CLASS ACTIVITY 4

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

1) Which of the following variables would most likely follow a Normal model?
   A) family income
   B) weights of adult male elephants
   C) scores on an easy test
   D) heights of singers in a co-ed choir
   E) all of these

2) Two sections of a class took the same quiz. Section A had 15 students who had a mean score of 80, and Section B had 20 students who had a mean score of 90. Overall, what was the approximate mean score for all of the students on the quiz?
   A) none of these  
   B) 84.3  
   C) 85.0  
   D) 85.7  
   E) It cannot be determined.

3) Your Stats teacher tells you your test score was the 3rd quartile for the class. Which is true?
   I. You got 75% on the test.
   II. You can’t really tell what this means without knowing the standard deviation.
   III. You can’t really tell what this means unless the class distribution is nearly Normal.
   A) II only  
   B) II and III  
   C) I only  
   D) III only  
   E) none

4) The SPCA has kept these data records for the past 20 years. If they want to show the trend in the number of dogs they have housed, what kind of plot should they make?
   A) pie chart  
   B) boxplot  
   C) timeplot  
   D) histogram  
   E) bar graph

5) The veterinary bills for the dogs are summarized in the ogive shown. Estimate the IQR of these expenses.

   ![Ogive graph]

   A) $100  
   B) $150  
   C) $200  
   D) $50  
   E) $75

6) Last weekend police ticketed 18 men whose mean speed was 72 miles per hour, and 30 women going an average of 64 mph. Overall, what was the mean speed of all the people ticketed?
   A) 69 mph  
   B) 68 mph  
   C) none of those  
   D) 67 mph  
   E) It cannot be determined.

7) Which is true of the data shown in the histogram?
I. The distribution is skewed to the right.
II. The mean is probably smaller than the median.
III. We should use median and IQR to summarize these data.

A) III only  
B) II and III only  
C) I, II, and III  
D) II only  
E) I only

8) Suppose that a Normal model describes fuel economy (miles per gallon) for automobiles and that a Toyota Corolla has a standardized score (z-score) of +2.2. This means that Corollas . . .

A) achieve fuel economy that is 2.2 standard deviations better than the average car.
B) get 2.2 mpg more than the average car.
C) get 2.2 miles per gallon.
D) get 2.2 times the gas mileage of the average car.
E) have a standard deviation of 2.2 mpg.

9) Which scatterplot shows a strong association between two variables even though the correlation is probably near zero?

A)  
B)  
C)  
D)  
E)  
10) The correlation between a family’s weekly income and the amount they spend on restaurant meals is found to be $r = 0.30$. Which must be true?

I. Families tend to spend about 30% of their incomes in restaurants.
II. In general, the higher the income, the more the family spends in restaurants.
III. The line of best fit passes through 30% of the $(\text{income, restaurant})$ data points.

A) III only  
B) II only  
C) II and III only  
D) I only  
E) I, II, and III

11) A least squares line of regression has been fitted to a scatterplot; the model’s residuals plot is shown.

Which is true?

A) The linear model is poor because the correlation is near 0.
B) A curved model would be better.
C) The linear model is appropriate.
D) none of these
E) The linear model is poor because some residuals are large.

**SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.**

12) Assembly line Your new job at Panasonic is to do the final assembly of camcorders. As you learn how, you get faster. The company tells you that you will qualify for a raise if after 13 weeks your assembly time averages under 20 minutes. The data shows your average assembly time during each of your first 10 weeks.

<table>
<thead>
<tr>
<th>Week</th>
<th>Time(min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td>2</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
</tr>
</tbody>
</table>

a. Which is the explanatory variable? ________________

b. What is the correlation between these variables? __________

c. You want to predict whether or not you will qualify for that raise. Would it be appropriate to use a linear model? Explain.

13) It takes a while for new factory workers to master a complex assembly process. During the first month new employees work, the company tracks the number of days they have been on the job and the length of time it takes them to complete an assembly. The correlation is most likely to be

A) near +0.6  
B) exactly +1.0  
C) near -0.6  
D) exactly -1.0  
E) near 0
14) All but one of these statements contain a mistake. Which could be true?
A) There is a high correlation (1.09) between height of a corn stalk and its age in weeks.
B) The correlation between a car’s length and its fuel efficiency is 0.71 miles per gallon.
C) The correlation between a football player’s weight and the position he plays is 0.54.
D) The correlation between the amount of fertilizer used and the yield of beans is 0.42.
E) There is a correlation of 0.63 between gender and political party.

15) Which statement about influential points is true?
I. Removal of an influential point changes the regression line.
II. Data points that are outliers in the horizontal direction are more likely to be influential than points that are outliers in the vertical direction.
III. Influential points have large residuals.
A) II and III  B) I only  C) I and II  D) I, II, and III  E) I and III

16) A regression analysis of students’ college grade point averages (GPAs) and their high school GPAs found $R^2 = 0.311$. Which of these is true?
I. High school GPA accounts for 31.1% of college GPA.
II. 31.1% of college GPAs can be correctly predicted with this model.
III. 31.1% of the variance in college GPA can be accounted for by the model
A) III only  B) II only  C) none  D) I only  E) I and II

17) Two variables that are actually not related to each other may nonetheless have a very high correlation because they both result from some other, possibly hidden, factor. This is an example of

18) If the point in the upper right corner of the scatterplot below is removed from the data set, then what will happen to the slope of the line of best fit ($b$) and to the correlation ($r$)?

A) $b$ will decrease, and $r$ will increase.
B) both will decrease.
C) $b$ will increase, and $r$ will decrease.
D) both will remain the same.
E) both will increase.
ANSWER KEY

1) B
2) D
3) E
4) C
5) A
6) D
7) B
8) A
9) C
10) B
11) C
12) a. weeks worked
   b. $r = -0.97$
   c. No. The residuals plot shows a distinct curve, and predictions about what will happen three weeks in the future are likely to be unreliable.
13) C
14) D
15) C
16) A
17) B
18) A