

Measures of Central Tendency

OBJECTIVES

- Find the mean, median, and mode of a set of data.
- Find measures of central tendency of data organized in a stem-and-leaf plot or a frequency distribution table.



ENTERTAINMENT For many years, people have attended Broadway plays in New York City.

Broadway became an important theatrical district in the mid-1800s. The theatrical activity of Broadway peaked in the 1920s. Today, there are fewer new shows, but Broadway still remains a major theatrical center. The numbers of new Broadway productions in recent seasons are listed below. What is the average number of new Broadway productions for these seasons?

This problem will be solved in Example 3.



Season	New Productions	Season	New Productions	Season	New Productions
1960–1961	48	1973–1974	43	1986–1987	41
1961–1962	53	1974–1975	54	1987–1988	32
1962–1963	54	1975–1976	55	1988–1989	30
1963–1964	63	1976–1977	54	1989–1990	35
1964–1965	67	1977–1978	42	1990–1991	28
1965–1966	68	1978–1979	50	1991–1992	37
1966–1967	69	1979–1980	61	1992–1993	33
1967–1968	74	1980–1981	60	1993–1994	37
1968–1969	67	1981–1982	48	1994–1995	29
1969–1970	62	1982–1983	50	1995–1996	38
1970–1971	49	1983–1984	36	1996–1997	37
1971–1972	55	1984–1985	33	1997–1998	33
1972–1973	55	1985–1986	33	1998–1999	20

Source: The League of American Theatres and Producers, Inc.

The average number of new Broadway productions is an ambiguous term. Loosely stated, the average means the center of the distribution or the most typical case. Measures of average are also called **measures of central tendency** and include the **mean, median, and mode**.

The **arithmetic mean** \bar{X} is often referred to as the mean. The mean is found by adding the values in a set of data and dividing the sum by the number of values in that set. Every number in a set of data affects the value of the mean. Consequently, the mean is generally a good representative measure of central tendency. However, the mean can be considerably influenced by extreme values.

Example 1 Find the mean of the set {19, 21, 18, 17, 18, 22, 46}.

$$\bar{X} = \frac{\text{sum of the values in the set of data}}{\text{number of values in the set}}$$

$$\bar{X} = \frac{19 + 21 + 18 + 17 + 18 + 22 + 46}{7}$$

$$\bar{X} = \frac{161}{7} \text{ or } 23$$

The mean of the set of data is 23.

Notice that the mean is not necessarily a member of the set of data.

The general formula for the mean of any set of data can be written using sigma notation. If X is a variable used to represent any value in a set of data containing n items, then the arithmetic mean \bar{X} of n values is given by the following formula.

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

The numerator of the fraction can be abbreviated using the summation symbol Σ . *Recall that Σ is the uppercase Greek letter sigma.*

$$\sum_{i=1}^n X_i = X_1 + X_2 + X_3 + \dots + X_n$$

The symbol X_i represents successive values of the set of data as i assumes successive integral values from 1 to n . Substitute the sigma notation into the formula for the mean to obtain the formula below.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \text{ or } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Arithmetic Mean

If a set of data has n values given by X_i such that i is an integer and $1 \leq i \leq n$, then the arithmetic mean \bar{X} can be found as follows.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Another measure of central tendency is the median, symbolized by M_d .

Median

The median of a set of data is the middle value. If there are two middle values, it is the mean of the two middle values.

Notice that the median is not necessarily a member of the set of data.

Before the median can be found, the data must be arranged in an ordered sequence, usually from least to greatest. The median of the set {5, 6, 8, 11, 14} is the middle value 8. The median of the set {3, 4, 6, 7, 8, 10} is $\frac{6+7}{2}$ or 6.5.

The median is preferable to the mean as a measure of central tendency when there are a few extreme values or when some of the values cannot be determined. Unlike the mean, the median is influenced very little by extreme values.



The mode of a set of data is the most frequent value. Some sets of data have multiple modes and others have no mode.

Data with two modes are **bimodal**. Sets have no mode when each item of the set has equal frequency. The value of the mode is not affected by extreme values. Unlike the mean and median, the mode, if it exists, is always a member of the set of data.

Example

2 IMMIGRATION The numbers of immigrants entering the United States in a recent year are given at the right.

- Find the mean of the data.
- Find the median of the data.
- Find the mode of the data.
- State which measures of central tendency seem most representative of the set of data. Explain.

**Immigrants Entering the United States**

Country	Number of Immigrants
China	41,700
Cuba	26,500
Dominican Republic	39,600
India	44,900
Jamaica	19,100
Mexico	163,600
Philippines	55,900
Russia	19,700

Source: U.S. Immigration and Naturalization Service

- Since there are 8 countries, $n = 8$.

$$\frac{1}{8} \sum_{i=1}^8 X_i = \frac{1}{8} (41,700 + 26,500 + 39,600 + 44,900 + 19,100 + 163,600 + 55,900 + 19,700) \text{ or } 51,375$$

The mean is 51,375.

- To find the median, order the data. Since all the numbers are multiples of 100, you can order the set by hundreds.

191 197 265 396 417 449 559 1636

Since there are an even number of data, the median is the mean of the two middle numbers, 39,600 and 41,700. The median number of immigrants is $\frac{39,600 + 41,700}{2}$ or 40,650.

- Since all elements in the set of data have the same frequency, there is no mode.
- Notice that the mean is affected by the extreme value 163,600 and does not accurately represent the data. The median is a more representative measure of central tendency in this case.

When you have a large number of data, it is often helpful to use a **stem-and-leaf plot** to organize your data. In a stem-and-leaf plot, each item of data is separated into two parts that are used to form a **stem** and a **leaf**. The parts are organized into two columns.

Stems: The column on the left shows the stems. Stems usually consist of the digits in the greatest common place value of all the data. For example, if the set of data includes the numbers 890 and 1160, the greatest common place value is hundreds. Therefore, the stem of 890 is 8, and the stem of 1160 is 11.

Leaves: The column at the right contains the leaves. The leaves are one-digit numbers, which are in the next greatest place value after the stem. The leaf of 890 is 9, and the leaf of 1160 is 6. The stems and leaves are usually arranged from least to greatest.

Example

3

ENTERTAINMENT Refer to the application at the beginning of the lesson.



- Make a stem-and-leaf plot of the number of new productions for the seasons listed.
- Find the mean of the data.
- Find the median of the data.
- Find the mode of the data.
- What is a good representative number for the average of the new Broadway productions for the seasons 1960-1999?



- Since the number of productions range from 20 to 74, we will use the tens place for the stems. List the stems and draw a vertical line to the right of the stems. Then list the leaves, which in this case will be the ones digit. As shown below, it is often helpful to list the leaves as you come to them and then rewrite the plot with the leaves in order from the least to greatest.

stem	leaf
2	8 9 0
3	6 3 3 2 0 5 7 3 7 8 7 3
4	8 9 3 2 8 1
5	3 4 5 5 4 5 4 0 0
6	3 7 8 9 7 2 1 0
7	4

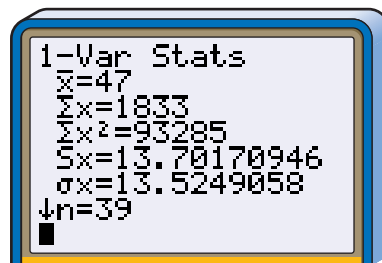
➔

stem	leaf
2	0 8 9
3	0 2 3 3 3 3 5 6 7 7 7 8
4	1 2 3 8 8 9
5	0 0 3 4 4 4 5 5 5
6	0 1 2 3 7 7 8 9
7	4

$$2|8 = 28$$

An annotation usually accompanies a stem-and-leaf plot to give meaning to the representation.

- Enter the data in the L1 list of a graphing calculator. Use the statistics mode of the calculator to find \bar{X} . The mean is 47.
- Since the median is the middle value, it is the 20th leaf on the plot. The median is 48.
- The stem-and-leaf plot shows the modes by repeated digits for a particular stem. There are four 3s with the stem 3. The mode is 33.



- e. Although 33 is the most common number in the data, it is not a central number for the data. In this case, the mean and median seem to be more representative of the data. Therefore, a representative average number of new Broadway productions could be either 47 or 48.

In a frequency distribution containing large amounts of data, each individual value in the set of data loses its identity. The data in each class are assumed to be uniformly distributed over the class. Thus, the class mark is assumed to be the mean of the data tallied in its class. For example, the mean of the data in the class with limits 17.5-22.5 is assumed to be 20, the class mark.

In the frequency distribution, the sum of the values in a class is found by multiplying the class mark X by the frequency f of that class. The sum of all the values in a given set of data is found by adding the sums of the values of each class in the frequency distribution. The sum of all values in the set can be

Remember that the measures of central tendency are only representations of the set of data.

represented by $\sum_{i=1}^k (f_i \cdot X_i)$, where k is the number of classes in the frequency distribution. Thus, the arithmetic mean of n values in a frequency distribution is found by dividing the sum of the values in the set by n or an expression equivalent to n , such as $\sum_{i=1}^k f_i$.

Mean of the Data in a Frequency Distribution

If X_1, X_2, \dots, X_k are the class marks in a frequency distribution with k classes and f_1, f_2, \dots, f_k are the corresponding frequencies, then the arithmetic mean \bar{X} can be approximated as follows.

$$\bar{X} = \frac{\sum_{i=1}^k [f_i \cdot X_i]}{\sum_{i=1}^k f_i}$$

Example



- 4 EDUCATION** Estimate the mean of the scores of 100 students on an algebra test given the following frequency distribution.

Class Limits	Class Marks (X)	Frequency (f)	fX
97.5-102.5	100	5	500
92.5-97.5	95	9	855
87.5-92.5	90	17	1530
82.5-87.5	85	26	2210
77.5-82.5	80	22	1760
72.5-77.5	75	10	750
67.5-72.5	70	7	490
62.5-67.5	65	4	260
		$\sum_{i=1}^8 f_i = 100$	$\sum_{i=1}^8 (f_i \cdot X_i) = 8355$

$$\bar{X} = \frac{8355}{100} \text{ or } 83.55$$

The mean is approximately 84.

The median M_d of the data in a frequency distribution is found from the **cumulative frequency distribution**. The cumulative frequency of each class is the sum of the frequency of the class and the frequencies of the previous classes. The chart shows the cumulative frequency for the data in Example 4.

It is often helpful to calculate the cumulative frequency from the last interval to the first.

Class Limits	Frequency f	Cumulative Frequency
97.5–102.5	5	100
92.5–97.5	9	95
87.5–92.5	17	86
82.5–87.5	26	69
77.5–82.5	22	43
72.5–77.5	10	21
67.5–72.5	7	11
62.5–67.5	4	4

For the class limit 77.5–82.5, the cumulative frequency equals $22 + 10 + 7 + 4$ or 43. This means that 43 algebra test scores fall below 82.5.

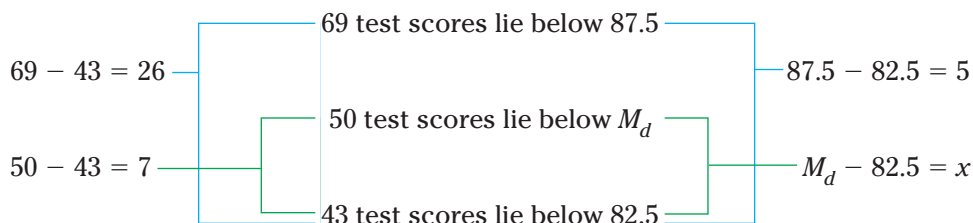
Since the median is the value below which 50% of the data lie, the class in which the median lies can be located. This class is called the **median class**. The median can be found by using an estimation technique called interpolation. *This method can also be used to find a score at any percent level.*

Example 5 **EDUCATION** Estimate the median of the data in the frequency distribution in Example 4.



Since there are 100 scores in this frequency distribution, 50 scores are below the median and 50 are above. From the chart above, find the least cumulative frequency that is greater than or equal to 50. That cumulative frequency is 69. So, the median class is 82.5–87.5.

You can use a proportion to find the value of M_d by finding the ratios of the differences in the cumulative frequencies and the upper limits of the classes.



$$\frac{26}{5} = \frac{7}{x}$$

$$x \approx 1.346153846 \quad \text{Use a calculator.}$$

$$M_d - 82.5 = x$$

$$M_d - 82.5 \approx 1.3 \quad x \approx 1.3$$

$$M_d \approx 83.8$$

The median of the data is approximately 83.8.



CHECK FOR UNDERSTANDING

Communicating Mathematics

Read and study the lesson to answer each question.

1. **Name** three measures of central tendency. Explain how to determine each measure.
2. **Write** a set of data that is bimodal.
3. **Explain** how to make a stem-and-leaf plot for a set of data whose greatest value is 1430 and least value is 970.
4. **You Decide** Eight people volunteer at Central City Recreation Center. The numbers of hours the volunteers spent at the center last week are 2, 3, 15, 2, 2, 23, 19, and 2. Omar says the median is the best representative for the number of volunteer hours. Tia disagrees and claims that the mean is a better representation. Casey says the most representative number is the mode. Who is correct? Explain.

Guided Practice

Find the mean, median, and mode of each set of data.

5. {10, 45, 58, 10}
6. {24, 28, 21, 37, 31, 29, 23, 22, 34, 31}

7. Find the mean, median, and mode of the data represented by the stem-and-leaf plot at the right.

stem	leaf
9	1 4 5 8 9
10	5 5 7 7 7
11	1 1 2
<i>9 1 = 9100</i>	

8. The Laketown Senior Center recorded how many times its members use the center's fitness facilities during a typical month. They organized their data into the following chart.

Visits	1-5	5-9	9-13	13-17	17-21	21-25	25-29	29-33
Members	2	8	15	6	38	31	13	7

- a. How many members used the center's fitness facilities during the month?
 - b. Estimate the mean of the data.
 - c. What is the median class of the data?
 - d. Estimate the median of the data.
9. **Football** Each December, the Liberty Bowl is played in Memphis, Tennessee. The winning scores of the first 40 Liberty Bowl games are listed below.

7, 41, 15, 6, 16, 32, 13, 14, 14, 34, 47, 17, 7, 31, 31, 7, 20, 36, 21, 20,
9, 28, 31, 21, 19, 21, 21, 21, 20, 34, 42, 23, 38, 13, 18, 30, 19, 41, 41, 23

- a. Make a stem-and-leaf plot of the winning scores.
- b. What is the mean of the data?
- c. What is the median of the data?
- d. Find the mode of the data.
- e. What is the most representative measure of central tendency for the number of points scored by the winning team at the Liberty Bowl? Explain.



EXERCISES

Practice

Find the mean, median, and mode of each set of data.

10. {140, 150, 160, 170} 11. {3, 3, 6, 12, 3}
12. {21, 19, 17, 19} 13. {5, 8, 18, 5, 3, 18, 14, 15}
14. {64, 87, 62, 87, 63, 98, 76, 54, 87, 58, 70, 76}
15. {6, 9, 11, 11, 12, 7, 6, 11, 5, 8, 10, 6}

16. Crates of books are being stored for later use. The weights of the crates in pounds are 142, 160, 151, 139, 145, 117, 172, 155, and 124.
- What is the mean of their weights?
 - Find the median of their weights.
 - If 5 pounds is added to each crate, how will the mean and median be affected?

Find the mean, median, and mode of the data represented by each stem-and-leaf plot.

17.

stem	leaf
3	5 8 8 9
4	4 5 5 5 8
5	7 7 9
3	5 = 35

18.

stem	leaf
5	2 4 6
6	0 1 7 8 9
7	1 6
8	0 2 6
9	1
5	2 = 5.2

19.

stem	leaf
9	0 1 7 8 9
10	5 6 9
11	3 8 8 8
12	0 5 5
9	0 = 900

20. Make a stem-and-leaf plot of the following ages of people attending a family picnic.

15, 55, 35, 46, 28, 35, 25, 17, 30, 30, 27, 35,
15, 25, 25, 20, 20, 15, 20, 17, 15, 25, 10

21. The store manager of a discount department store is studying the weekly wages of the part-time employees. The table profiles the employees.

Weekly Wages	Frequency
\$130–\$140	11
\$140–\$150	24
\$150–\$160	30
\$160–\$170	10
\$170–\$180	13
\$180–\$190	8
\$190–\$200	4

- Find the sum of the wages in each class.
- What is the sum of all of the wages in the frequency distribution?
- Find the number of employees in the frequency distribution.
- What is the mean weekly wage in the frequency distribution?
- Find the median class of the frequency distribution.
- Estimate the median weekly wage in the frequency distribution.
- Explain why both the mean and median are good measures of central tendency in this situation.

22. Find the value of x so that the mean of {2, 4, 5, 8, x } is 7.5.
23. What is the value of x so that the mean of { x , $2x - 1$, $2x$, $3x + 1$ } is 6?
24. Find the value of x so that the median of {11, 2, 3, 3.2, 13, 14, 8, x } is 8.



25. The frequency distribution of the verbal scores on the SAT test for students at Kennedy High School is shown below.

Scores	Number of Students	Scores	Number of Students
200–250	9	500–550	18
250–300	14	550–600	12
300–350	23	600–650	7
350–400	30	650–700	3
400–450	33	700–750	1
450–500	28	750–800	1

- What is the mean of the verbal scores at Kennedy High School?
- What is the median class of the frequency distribution?
- Estimate the median of the verbal scores at Kennedy High School.

**Applications
and Problem
Solving**



26. **Weather** The growing season in Tennessee is the period from May to September. The table at the right shows the normal rainfall for those months.

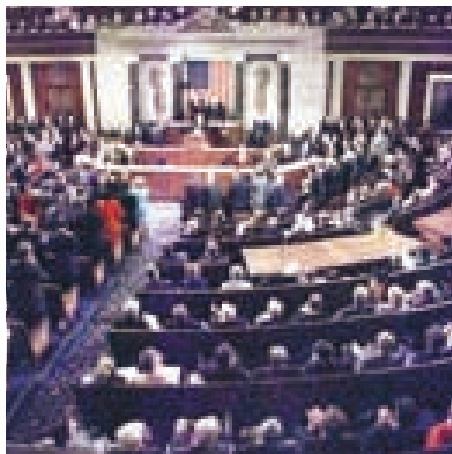
Normal Rainfall for Tennessee (inches)	
May	4.8
June	3.6
July	3.9
August	3.6
September	3.7

- Find the mean, median, and mode of this data.
- Suppose Tennessee received heavy rain in May totaling 8.2 inches. If this figure were used for May, how would the measures of central tendency be affected?
- If September were eliminated from the period, how would this affect the measures of central tendency?

27. **Critical Thinking** Find a set of numbers that satisfies each list of conditions.

- The mean, median, and mode are all the same number.
- The mean is greater than the median.
- The mode is 10 and the median is greater than the mean.
- The mean is 6, the median is $5\frac{1}{2}$, and the mode is 9.

28. **Government** As of 1999, the number of members in the House of Representatives for each state is given below.



AL	7	HI	2	MA	10	NM	3	SD	1
AK	1	ID	2	MI	16	NY	31	TN	9
AZ	6	IL	20	MN	8	NC	12	TX	30
AR	4	IN	10	MS	5	ND	1	UT	3
CA	52	IA	5	MO	9	OH	19	VT	1
CO	6	KS	4	MT	1	OK	6	VA	11
CT	6	KY	6	NE	3	OR	5	WA	9
DE	1	LA	7	NV	2	PA	21	WV	3
FL	23	ME	2	NH	2	RI	2	WI	9
GA	11	MD	8	NJ	13	SC	6	WY	1

- Make a stem-and-leaf plot of the number of representatives.
- Find the mean of the data.
- What is the median of the data?
- Find the mode of the data.
- What is a representative average for the number of members in the House of Representatives per state? Explain.

Data Update

For the latest information about the number of goals scored in hockey, visit www.amc.glencoe.com



29. Hockey A frequency distribution for the number of goals scored by teams in the National Hockey League during a recent season are given at the right.

- Use the frequency chart to estimate the mean of the number of goals scored by a team.
- What is the median class of the frequency distribution?
- Use the frequency chart to estimate the median of the number of goals scored by a team.



National Hockey League Goals

Goals	Number of Teams
160–180	1
180–200	6
200–220	10
220–240	6
240–260	3
260–280	1

Source: National Hockey League

- The actual numbers of goals scored are listed below. Find the mean and median of the data.

268, 248, 245, 242, 239, 239, 237, 236, 231, 230, 217, 215, 214, 211, 210, 210, 207, 205, 202, 200, 196, 194, 192, 190, 189, 184, 179

- How do the measures of central tendency found by using the frequency chart compare with the measures of central tendency found by using the actual data?

30. Critical Thinking A one-meter rod is suspended at its middle so that it balances. Suppose one-gram weights are hung on the rod at the following distances from one end.

5 cm 20 cm 37 cm 44 cm 52 cm 68 cm 71 cm 85 cm

The rod does not balance at the 50-centimeter mark.

- Where must a one-gram weight be hung so that the rod will balance at the 50-centimeter mark?
- Where must a two-gram weight be hung so that the rod will balance at the 50-centimeter mark?

31. Salaries The salaries of the ten employees at the XYZ Corporation are listed below.

\$54,000, \$75,000, \$55,000, \$62,000, \$226,000,
\$65,000, \$59,000, \$61,000, \$162,000, \$59,000

- What is the mean of the salaries?
- Find the median of the salaries.
- Find the mode of the salaries.
- What measure of central tendency might an employee use when asking for a raise?
- What measure of central tendency might management use to argue against a raise for an employee?
- What measure of central tendency do you think is most representative of the data? Why?
- Suppose you are an employee of the company making \$75,000. Write a convincing argument that you deserve a raise.

- 32. Education** The grade point averages for a graduating class are listed in the frequency table below.

Grade Point Averages	1.75–2.25	2.25–2.75	2.75–3.25	3.25–3.75	3.75–4.25
Frequency	12	15	31	37	5

- a. What is the estimated mean of the data?
 b. Estimate the median of the data.
- 33. Basketball** Jackson High School just announced the members of its varsity basketball team for the year. Kwan, who is 5' 9" tall, is the only sophomore to make the team. The other basketball team members are 5' 11", 6' 0", 5' 7", 6' 3", 6' 1", 6' 6", 5' 8", 5' 9" and 6' 2". How does Kwan compare with the other team members?

Mixed Review

- 34. Highway Safety** The maximum speed limits in miles per hour for interstate highways for the fifty states are given below. Construct a frequency polygon of the data. (Lesson 14-1)
- 70, 65, 75, 70, 70, 75, 65, 65, 70, 70, 55, 75, 65, 65, 65, 70, 65, 70, 65, 65, 65, 70, 70, 70, 70, 65, 75, 75, 65, 65, 75, 65, 70, 70, 65, 75, 65, 65, 65, 65, 75, 65, 70, 75, 65, 65, 70, 70, 65, 75



Source: National Motorists Association

- 35.** Determine if the following event is *independent* or *dependent*. Then determine the probability. (Lesson 13-4)
- the probability of randomly selecting two fitness magazines at one time from a basket containing 6 news magazines, 3 fitness magazines, and 2 sports magazines
- 36.** Use the ratio test to determine if the series $\frac{1}{3} + \frac{2}{3^2} + \frac{3}{3^3} + \dots + \frac{n}{3^n} + \dots$ is *convergent* or *divergent*. (Lesson 12-4)
- 37. Investments** An annuity pays 6%. What is the future value of the annuity if \$1500 is deposited into the account every 6 months for 10 years? (Lesson 11-2)
- 38.** Graph the system of inequalities. (Lesson 10-8)
- $$3x + y^2 \leq 18$$
- $$x^2 + y^2 \geq 9$$
- 39. SAT Practice** The area of $\triangle ABC$ is between which pair of numbers?
- A 16 and 17
 B 15 and 16
 C 12 and 13
 D 10 and 11
 E 9 and 10

