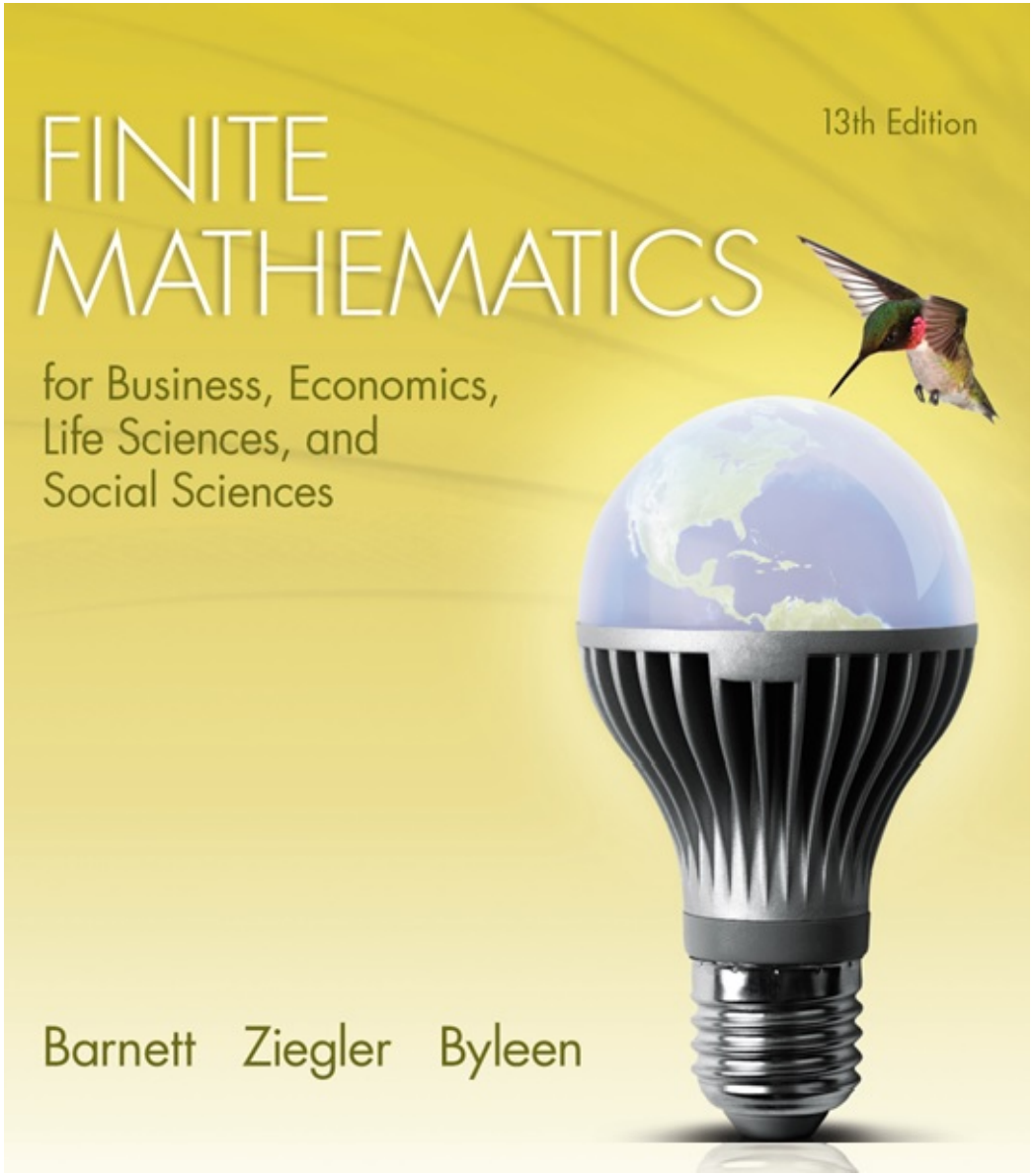


Chapter 11

Data Descriptions and Probability Distributions

Section 2

Measures of Central Tendency

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

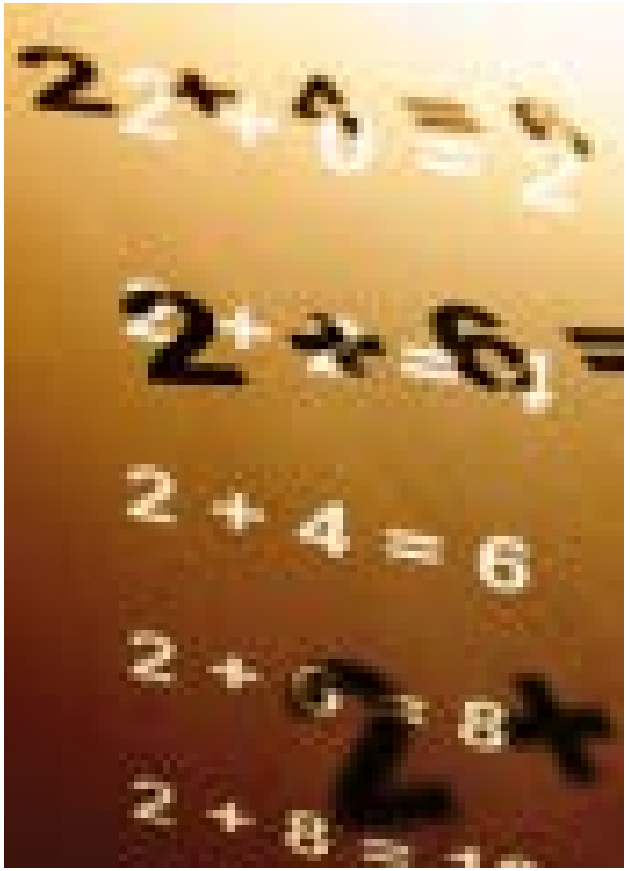
Learning Objectives for Section 11.2 Measures of Central Tendency

- The student will be able to calculate the mean of a distribution.
- The student will be able to identify the median of a distribution.
- The student will be able to identify the mode of a distribution.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

11.2 Measures of Central Tendency

In this section, we will study three measures of central tendency: the **mean**, the **median** and the **mode**. Each of these values determines the “center” or middle of a set of data.



Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean

- Most common
- Definition: the **mean** is the sum of the numbers divided by number of numbers
- Notation:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- **Example:** The salaries of 5 employees (in thousands) are 14, 17, 21, 18, 15. Find the mean.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean

- Most common
- Definition: the **mean** is the sum of the numbers divided by number of numbers
- Notation: $$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$
- **Example:** The salaries of 5 employees (in thousands) are 14, 17, 21, 18, 15. Find the mean.
- **Solution:** Sum = (14 + 17 + 21 + 18 + 15) = 85

  Number of numbers: 5

  Divide 85 by 5 = 17. The average salary is $17,000

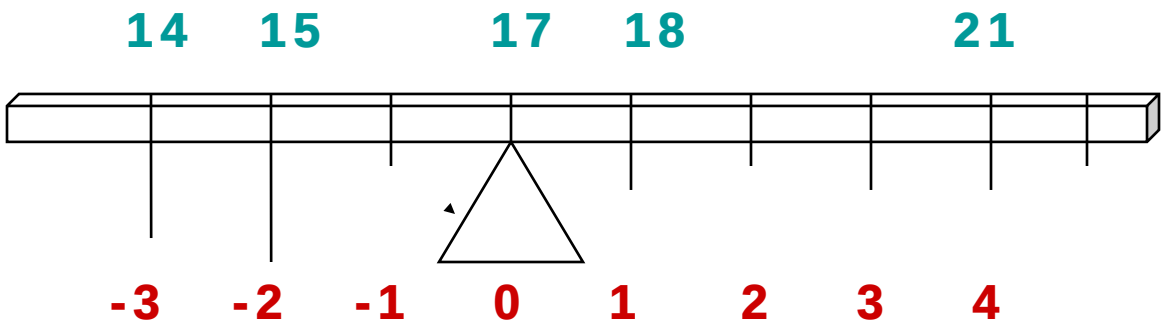
Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean as Center of Gravity

- We will represent each data value on a “teeter-totter” in the following slide. The teeter-totter serves as a number line.
- You can think of each point's deviation from the mean as the influence the point exerts on the tilt of the teeter-totter. Positive values push down on the right side; negative values push down on the left side. The farther a point is from the fulcrum, the more influence it has.
- Note that the mean deviation of the scores from the mean is always zero. That is why the teeter-totter is in balance when the fulcrum is at the mean. This makes the mean the center of gravity for all the data points.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean as Center of Gravity (continued)



Data balances at 17. Sum of the deviations from mean equals zero.

$$(-3 + -2 + 0 + 1 + 4 = 0) .$$

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean for Grouped Data

- To find the mean for grouped data, find the midpoint of each class by adding the lower class limit to the upper class limit and dividing by 2. For example, the midpoint of the class [0,7) is (0 + 7)/2 = 3.5.
- Multiply the midpoint value by the frequency of the class.
- Find the sum of the products of x and f.
- Divide this sum by the total frequency.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean for Grouped Data Example

Find the mean of the distribution given in the following table:

class	**midpoint *x***	**frequency *f***
[0,7)	3.5	0
[7,14)	10.5	2
[14,21)	17.5	10
[21,28)	24.5	21
[28,35)	31.5	23
[35,42)	38.5	14
[42,49)	45.5	5

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mean for Grouped Data Example (continued)

class	midpoint x	frequency f	$x \cdot f$
[0,7)	3.5	0	0
[7,14)	10.5	2	21
[14,21)	17.5	10	175
[21,28)	24.5	21	514.5
[28,35)	31.5	23	724.5
[35,42)	38.5	14	539
[42,49)	45.5	5	227.5
total		75	2201.5
mean			**29.3533**

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median

- The mean is not always the best measure of central tendency especially when the data has one or more “outliers” (numbers which are unusually large or unusually small and not representative of the data as a whole).
- Definition: The **median** of a data set is the number that divides the bottom 50% of data from top 50% of data.
- To obtain the median:
  - Arrange the data in ascending order
  - With an odd number of data points, it is the middle measurement
  - With an even number of data points, it is the mean to middle two measurements.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median
Solution

- **Example:** Find the median of the data set 14, 17, 21, 18, 15
- **Solution:**

  1. Arrange data in order: 14, 15, 17, 18, 21
  2. Determine the location of the median: (5+1)/2 = 3.
  3. Count from the left until you reach the number in the third position (17) .
  4. The value of the median is 17.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median
Example

- This example illustrates the case when the number of observations is an even number. The value of the median in this case will **not** be one of the original pieces of data.
- **Example:** Determine the median of the data set 14, 15, 17, 19, 23, 25

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median
Solution

- This example illustrates the case when the number of observations is an even number. The value of the median in this case will **not** be one of the original pieces of data.
- **Example:** Determine the median of the data set
  14, 15, 17, 19, 23, 25
- **Solution:**
  1. Data is already arranged in order.
  2. Determine the position of the median: (6+1)/2 = 3.5
  3. Take the average of the 3rd and 4th data value.
  4. (17+19)/2 = 18. Thus, the median is 18.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Which is better? Median or Mean?

The yearly salaries of 5 employees of a small company are 19, 23, 25, 26, and 57 (in thousands).

1. Find the mean salary
2. Find the median salary
3. Which measure is more appropriate and why?

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Which is better? Median or Mean?

The yearly salaries of 5 employees of a small company are 19, 23, 25, 26, and 57 (in thousands).

1. Find the mean salary 30
2. Find the median salary 25
3. Which measure is more appropriate and why?

Answer: The median is better since the mean is skewed (affected) by the outlier 57.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Properties of the Mean

1. Mean takes into account all values
2. Mean is sensitive to extreme values (outliers)
3. Mean is called a **non-resistant measure** of central tendency since it is affected by extreme values.
4. Population mean = mean of all values of the population
5. Sample mean = mean of sample data
6. Mean of a representative sample tends to best estimate the mean of population (for repeated sampling)

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Properties of the Median

1. Not sensitive to extreme values; **resistant** measure of central tendency
2. Takes into account only the middle value of a data set or the average of the two middle values.
3. Should be used for data sets that have outliers, such as personal income, or prices of homes in a city.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Mode

- Definition: the **mode** is the most frequently occurring value in a data set.
- To obtain mode, find the frequency of occurrence of each value and then note the value that has the greatest frequency.
- If the greatest frequency is 1, then the data set has no mode.
- If two values occur with the same greatest frequency, then we say the data set is **bimodal**.

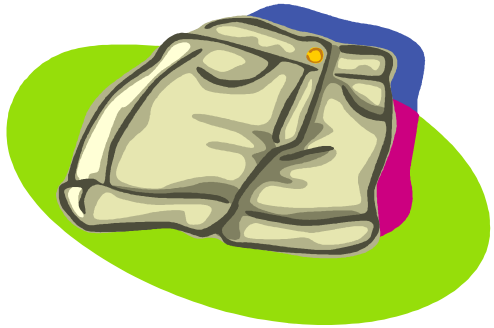
Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Example of Mode Solution

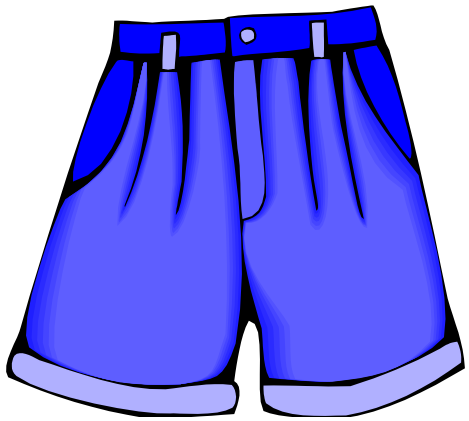
- **Example 1:** Find the mode of the following data set:

  45, 47, 68, 70, 72, 72, 73, 75, 98, 100

- **Answer:** The mode is 72.



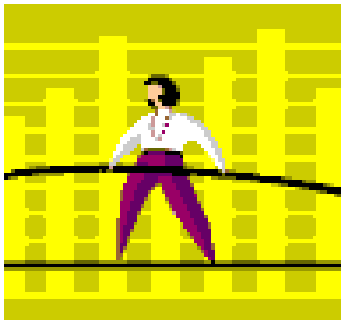
- **Example 2:** Shorts are classified as small, medium, large, and extra large. A store has on hand 12 small, 15 medium, 17 large and 8 extra large pairs of shorts. Find the mode.
- **Solution:** The mode is *Large*. This is the modal class (the class with the greatest frequency. It would not make sense to find the mean or median for nominal data.

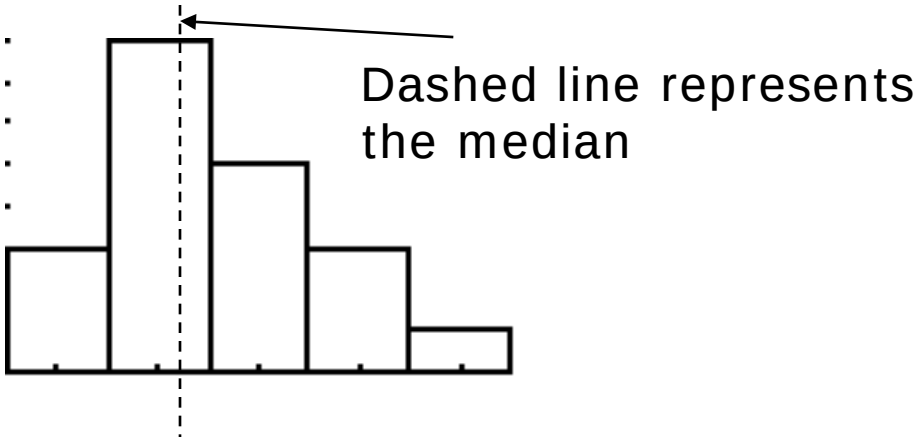


Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median for Grouped Data

The **median for grouped data** with no classes of frequency 0 is the number such that the histogram has the same area to the left of the median as to the right of the median.





Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median for Grouped Data Example

Example: Compute the median for the grouped data in the table.

Class Interval	Frequency
3.5-4.5	3
4.5-5.5	1
5.5-6.5	2
6.5-7.5	4
7.5-8.5	3
8.5-9.5	2

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median for Grouped Data Example

Solution: First, add a cumulative frequency column.

Class Interval	Frequency	Cumulative Frequency
3.5-4.5	3	3
4.5-5.5	1	4
5.5-6.5	2	6
6.5-7.5	4	10
7.5-8.5	3	13
8.5-9.5	2	15

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median for Grouped Data Example (continued)

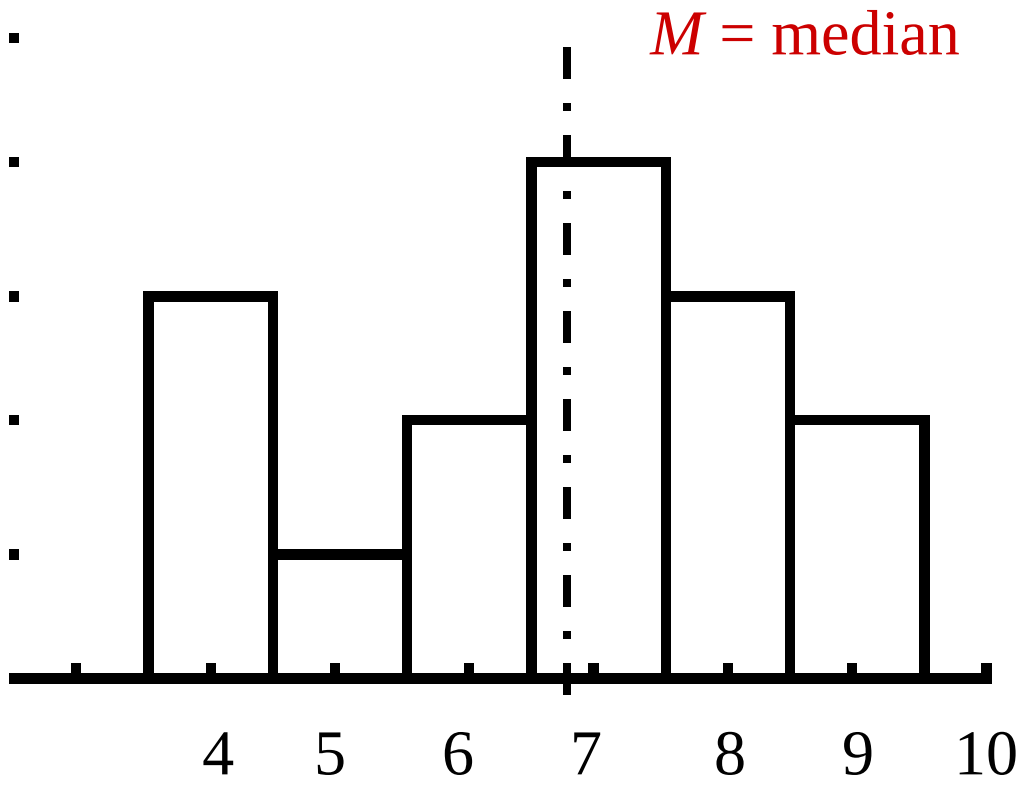
The total area is 15. Half the area is 15/2 = 7.5.
The half-way point is somewhere in the interval [6.5,7.5], since the total area goes from 6 to 10 in that interval.
If the median is M, the area to the left of the median M must be 7.5:

$$(1)(3) + (1)(1) + (1)(2) + (M\text{-}6.5)(4) = 7.5$$
$$3 + 1 + 2 + 4M - 26 = 7.5$$
$$4M - 20 = 7.5$$
$$4M = 27.5$$
$$M = 6.875$$

The median for the grouped data in the table is 6.875.

Copyright © 2015, 2011, and 2008 Pearson Education, Inc.

Median for Grouped Data Example (continued)



Copyright © 2015, 2011, and 2008 Pearson Education, Inc.