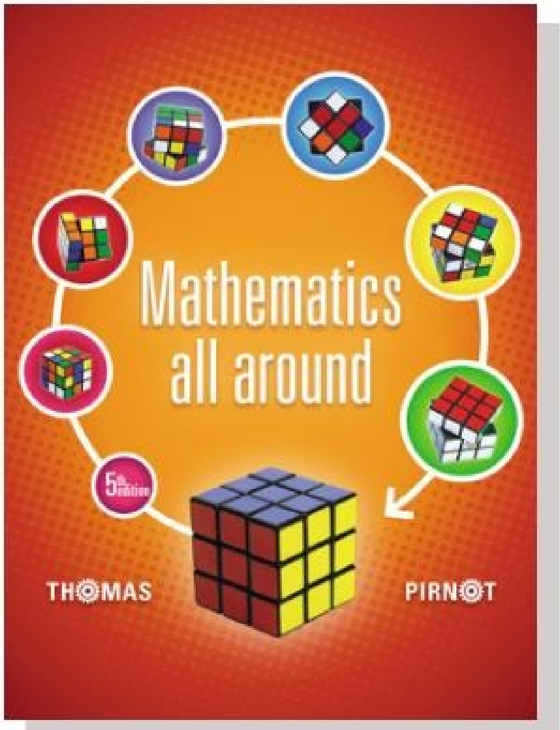
14

# Descriptive Statistics




## *What a Data Set Tells Us*

# 14.3 Measures of Dispersion

- Compute the range of a data set.
- Understand how the standard deviation measures the spread of a distribution.
- Use the coefficient of variation to compare the standard deviations of different distributions.

# The Range of a Data Set

**DEFINITION** The **range** of a data set is the difference between the largest and smallest data values in the set.

**KEY POINT**

The range is a crude measure of the spread of a data set.

# The Range of a Data Set

- Example: Find the range of the heights of the people listed in the accompanying table.

| Person | Height | Height in Inches |
|---|---|---|
| Leonid Stadynk (World's Tallest Person) | 8 feet, 6 inches | 102 inches |
| LaDainian Tomlinson | 5 feet, 10 inches | 70 inches |
| Madge Bester (World's Shortest Person) | 2 feet, 2 inches | 26 inches |
| LeBron James | 6 feet, 8 inches | 80 inches |

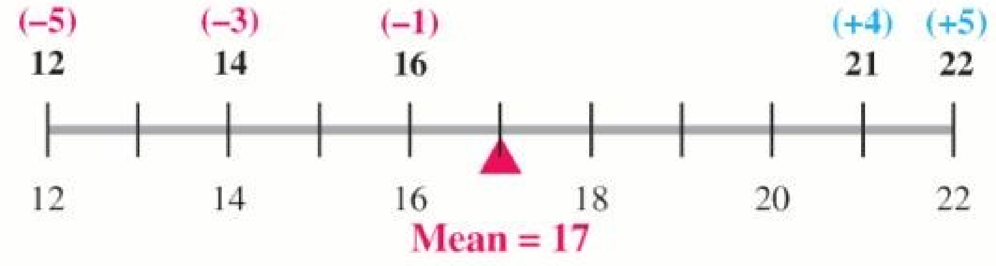
- Solution:

range = largest data value − smallest data value = 102 − 26 = 76 inches = 6 feet, 4 inches

# Standard Deviation

**DEFINITION** If $x$ is a data value in a set whose mean is $\bar{x}$, then $x - \bar{x}$ is called $x$'s *deviation from the mean.*

| Data Value, $x$ | Deviation from Mean, $x - \bar{x}$ |
|---|---|
| 16 | −1 |
| 14 | −3 |
| 12 | −5 |
| 21 | 4 |
| 22 | 5 |
| Total | 0 |







PEARSON

# Standard Deviation

**DEFINITION** We denote the **standard deviation** of a *sample* of $n$ data values by $s$, which is defined as follows:

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

| Data Value, $x$ | Deviation from Mean, $x-\bar{x}$ | Deviation Squared, $(x-\bar{x})^2$ |
|---|---|---|
| 16 | −1 | 1 |
| 14 | −3 | 9 |
| 12 | −5 | 25 |
| 21 | 4 | 16 |
| 22 | 5 | 25 |

$$\frac{1+9+25+16+25}{\underbrace{5-1}_{n-1}}=\frac{76}{4}=19 \longrightarrow \sqrt{19}\approx 4.36$$




PEARSON

# Standard Deviation

**COMPUTING THE STANDARD DEVIATION** To compute the standard deviation for a sample consisting of $n$ data values, do the following:

1. Compute the mean of the data set; call it $\bar{x}$.
2. Find $(x - \bar{x})^2$ for each score $x$ in the data set.
3. Add the squares found in step 2 and divide this sum by $n - 1$; that is, find

$$\frac{\sum (x - \bar{x})^2}{n - 1},$$

which is called the **variance**.

4. Compute the square root of the number found in step 3.

# Standard Deviation

- Example: A company has hired six interns. After 4 months, their work records show the following number of work days missed for each worker:

0, 2, 1, 4, 2, 3

Find the standard deviation of this data set.

- Solution:

**Mean:** $\dfrac{0+2+1+4+2+3}{6}=\dfrac{12}{6}=2$

(continued on next slide)

# Standard Deviation

We calculate the squares of the deviations of the data values from the mean.

| Number of Days Missed | Deviation from Mean | Square of Deviation from the Mean |
|---|---|---|
| 0 | −2 | 4 |
| 2 | 0 | 0 |
| 1 | −1 | 1 |
| 4 | 2 | 4 |
| 2 | 0 | 0 |
| 3 | 1 | 1 |
| | | $\Sigma(x-2)^2 = 10$ |

**Standard Deviation:** $$s = \sqrt{\frac{\sum(x-2)^2}{\underbrace{6-1}_{n-1}}} = \sqrt{\frac{10}{5}} \approx 1.41$$

# Standard Deviation

**FORMULA FOR COMPUTING THE SAMPLE STANDARD DEVIATION FOR A FREQUENCY DISTRIBUTION** We calculate the standard deviation, $s$, of a sample that is given as a frequency distribution as follows:

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

where $\bar{x}$ is the mean of the distribution, $f$ is the frequency of data value $x$, and $n = \Sigma f$, the number of data values in the distribution. 10

# Standard Deviation

- Example: The following are the closing prices for a stock for the past 20 trading sessions:

37, 39, 39, 40, 40, 38, 38, 39, 40, 41,
41, 39, 41, 42, 42, 44, 39, 40, 40, 41

What is the standard deviation for this data set?

- Solution:

**Mean:** $\frac{800}{20} = 40$ (sum of the closing prices is 800)

(continued on next slide)

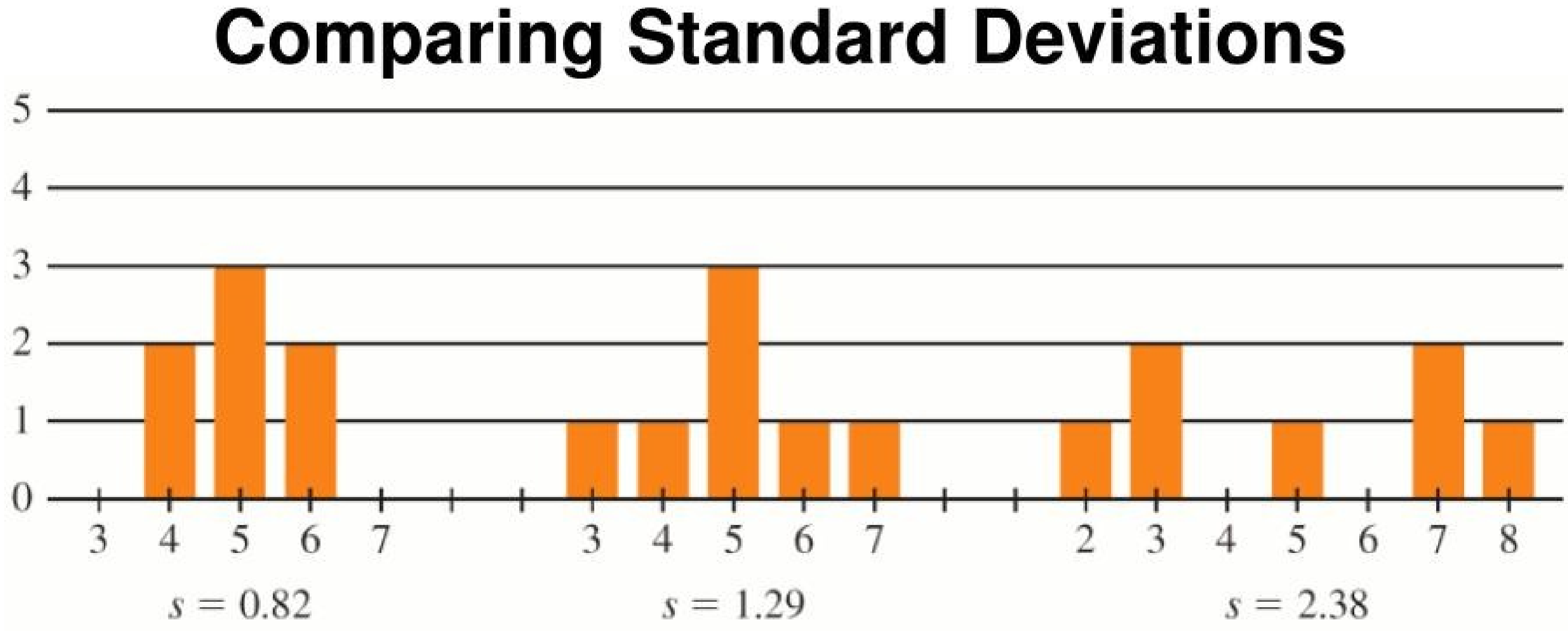
# Standard Deviation

We create a table with values that will facilitate computing the standard deviation.

| Closing Price, $x$ | Frequency, $f$ | Product, $x \cdot f$ | Deviation, $(x - 40)$ | Deviation Squared, $(x - 40)^2$ | Product, $(x - 40)^2 \cdot f$ |
|---|---|---|---|---|---|
| 37 | 1 | 37 | −3 | 9 | 9 |
| 38 | 2 | 76 | −2 | 4 | 8 |
| 39 | 5 | 195 | −1 | 1 | 5 |
| 40 | 5 | 200 | 0 | 0 | 0 |
| 41 | 4 | 164 | 1 | 1 | 4 |
| 42 | 2 | 84 | 2 | 4 | 8 |
| 44 | 1 | 44 | 4 | 16 | 16 |
| | $\Sigma f = 20$ | $\Sigma(x \cdot f) = 800$ | | | $\Sigma(x - 40)^2 \cdot f = 50$ |

**Standard Deviation:** $s = \sqrt{\dfrac{\sum(x-40)^2 \cdot f}{n-1}} = \sqrt{\dfrac{50}{19}} \approx 1.62$

# Standard Deviation




All three distributions have a mean and median of 5; however, as the spread of the distribution increases, so does the standard deviation.