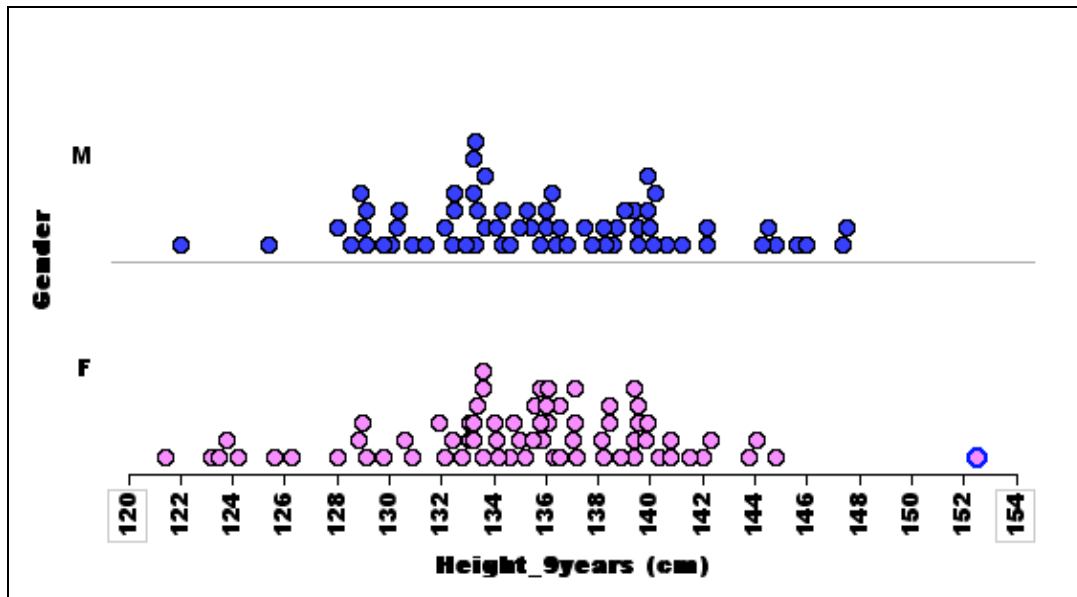


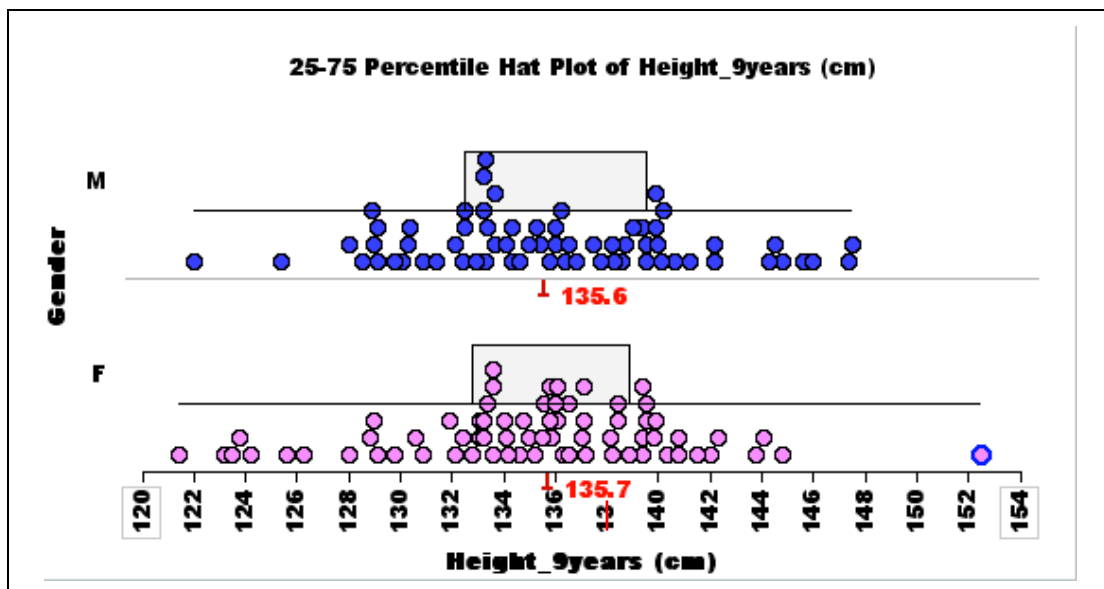
Examples of misjudgements of difference for two samples/populations

The graph below shows the heights of 136 students at the age of 9 years. Who is taller at 9 years, boys or girls?

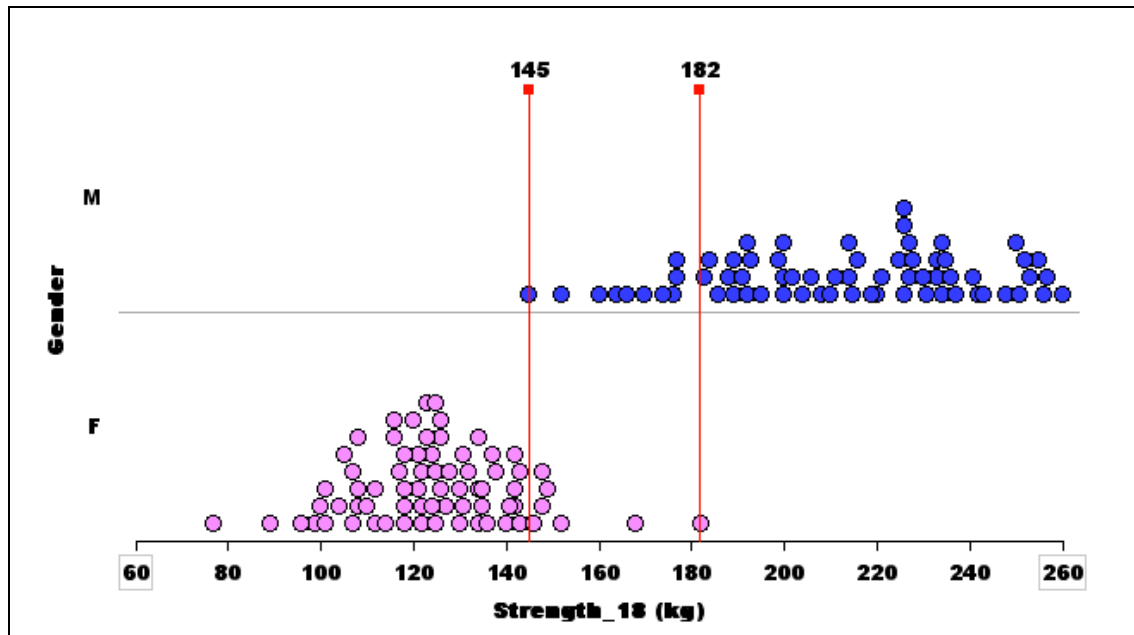


Student Answer: Girls, because there is a girl (152.5 cm) taller than all of the boys.

This response illustrates a student focussing on individual values rather than looking at the distribution. There is also a girl shorter than all of the boys but this value does not stand out as much. In fact looking at the same graph below we see that the medians are virtually the same, as are the hats showing the middle 50% of the data and the brims of the hats showing the highest and lowest 25% of the data (except for that one female value).



The graph below shows the strength (measured in kg) of the same 136 students at the age of 18. Who is stronger at 18 years, boys or girls?



Student Answer: You can't say because there are some boys (about 10) whose strengths overlap with some girls (about 9).

This response shows the reluctance to claim any difference in the two samples (and population of 18 year-olds) if there is any overlap in the data sets. In fact looking at the same data set below we see that the medians of the two data sets are very different (214.5 kg for Males and 124.5 kg for Females, a difference of 90 kg) and the entire hat plot (100% of data) for Males is greater than the crown and left brim of the hat plot (75% of the data) for females. Such a difference leads us to be quite confident that at 18 years generally Males are stronger than Females.

