



Statistics & Probability

Creating a Box Plot

Grade 6

Rationale

- Graphical displays of data are helpful when interpreting information. Box plots are frequently found in statistics and are one way to view data. The ability to understand box plots can help students interpret trends and information in the data easily.

Goal

- To understand the relationship between data points and box plots

Standards

- 6.SP.4** Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
- 6.SP.5** Summarize numerical data sets in relation to their context, such as by:
 - Reporting the number of observations.
 - Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.
 - Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
 - Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

Objectives

- Students will understand how to create box plots.
- Students will identify mean, median, and interquartile range.
- Students will find outliers in a box plot.

Materials

- August Temperatures Box Plot* Activity Sheet (one per group)



Procedures

- ✚ Initiate a discussion about box plots. Ask if they have they seen box plots, where they've seen them, and for what purposes they are used. This is intended to be an introduction to box plots, so it is ok if they don't yet have answers for this question, but this will help you to assess their prior knowledge.
- ✚ Explain to students that a box plot can be used to display data points. Divide students into groups of 2 or 3. Give each group a copy of the *August Temperatures Box Plot* activity sheet.
- ✚ Explain to students that they will work in groups to determine how a box plot is created. Groups will use the data points and box plot from the activity sheet to determine the relationship between the box plot and the data points. Groups may use guess and check, working backwards, or other strategies.
- ✚ As groups work, pose questions that will help guide students toward finding the relationship. For example, ask students to locate the data points on the box plot. Check that groups order the data points. Have students find the median (74.6) of the data points.
- ✚ Once groups have had a chance to discover the relationship on their own, discuss their findings. Students may have found that the median, the minimum, and the maximum are all identified on the box plot. Discuss how the median splits the data points into two groups. Students may also have found that the medians of each of the two groups (68.1 and 79.95) are identified on the box plot.
- ✚ Discuss with groups how the data points are divided on the box plot. Students should notice that the data points are divided into four sections on the box plot. Explain that these sections are called quartiles. The greatest point in each quartile is given a name. Ask groups to label each quartile point with Q1, Q2, Q3, or Q4 as it is discussed.
- ✚ Q2 is the median of all data points. Have students label Q2 (74.6) by the line dividing the data points into two groups.
- ✚ Q1 is the median of the lower half of data points. Have students label Q1 (68.1) by the line on the bottom of the box, which divides the lower half of data points into two groups.
- ✚ Q3 is the median of the higher half of data points. Have students label Q3 (79.95) by the line on the top of the box, which divides the higher half of data points into two groups.
- ✚ Q4 is the greatest data point. Have students label Q4 (81.4) by the line at the top of the box plot.
- ✚ Have groups calculate the range of the box. This can be found by subtracting Q1 from Q3 (11.85). Explain that the range of the box is called the interquartile range (IQR). Any data points that lie outside the box a distance greater than one and a half times the IQR are considered outliers.
- ✚ Explain to students that outliers are labeled on a box plot by a dot. They are not included in the calculations for the box plot and thus lie beyond the whiskers. Have groups calculate these values to create an outlier data point

that could be added to the box plot. $Q1 - 1.5 * IQR$ and $Q3 + 1.5 * IQR$. For example, 98 and 50 would be outliers.

- ✚ Explain to students that some box plots have the mean indicated on the box plot by an "x". Have groups calculate the mean of the data points (74.48) and mark the mean value on the box plot with an "x."
- ✚ **Closing:** Ask students questions about box plots to test their understanding. For example, what would it mean if a box plot had a short box? How about a long whisker?

Teacher Tips

- ✚ When discussing a median dividing data points into two groups, you may wish to draw a physical line in the data list show how it splits the data points
- ✚ Explain to students that in the case of an odd number of data points, the median would not be included in either of the two groups when finding the first and third quartiles.
- ✚ Explain the effects of including an outlier in the data list and calculations when creating a box plot.

Extension Activities

- ✚ Have students find or collect data and create their own box plots. Students may wish to include one outlier in their data lists.
- ✚ Students find examples of box plots online or in newspapers to see how they are used in real world situations. Discuss how the examples differ and what information can be found by viewing the box plot without the data list.
- ✚ Have students work in groups to create a data list and box plot. Groups make one purposeful error when creating their box plots. Groups trade and identify the error that was made when creating the box plot. For example, the median was calculated incorrectly, thus $Q2$ is incorrect or an outlier was included in the box plot.

August Temperatures Box Plot

The following data points are temperatures recorded in degrees Fahrenheit on days in the month of August. The points are graphed onto a box plot.

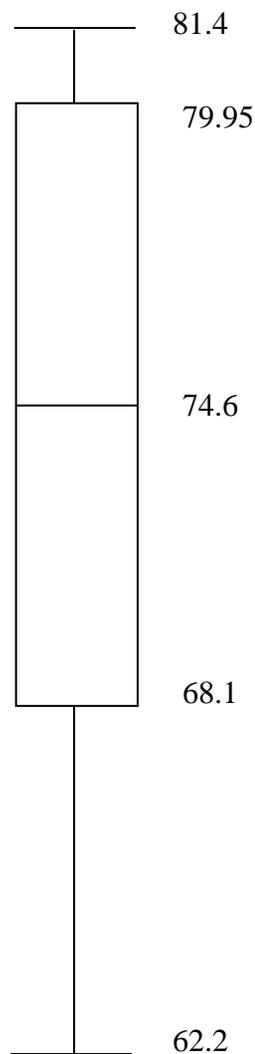
Part A-

Data Points:

69.2, 75.5, 80.1, 64.5, 71.4, 62.2, 78.7, 79.8, 80.3, 81.4, 67.0, 73.7

Part B-

Box Plot:



Part C-

Directions: Complete the questions below as you determine the relationship between the data points and the box plot. When you have finished, discuss your results with the class.

1. How many sections do you see on the box plot?

2. How many data points are found in each section of the box plot?

3. What do you think the box represents?

4. The extensions from the box are called “whiskers,” what do you think those represent?

5. Why do you think the two whiskers different lengths?
