3.2: Measures of Spread:

Consider the home values in these two neighborhoods:

Orchard hill: $125k, $132k, $144k, $156k, $452k

River district: $178k, $189k, $204k, $213k, $227k

- We prefer to use the mean as a measure of center since it considers every number …
- Find the mean for each neighborhood.

  Orchard hill:

  River district:
- By itself the mean leaves out important information

Range

Definition: Range

Range = max – min

One way to measure the spread is to calculate the range. The range is the difference between the largest and smallest values in the data set;

- The range is strongly affected by outliers
- Find the range for the two neighborhoods

  Orchard hill:

  River district:

Average Deviation

- Each data value has an associated deviation from the mean,

Definition: Deviation from the Mean

$$\boldsymbol{x - \bar{x}}$$

Example The deviation of the first Orchard Hill house is
$x - \bar{x} = \$125{,}000 - \$201{,}800 = -\$76{,}800$

- *Note:* A deviation is positive if it falls above the mean and negative if it falls below the mean

- *The Average Deviation is the sum of the absolute value of deviations divided by the number of items.*

Definition: Average Deviation

$$\boldsymbol{\frac{\Sigma|x - \bar{x}|}{n}}$$

- Orchard hill: $125k, $132k, $144k, $156k, $452k;

  mean = $201,800

  River district: $178k, $189k, $204k, $213k, $227k

  mean = $202,200

- Find the average deviation from the mean for each neighborhood

Orchard Hill

Home Value	$x - \bar{x}$	$\lvert x - \bar{x} \rvert$
$125,000	-$76,800	$76,800
$132,000	-$69,800	$69,800
$144,000	-$57,800	$57,800
$156,000	-$45,800	$45,800
$452,000	$250,200	$250,200

Average (mean) deviation = $100,080

A large deviation from the mean!

River District

Home Value	$x - \bar{x}$	$\lvert x - \bar{x} \rvert$
$178,000	-$24,200	$24,200
$189,000	-$13,200	$13,200
$204,000	$1,800	$1,800
$213,000	$10,800	$10,800
$227,000	$24,800	$24,800

Average (mean) deviation = $14,960

A small deviation from the mean!

Standard Deviation

- The Standard Deviation Gives a measure of variation by summarizing the *deviations* of each observation from the mean and calculating an *adjusted average* of these deviations
- Standard Deviation for populations is represented by the lower case Greek letter σ (“sigma”)

Formula:

Standard Deviation:

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

$$\sigma = \sqrt{\frac{\sum (x - \overline{x})^2}{n}}$$

Steps for finding Standard Deviation:

1. Find the mean
2. Find the deviation of each value from the mean
3. Square the deviations
4. Add the squared deviations
5. Divide the sum by *n*
6. Find the square root of that sum

Example: Find the **Stand**ard Deviation

Metabolic rates of 7 men (cal./24hr.) :1792 1666 1362 1614 1460 1867 1439

First Find Mean: $\bar{x} = 1600$

Now find the deviations and squared deviations

Observations	Deviations $(x - \bar{x})$	Squared deviations $(x - \bar{x})^2$
1792	1792–1600 = 192	$(192)^2$ = 36,864
1666	1666 –1600 = 66	$(66)^2$ = 4,356
1362	1362 –1600 = -238	$(-238)^2$ = 56,644
1614	1614 –1600 = 14	$(14)^2$ = 196
1460	1460 –1600 = -140	$(-140)^2$ = 19,600
1867	1867 –1600 = 267	$(267)^2$ = 71,289
1439	1439 –1600 = -161	$(-161)^2$ = 25,921
	sum = 0	sum = 214,870

Now apply the formula

$$\sigma^2 = \frac{214{,}870}{7} \approx 30{,}695.71$$

$$\sigma = \sqrt{30{,}695.71} = 175.2\ calories$$

You Try It! Find the standard deviation of the data collected in class.

Find the mean:

Now find the deviations and squared deviations

Observations	**Deviations** $(x - \bar{x})$	**Squared deviations** $(x - \bar{x})^2$
	sum = $\Sigma(x - \bar{x})$ =	sum = $\Sigma(x - \bar{x})^2$=

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

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

Properties of the Standard Deviation

- σ measures the spread of the data
- σ = 0 only when all observations have the same value, otherwise σ > 0. As the spread of the data increases, σ gets larger.
- σ has the same units of measurement as the original observations. σ is typically rounded to one more place than the original data.
- σ is not resistant. Outliers (skewed data) can greatly increase σ.

Standard Deviation: 2 types

- Population Standard Deviation:

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

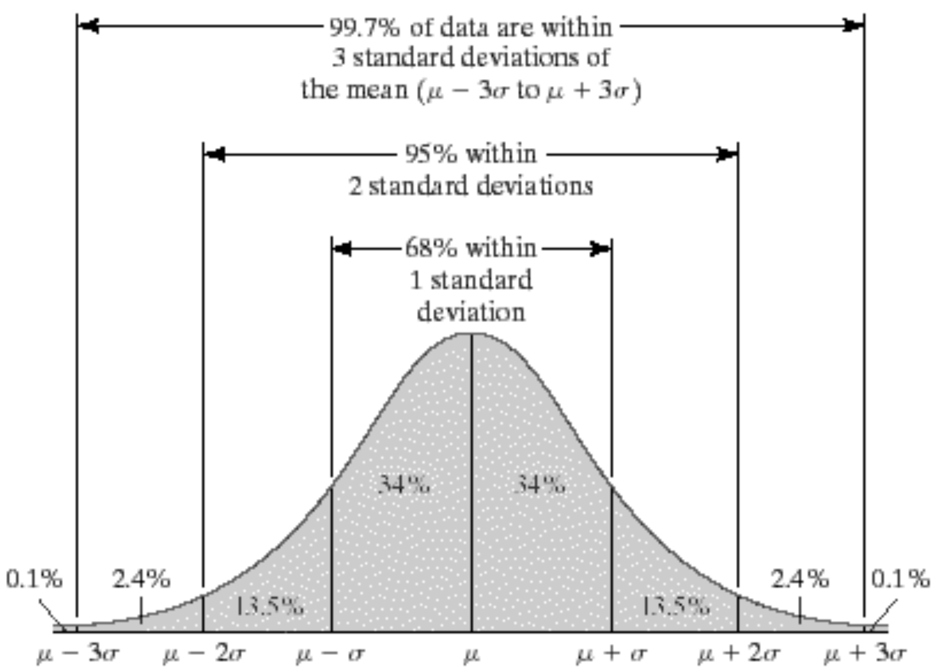
- Sample Standard Deviation

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

Empirical Rule

If a population is "normal" (meaning the graph is symmetric), then

- 68% of the population falls within 1 standard deviation of the mean
- 95% of the population falls within 2 standard deviation of the mean
- 99.7% of the population falls within 3 standard deviation of the mean



Note: the Greek Letter μ (pronounced "me-you" or "mew") represents the mean of the population.