

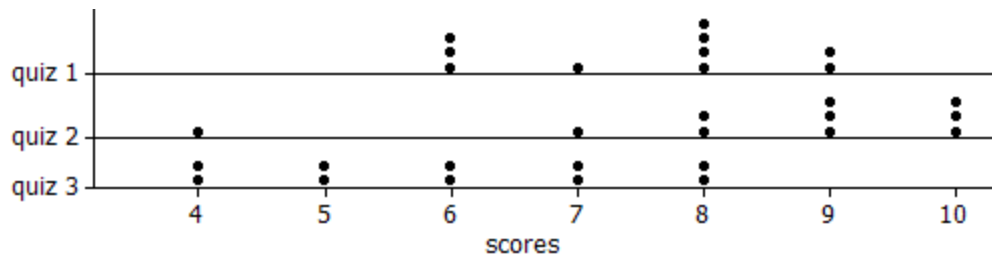
Name: \_\_\_\_\_

Date: \_\_\_\_\_

## Unit 6: Data and Statistics

## Review

- 1) The scores of three quizzes are shown in the following data plot for a class of 10 students. Each quiz has a maximum possible score of 10. Possible dot plots of the data are shown below.



- On which quiz did students tend to score the lowest? Justify your choice.
- Without performing any calculations, which quiz tended to have the most variability in the students' scores? Justify your choice based on the graphs.
- If you were to calculate a measure of variability for Quiz 2, would you recommend using the interquartile range or the standard deviation? Explain your choice.
- For Quiz 3, move one dot to a new location so that the modified data set will have a larger standard deviation than before you moved the dot. Be clear which point you decide to move, where you decide to move it, and explain why. State the standard deviation for the original data set, and for your new data set.

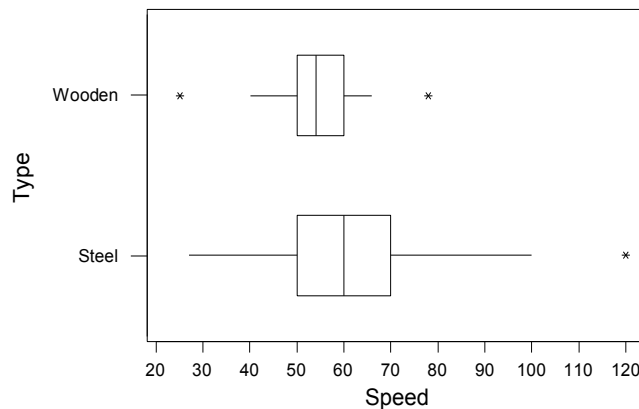
Name: \_\_\_\_\_

Date: \_\_\_\_\_

**Unit 6: Data and Statistics**

**Review**

2) The box plots below display the distributions of maximum speed for 145 roller coasters in the United States, separated by whether they are wooden coasters or steel coasters.



Based on the box plots, answer the following questions or indicate that you do not have enough information.

Which type of coaster has more observations?

Explain your choice:

- A. Wooden
- B. Steel
- C. About the same
- D. Cannot be determined

a. Which type of coaster has a higher percentage of coasters that go faster than 60 mph?

Explain your choice:

- A. Wooden
- B. Steel
- C. About the same
- D. Cannot be determined

b. Which type of coaster has a higher percentage of coasters that go faster than 50 mph?

Explain your choice:

- A. Wooden
  - B. Steel
  - C. About the same
  - D. Cannot be determined
- Explain your choice:

c. Which type of coaster has a higher percentage of coasters that go faster than 48 mph?

Explain your choice:

- A. Wooden
- B. Steel
- C. About the same
- D. Cannot be determined

d. Write 2–3 sentences comparing the two types of coasters with respect to which type of coaster normally goes faster.

Name: \_\_\_\_\_

Date: \_\_\_\_\_

**Unit 6: Data and Statistics**

**Review**

3) Fifty moviegoers were surveyed about their favorite movie types.

- 15 men and 6 women chose "Action" as their favorite type
- 9 men and 10 women chose "Drama" as their favorite type
- 6 men and 4 women chose "Comedy" as their favorite type

a) Use the table below to construct a two-way frequency table.

<b>Favorite Movie Types</b>				
	<b>Action</b>	<b>Drama</b>	<b>Comedy</b>	<b>Total</b>
<b>Men</b>				
<b>Women</b>				
<b>Total</b>				

b) Find the relative frequencies to compare and describe the survey.

<b>Favorite Movie Types</b>				
	<b>Action</b>	<b>Drama</b>	<b>Comedy</b>	<b>Total</b>
<b>Men</b>				
<b>Women</b>				
<b>Total</b>				

c) Find the row conditional relative frequencies to compare and describe the survey.

<b>Favorite Movie Types</b>				
	<b>Action</b>	<b>Drama</b>	<b>Comedy</b>	<b>Total</b>
<b>Men</b>				
<b>Women</b>				
<b>Total</b>				

d) Find the column conditional relative frequencies to compare and describe the survey.

<b>Favorite Movie Types</b>				
	<b>Action</b>	<b>Drama</b>	<b>Comedy</b>	<b>Total</b>
<b>Men</b>				
<b>Women</b>				
<b>Total</b>				

e) What movie type do the respondents prefer? How do you know?

f) Is there an association between movie type and gender? How do you know? Does gender cause an individual to have a particular movie preference? Explain.

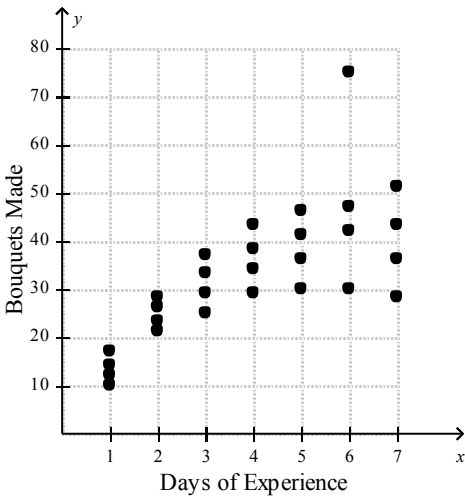
Name: \_\_\_\_\_

Date: \_\_\_\_\_

**Unit 6: Data and Statistics**

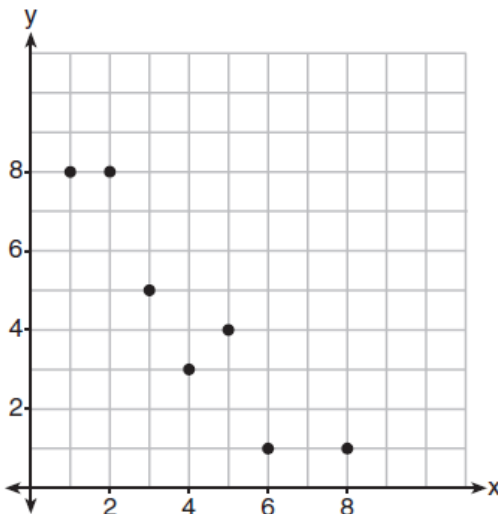
**Review**

- 4) A floral delivery company conducts a study to measure the effect of worker experience on productivity. Tell whether the scatter plot appears to have a linear or non-linear pattern of association.



- 5) Multiple Choice – Pick the BEST answer.

What is the correlation coefficient of the linear fit of the data shown below, to the *nearest hundredth*?



- (1) 1.00                      (3) -0.93  
(2) 0.93                      (4) -1.00

Name: \_\_\_\_\_

Date: \_\_\_\_\_

**Unit 6: Data and Statistics**

**Review**

- 6) Twenty-two students were surveyed about the number of days they played outside in one month. The results of this survey are shown below.

3, 3, 4, 4, 4, 4, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 8, 9, 9, 10, 11

- a. Use your calculator to find the Mean, Median, Q1 and Q3. On the grid below, create a **dot plot** based on the data. Be sure to check for outliers.

- b. Identify the typical number of days spent outside by the twenty-five students.

- c. Use the statistical features of your calculator to find the standard deviation of the data set (round to the nearest hundredth).

Standard Deviation: \_\_\_\_\_

- d. Write a sentence Interpreting Standard Deviation in the context of the question.

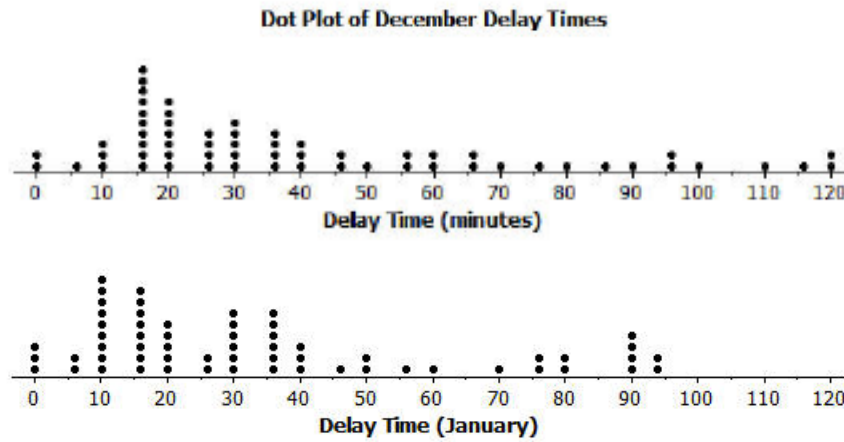
Name: \_\_\_\_\_

Date: \_\_\_\_\_

**Unit 6: Data and Statistics**

**Review**

- 7) Transportation officials collected data on sixty flight delays in the month of December and sixty flight delays in the month of January.



Which measure of variability would be appropriate in describing the typical flight delays in December and January (pick from the Standard Deviation of the Interquartile Range)? Why? Explain the meaning of this measure of variability in the context of the question.

- 8) Find the standard deviation BY hand for the data given below.

10, 15, 18, 19, 21, 26, 28

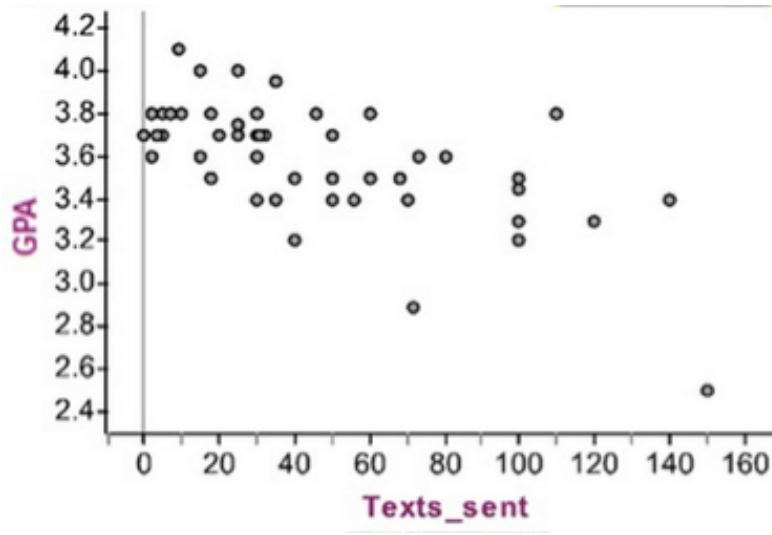
Name: \_\_\_\_\_

Date: \_\_\_\_\_

**Unit 6: Data and Statistics**

**Review**

9) The following scatterplot displays data for GPA and Texts sent by a sample of Newton high school students:



- a) There appears to be a linear relationship between the variables. Describe the relationship (strong positive, strong negative, etc.)
  
- b) Draw what you think may be a good line of best fit on the graph above AND find the equation of your line BY HAND.
  
  
  
  
  
  
  
  
  
  
  
  
  
- c) What would you estimate is the correlation coefficient for the data set. Explain why you picked your estimate.
  
  
  
  
  
  
  
  
  
  
  
  
  
- d) Explain what a residual is and what it means. What does the sign (positive or negative) of the residual tell you about the actual versus the predicted value? What does the sign tell you about the location of the actual value relative to the regression line?

## Unit 6: Data and Statistics

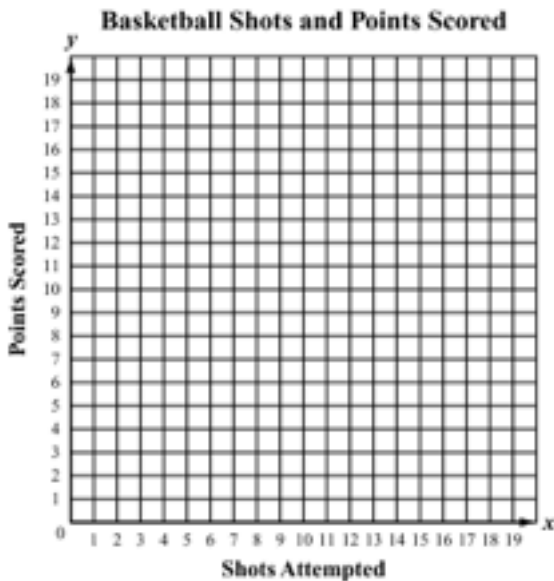
## Review

10)

The table below shows the number of shots attempted by a basketball player in ten games and the number of points scored as a result of the attempted shots.

<b>Shots Attempted (<math>x</math>)</b>	4	7	12	5	2	5	10	6	9	5
<b>Points Scored (<math>y</math>)</b>	2	4	7	3	2	3	6	4	6	4

a) Make a scatterplot of the data.



b) Use your calculator to find the linear regression equation for this data set. Write the equation below.

c) Calculate the residual for the least squares regression line for an  $x$  value of 7.

c) What is the slope of the linear regression equation? Interpret the slope in context.

d) Use your calculator to find the correlation coefficient. Interpret the correlation coefficient.

e) Does the number of shots attempted have a causal relationship with the number of points earned? Explain.