Math 2201 **Date:**____________________

5.3 Standard Deviation

Standard Deviation
We looked at **range** as a measure of **dispersion**, or spread of a data set. The problem with using range is that it is only a measure of how spread out the extreme values, the smallest and largest, are. It doesn't provide any information about the variation within the other data values. Thus, we will now look at a different measure of dispersion called **standard deviation**.

Standard deviation: a measure of the dispersion or scatter of data values in relation to the mean; a low standard deviation indicates that most data values are close to the mean, and a high standard deviation indicates that most data values are scattered farther from the mean. The symbol used to represent standard deviation is the Greek letter σ, sigma.

Review of Mean or Average
The mean, $\bar{x}$, can be expressed using symbols:

"x bar"

sum of all data points

$$\bar{x} = \frac{\Sigma x}{n}$$

mean

total number of data points

Thus, the mean is obtained by summing up all the data points and dividing by the total number of points.

Example 1:
Calculate the mean of the following data set: 3, 5, 6, 6, 7, 9, 9, 10, 12

$$\bar{x} = \frac{\Sigma x}{n} = \frac{3+5+6+6+7+9+9+10+12}{9}$$

$$= \frac{67}{9}$$

$$= 7.4$$

Calculating Standard Deviation

The standard deviation, σ, can be found using the following formula:

Step 1: Subtract the mean from each data point. Square what you get and sum up all the results

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

Step 3: take the square root

Step 2: divide by the total number of data points

Example 2:

Recall this example from section 5.1. The following table shows the test scores that Tim and Mary receive in Math 2201.

	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5
Tim	60	65	70	75	80
Mary	60	69	70	71	80

(A) Calculate the standard deviation for both Tim's and Mary's test scores.

Tim

$$\bar{x} = \frac{350}{5} = 70$$

Mary

$$\bar{x} = \frac{350}{5} = 70$$

Step 1: Subtract the mean from each data point. Square what you get and sum up all the results.

The following table will help organize the data and calculations for step 1.

Tim:

Data Point (x)	Data Point – Mean $(x - \bar{x})$	(Data Point – Mean)^2 $(x - \bar{x})^2$
60	$60 - 70 = -10$	$(-10)^2 = 100$
65	$65 - 70 = -5$	$(-5)^2 = 25$
70	$70 - 70 = 0$	$(0)^2 = 0$
75	$75 - 70 = 5$	$(5)^2 = 25$
80	$80 - 70 = 10$	$(10)^2 = 100$
		$\Sigma(x - \bar{x})^2 = 250$

Step 2: Divide sum by the total number of data points.

$$\frac{250}{5} = 50$$

Step 3: Take the square root of the total, divided by the sum.

$$\sigma = \sqrt{50}$$
$$= 7.1$$

Now repeat the process for **Mary**:

Data Point (x)	Data Point – Mean $(x-\bar{x})$	(Data Point – Mean)^2 $(x-\bar{x})^2$
60	$60-70=-10$	$(-10)^2=100$
69	$69-70=-1$	$(-1)^2=1$
70	$70-70=0$	$(0)^2=0$
71	$71-70=1$	$(1)^2=1$
80	$80-70=10$	$(10)^2=100$
		$\Sigma(x-\bar{x})^2=$ 202

$$\sigma=\sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}=\sqrt{\frac{202}{5}}=6.4$$

(B) Whose marks are more dispersed? What does this mean in terms of standard deviation?

Tim's. The σ is larger.

(C) If the data is clustered around the mean, what does this mean about the value of the standard deviation?

Smaller σ.

(D) Who was more consistent over the five unit tests?

Mary. A smaller σ means her test scores were closer together.

Example 3:

(A) Is it possible for a data set to have a standard deviation of 0? Use an example.

Yes. When all data values are equal.
for example, 65, 65, 65, 65, 65.

(B) Can the standard deviation ever be negative? Explain why or why not.

No, because you square the difference of the data values subtracted by the mean.

(C) If 5 were added to each number in a set of data, what effect would it have on the mean?

Original data	add 5
72	77
78	83
74	84
$\bar{x} = \frac{234}{3}$	$\bar{x} = \frac{249}{3}$
$\bar{x} = 78$	$\bar{x} = 83$

When you add a number to all data points, the mean will increase by the same number.

(D) If 5 were added to each number in a set of data, what effect would it have on the standard deviation?

Adding a value to all points does not change the spread, or the σ of a data set.

(E) What effect would multiplying each number by -3 have on the standard deviation?

Ignoring the negative, the standard deviation will get 3 times larger.

Whenever we multiply all the numbers in a data set by a common value, the standard deviation gets multiplied by that value.

Example 4:
Two high schools kept a record of the number of students sent to the office for smoking on school grounds. Over a 5 day period, the following results were obtained:

School A: 4, 8, 13, 2, 5
School B: 9, 6, 11, 10, 8

Determine the standard deviation for each school. Which school has the greatest variation? Why?

School A

$$\bar{x} = \frac{4+8+13+2+5}{5}$$

$$\bar{x} = 6.4$$

$4 - 6.4 = (-2.4)^2 = 5.76$

$8 - 6.4 = (1.6)^2 = 2.56$

$13 - 6.4 = (6.6)^2 = 43.56$

$2 - 6.4 = (-4.4)^2 = 19.36$

$5 - 6.4 = (-1.4)^2 = 1.96$

73.2

$$\sigma = \sqrt{\frac{73.2}{5}} = 3.8$$

School B

$$\bar{x} = \frac{9+6+11+10+8}{5}$$

$$\bar{x} = 8.8$$

$9 - 8.8 = (0.2)^2 = 0.4$

$6 - 8.8 = (-2.8)^2 = 7.84$

$11 - 8.8 = (2.2)^2 = 4.84$

$10 - 8.8 = (1.2)^2 = 1.44$

$8 - 8.8 = (-0.8)^2 = 0.64$

15.16

$$\sigma = \sqrt{\frac{15.16}{5}} = 1.7$$

Example 5:
Sports Illustrated is doing a story on the variation of player heights on NBA basketball teams. The heights, in cm, of the players on the starting line ups for two basketball teams are given in the table below.

Lakers	195	195	210	182	205
Celtics	193	208	195	182	180

The team with the most variation in height will be selected for the cover of Sports Illustrated magazine. Which team will appear on the cover.

Lakers

$$\bar{x} = \frac{195+195+210+182+205}{5}$$

$$\bar{x} = \frac{987}{5}$$

$$\bar{x} = 197.4$$

$195-197.4=(-2.4)^2=5.76$

$195-197.4=(-2.4)^2=5.76$

$210-197.4=(12.6)^2=158.76$

$182-197.4=(-15.4)^2=237.16$

$205-197.4=(7.6)^2=57.76$

465.2

$$\sigma = \sqrt{\frac{465.2}{5}} = 9.6$$

Celtics

$$\bar{x} = \frac{193+208+195+182+180}{5}$$

$$\bar{x} = \frac{985}{5}$$

$$\bar{x} = 191.6$$

$193-191.6=(1.4)^2=1.96$

$208-191.6=(16.4)^2=268.96$

$195-191.6=(3.4)^2=11.56$

$182-191.6=(-9.6)^2=92.16$

$180-191.6=(-11.6)^2=134.56$

509.20

$$\sigma = \sqrt{\frac{509.20}{5}} = 10.1$$

Celtics have more variation so they will appear on cover.

Example 6:
Two obedience schools for dogs monitor the number of trials required for twenty puppies to learn the **sit** *and* **stay** command.

True Companion Dog School	
Number of Trials	Number of Puppies
7 ×	1 = 7
8 ×	2 = 16
9 ×	5 = 45
10 ×	4 = 40
11 ×	4 = 44
12 ×	4 = 48

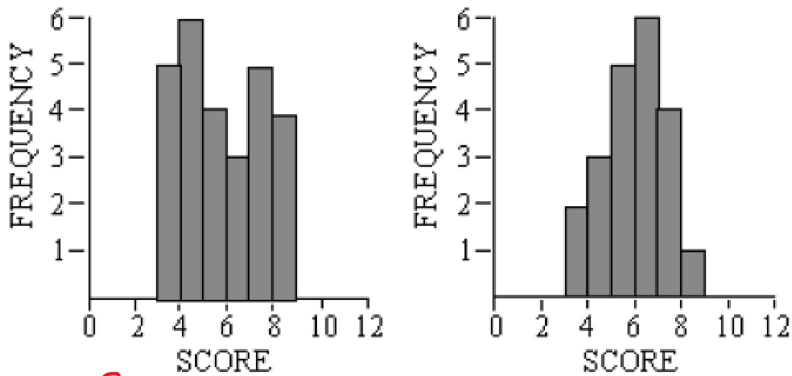
Dog Top School	
Number of Trials	Number of Puppies
7 ×	4 = 28
8 ×	3 = 24
9	2 = 18
10	3 = 30
11	4 = 44
12	4 = 48

(A) Determine the mean and standard deviation of the number of trials required to learn the *sit and stay* command.

(B) Which school is more consistent at teaching puppies to learn the *sit and stay* command. Justify your reasoning.

Example 7:

(A) Indicate whether one of the graphs has a larger standard deviation than the other or if the two graphs have the same standard deviation. What approaches did you use to decide which graph had the larger standard deviation?



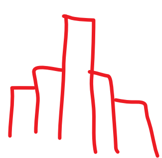
larger σ

smaller σ

- data is more clustered around the center.

(B) Identify the characteristics of the graphs that make the standard deviation larger or smaller. What features of a histogram seem to have no bearing on the standard deviation? What features do appear to affect the standard deviation?

Small σ



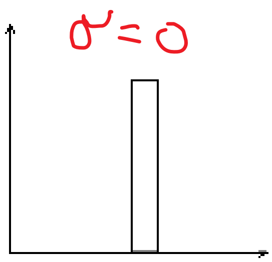
high in the middle, lower on the ends.

Example 8:
For each of the following histograms, match the following standard deviations: 0, 1, 3, 10

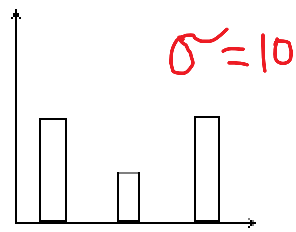
(A)

(B)

(C)



(D)



Textbook Questions: page 261, 262 #1a,c, 4, 6