

4.1 Measures of Location

Arithmetic Mean

Suppose there are n observations in a data set, consisting of the observations x_1, x_2, \dots, x_n ; then the **arithmetic mean** is defined to be

$$\frac{1}{n}(x_1 + x_2 + \dots + x_n).$$

DEFINITION

Calculate the sample mean of the following sample data values: 4, 10, 7, 15.

$$\frac{4 + 10 + 7 + 15}{4} = \frac{36}{4} = \boxed{9}$$

Weighted Mean

The weighted mean of a data set with values $x_1, x_2, x_3, \dots, x_n$ is given by

$$\bar{x} = \frac{w_1x_1 + w_2x_2 + \dots + w_nx_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum (w_ix_i)}{\sum w_i}$$

where w_i is the weight of observation x_i .

FORMULA

Meghan is a freshman in college and she received the following grades for her first semester.

Meghan's Grades		
Course	Grade	Credit Hours
Psychology 101	B <i>3</i>	3
Probability and Statistics	A <i>4</i>	4
Anatomy I	C <i>2</i>	5
English 101	A <i>4</i>	3

$$\begin{aligned} 3(3) &= 9 \\ 4(4) &= 16 \\ 2(5) &= 10 \\ 4(3) &= 12 \end{aligned} + = \boxed{47}$$

A grade of A is worth 4 points on a 4-point scale. A grade of B is worth 3 points and a grade of C is worth 2 points.

$$\frac{47}{15} = 3.13$$

- Calculate Meghan's GPA using the credit hours as weights. Round to two decimal places.
- If Meghan's goal was to have a GPA of 3.4, what grade did she need to make in the Anatomy I class to reach her goal?

Meghan is a freshman in college and she received the following grades for her first semester.

Meghan's Grades		
Course	Grade	Credit Hours
Psychology 101	B 3	3
Probability and Statistics	A 4	4
Anatomy I	X	5
English 101	A 4	3

9

$$\frac{37 + 5x}{15} = 3.4$$

16

5x

$$37 + 5x = 51$$

-37

-37

A grade of A is worth 4 points on a 4-point scale. A grade of B is worth 3 points and a grade of C is worth 2 points.

If Meghan's goal was to have a GPA of 3.4, what grade did she need to make in the Anatomy I class to reach her goal?

B

2

$$\frac{5x}{5} = \frac{14}{5}$$

x ≈ 3

Average Monthly Balances for a Bank Customer (January through March)

Month	Average Monthly Balance	
January	(31)	\$1885.67
February	(28)	\$1312.92
March	(31)	\$2001.53

$= 58455,77$
 $= 36761,76$
 $= 62047,43$

90

157264.96

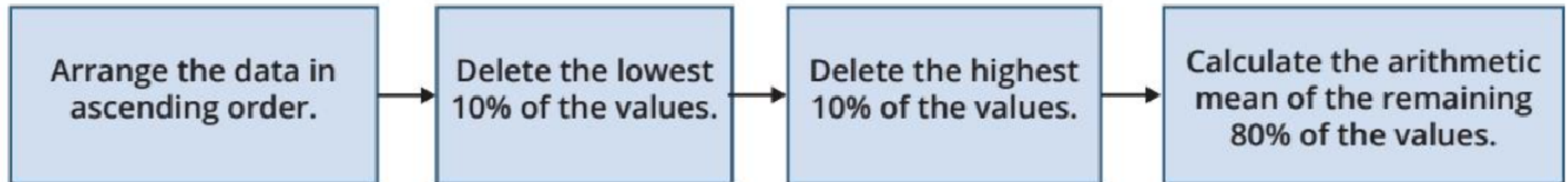
$\frac{157264,96}{90} = \$1,747.39$

Trimmed Mean

The **trimmed mean** is a modification of the arithmetic mean which ignores an equal percentage of the highest and lowest data values in calculating the mean.

DEFINITION

Finding the 10% Trimmed Mean



Consider the following data taken from a poll on how many text messages a person sent in a day.

~~16~~ 18 20 21 23 23 24 32 36 ~~42~~

mean = 25.5

Find the 10% trimmed mean.

$$\frac{197}{8} = 24.6$$

Consider the same data set, except the last data value is replaced with an outlier.

~~16~~ 18 20 21 23 23 24 32 36 ~~490~~

mean = 70.3

Find the 10% trimmed mean.

$$\frac{197}{8} = 24.6$$

Outliers and Resistant Measures

An **outlier** is a data value that is extremely different from other measurements in the data set. Statistical measures which are not affected by outliers are said to be **resistant**.

DEFINITION

Daily Production

Day	1	2	3	4	5	6	7	8	9	10
Units	100	104	117	20	20	111	105	106	115	101
Day	11	12	13	14	15	16	17	18	19	20
Units	101	102	115	116	113	103	104	119	118	108

- What level of measurement does the data possess?
- Compute the mean, 10% trimmed mean and the 20% trimmed mean.
- Considering the worker's illness, which measure computed in part **b.** best describes the production capability of the machine? Discuss.

$$10\% \quad \frac{1721}{16} = 108$$

$$20\% \quad \frac{1287}{12} = 107$$

Median

The **median** of a set of observations is the measure of center that is the middle value of the data when it is arranged in ascending order. The same number of data values lie on either side of the median.

DEFINITION

To determine the median of a set of data, we use the following steps.

Finding the Median of a Data Set

1. Arrange the data in ascending order.
2. Determine the number of values in the data.
3. Find the data value in the middle of the data set.
4. If the number of data values is odd, then the median is the data value that is exactly in the middle of the data set.
5. If the number of data values is even, then the median is the mean of the two middle observations in the data set.

PROCEDURE

Consider the following ten test scores from a student taking a high school calculus class.

65, 98, 76, 83, 94, 79, 88, 72, 90, 85

Find the median.

~~65~~, ~~72~~, ~~76~~, ~~79~~, 83, 85, ~~88~~, ~~90~~, ~~94~~, ~~98~~

$$\frac{83 + 85}{2} = 84$$

Mode

The **mode** of a data set is the most frequently occurring value.

DEFINITION

Find the mode of the following data regarding the number of power outages reported over a period of eleven days.

0, 1, 4, 3, 9, 8, 10, 0, 1, 3, 0

Bell-Shaped Distribution

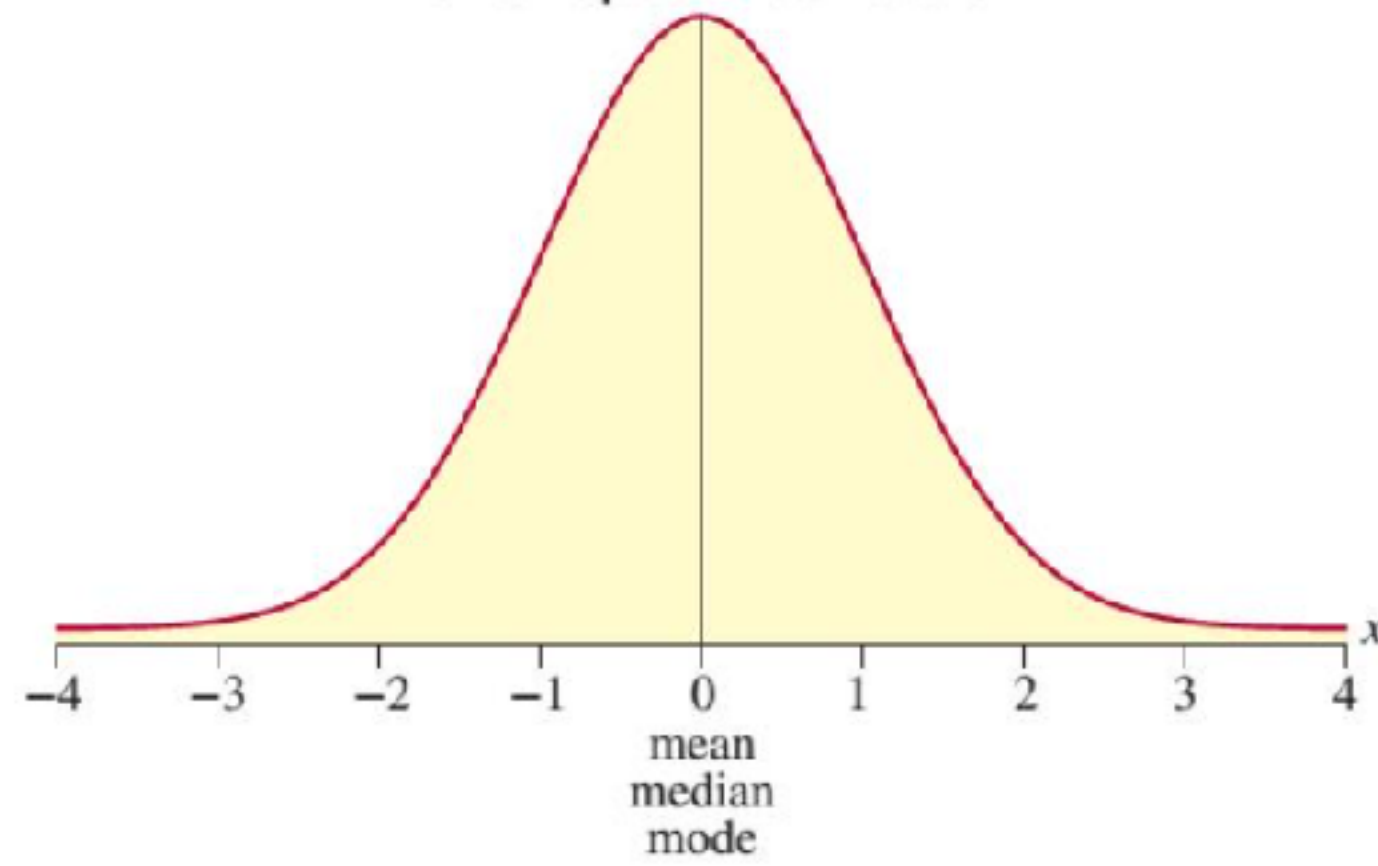
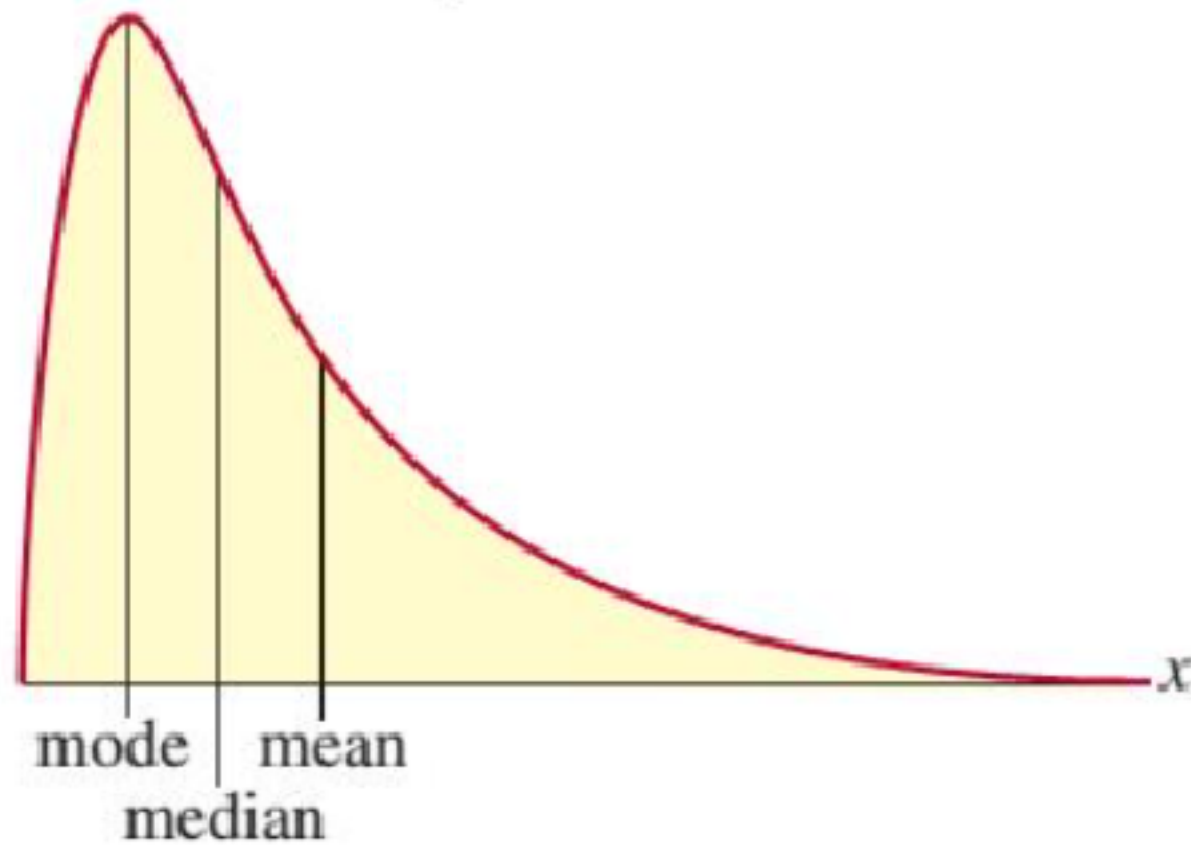


Figure 4.1.3

Positively Skewed Curve



Negatively Skewed Curve

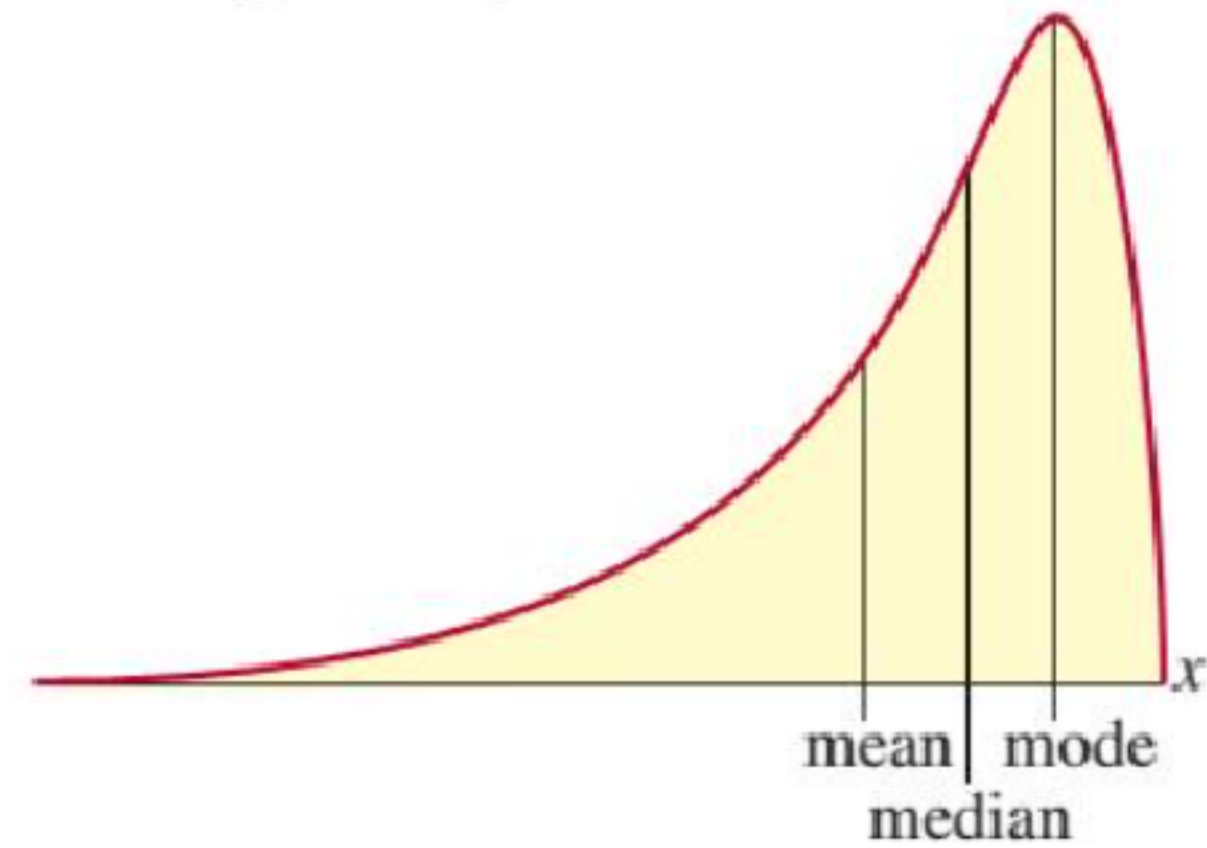


Table 4.1.3 – Applicable Level of Measurement

	Qualitative		Quantitative	
	Nominal	Ordinal	Interval	Ratio
Mean			✓	✓
Median		✓	✓	✓
Mode	✓	✓	✓	✓
Trimmed Mean			✓	✓

- a.** A doctor is interested in analyzing the increase in systolic blood pressure caused by a certain antibiotic. Would the doctor be more interested in studying the mean, median or mode of the systolic blood pressures?
- b.** A car manufacturer is trying to decide in what colors it should offer its new sports coupe. In analyzing the preferred colors of other sports coupes, would the manufacturer be more interested in the mean, median, or mode of the colors?
- c.** A manufacturer of chocolate bars is interested in knowing how people rate its chocolate: the best, above average, average, below average, or the worst. Would the company be more interested in the mean, median, or mode of the ratings?
- d.** A realtor is interested in studying the prices of recent home sales in an area which has many diverse neighborhoods. Would the mean, median, or mode of the prices of recent home sales be the best measure of central tendency?