

## 5.1 Measures of Central Tendency

---

In Grade 7, you described data using the measures of central tendency and range. Mean, median and mode were classified as measures of central tendency. When deciding which measure of central tendency best represented a set of data, the effect of outliers were considered.

**Mean:** a measure of central tendency determined by dividing the sum of all the values in a data set by the number of values in the set.

**Median:** a measure of central tendency represented by the middle value of an ordered set.

**Mode:** A measure of central tendency represented by the value that occurs most often in a data set.

**Example 1:**

Find the mean, median and mode for the following data set:  $\{\cancel{6}, \cancel{3}, \cancel{3}, \cancel{5}, \cancel{3}, \cancel{5}, \cancel{4}\}$

Always order the data: 3, 3, 3, 4, 5, 5, 6

$$\text{mean} = \frac{3+3+3+4+5+5+6}{7} = \frac{29}{7} = 4.1$$

$$\text{median} = 4$$

$$\text{mode} = 3$$

## Outliers and Measures of Central Tendency

**Outlier:** a value in a data set that is very different from the other values in the set

### Example 2:

Jim and Glen received Math 2201 test scores as follows:

Jim: 86%, 78%, 91%, 89%, 23%

23%, 78%, 86%, 89%, 91%

Glen: 84%, 77%, 95%, 91%, 82%

77%, 82%, 84%, 91%, 95%

(A) Determine the mean, median and mode of each students test scores.

Jim

$$\text{mean} = \frac{367}{5} = 73.4\%$$

$$\text{median} = 86\%$$

$$\text{mode} = \text{none}$$

Glen

$$\text{mean} = \frac{429}{5} = 85.8\%$$

$$\text{median} = 84\%$$

$$\text{mode} = \text{none}$$

(B) Would you say the two students' marks were similar or different for the most part?

The students' marks are similar, except for Jim's outlier of 23%.

(C) Which measure of central tendency is very different for the two students? Explain why this is the case.

mean. This is because Jim scored 23% on a test.

## Range and Dispersion

The range, based on the two extreme values of the data set, is one measure of dispersion. It gives a general idea about the total spread of the data but gives no weight to the central values of the data.

Range is commonly used as a preliminary indicator of dispersion. However, because it takes into account only the scores that lie at the two extremes, it is of limited use. Later in this unit, a more complete measure of dispersion known as standard deviation will be introduced. It takes into account every score in a distribution.

### Example 3:

Tim and Luke are both enrolled in Mathematics 2201 and scored the following marks on the last five unit tests.

	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5
Tim	60	65	70	75	80
Luke	68	69	70	71	72

Calculate the mean, median and mode for each student.

Tim

$$\text{mean} = \frac{350}{5} = 70$$

$$\text{median} = 70$$

$$\text{mode} = \text{none}$$

Luke

$$\text{mean} = \frac{350}{5} = 70$$

$$\text{median} = 70$$

$$\text{mode} = \text{none}$$

Even though the measures of central tendency are very similar for both people, the data points themselves are actually different. One set of data, Tim's test scores, is actually more spread out than Luke's. We often use the term **dispersion** to refer to the "spread" of a set of data.

**Dispersion:** a measure that varies by the spread among the data in a set; dispersion has a value of zero if all the data in a set is identical, and it increases in value as the data becomes more spread out.

One simple, preliminary way to examine the dispersion of a set of data is to calculate the **range**. We will look at more complete measures of dispersion later in the unit.

**Range:** the difference between the maximum value and the minimum value in a data set.

**Example 4:**

(A) Calculate the range for Tim and Luke's test scores.

$$\text{Tim: } 80 - 60 = 20$$

$$\text{Luke: } 72 - 68 = 4$$

(B) Are there any limitations in using range as a measure of central tendency?

Yes. Range is an indication of dispersion, but not central tendency. It tells us nothing about individual scores.

**Your turn:**

1. In a science experiment, students tested whether compost helped plants grow faster by counting the number of leaves on each plant. The following results were obtained:

Plant Growth without Compost (# of leaves per plant)	Plant Growth with Compost (# of leaves per plant)
<del>6</del>	6
<del>4</del>	11
<del>5</del>	1
<del>4</del>	6
<del>8</del>	2
<del>3</del>	4

(A) Calculate the mean, median and mode for each group.

With Compost  
mean = 5  
median = 4.5  
mode = 4

Without compost  
mean = 5  
median = 5  
mode = 6

(B) Calculate the range for each group.

$$\text{Range} = 8 - 3 = 5$$

$$\text{Range} = 11 - 1 = 10$$

(C) Describe the dispersion in the data for each group.

without compost is more disperse.  
(larger range)

(D) Which group of plants grew better? Justify your decision.

Plants with compost grew better.  
Median and mode are larger.

2. Calculate the range of each group. Explain why the range, by itself, can be a misleading measure of dispersion.

Group A: 8, 13, 13, 14, 14, 14, 15, 15, 20

Group B: 7, 7, 8, 9, 11, 13, 15, 15, 17, 18

$$\text{Group A: Range} = 20 - 8 = 12$$

$$\text{Group B: Range} = 18 - 7 = 11$$

Because it only reflects the difference between lowest and highest value. It tells us nothing about the other numbers.