

CHAPTER 5

Statistical Reasoning



Section 5.1 Exploring Data

Mean

Median

Mode

Range

Outlier



Section 5.1 Exploring Data

Mean

- ↳ a measure of central tendency found 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 data set, if there are two middle values the median is the mean of the two.

Hint: arrange the numbers in order!

Mode

- ↳ a measure of central tendency represented by the value that occurs most often in the data set

Range

- ↳ a measure of dispersion that represents the difference between the maximum value and the minimum value in a data set

Outlier

- ↳ a value in a data set that is very different from the other values in the set.

Note:

Mean, median and mode are three measures of central tendency. A single number that tends to fall around the center of the data.

Example 1:

Tim and Luke are both enrolled in Mathematics 2201 and scored the following marks on their last 5 unit tests.

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

- a) Determine the measures of central tendency (**mean**, **median**, and **mode**) for the test marks for each student.

- b) Calculate the range values for the test marks for each student.

- c) Describe the similarities and differences in the student's marks.



WHAT DO You Think?

Decide whether you agree or disagree with each statement.
Explain your decision.

- To compare two sets of data, you need only the mean, the median, and the mode.

- Most sets of data are evenly distributed about their mean.



Example 2:

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)
6	6
4	11
5	1
4	6
8	2
3	4

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

b) Calculate the range for each group.

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



Example 3:

a) Calculate the range of each group below.

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

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

b) Explain why the range, by itself, can be a misleading measure of how spread out the data values are.



Example 4:

Paulo needs a new battery for his car. He is trying to decide between two different brands. Both brands are the same price. He obtains data for the lifespan, in years, of 30 batteries of each brand, as shown below.

Measured Lifespans of 30 Car Batteries (years)									
Brand X					Brand Y				
5.1	7.3	6.9	4.7	5.0	5.4	6.3	4.8	5.9	5.5
6.2	6.4	5.5	5.7	6.8	4.7	6.0	4.5	6.6	6.0
6.0	4.8	4.1	5.2	8.1	5.0	6.5	5.8	5.4	5.1
6.3	7.5	5.0	5.7	8.2	5.7	6.8	5.6	4.9	6.1
3.3	3.1	4.3	5.9	6.6	4.9	5.7	6.2	7.0	5.8
5.8	6.4	6.1	4.6	5.7	6.8	5.9	5.3	5.6	5.9

- a) Find the measures of central tendency and the range for each set of data.

<http://easycalculation.com/statistics/mean-median-mode.php>

Brand X

Brand Y

→

b) Describe any similarities and differences between the two sets of data. Explain what additional information can be learned from the range of the data.

c) Is the mode useful to compare in this situation? Explain.

d) Suppose that one battery included in the set of data for Brand Y is defective, and its lifespan is 0.5 years instead of 5.9 years.

Discuss how this would or would not affect Paulo's decision.

To Complete: page 239-240 #1-3