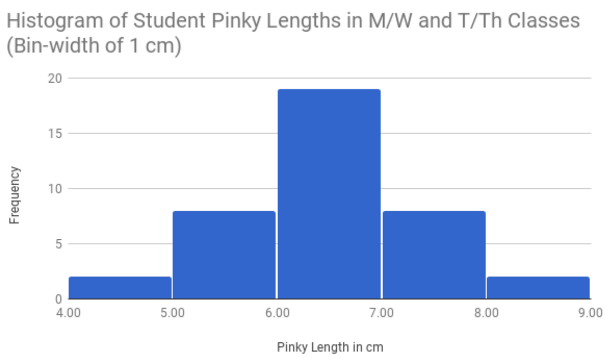
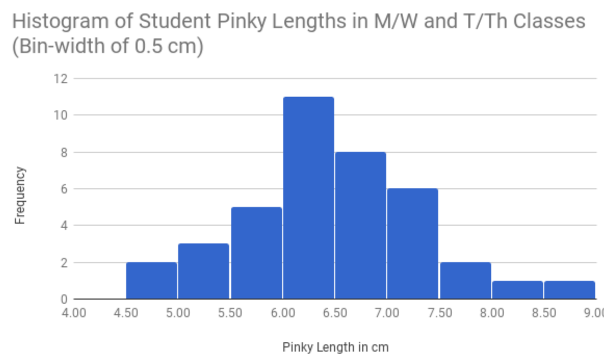
6B, C: Standard Deviation and the Normal Distribution

Group Activity

1. Here are two different histograms for the same data - the pinky length measurements for two Math 105 classes. Discuss the differences between the two graphs with your group. Which histogram do you think represents the data better?





2. The pinky length data is on the next page. Calculate the following statistics for our sample. Add units to all your statistics.

Mean (round to one decimal place): **254.5/41 = 6.2 cm**

Median: **6 cm**

Mode (remember to look at the histograms and the data): **6 cm, 6.5 cm or 6.25 cm**

5-number summary: **(4.5, 5.5, 6, 6.75, 8.5) cm**

Range: 8.5 – 4.5 = **4 cm**

IQR: Q3 – Q1 = **6.75 – 5.5 = 1.25 cm**

Range/4: **4/4 = 1 cm**

Standard deviation (calculate on the next page and round to two decimal places): **0.87 cm**

3. Once you have all your statistics, write a few sentences to describe the shape, center and spread and any unusual features of this data.

The data is unimodal and approximately symmetric. The mean is 6.2 cm and the standard deviation is 0.87 cm. There do not appear to be any outliers or unusual features.

Or

The data is unimodal and slightly skewed to the right. The median is 6 cm and the IQR is 1.25 cm. There do not appear to be any outliers or unusual features.

Check your statistics with Cara and/or the key before you move on to page 3.

Pinky Length Data. Calculate the sample standard deviation, rounded to two decimal places:

Pinky Length (cm)	Deviation from the mean	Squared deviation
4.5 − 6.2	-1.7	2.89
4.5	-1.7	2.89
5 − 6.2	-1.2	1.44
5	-1.2	1.44
5	-1.2	1.44
5.5	-0.7	0.49
5.5	-0.7	0.49
5.5	-0.7	0.49
5.5	-0.7	0.49
5.5	-0.7	0.49
5.5	-0.7	0.49
6 − 6.2	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6	-0.2	0.04
6.5 − 6.2	0.3	0.09
6.5	0.3	0.09
6.5	0.3	0.09
6.5	0.3	0.09
6.5	0.3	0.09
6.5	0.3	0.09
6.5	0.3	0.09
6.5	0.3	0.09
7 − 6.2	0.8	0.64
7	0.8	0.64
7	0.8	0.64
7	0.8	0.64
7	0.8	0.64
7	0.8	0.64
7.5 − 6.2	1.3	1.69
7.5	1.3	1.69
8 − 6.2	1.8	3.24
8.5 − 6.2	2.3	5.29
Sum of the Squared Deviations		29.99

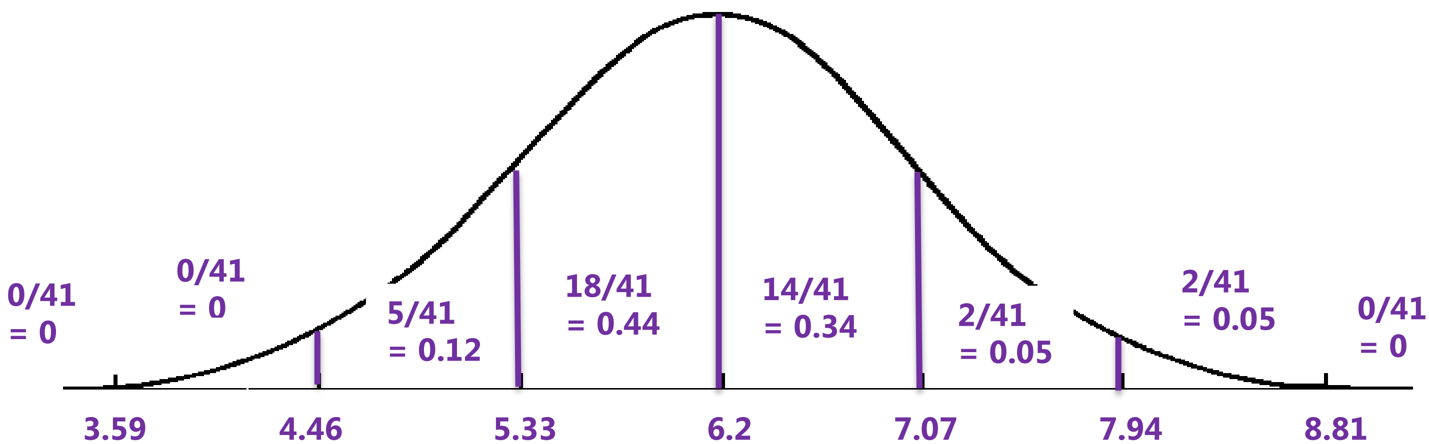
$$s = \sqrt{\frac{\sum (x - mean)^2}{n-1}}$$

$$= \sqrt{\frac{29.99}{41-1}}$$

$$\approx 0.87$$

We want to see if a Normal curve fits our data.

4. Label the sample mean and three standard deviations on each side.



5. Divide the Normal curve into subsections, like we did with the Empirical rule. Then calculate the percentage of students whose pinky length falls within each region. Write the percentage in each region.

6. Compare your percentages with the Empirical rule. How do they compare?

The percentages are close, but not the same. There are more students just to the left of the mean.

Shapes (Distributions) of Data

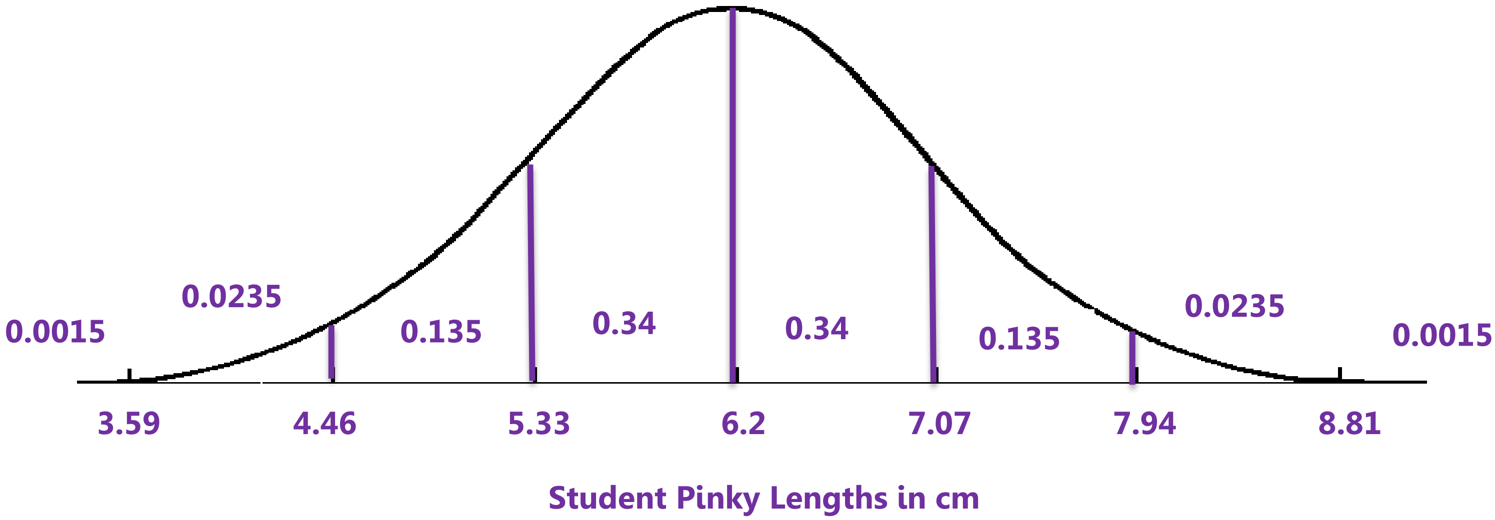
Measurements like heights and weights of people tend to follow a Normal distribution. Measurements of machined parts or test scores on very complex tests, like the SAT's, also follow a Normal distribution.

Variables that have a limit on one side tend to be skewed away from that side. For example, the number of tattoos people have tends to be skewed to the right.

7. Match the description of the data with its most likely shape:

a. The amount of liquid in soda cans (ml)	i. Normally distributed
b. Lengths of newborn babies (in)	ii. Skewed to the right
c. Household incomes in the U.S. (dollars)	iii. Skewed to the left

8. **Theoretically**, pinky measurements should follow a Normal distribution if we had a larger sample size (we may need to separate the data by gender identity). Let's use our Normal model to represent PCC students. Label the mean and standard deviation and the probabilities from the Empirical rule.



9. Using the **Empirical rule**, find the percentage of students who have pinky lengths

a. Less than 4.46 cm **0.0015 + 0.0235 = 0.025**

b. Between 6.2 and 7.94 cm **0.34 + 0.135 = 0.475**

c. Greater than 7.07 cm **0.135 + 0.0235 + 0.0015 = 0.16**

Z-Scores

10. Calculate the Z-score for a person with a pinky length of 4 cm. What does that mean?

$Z = \frac{4-6.2}{0.87} \approx$ **-2.5 standard deviations. This person is 2.5 standard deviations below the mean, which is rare.**

11. Calculate the Z-score for a person with a pinky length of 7 cm. What does that mean?

$Z = \frac{7-6.2}{0.87} \approx$ **0.9 standard deviations. This person is almost one standard deviation above the mean, which is the average deviation.**

Percentiles

12. In what percentile is a student with a pinky length of 5.33 cm?

0.135 + 0.0235 + 0.0015 = 0.16. This person is in the 16th percentile, which means their pinky length is longer than 16% of the population.

13. What pinky length is the 84th percentile?

The 84th percentile corresponds to one standard deviation above the mean, so the 84th percentile is 7.07 cm.