

# ANALYZING NUMERICAL DATA: CENTRAL TENDENCY



The student is expected to summarize numeric data with numerical summaries, including the mean and median (measures of center) and the range and interquartile range (IQR) (measures of spread), and use these summaries to describe the center, spread, and shape of the data distribution.



## TELL ME MORE...

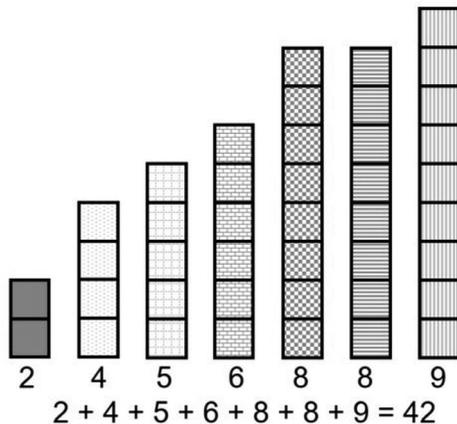
There are two characteristics of a data set that can help you to better understand what the data in the set means: center and spread. The **center** of a data set is a way to describe a typical value in the data set. The **spread** of a data set is how spread out the data values are in the set.

There are two numbers you can use to identify the center of a data set. Consider the data set shown:

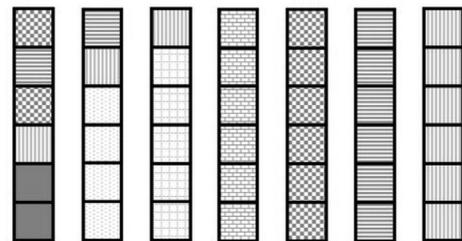
2, 4, 5, 6, 8, 8, 9

### MEAN

**Mean** is the average value in the data set. Use tiles to represent each number in the data set. 2 is represented with 2 tiles, 4 with 4 tiles, etc. The total number of tiles used is the sum of all of the numbers in the data set.



Rearrange the tiles so that each stack has the same number. In other words, divide the 42 tiles into 7 equal stacks.



$$2 + 4 + 5 + 6 + 8 + 8 + 9 = 42$$

$$42 \div 7 = 6, \text{ so the mean is } 6.$$

### MEDIAN

The **median** is the data value that, when listed in order from least to greatest, is the middle value in the data set.

The data set shown has 7 data values, which is an odd number. The median is the middle number in the data set.

2, 4, 5, **6**, 8, 8, 9

Half the data values are less than 6.

Half the data values are greater than 6.

The data set shown has 8 data values, which is an even number. The median is the mean of the two middle numbers in the data set:  
 $6 + 8 = 14$  and  $14 \div 2 = 7$ .

2, 4, 5, **6, 8**, 8, 9, 12

Half the data values are less than 7.

Half the data values are greater than 7.



## EXAMPLES

**EXAMPLE 1:** The table shows the population estimates in 2017 of the six least populated counties in Texas. What is the mean and median of this data set.

County	Population (2017)
Borden	673
Kenedy	417
Kent	763
King	296
Loving	134
McMullen	778

**STEP 1** Determine the mean of the data set.

- Determine the sum of the six populations.
  - $673 + 417 + 763 + 296 + 134 + 778 = 3,061$
- Divide the sum by the number of counties, which is 6. Round to the nearest tenth if necessary.
  - $3,061 \div 6 \approx 510.2$

**The mean of the data set is 510.2.**

**STEP 2** List the data set in order from least to greatest.

**134, 296, 417, 673, 763, 778**

**STEP 3** Determine the median of the data set.

- There are six data values, so identify the two middle numbers.

134, 296, 417, 673, 763, 778

- Calculate the mean of these two numbers. Add them together and divide the sum by 2.
  - $417 + 673 = 1,090$
  - $1,090 \div 2 = 545$

**The median of the data set is 545.**

**EXAMPLE 2:** The list shows the price of gasoline on a day in 2018 for several cities in Texas. Determine the mean of the data set. Record your answer and fill in the bubbles. Be sure to use the correct place value.

\$2.58, \$2.69, \$2.79, \$2.72, \$3.07

**STEP 1** Determine the sum of the five data values.

- $2.58 + 2.69 + 2.79 + 2.72 + 3.07 = 13.85$

**13.85**

**STEP 2** Divide the sum by the number of data values in the set.

- • There are 5 data values in the set.
- •  $13.85 \div 5 = 2.77$

**The mean is \$2.77.**

### YOU TRY IT!

Margie bowls in a league. The list shows her five most recent scores. What are the mean and median of this data set?

147, 190, 205, 184, 169

List the values in order from least to greatest.

\_\_\_\_/\_\_\_\_/\_\_\_\_/\_\_\_\_/\_\_\_\_

Identify the median: \_\_\_\_\_

Add the data values: \_\_\_\_\_

How many data values are there? \_\_\_\_

Determine the mean: \_\_\_\_\_  $\div$  \_\_\_\_\_

The mean is \_\_\_\_\_.

**STEP 3** Since the question is a gridded response question, enter your response on the grid provided. Practice using the grid with the instructions.

- Record a 2 in the ones column.  
Record a 7 in the tenths column.  
Record a 7 in the hundredths column.
- Bubble 2 beneath the numeral 2.  
Bubble 7 beneath the numeral 7.  
Bubble 7 beneath the numeral 7.

				2	.	7	7										
+	0	0	0	0		0	0			0	0			0	0		
-	1	1	1	1		1	1			1	1			1	1		
	2	2	2	2		2	2			2	2			2	2		
	3	3	3	3		3	3			3	3			3	3		
	4	4	4	4		4	4			4	4			4	4		
	5	5	5	5		5	5			5	5			5	5		
	6	6	6	6		6	6			6	6			6	6		
	7	7	7	7		7	7			7	7			7	7		
	8	8	8	8		8	8			8	8			8	8		
	9	9	9	9		9	9			9	9			9	9		

**EXAMPLE 3:** The histogram shows the ages of 20 actresses who recently won the Academy Award for Best Actress. Based on the data in the histogram, in which interval is the mean and median most likely located?

**STEP 1** Determine the location of the median.

- The median is the middle number in the data set.
- There are 20 numbers in the data set, which is an even number, so data values 10 and 11 are the two middle numbers.
- Count from left to right in the histogram to determine the interval containing data values 10 and 11.

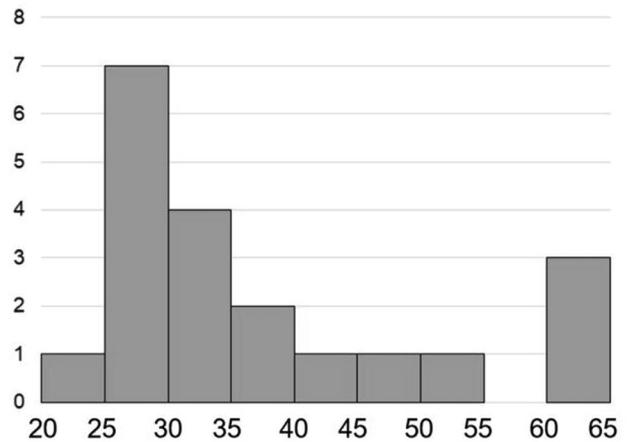
**The median lies in the 30-35 interval.**

**STEP 2** Estimate the location of the mean.

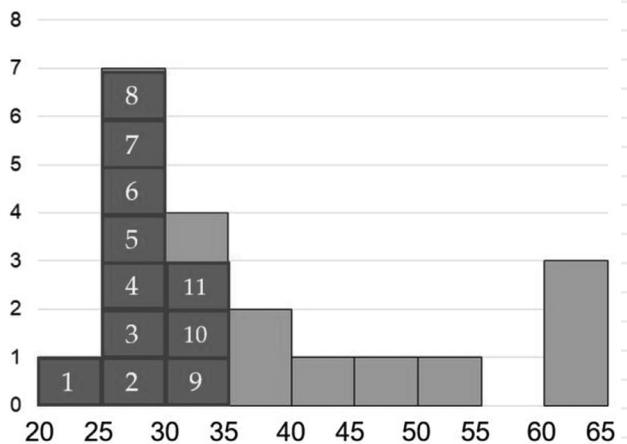
- The data set is not symmetrical, so the mean and median may not be in the same interval.
- The data set is skewed right, so the mean will be located to the right of the median.

**The mean is most likely in the 35-40 interval.**

Ages of Best Actresses



Ages of Best Actresses





# PRACTICE

Use the following information for questions 1-3.

A small clothing store kept recorded the number of customers visiting the store each day during the week using the table below.

Day	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
Customers	13	39	25	39	76	71	38

1. What is the mean number of customers during the week?
2. What is the median number of customers during the week?
3. Using the mean and median information, how can the shape of the data distribution best be described?

Use the following information for questions 4-5.

The table below displays the number of units sold for a product during the second half of the year.

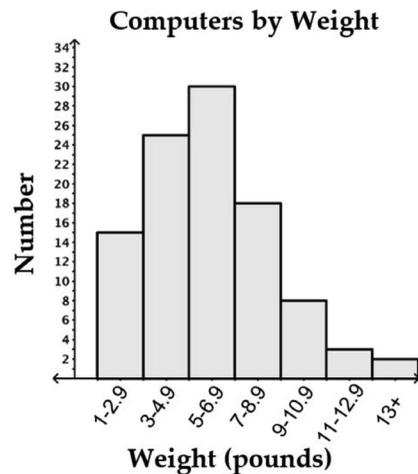
Monthly Sales (units)

July	150
August	126
September	214
October	380
November	416
December	400

4. Describe the data using the mean and median.
5. How do the mean and median best describe the shape of the data distribution?

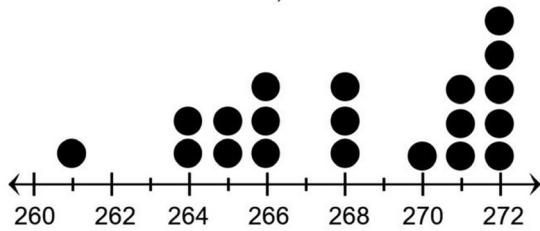
Use the following information for questions 6-7.

The histogram shows the numbers of computers being sold by weight, measured to the nearest tenth of a pound.



6. In which interval is the median likely to fall?
7. Based on the shape of the distribution, is the mean less than or greater than the median?

8. The dot plot shows the scores of the top 20 golfers in the 2018 Byron Nelson Tournament in Dallas, Texas.



What is the median of this data set? Record your answer and fill in the bubbles. Be sure to use the correct place value.

					.		
+	0	0	0	0		0	0
-	1	1	1	1		1	1
	2	2	2	2		2	2
	3	3	3	3		3	3
	4	4	4	4		4	4
	5	5	5	5		5	5
	6	6	6	6		6	6
	7	7	7	7		7	7
	8	8	8	8		8	8
	9	9	9	9		9	9

9. The median of the list of data shown is 34.

18, 32, 65, , 36, 68, 24, 30, 34

One value in the data is missing. Which could be the missing value that belongs in the box?

- A 14
- B 42
- C 30
- D 26

10. The list shows the number of telemarketing calls 10 people reported receiving during the past week.

43, 37, 35, 23, 41, 30, 33, 31, 16, 21

Which statement is best supported by the data?

- F The mean is 31 and the median is 32 making the distribution symmetrical.
- G The mean is 32 and the median is 35.5 making the distribution skewed left.
- H The mean is 35.5 and the median is 32 making the distribution skewed right.
- J The mean is 31 and the median is 35.5 making the distribution symmetrical.

11. Shawn's scores from his last 4 assignments are listed.

85, 90, 78, 89

Shawn has 1 more assignment and wants to have a mean score of 85. What score does Shawn need to earn on the fifth assignment?

- A 100
- B 86
- C 68
- D 83