

Now that we understand the mean as a measure of center and standard deviation as a measure of spread, our next data analysis will combine the two using the median and quartiles. When we think of spread we should look at three major measures to determine consistency. The range and standard deviation can tell us how spread out the whole data set is, but what if we want more specific information? One way to analyze the data is to break it up into four equal sections and see how spread out each section is. Even though these sections will have the same number of values in them, they will not necessarily be the same length. Let's see how this works.

Measure the length of each complete Cheetos to the nearest tenth of a centimeter using a ruler and write these measurements in the table below.

[illegible]

For the next part, finding the five number summary, you must SHOW ALL THE WORK to get credit. To find our quartiles without the calculator we must first line up all of our lengths in order from least to greatest.

We find the median by finding the middle number of the list (or the average of the two middle numbers) and mark it.

Next, find the middle number of each side of the median and label the smaller number as quartile one ( $Q_1$ ) and the larger one as quartile three ( $Q_3$ ).

The last two numbers in this five number summary will be the minimum and maximum values which should be labeled min and max.

To graph this information we must draw a vertical line ( | ) at each of these numbers.  
Connect the minimum and  $Q_1$  and the maximum and  $Q_3$  with a horizontal line called a whisker.  
Connect  $Q_1$  and  $Q_3$  with two horizontal lines to create the box. Be sure to label the axis and title the graph.

[illegible]

Using these plots we can determine our SOCS with the information below.

### Shape

If the sections are all about the same length, then the data can be considered **symmetric**

If the sections on the right are longer than we can say the data is **skewed right**

If the sections on the left are longer than we can say the data is **skewed left**

### Outliers

We can now use a formula to determine whether there are any outliers. If we have any numbers that are less than  $Q_1 - 1.5$  (IQR) or greater than  $Q_3 + 1.5$  (IQR) then it is an outlier. The IQR is the interquartile range which equals  $Q_3 - Q_1$ .

### Center

List just the median to describe this type of data

### Spread

List all of the five number summary values to describe the spread of the data and decide on the consistency based on the skew

Describe your data using your SOCS below.

Put your data into  $L_1$  in the calculator and perform a 1 varstat to check all your five number summary values.

Graphing box and whisker plots in the calculator.

Press **2<sup>nd</sup>** then **y =** to open the statplot menu. Press **enter** to set up the graph as a boxplot and press **graph** then **zoom 9**.

Next you will check your outliers by creating a modified box and whisker plot. Go through the same graphing steps but choose the box and whisker plot with the points outside the plot. If you get points below or above the whiskers then these values are outliers. You can find these numbers by pressing **trace** and highlighting the points.

Now let's compare your box plot to another.

Keeping your data in  $L_1$  type the data of a classmates into  $L_2$ . Set up a statplot using  $L_2$  and make sure that both plots are on. Press graph and if necessary, zoom 9.

Create a quick sketch of the two box and whisker plots below.



What similarities and differences do you notice about the plots?

Write the five number summary for the other group's data.

What similarities and differences do you notice about the five number summaries?

What does this say about the lengths of your Cheetos compared to your classmate's?