STAT200 Problem 3.2.32

a) Add a constant to data.

Let's assume we have a data set with three numbers. Call them a, b, c.

The *mean* is $\bar{x} = (a+b+c)/3$. Let's add a constant k to each, so now we have the data set $a+k, b+k, c+k$. The mean, by definition of mean, is $[(a+k)+(b+k)+(c+k)]/3$, so by simple algebra, $(a+b+c)/3+(3k)/3$, and this is $\bar{x}+k$; **If k is added to each datum, the mean is increased by k.**

For the *median*, let's arrange the data in order, so $a < b < c$ and b is the median. Add k to each and get $a+k < b+k < c+k$. **The median is not affected: It's the same datum with k added.** It's **likewise for the mode**, if one exists.

The *standard deviation* is more work. Part of it is done above with the arguments for $\bar{x}$. The critical part of the calculation, using the formula on page 93 of Triola, is

$$3(a^2+b^2+c^2)-(a+b+c)^2$$

Doing the k stuff on that gives

$$3[(a+k)^2+(b+k)^2+(c+k)^2]-([a+k]+[b+k]^2+[c+k])^2$$

Then there is a mess of algebra to do. The result is that *all of the terms with k cancel!*, giving

$$3(a^2+b^2+c^2)-(a+b+c)^2$$

That is **the standard deviation is unchanged**.

Though I have proven this for a data set of size 3, it's easy to see that it holds for a data set of any size. It's also easy to see by drawing pictures.

b) Multiply the data by a constant.

The same sort of arguments hold. In fact, the algebra is simpler.