the data points into fourths and will identify the median, the quartiles, and the extremes. The median is the number that separates the upper half and the lower half. When each of the halves is separated into two equal parts, the upper quartile is the number that separates the upper half and the lower quartile is the number that separates the lower half. The word *quartile* refers to the points that divide the data points into four groups. The word *quartile* is also sometimes used to refer to the four groups. When used this way, the first and second quartiles are the lower half of the set and the third and fourth quartiles are the upper half. From left to right, they are the first, second, third, and fourth quartiles. The lower extreme is the least data value. The upper extreme is the greatest data value.

**Example 1:** Display the data from the following chart using a box-and-whisker plot.

**Time spent practicing an instrument during the week (in minutes)**

| | M | T | W | Th | F |
|---|---|---|---|---|---|
| **Jerry** | 25 | 15 | 13 | 36 | 27 |
| **Jeannie** | 22 | 10 | 15 | 17 | 20 |
| **Judy** | 10 | 12 | 18 | 19 | 20 |

Step 1: Order the data points from least to greatest.

10 10 12 13 15 15 17 18 19 20 20 22 25 27 36

Step 2: Find the median in the set of data points. Since this set of data points has 15 values, the median (middle) value is 18. There are seven values to the left of 18 and seven values to the right of 18.

10 10 12 13 15 15 17 18 19 20 20 22 25 27 36

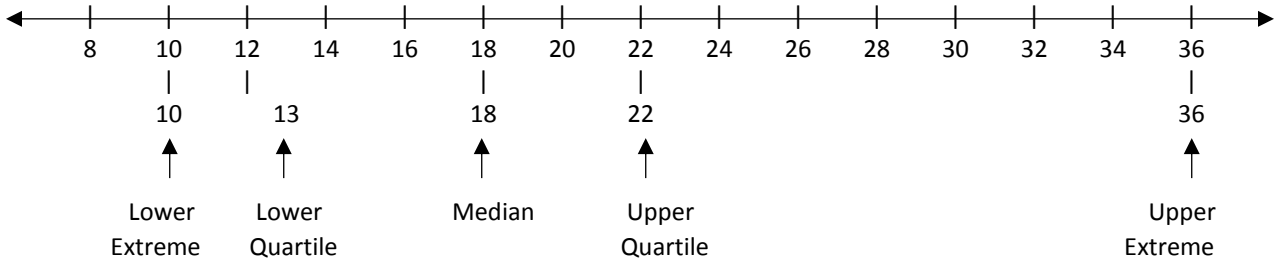
Median

Step 3: Find the median values in the lower half and the upper half of the data set. The median value of the upper half is called the upper quartile. The median value of the lower half is called the lower quartile. In this set of data points, the lower quartile (middle value in the lower half) is 13 and the upper quartile (middle value in the upper half) is 22. Note that these values separate the data points into four groups. The values at the lower end and the upper end are the *lower extreme* and the *upper extreme.*

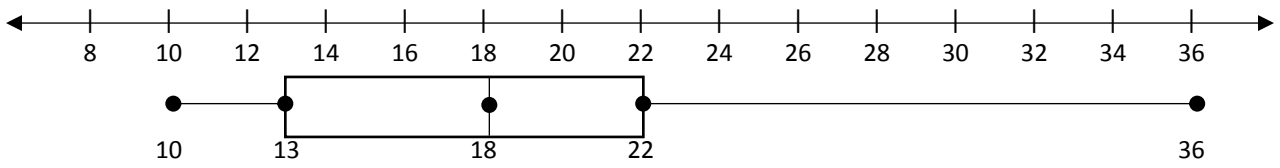
10 10 12 13 15 15 17 18 19 20 20 22 25 27 36

Lower Extreme — Lower Quartile — Median — Upper Quartile — Upper Extreme

Step 4: Plot the five values below a number line.



Step 5: Draw a box with sides at the upper quartile and lower quartile. Draw a vertical line through the median. Draw *whiskers* from the quartiles to both extremes.




**Q:** Does the number 12 fall in the 1st, 2nd, 3rd, or 4th quartile?
**A:** 1st.

**Extended Reasoning**

**Q:** Notice that this box is rather small compared to the length of the whiskers. What does this tell you about the values for the middle half of this set of numbers?
**A:** They are all very close to each other and very close to the median. There is not a lot of "difference" between the middle numbers.

**Q:** Notice that the whisker from the upper quartile to the upper extreme is pretty long compared to the box. What does this tell you about the values from the upper fourth of the set?
**A:** They were spread pretty far apart. There was a big difference between the middle or the median and the upper extreme.

**Example 2:** Display the data from the following set using a box-and-whisker plot

Miss Ortiz decided to have her students keep track of how much money they spent at the school's concession stand for meals and snacks during basketball season. Each of her 19 students rounded their expenses to the nearest dollar and reported at the end of the season. Their reported amounts follow.

6 14 15 21 11 37 12 19 22 20 11 17 29 15 24 16 26 15 7

Step 1: Order the data points from least to greatest.

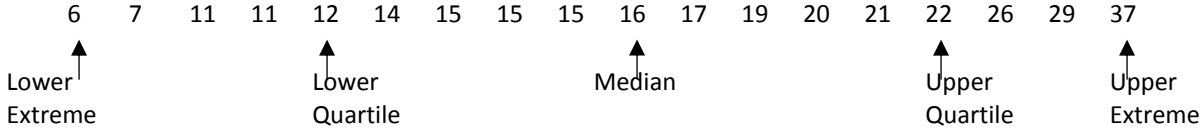
6 7 11 11 12 14 15 15 15 16 17 19 20 21 22 24 26 29 37

Step 2: Find the median in the set of data points. Since this set of data points has 19 values, the median (middle) value is 16. There are nine values to the left of 16 and nine values to the right of 16.

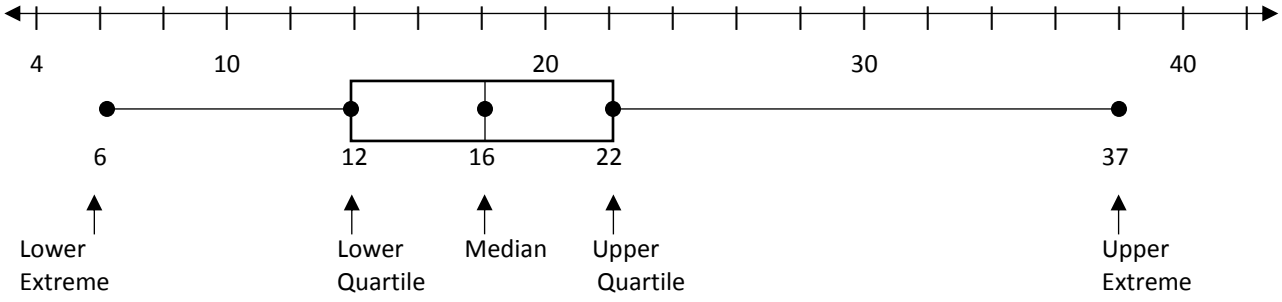
6 7 11 11 12 14 15 15 15 16 17 19 20 21 22 26 29 37

Median

Step 3: Find the median (middle) values in the lower half and the upper half of the data set. In this set of data points, the middle value in the lower half is 12 and the middle value in the upper half is 22. Twelve is the lower quartile and 22 is the upper quartile. The values at the lower end and the upper end are the *lower extreme* and the *upper extreme.*




Step 4: Plot the five values below a number line.



Step 5: Draw a box with sides at the upper and lower quartiles. Draw a vertical line through the median. Draw *whiskers* from the quartiles to both extremes.

**Question/Answer**

**Q:** What is the mean for the set of data? (the average amount spent by a student in Miss Ortiz's class). Round your answer to the nearest dollar.
**A:** $18

**Q:** How does the mean for this set of data compare to the median for this set of data?
**A:** They are close but not exactly the same. There is a $2 difference between the median and the mean. [Note that mean and median numbers are most often close in value, but they are not the same and represent different ways of thinking about the middle of a set of data points.]

**Q:** What is the range of the data for Miss Ortiz's class?
**A:** $31

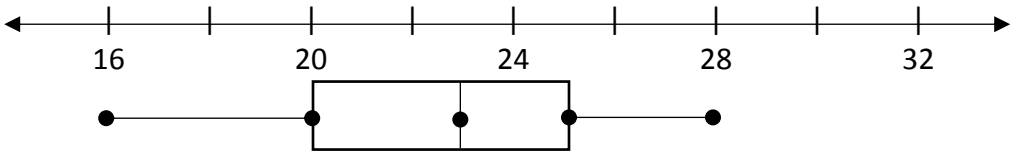
**Extended Reasoning**

**Q:** Can we infer from this range that one student has more money than another student?
**A:** Not really. The data only tells us that one student spent $31 more than another student.

**Q:** What other reasons might explain the range?
**A:** The student who spent more money might have attended a lot more games. The student who spent more money might have purchased things for friends. The student who spent less money might be saving her money for something else. [We must be careful about the conclusions we draw from data. The only valid conclusions are those actually supported by the data.]

## Interpreting Box-and-Whisker Plots

Miss Ortiz's class planted bean seeds in science class. After four weeks, they measured the height of their plants. The box-and-whisker plot shows the height, in centimeters, of the plants.




**Question/Answer:**

**Q:** What is the median height for the plants?
**A:** 23 cm

**Q:** What is the range of height for the plants?
**A:** 12 cm

**Q:** What is the upper extreme for the plants?
**A:** 28 cm

**Q:** What is the lower extreme for the plants?
**A:** 16 cm

**Q:** What is the upper quartile for the plants?
**A:** 25 cm

**Q:** What is the lower quartile for the plants?
**A:** 20 cm

**Extended Reasoning**

**Q:** Can you calculate the mean (average height of a plant) for the set of data?
**A:** No, the mean can't be calculated from the data provided.

**Q:** Why can't the mean be calculated?
**A:** The plot does not indicate the number of plants that were measured or how many plants were at each height. The number of data points is not included in a box-and-whisker plot. [Note that a box-and-whisker plot only provides some information about the data—the median, the quartiles, and the extremes.]

**Q:** Is the mean plant height *likely* to be in the box or in one of the whiskers? Explain your answer.
**A:** In the box. The *mean* value and the *median* value are most often close in value, and the median is in the box.