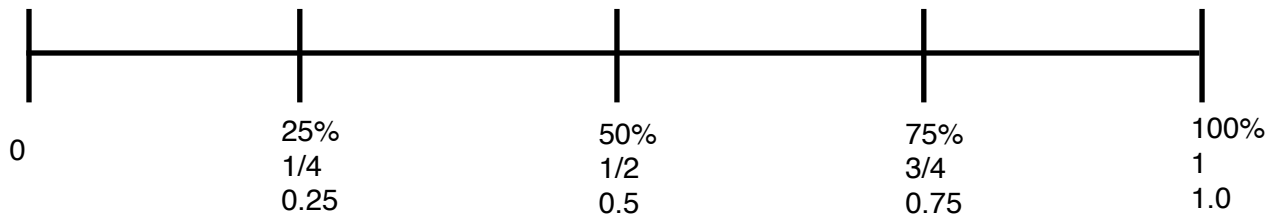


# Unit 6 End of Unit Assessment

## Study Guide **ANSWERS**

1) Mark the following scenarios on the Likelihood Line:



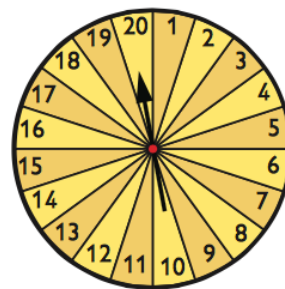
a) From a group of randomly chosen 9th graders, at least one will like ice cream

Place "a" at or between 3/4 and 1 (or 0.75 - 1.0, 75% - 100%)

b) A randomly chosen American 7-year-old is male

Place "b" at 1/2 marker (or 0.50, 50%)

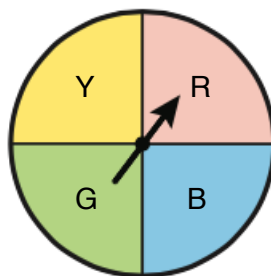
2) If you spin this spinner, what is the probability that you will land on a multiple of 3?



Multiples of 3: 3, 6, 9, 12, 15, 18 → 6 multiples of 3

Answer:  $P(\text{multiple of 3}) = \frac{6}{20}$ , or  $\frac{3}{10}$ , or 30%

3)



a) Make a list of all the possible combined outcomes of flipping this coin and spinning the spinner

**HY TY HR TR HG TG HB TB**

b) What is the probability you will get Tails and Blue on the first try?

Possible outcomes: HY TY HR TR HG TG HB **TB**

Answer:  $P(TB) = 1/8 = 0.125 = 12.5\%$

- 4) Jack and Sophie are doing a survey about lunch choices at their school. They can't ask everybody—there are 1,000 students—so they decide to ask 50 students.

Here are some plans for finding the 50 students. Which is best?

- A** Pick 50 locker numbers using a random-number table and survey the students who have those lockers.
- B** At the front door in the morning survey every 20th student until 50 students have been surveyed.
- C** At the cafeteria door at lunch survey every 20th student until 50 students have been surveyed.
- D** At the front door in the morning have each student roll a 20-sided die. If a student rolls a 0 survey that student. Continue until they have 50 responses.

Answer: A

- 5) Five samples of 100 marbles show the following number of blue marbles:

45, 42, 44, 47, 47

Based on these samples, how many blue marbles are in the jar if there are a total of 600 marbles in the jar?

Answer: About 270 marbles

The sum of the samples is 225 ( $45 + 42 + 44 + 47 + 47 = 225$ ) and there are five data values:  $225 \div 5 = 45$ . So, about 45 100 marbles are blue. The jar contains six times this amount ( $100 \cdot 6 = 600$ ); so, about  $45 \cdot 6$ , or 270, marbles are blue.

- 6) What is the median for this set of data?

3, 5, 6, 7, 8, 8, 9, 10

- A** 6
- B** 7
- C** 7.5
- D** 8

Answer: C

- 7) Which set of data has the highest mean?

- A** 5, 7, 9, 11, 13
- B** 3, 6, 6, 8, 11, 14
- C** 9, 9, 9, 9, 10, 10
- D** 3, 7, 9, 9, 12, 14

- 8) Which set of data has the highest mean absolute deviation?

- A** 5, 7, 9, 11, 13
- B** 4, 7, 7, 9, 12, 15
- C** 8, 8, 9, 9, 10, 10
- D** 3, 7, 9, 9, 12, 14

- 7) A:  $5+7+9+11+13 = 45 \rightarrow 45/5 = 9$   
 B:  $3+6+6+8+11+14 = 48 \rightarrow 48/6 = 8$   
 C:  $9+9+9+9+10+10 = 56 \rightarrow 56/6 = 9.33$   
 D:  $3+7+9+9+12+14 = 54 \rightarrow 54/6 = 9$

Answer: C has the highest mean

8)

B: Mean = 9

$9 - 4 = 5$  ;  $9 - 7 = 2$  ;  $9 - 7 = 2$  ;  $9 - 9 = 0$  ;  $9 - 12 = |-3| = 3$  ;  $9 - 15 = |-6| = 6$

MAD:  $5+2+2+0+3+6 = 18 \rightarrow 18/6 = 3$

Answer = B

9)

Class A and Class B measured the height of all the students (both classes have 32 students) and made these two box plots (using the same scale) to compare the classes. With the given information, is it possible to determine which class has the taller typical student?



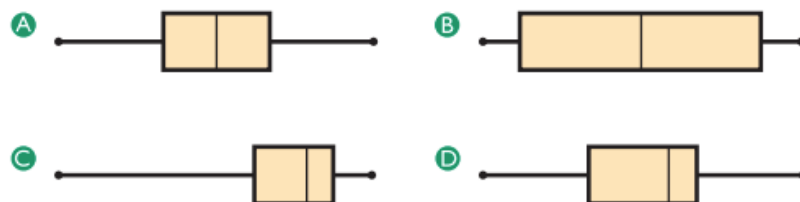
Answer:

Answers will vary. Possible answer: No, because an argument could be made for either class. Although Class A has a slightly higher median it also has a wider range and a wide interquartile range. 50% of Class A's students are taller than 75% of Class B's students. However, Class B has 50% of its students clustered close to the median.

This is a situation in which knowing the mean or the MAD (mean absolute deviation) would be very helpful, because it is not known how the data are spread among each quartile.

10)

Each box plot summarizes one class's scores on a district math test. Which class scored highest?



Answer = C

11) Jim is doing a survey to figure out how his customers are liking his company's new Lego set. He needs to determine the following for his survey:

1. *The target population*
2. *Whether he should survey a sample, or the population*
3. *What the method for sampling should be*

Here is what Jim decided:

1. *The target population would be: all kids ages 6 -17*
2. *Survey a sample*
3. *Put a link to the survey in the Lego box offering a 10% discount on their next Legos purchase if they fill out the survey.*

a. Which target population is correct?

- |                         |                                |
|-------------------------|--------------------------------|
| a) All kids ages 6 - 17 | c) All kids who play Legos     |
| b) Boys ages 6 - 17     | d) Anyone who plays with Legos |

Answer = D

b. What do you think of Jim's sampling method? Do you think it's biased or unbiased, and why?

Answer = Could be biased because not everyone can afford a computer and have access to a computer. Also, only those who have time to answer the survey will answer it.

12) a) Create a box plot for this data set:

3, 5, 7, 8, 9, 14, 14, 18, 20, 22, 23, 23, 23, 35

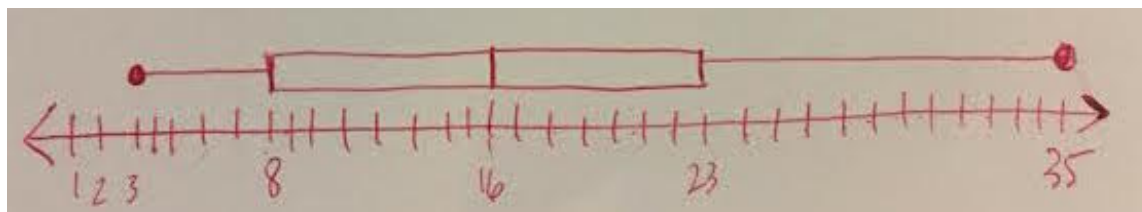
Lowest Value: 3

Highest Value: 35

Median: 16

Lower Quartile: 8

Upper Quartile: 23



b) Calculate the mean of this data set.

$$\text{Mean} = 3+5+7+8+9+14+14+18+20+22+23+23+23+35 = 224 \rightarrow 224/14 = 16$$

c) Calculate the Mean Absolute Deviation (MAD) of this data set (Round to two decimal places).

$$\text{Mean} = 16$$

$$\begin{aligned} 16 - 3 &= \underline{13} ; 16 - 5 = \underline{11} ; 16 - 7 = \underline{9} ; 16 - 8 = \underline{8} ; 16 - 9 = \underline{7} ; 16 - 14 = \underline{2} ; 16 - 14 = \underline{2} ; \\ 16 - 18 &= |-2| = \underline{2} ; 16 - 20 = |-4| = \underline{4} ; 16 - 22 = |-5| = \underline{5} ; 16 - 23 = |-7| = \underline{7} ; 16 - 23 = |-7| = \underline{7} ; \\ 16 - 23 &= |-7| = \underline{7} ; 16 - 35 = |-19| = \underline{19} \end{aligned}$$

$$13+11+9+8+7+2+2+2+4+5+7+7+7+19 = 139 \rightarrow 139/14 = \underline{9.93}$$

$$\text{MAD} = 9.93$$

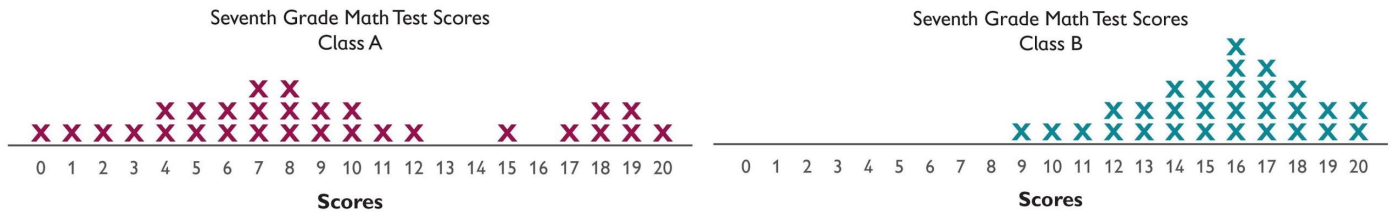
d) What does the MAD tell you about this data set?

It tells you that much of your data is far from the mean, and that mean is probably not the best way to analyze your data. It also means that your data has a high range of values (distance between lowest and highest values).

## DEFINITIONS!

Distribution: How the data sets are spread out.

Ex. How is the data distributed in these two data sets:



Variability: Refers to how spread out the data is (think about the range).

Ex. Using the two data sets above, which class has a higher variability? (Or, which has a bigger range of numbers?)

Mean: Add up all the numbers in the data set (even if some numbers repeat), and divide that sum by the total number of numbers in your data set.

Ex. Data set: {2, 4, 4, 6, 7}

$$\text{Mean} = 2 + 4 + 4 + 6 + 7 = 23 \rightarrow 23 / 5 = 4.6$$

Median: Make a list of all the numbers in your data set from lowest to highest (include all repeated numbers), and find the number in the middle. If there are two numbers that share the middle, then find the mean of those two numbers.

Ex. Data set: {2, 4, 4, 6, 7}

Median = 2, 4, 4, 6, 7

Median = 4

Ex. Data set: {2, 2, 4, 6, 7, 8}

Median = 2, 2, 4, 6, 7, 8

$$\text{Median} = 4 + 6 = 10 \rightarrow 10 / 2 = 5$$

Mean Absolute Deviation (MAD): The average distance between each individual data value and the mean.

1) Find the mean of your data set

2) Take each number in your data set, and subtract the mean - each of those numbers. Take the absolute value of each of the results (they're always positive).

3) Add up all those results from step 2, and divide them by the total number of values in your data set

Ex. Data set: {2, 4, 4, 6, 7}

1) Mean:  $2 + 4 + 4 + 6 + 7 = 23 \rightarrow 23 / 5 = 4.6$

2)  $4.6 - 2 = \mathbf{2.6}$

$4.6 - 4 = \mathbf{0.6}$

$4.6 - 4 = \mathbf{0.6}$

$4.6 - 6 = \mathbf{1.4}$

$4.6 - 7 = \mathbf{2.4}$

3)  $2.6 + 0.6 + 0.6 + 1.4 + 2.4 = 7.6 \rightarrow 7.6 / 5 = \mathbf{1.52}$

What does the MAD tell you? It tells you how useful the mean is for analyzing your data.

If your MAD is really **small** (like 0.1 - 4), then that shows you that your data values are pretty similar (or close) to what the mean is.

If the MAD is a **big** number (Like 10), then that means many of your data values are far from the mean, such as outliers, and mean is probably not a great value to use to analyze your data.