# Box-and-Whisker Plots Key

Suppose you have a large set of data and you want a display that gives a general idea of how the data cluster together. A box-and-whisker plot displays the median, the quartiles, and outliers of a set of data, but does not display any other specific values.

In box-and-whisker plots (box plots), data are grouped by first being reduced to a five-number summary, consisting of the minimum value, the lower quartile, the median, the upper quartile, and the maximum value. The shape of the box plot and the five-number summary provide benchmarks for dividing the data into quartiles. It is not possible to identify individual data items from a box plot; however, it is easy to see the spread of the data and to identify the median, as it is one of the five summary numbers.
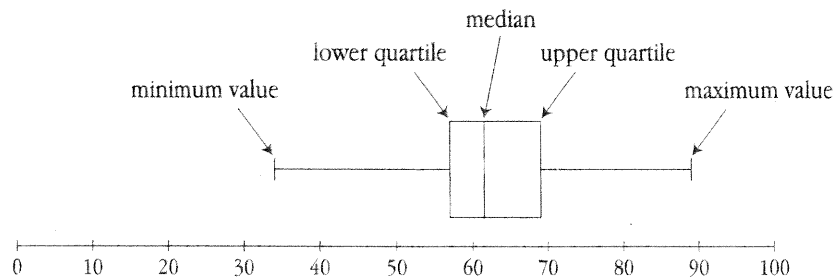
A box plot is constructed from the five-number summary of the data. You know how to find the minimum value, maximum value, and median in a set of data. The lower quartile is the median of the data values below the median. The upper quartile is the median of the data values above the median. The diagram below illustrates the five-number summary for the quality ratings of the natural brands of peanut butter. (Note: we will be discussing that information in an upcoming class activity.)
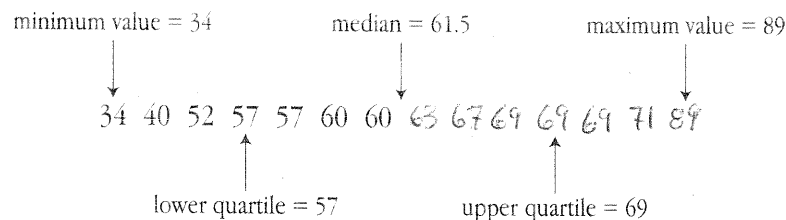

The box plot below shows how the numbers in the five-number summary correspond to the features of the box plot.
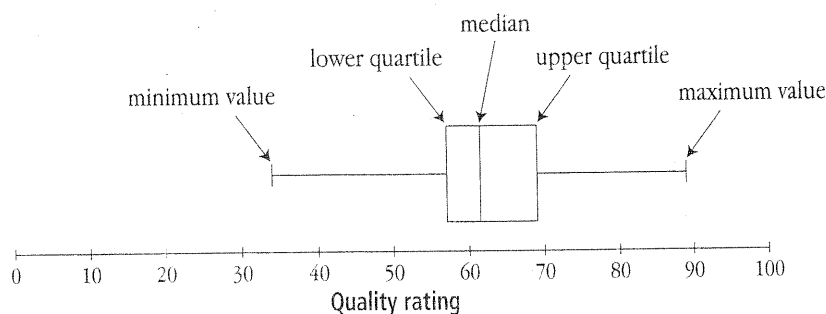
At the spring parade, the local dentists' organization was tossing packets of sugarless gum balls to all the children. Sook Leng, who was watching the parade, wondered how the gum balls had been put into the packets. A dentist told her that each packet contained 8 gum balls, which came in green, red, and yellow. A total of 40,000 gum balls had been mixed thoroughly and put into the packets. However, the dentist did not know the mix of colors in the batch of 40,000.

Sook Leng asked the children near her to tell her the number of each color of gum balls in their packets. When she went home, she made three line plots from the data she had collected from 30 packets. The line plots are shown below:
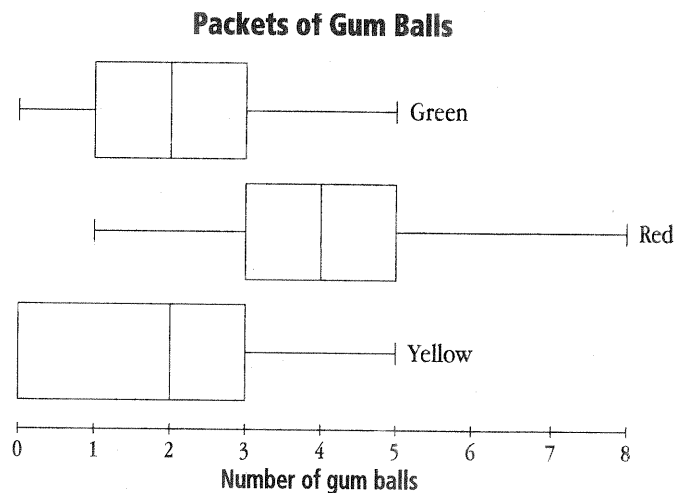
Colors of Gum Balls

Green gum balls   Red gum balls   Yellow gum balls
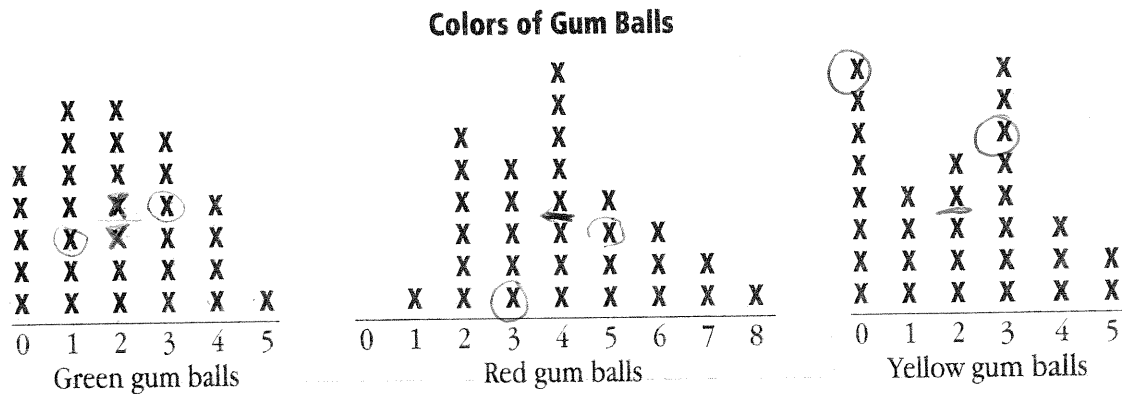
a. Make a box plot of the distribution of the number of gum balls of each color: Draw your three box plots on the same scale.

five number summaries

	G	R	Y
min value	0	1	0
lower quartile	1	3	0
median	2	4	2
upper quartile	3	5	3
max value	5	8	5

BOX PLOTS

Packets of Gum Balls

Number of gum balls

This gives about 10,000 green, 20,000 red, + 10,000 yellow, or 25% green, 50% red, and 25% yellow.

b. Using your box plots, estimate how many – or about what percent – of the 40,000 gum balls were green, how many were red, and how many were yellow. Explain your reasoning. Possible answer: Using the median values (2 G, 4 R, + 2 Y out of 8 gum balls) is one way to predict the proportion of each color in 40,000 gum balls.