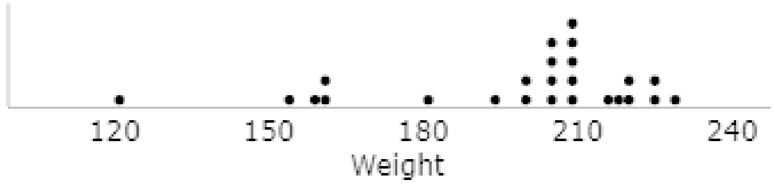
Introduction to Statistics
Activity 3.1b **SOLUTIONS**

Rowers' Weights



a) Based on this dotplot, estimate the values of the mean and median of the rowers' weights.
Mean (estimate): 198 Median (estimate): 205

b) Use technology to calculate the mean and median. You can find the data on the class Google Doc under Data: Rower. It's a spreadsheet. Make a copy of it to do your calculations ("=AVERAGE()" and "=MEDIAN()"). Record them in the first column of the following table. Why does it make sense that the mean and median differ in this direction? [*Hint:* Think about more than the shape of the distribution.]
Explanation: The graph is skewed left, this means the tail on the left drags the mean to the left more than the median. Therefore the mean should be smaller than the median.

	Whole Team	w/out Coxswain	changing max weight to 329	changing max weight to 2229
Mean	197.96	201	205	284.5
Median	205	207	207	207

c) Predict what will happen to the mean if you remove the coxswain (McElhenney, the outlier whose job is to call out rowing instructions and keep the rowers in sync) from this analysis. Also predict what will happen to the median. Explain.
Since the coxswain is an outlier, and very small, it is affecting the mean more than other values, dragging the mean to the left. Therefore, we should see the mean increase noticeably, while the median (being less influenced by outliers) should increase to a lesser degree.

d) Remove the coxswain McElhenney and recalculate the mean and median. Record the values in the second column of the table. Did these values change as you predicted?
Yes, the mean increased by about 3, while the median increase by only 2.

e) Now (putting the coxswain back in) consider what would happen if you increased the weight of the heaviest rower. How would this change the mean (if at all)? How would it affect the median? Explain.
Similar to how the mean changed without the lowest weight rower, in this situation the mean would increase noticeably, while the median would increase less so.

f) Increase Newlin's weight from 229 to 329. Recalculate the mean and median, and record them in the third column of the previous table. Did they change (or not) as you predicted?
Yes, the mean increased by about 5 while the median increased by only 2.

g) Suppose the typist's finger slipped when recording Newlin's weight and entered 2229. Predict how this error will affect the mean and median.
Since the median doesn't care about the value of the extreme data points, the median would remain the same. However, the mean is very much affected by these values, so it would increase significantly.

h) Increase Newlin's weight to 2229 and recalculate the mean and median. Record them in the last column of the table.

> A measure whose value is relatively unaffected by the presence of outliers in a distribution is said to be **resistant**.

i) Based on these calculations, is the mean and/or the median resistant? Explain why this conclusion makes sense; base your argument on how each measure is calculated.
The median is resistant. This is because the mean incorporates all of the magnitudes of the data, while the median throws out the biggest and smallest numbers.

j) In principle, is there any limit to how much the mean can increase or decrease simply by changing *one* of the values in the distribution?
No.