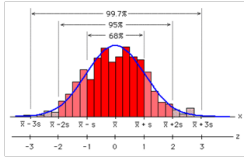


Exploring Data with Graphs and Numerical Summaries | 2

2.4 Describe the Spread of Quantitative Data

Aug 8-9:16 AM

Learning Objectives

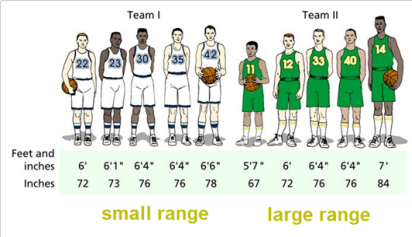


- Calculate the **range**
- Calculate the **standard deviation**
- Know the properties of the standard deviation
- Know how to interpret the magnitude of s: **The Empirical Rule**

Aug 8-9:16 AM

Range

Range = max – min
The range is strongly affected by outliers.



	Team I					Team II				
Feet and inches	6'	6'1"	6'4"	6'4"	6'6"	5'7"	6'	6'4"	6'4"	7'
Inches	72	73	76	76	78	67	72	76	76	84

small range large range

Aug 8-9:16 AM

Deviation

- Each data value has an associated **deviation** from the mean, $x - \bar{x}$
- A deviation is **positive** if it falls **above** the mean and **negative** if it falls **below** the mean
- The sum of the deviations is always **zero**

Aug 8-9:16 AM

Standard Deviation

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:

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

- Find mean
- Find each deviation
- Square deviations
- Sum squared deviations
- Divide sum by $n-1$
- Take square root

Aug 8-9:16 AM

Standard Deviation

Metabolic rates of 7 men (calories/24 hours)

Observations	Deviations	Squared deviations
1792		
1666		
1362		
1614		
1460		
1867		
1439		

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

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

Aug 8-9:16 AM

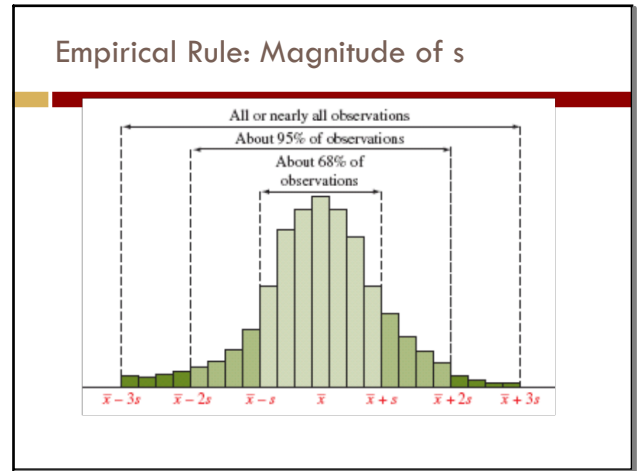
Properties of Sample Standard Deviation

S root sample mean square standard deviation

observation \bar{x}

- Measures spread of data
- Only zero when all observations are same; otherwise, $s > 0$
- As the spread increases, s gets larger
- Same units as observations
- Not resistant
- Strong skewness or outliers greatly increase s

Aug 8-9:16 AM



Aug 8-9:16 AM

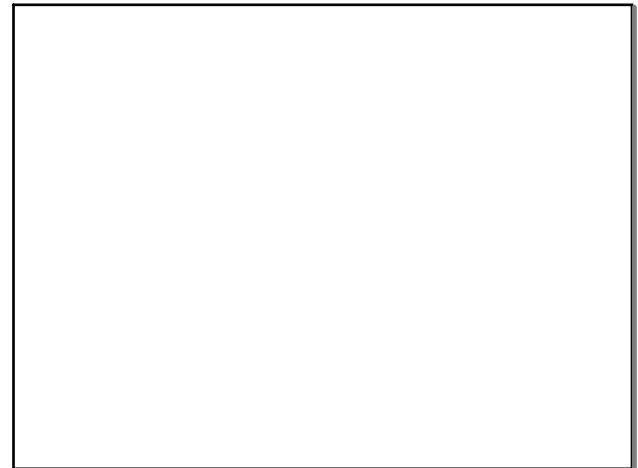
2.47 **Sick leave:** A company decides to investigate the amount of sick leave taken by its employees. A sample of eight employees yields the following numbers of days of sick leave taken in the past year:

TRY

0 0 4 0 0 0 6 0

- Find and interpret the range.
- Find and interpret the standard deviation s .
- Suppose the 6 was incorrectly recorded and is supposed to be 60. Redo (a)-(b) with the correct data and describe the effect of this outlier.

Aug 8-9:16 AM



Aug 14-3:15 PM