

- 12 Calculate the standard deviation of the numbers 4, 6, 8 (the mean is 6).  
When you are done, use StatCrunch to check your work.

This one is for you to do. Follow the procedure in the text.

The mean is  $(4+6+8)/3=6$ :

Value	minus mean	squared
4	-2	4
6	0	0
8	2	4

Add up the last column to get 8, and divide by  $n-1$ . Here  $n=3$ , so divide by 2 to get a variance of 4. The SD is the square root of this, 2.

For StatCrunch, I typed the values 4, 6, 8 into a column and called it x. Then Stat and Summary Stats gave me:

### Summary statistics:

Column	n	Variance	Std. Dev.
x	3	4	2

so the variance of 4 and the SD of 2 were correct.

14 In the cereal data, we found that the IQR for potassium was bigger than the IQR for calories. Use StatCrunch to find out whether the same is true for SD. Does your result make sense?

You do this one. The IQR and SD values show a consistent picture.

Stats and Summary Stats again. Just asking for the IQR and SD (unselecting the others) gives you:

### Summary statistics:

Column	Std. Dev.	IQR
calories	19.48412	10
potassium	71.28681	80

Calories has less spread than potassium, measured both ways. You can't really compare SD and IQR for one variable (it doesn't make much sense to ask which one is bigger for, say, calories) but IQR being smaller for calories means that the SD is likely to be smaller for calories too.

- 15 Start with set of numbers 10, 11, 12, 13, 15, 16, 17.
- Using StatCrunch, find the mean, median, IQR and SD.
  - Change the 17 to something bigger. What happens to the mean and SD? What happens to the median and IQR?
  - Repeat the previous question, replacing 17 by something bigger still.
  - What does this tell you about when to use SD and when to use IQR?

In my table below, x is the original data, y has 17 replaced by 37, and z has it replaced by 107:

### Summary statistics:

Column	Mean	Median	Std. Dev.	IQR
x	13.428572	13	2.6367369	5
y	16.285715	13	9.375754	5
z	26.285715	13	35.654427	5

The median and IQR don't change at all, while the mean increases, and the SD goes through the roof.

This indicates the danger of relying on mean and SD: they can be heavily influenced by a single outlier.