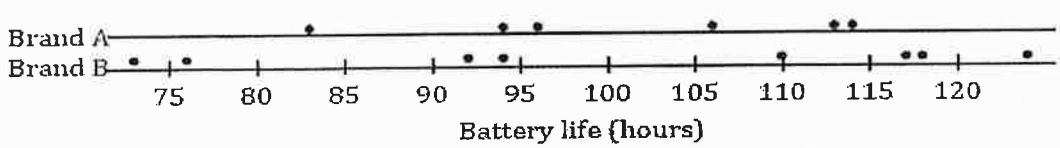
Unit 10: Statistics **Name:** KEY

Standard Deviation **Date:** March 13, 2017

Example 1:

A consumers' organization is planning a study of the various brands of batteries that are available. As part of its planning, it measures lifetime (how long a battery can be used before it must be replaced) for each of six batteries of Brand A and eight batteries of Brand B. Dot plots showing the battery lives for each brand are shown below.



1. Does one brand of battery tend to last longer, or are they roughly the same? What calculations could you do in order to compare battery life for the two brands?

Find the mean of each data set.

$\bar{X}_A = 101$ hr

$\bar{X}_B = 100.5$ hr

There is little difference in the average battery life.

2. Do the battery lives tend to differ more from battery to battery for Brand A or for Brand B?

Brand B looks to have a wider range than Brand A.

3. Would you prefer a battery brand that has battery lives that do not vary much from battery to battery? Why or why not?

Smaller variability is more consistent. So with roughly the same mean, Brand A would be a better choice.

The table below shows the lives (in hours) of the Brand B batteries and the deviations from the mean.

Life (Hours)	73	76	92	94	110	117	118	124
Deviation from the Mean	−27.5	−24.5	−8.5	−6.5	9.5	16.5	17.5	23.5

4. Calculate the deviations from the mean for Brand A, and write your answers in the appropriate places in the table.

- For any given value in a data set, the deviation from the mean is the value minus the mean. Written algebraically, this is $x - \bar{x}$.

Life (Hours)	83	94	96	106	113	114
Deviation from the Mean	−18	−7	−5	5	12	13

101

- The greater the variability (spread) of the distribution, the greater the deviations from the mean (ignoring the signs of the deviations).

5. Does Brand A or Brand B have greater variability?

Brand B

The **standard deviation** measures a typical deviation from the mean.

To calculate the standard deviation:
1. Find the mean of the data set.
2. Calculate the deviations from the mean.
3. Square the deviations from the mean.
4. Add up the squared deviations.

5a. Divide by 'n – 1',if you are working with a data from a sample, which is the most common case.
5b. Divide by 'n', if the data is from the entire population.
6. Take the square root.

The unit of the standard deviation is always the same as the unit of the original data set.

The larger the standard deviation, the greater the spread (variability) of the data set.

Use the calculator to find the standard deviation for Brand A and Brand B batteries. Since this is from a sample, we will divide by 'n-1'.

- Brand A $(-18)^2+(-7)^2+(-5)^2+(5)^2+(12)^2+(13)^2=736$

  $S_x = \sqrt{\frac{736}{6-1}} \approx 12.13$ hr

- Brand B $(-27.5)^2+(-24.5)^2+(-8.5)^2+(-6.5)^2+(9.5)^2+(16.5)^2+(17.5)^2+(23.5)^2=2692$

  $S_x = \sqrt{\frac{2692}{8-1}} \approx 19.61$ hr

6. Does Brand A or $\boxed{\text{Brand B}}$ have greater variability?

Try-It!

A set of eight men had heights (in inches) as shown below.

67.0 70.9 67.6 69.8 69.7 70.9 68.7 67.2

Deviations: 2 1.9 1.4 .8 .7 1.9 .3 1.8

1. Indicate the mean and standard deviation to the nearest hundredth.

Mean: 69

$S_x = \sqrt{\frac{17.64}{7}}$

Standard Deviation: 1.59

$(2)^2 + (1.9)^2 + (1.4)^2 + (.8)^2 + (.7)^2 + (1.9)^2 + (.3)^2 + (1.8)^2 = 17.64$

The heights (in inches) of 9 women were as shown below.

68.4 70.9 67.4 67.7 67.1 69.2 66.0 70.3 67.6

.1 2.6 .9 .6 1.2 .9 2.3 2 .7

2. Indicate the mean and standard deviation to the nearest hundredth.

Mean: 68.3

$S_x = \sqrt{\frac{19.97}{8}}$

Standard Deviation: 1.58

$(.1)^2 + (2.6)^2 + (.9)^2 + (.6)^2 + (1.2)^2 + (.9)^2 + (2.3)^2 + (2)^2 + (.7)^2 = \cancel{20.08}$ 19.97

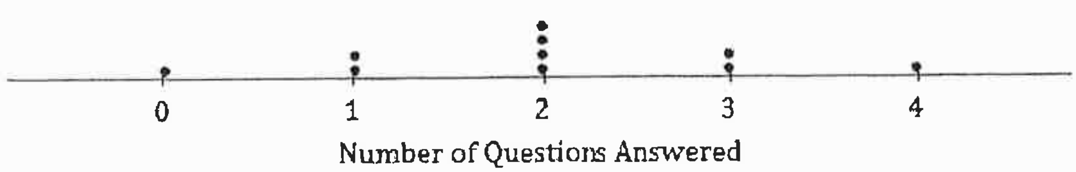
3. Write a few sentences to compare the typical heights of men and women and their variability.

In general, the men are typically taller than the women, but not by much. Their variabilities are virtually the same.

Example 2:

Ten people attended a talk at a conference. At the end of the talk, the attendees were given a questionnaire that consisted of four questions. The questions were optional, so it was possible that some attendees might answer none of the questions while others might answer 1, 2, 3, or all 4 of the questions (so the possible numbers of questions answered are 0, 1, 2, 3, and 4).

Suppose that the numbers of questions answered by each of the ten people were as shown in the dot plot below.



a. Find the mean and the standard deviation of the data set.

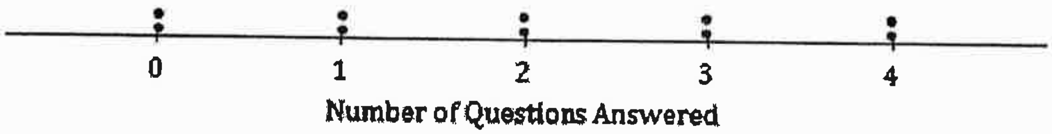
Deviations: 2, 1, 1, 0, 0, 0, 0, 1, 1, 2

Mean: 2

Standard Deviation: 1.15

$S_x = \sqrt{\frac{12}{9}}$

b. Suppose the dot plot looked like this:



Find the mean and the standard deviation of this distribution.

Deviations: 2, 2, 1, 1, 0, 0, 1, 1, 2, 2

Mean: 2

Standard Deviation: 1.49

$S_x = \sqrt{\frac{20}{9}}$

Remember that the size of the standard deviation is related to the size of the deviations from the mean.

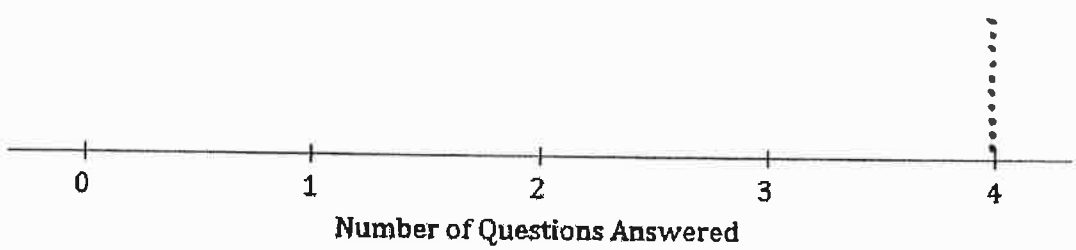
c. Explain why the standard deviation of part b is greater than the standard deviation in part a.

More data points are farther away from the mean in part b.

Try-It!

Suppose that every person answers all four questions on the questionnaire.

a) What would the dot plot look like?



b) What is the mean number of questions answered? (You should be able to answer without doing any calculations!)

4

c) What is the standard deviation? (Again, don't do any calculations!)

0 (no numbers differ from the mean)