How to Plot a Box Plot

When you want to compare sets of data, box and whiskers plots are very useful. The box shows the middle half of the data between the upper and lower quartiles with the median marked as a solid line across the box. The whiskers show the range of the data. However, if there are any *outliers* these are shown as points.

To identify an outlier

An outlier is a value that falls outside the overall pattern of the other observations. As it is not always easy to identify outliers by looking at the data we can use the following rule.

If a value is less than the lower limit, we consider this to be an outlier.

Lower limit = Lower quartile – 1.5 * interquartile range

If a value is greater than the upper limit, we consider this to be an outlier.

Upper limit = Upper quartile + 1.5 * interquartile range

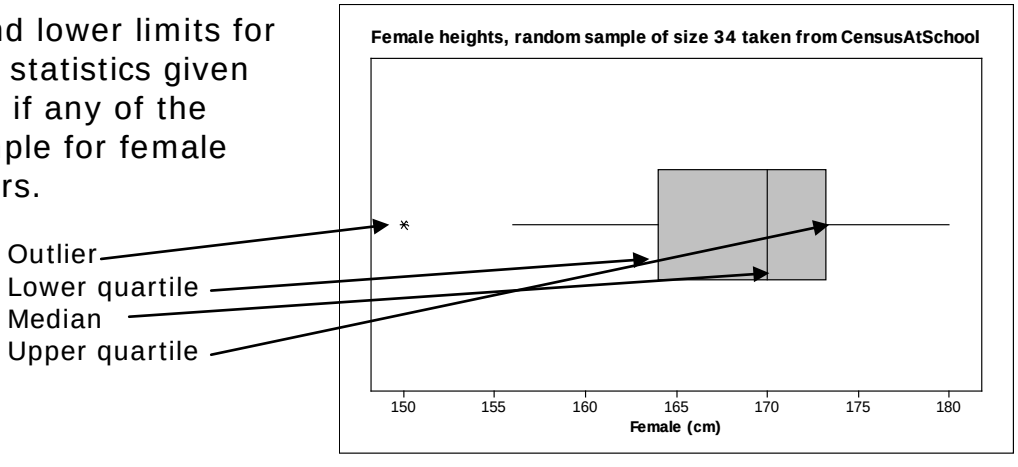
Below find two sets of data from *CensusAtSchool*. Both are from one class of year 11 pupils giving their height to the nearest centimetre.

Females	156	166	164	175	172	177	168	150	165	168	172	
	172	173	180	160	173	173	158	170	173	170	156	
	174	164	167	157	177	174	170	177	163	174	170	165
Males	173	193	176	183	174	180	187	185	180	179	178	165
	168	170	180	166	173	176	183	178	176	170	186	182
	174	175	173	179	190	174						

For females:

Lower Quartile	= 164 cm;
Median	= 170 cm;
Upper quartile	= 173.25 cm;
Range	= 150 – 180 = 30 cm.

Find the upper and lower limits for outliers using the statistics given above and decide if any of the values in the sample for female heights are outliers.



Now construct box plots for the heights of males and females given in the tables above presented so comparison of the distributions is easy. To help the box plot for females is shown above.

What are your conclusions?