

Section 6.1 Measures of Center

Objective: Compute a mean

In this lesson we are going to continue summarizing data. Instead of using tables and graphs we are going to make some numerical calculations that will aid in our understanding of data.

Definition. The *mean* of a set of data is the measure of center found by adding the data values and dividing the total by the number of data values.

Mean:

$$\mu = \frac{\text{sum of all the data values}}{\text{number of data values}}$$

Example 1. Shoplifters

The following data represents the number of shoplifters per week at a large electronics store for thirteen weeks. Calculate the mean number of shoplifters/week.

1	7	0	2	4	5	2	5	4	1	3	1	6
---	---	---	---	---	---	---	---	---	---	---	---	---

$$\mu = \frac{1 + 7 + 0 + 2 + 4 + 5 + 2 + 5 + 4 + 1 + 3 + 1 + 6}{13}$$
$$\mu = \frac{41}{13}$$

$$\mu = 3.2 \text{ shoplifters per week}$$

The mean is a one number summary that could describe the number of shoplifters for a typical week at the store.

Objective: Compute a median

Definition. The *median* of a data set is the measure of center that is the middle value when the data values are sorted from smallest to largest. This measure splits the data in 2 equal parts. Half of the data values are below the median and half of the data values are above the median.

Formula to help us locate the median:

$$\text{Position of median} = \frac{n+1}{2}$$

Note: This is NOT the median. It is just its location in our sorted data.

Example 2. Shoplifters

We will revisit the shoplifter example. This is the same data except that it is sorted from smallest to largest.

1	7	0	2	4	5	2	5	4	1	3	1	6
---	---	---	---	---	---	---	---	---	---	---	---	---

$$\text{Position of median} = \frac{13+1}{2} = 7$$

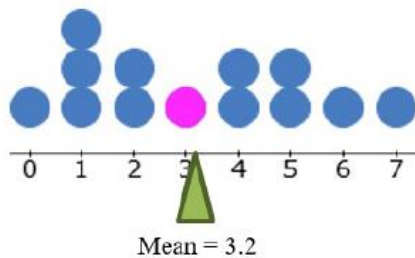
Our median is the 7th value in our sorted data!

0	1	1	1	2	2	3	4	4	5	5	6	7
1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	10 th	11 th	12 th	13 th

Median = 3 shoplifters/week

The median does slice the data in *half*. Six values are below it and above it.

Another way to think of the **mean** is that it is the balance point of the data. The fulcrum (triangle under dotplot) shows where this dot plot would balance. Notice that its' location is the same as our mean. The purple dot is our median. It is a colored dot because it is an actual observation in the data.



Example 3. Shoplifters again

Suppose a 14th week was added. Recalculate the mean and median.

Mean:

1	7	0	2	4	5	2	5	4	1	3	1	6	30
---	---	---	---	---	---	---	---	---	---	---	---	---	----

$$\mu = \frac{1 + 7 + 0 + 2 + 4 + 5 + 3 + 5 + 4 + 1 + 3 + 1 + 6 + 30}{14}$$

$$\mu = \frac{71}{14}$$

$\mu = 5.07$ shoplifters/week

Median:

$$\text{Position of median} = \frac{14+1}{2} = 7.5$$

This tells us that our median is between the 7th and 8th value in our sorted data! When there is an even number of data values in a dataset the median will always be between two data values.

0	1	1	1	2	2	3	4	4	5	5	6	7	30
1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	10 th	11 th	12 th	13 th	14 th

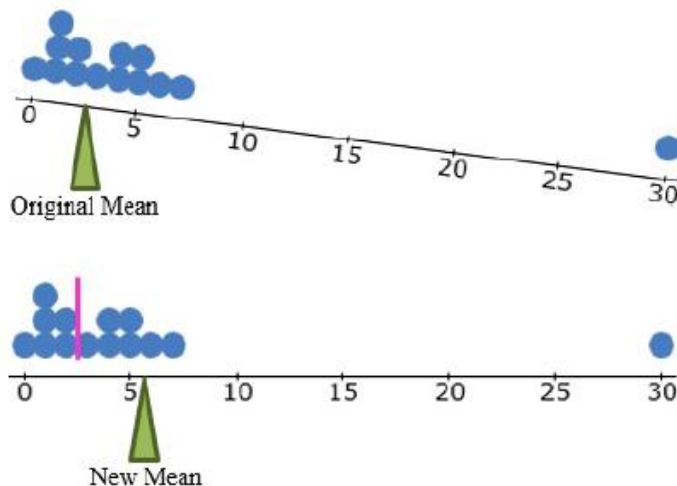
We will need to determine what number is exactly between the 7th and 8th data value. The midpoint formula will help us determine this.

$$\text{midpoint} = \frac{3+4}{2}$$

$$\text{midpoint} = \frac{7}{2}$$

Median = 3.5 shoplifters/week

If we take a look at the dot plot again you can see that the 30 is very different from the other weeks and it creates an imbalance.

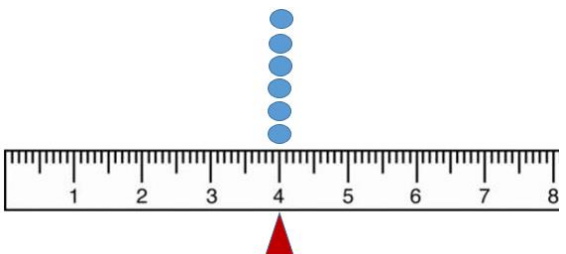
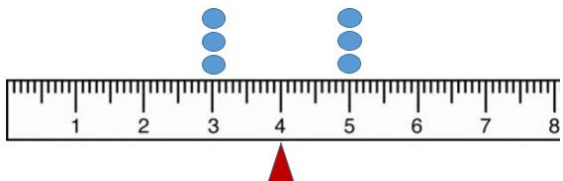


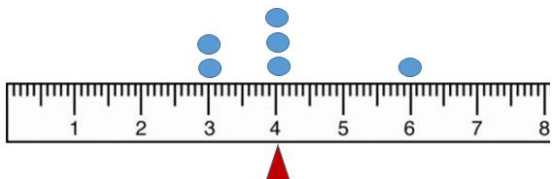
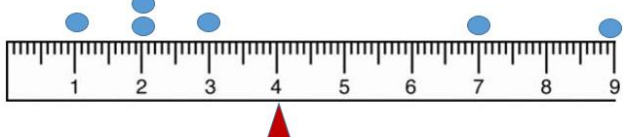
The new mean with the 30 included has shifted to the right to restore the balance. The median is again in purple. It is a line because it is not an actual value in our data this time. It still divides the data into two equal parts. Did the median react much to the value of 30? No, medians are not affected by extreme values in data sets. Such extreme values are called outliers.

The mean and median measure the center of data in different ways. The mean is the balance point. The median is the middle number (half the data values are less than the median and half are greater).

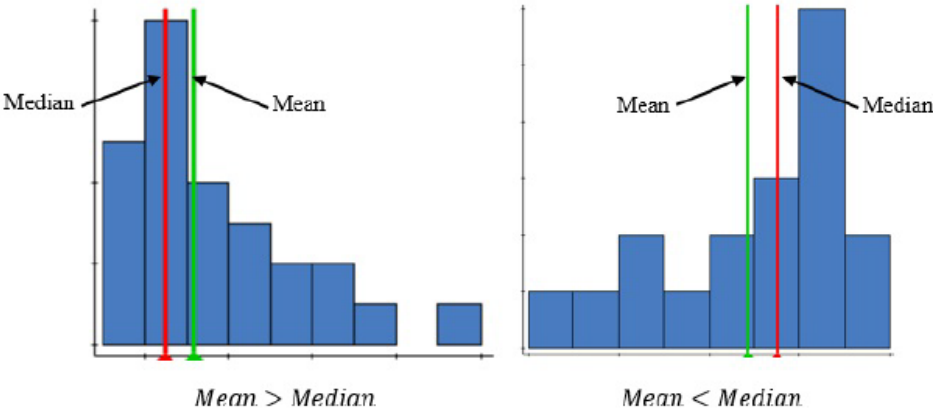
Example 3. Balance point is the mean.

Not all data are created equal. Imagine a dataset that has six values and a mean of 4. Here are just a few below. The red triangle represents the mean as well as the balance point.

Data 1: 4, 4, 4, 4, 4, 4	Data 2: 3, 3, 3, 5, 5, 5
$\mu = \frac{4 + 4 + 4 + 4 + 4 + 4}{6}$ $\mu = \frac{24}{6}$ $\mu = 4$	$\mu = \frac{3 + 3 + 3 + 5 + 5 + 5}{6}$ $\mu = \frac{24}{6}$ $\mu = 4$
	

Data 3: 3, 3, 4, 4, 4, 6	Data 4: 1, 2, 2, 3, 6, 9
$\mu = \frac{3 + 3 + 4 + 4 + 4 + 6}{6}$ $\mu = \frac{24}{6}$ $\mu = 4$	$\mu = \frac{1 + 2 + 2 + 3 + 6 + 9}{6}$ $\mu = \frac{24}{6}$ $\mu = 4$
	

When data is skewed to the right, the mean is generally larger than the median. When the data is skewed to the left the mean is generally smaller than the median.

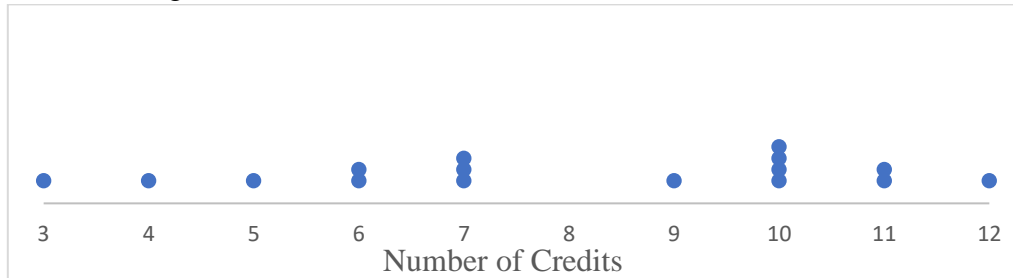


6.1 Practice

1. The following data represent the number of pop-up advertisements received by 11 families during the past month. Calculate the mean and median number of advertisements received by each family during the month.

43	37	35	30	41	23	33	31	16	21	39
----	----	----	----	----	----	----	----	----	----	----

2. The following data represents the number credits of students in a College Algebra class were taking in fall 2014.



- a. How many students were in the class?
 - b. What is the mean number of credits?
 - c. What is the median number of credits?
3. The following are daily dollar amounts collected at a neighborhood lemonade stand:

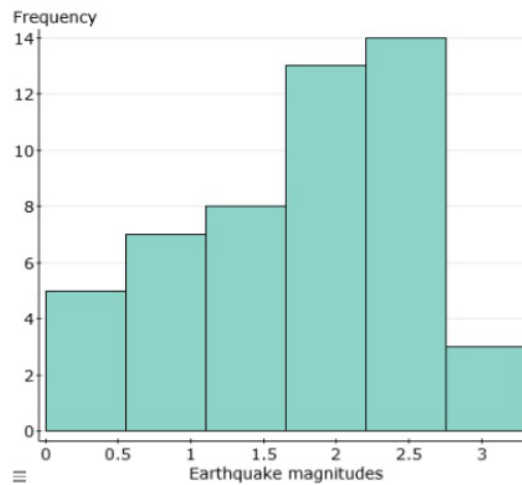
19	55	25	37	32	28	22	23	29	34	39	31	26	17
----	----	----	----	----	----	----	----	----	----	----	----	----	----

- a. Compute the mean.
 - b. Compute the median.
4. The times spent waiting in line (in minutes) for 21 randomly selected customers during the lunch rush hour at a local fast food restaurant are below. Use these data to answer the questions below.

2.0	1.8	4.0	1.5	1.0	3.4	2.3	3.2	4.5	3.2	2.5
3.1	2.5	3.4	5.1	3.5	3.2	3.5	2.9	4.2	2.7	

- a. Compute the mean
- b. Compute the median.

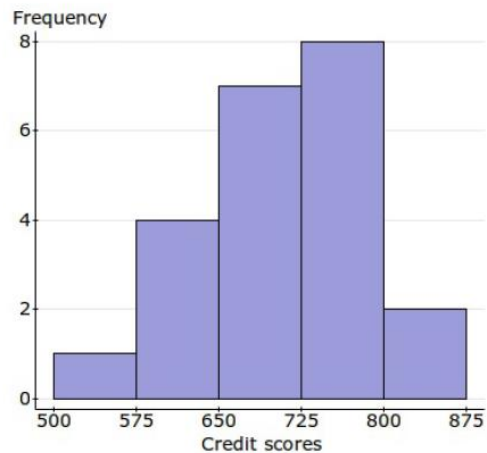
5. The histogram summarizes the magnitudes (Richter scale) of a sample of earthquakes. Refer to the histogram and comment on the relation between mean and median.



6. The amount spent on textbooks for one semester was recorded for a group of 200 students at your college. The mean expenditure was calculated to be 350 dollars. Suppose the textbook expenditure distribution is skewed to the right, which of the following values is most likely the value, in dollars, of the median of the textbook costs?

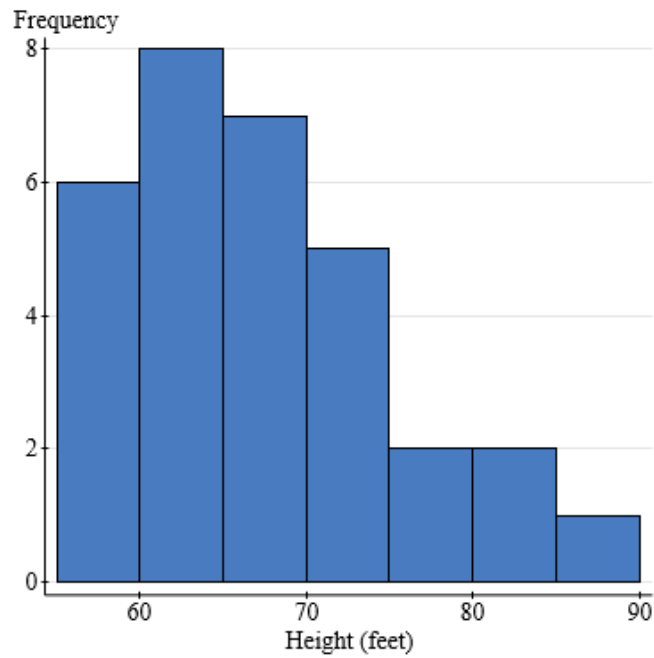
a. 450 b. 400 c. 350 d. 275

7. The histogram illustrates the credit scores of a group of people. Approximate the value of the mean.



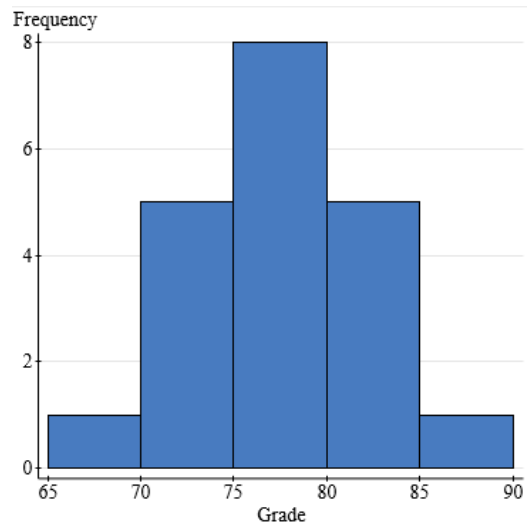
a. 715 b. 825 c. 600 d. 450

8. The histogram illustrates the heights of 31 black cherry trees. Approximate the value of the mean.



- a. 60 b. 67 c. 75 d. 4.4

9. The histogram illustrates the test grades of 20 students. Which statement below is true?



- The mean is larger than the median.
- The mean is smaller than the median.
- The mean is the same as the median.

6.1 Answers

1. Mean = 31.7 pop ups, median = 33 pop ups
2.
 - a. 16 students
 - b. Mean = 8 students
 - c. Median = 8 students
3.
 - a. Mean = 29.79 dollars
 - b. Median = 28.50 dollars
4.
 - a. Mean = 3.02 min
 - b. Median = 3.2 min
5. The mean will be smaller than the median because the data is skewed to the left.
6. d. 275
7. a. 715
8. b. 67
9. The mean is the same as the median.