Name: _____ Date: _____

So far in this unit we have explored:

Measures of central tendancy - Mean, Median, Mode.
Frequency Diagrams - line plot, histogram, polygon

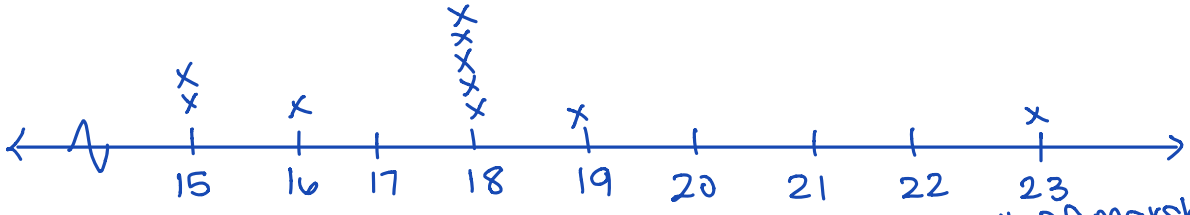
Today we are going to look at another way of describing dispersion, what is the spread in the data.

Let's consider the following situation:

Juliana is going into business selling bags of chocolate covered marshmallows. She contacts two different companies that are offering to do the packaging for her. She receives 10 samples from each company and counts the number of marshmallows in each bag:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Company #1 | Wrap 'em Up | 15 | 15 | 16 | 16 | 17 | 18 | 18 | 19 | 21 | 23 |
| Company #2 | Bags of Fun | 15 | 15 | 16 | 18 | 18 | 18 | 18 | 18 | 19 | 23 |

1. Make a line plot of the number of marshmallows in a bag for the two companies




2. Calculate the three measures of central tendency and the range for each company

| Company | Mean | Median | Mode | Range |
|---|---|---|---|---|
| Wrap 'em Up | 17.8 | 17.5 | 15, 16, 18 | 15 - 23 |
| Bags of Fun | 17.8 | 18 | 18 | 15 - 23 |

Based on your measures calculated above, which company is more consistent? Did any of our summary statistics capture this?

The second - mode

Assignment p. 233 #1-9, 13 Quiz Next Day!

Standard Deviation

a measure of how consistent the data set is.

$\sigma$ (sigma - lowercase)

Let's calculate the standard deviation for the marshmallow packages.

| Wrap 'em up $\bar{x} = 17.8$ | | | Bags of Fun $\bar{x} = 17.8$ | | |
|---|---|---|---|---|---|
| Observation | Distance from mean | distance$^2$ | Observation | Distance from mean | distance$^2$ |
| 15 | +2.8 | 7.84 | 15 | +2.8 | 7.84 |
| 15 | +2.8 | 7.84 | 15 | +2.8 | 7.84 |
| 16 | +1.8 | 3.24 | 16 | +1.8 | 3.24 |
| 16 | +1.8 | 3.24 | 18 | -0.2 | 0.04 |
| 17 | +0.8 | 0.64 | 18 | -0.2 | 0.04 |
| 18 | -0.2 | 0.04 | 18 | -0.2 | 0.04 |
| 18 | -0.2 | 0.04 | 18 | -0.2 | 0.04 |
| 19 | -1.2 | 1.44 | 18 | -0.2 | 0.04 |
| 21 | -3.2 | 10.24 | 19 | -1.2 | 1.44 |
| 23 | -5.2 | 27.04 | 23 | -5.2 | 27.04 |
| Total | ~~ | 61.6 | Total | ~~ | 47.6 |

Which data was more spread out?

bigger ⇒ more spread out.

How does this relate to standard deviation?

kind of take the average

Assignment p. 233 #1-9, 13 Quiz Next Day!

Standard Deviation: $\sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}}$

| | Wrap 'em Up ($x$) | $(x - \bar{x})^2$ | Bags of Fun ($x$) | $(x - \bar{x})^2$ |
|---|---|---|---|---|
| | 15 | 7.84 | 15 | 7.84 |
| | 15 | 7.84 | 15 | 7.84 |
| | 16 | 3.24 | 16 | 3.24 |
| | 16 | 3.24 | 18 | 0.04 |
| | 17 | 0.64 | 18 | 0.04 |
| | 18 | 0.04 | 18 | 0.04 |
| | 18 | 0.04 | 18 | 0.04 |
| | 19 | 1.44 | 18 | 0.04 |
| | 21 | 10.24 | 19 | 1.44 |
| | 23 | 27.04 | 23 | 27.04 |
| $\sum x$ | | 61.6 | | 47.6 |
| mean → $\bar{x}$ | 17.8 | --- | 17.8 | --- |
| $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$ | --- | $\sqrt{\frac{61.6}{10}}$ | --- | $\sqrt{\frac{47.6}{10}}$ |
| | | $= 0.78$ | | $= 0.69$ |