Math 140   Extra Credit   Name: _____

1. In a survey people are asked "Which brand of toothpaste do you prefer?" The data gathered from this question would be what type of data?
    a. categorical
    b. quantitative
    c. continuous

Items 2 and 3 refer to the following situation:
A student is gathering data on the driving experiences of other college students. One of the variables measured is the type of car the student drives. These data are coded using the following method: 1 = subcompact, 2 = compact, 3 = standard size, 4 = full size, 5 = premium, 6 = mini van, 7 = SUV, and 8 = truck.

2. What type of variable is this?
    a. categorical
    b. quantitative
    c. continuous

3. The student plans to see if there is a relationship between the number of speeding tickets a student gets in a year and the type of vehicle he or she drives. Identify the response variable in this study.
    a. College students
    b. type of car
    c. number of speeding tickets
    d. average number of speeding tickets last year

4.   A researcher is studying the relationship between a vitamin supplement and cholesterol level. What type of study needs to be done in order to establish that the amount of vitamin supplement causes a change in cholesterol level?
    a. Correlational study
    b. Randomized experiment
    c. Time Series study
    d. Survey

5. An instructor is going to model an experiment in his statistics class by comparing the effect of 4 different treatments on student responses. There are 40 students in the class. Which is the best way for the instructor to distribute the students to the 4 treatments for this experiment?
    a. Assign the first treatment to the first 10 students on the class list, the second treatment to the next 10 students, and so on.
    b. Assign a unique number to each student, then use random numbers to assign 10 students to the first treatment, 10 students to the second treatment, and so on.
    c. Assign the treatment as students walk into class, giving the first treatment to the first 10 students and the second treatment to the next 10 student, and so on.
    d. All of these are equally appropriate methods.
    e. None of these is an appropriate method.

Suppose two researchers wanted to determine if aspirin reduces the chance of a heart attack.

6. Researcher 1 studied the medical records of 500 randomly selected patients. For each patient, he recorded whether the person took aspirin every day and if the person had ever had a heart attack. Then he reported the percentage of heart attacks for the patients who took aspirin every day and for those who did not take aspirin every day.
   What type of study did Researcher 1 conduct?
       a. Observational
       b. Experimental
   c. Survey
       d. None of the above

7. Researcher 2 also studied 500 patients that visited a regional hospital in the last year.
   He randomly assigned half (250) of the patients to take aspirin every day and the other half to take a placebo everyday. Then after a certain length of time he reported the percentage of heart attacks for the patients who took aspirin every day and for those who did not take aspirin every day. What type of study did Researcher 2 conduct?
       a. Observational
       b.Experimental
       c. Survey
       d. None of the above

8. A team in the Department of Institutional Review at a large university wanted to study the relationship between completing an internship during college and students' future earning potential. From the same graduating class, they selected a random sample of 80 students who completed an internship and 100 students who did not complete an internship and examined their salaries 5 years past graduation. They found that there was a statistically higher mean salary for the internship group than for the non-internship group. Which of the following interpretations do you think is the most appropriate?
       a. More students should take internships because having an internship produces a higher salary.
       b. There could be a confounding variable, such as student major, that explains the difference in mean salary between the internship and no intership groups.
       c. You cannot draw any valid conclusions because the samples are not the same size.

9. Scores for a quiz were calculated as the number of correct responses. Below is a graphical display of the quiz scores. How many students have scores above 15? (Note: all scores are integers and bars begin at left endpoints)

Quiz Scores
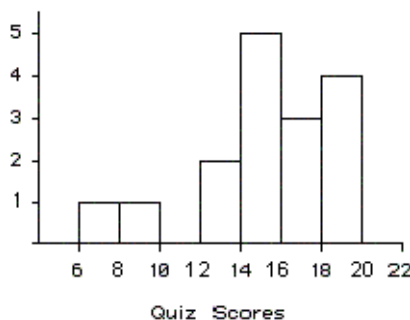
a. 6
b. 7
c. 12
d. 13
e. Can't be determined.

10. In order to determine which kind of data display (e.g., histogram versus bar graph) is appropriate for a given variable, one should consider which of the following:
    a. whether the relevant variable is quantitative or categorical
    b. whether the study is observational or experimental
    c. the range of the data

11. A class survey asked students to indicate if they are MAC or PC users. Of the following graphs, which is most appropriate to display their results?
    a. Pie chart
    b. Histogram
    c. Either a pie chart or a histogram
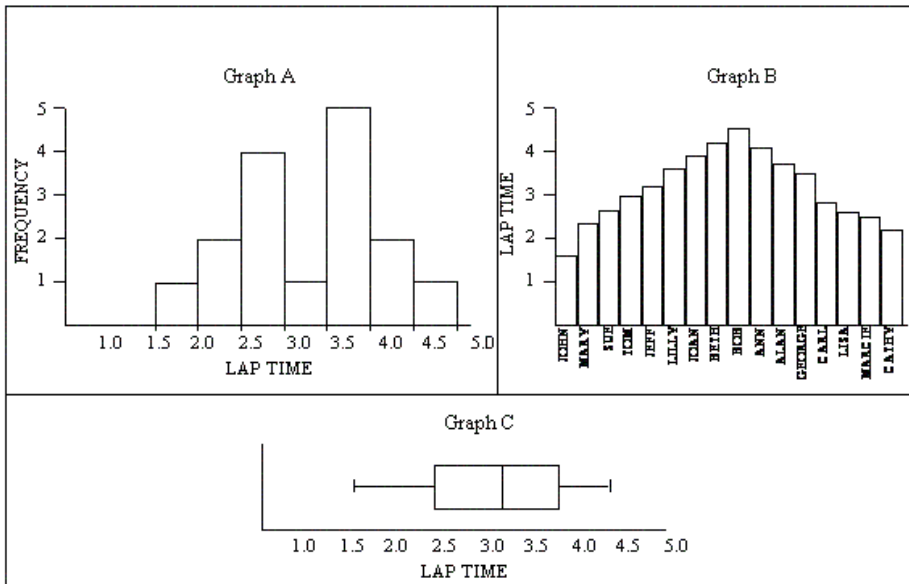   d. None of the above

12. A college statistics class conducted a survey. They gathered data from a large random sample of students who estimated how much money they typically spent each week in different categories (e.g., food, entertainment, etc.). A distribution of the survey results is presented below. One student claims the distribution of food costs basically looks bell-shaped, with one outlier. How would you respond?
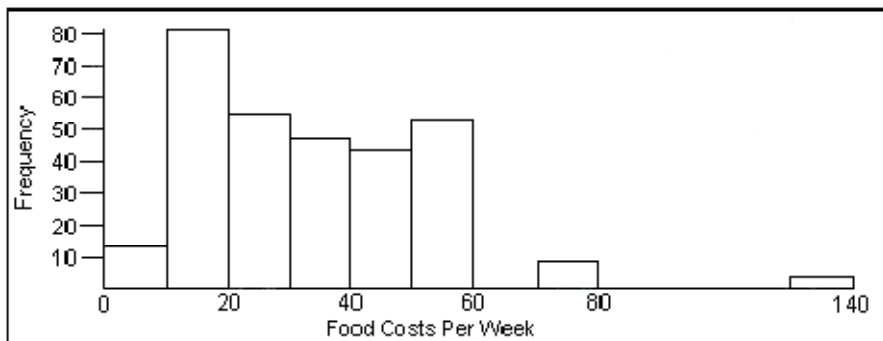
    a. Agree, it looks pretty symmetric if you ignore the outlier.
    b. Agree, most distributions are bell-shaped.
    c. Disagree, it looks more skewed to the left.
    d. Disagree, it looks more skewed to the right.
    e. Disagree, it looks more bimodal.

**Items 13 and 14 refer to the following situation:**
A local running club has its own track and keeps accurate records of each member's individual best lap time around the track, so members can make comparisons with their peers. Here are graphs of these data.

Graph A — FREQUENCY vs LAP TIME

Graph B — LAP TIME (JOHN, MARY, SUE, TOM, JEFF, LILLY, JOAN, BETH, BOB, ANN, ALAN, GEORGE, CARL, LISA, MARIE, CATHY)

Graph C — box plot, LAP TIME

13. Which of the above graphs allows you to most easily estimate the median running time.
    a. Graph A.
    b. Graph B.
    c. Graph C.
    d. All of the above.

14. Which of the above graphs allows you to most easily see the shape of the distribution of running times?
    a. Graph A.
    b. Graph B.
    c. Graph C.
    d. All of the above.

**Items 15 to 17 refer to the following situation:**
Here is a histogram for a set of test scores from a 10-item makeup quiz given to a group of students who were absent on the day the quiz was given.
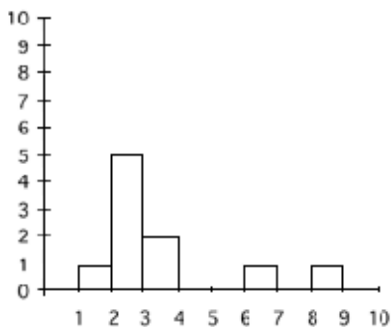
15. What do the numbers on the vertical axis represent?
    a. Independent variable
    b. Scores on the test
   c. Dependent variable
    d. Number of Students

16. How many people received scores higher than 4?
     a. 1
     b. 2
     c. 3
     d. 4
17. How many people took the test and have scores represented in the graph?
     a. 5
     b. 10
     c. 20

18. Select the description that best represents the shape of the following distribution.
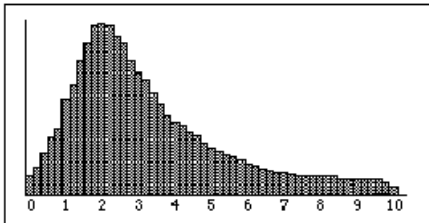
     a. Left (negatively) skewed
     b. Right (positively) skewed
   c. Normal leaning right
     d. Normal leaning left

19. Select the description that best represents the shape of the following distribution.
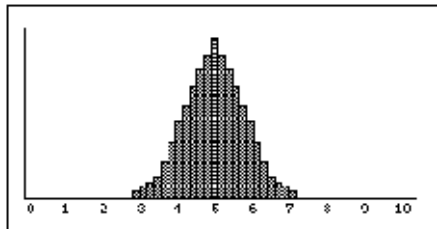
     a. Normal
   b. Skewed
     c. Bimodal
     d. Uniform

Items 20 and 21 refer to the following situation:
One of the items on the student survey for an introductory statistics course was "Rate your aptitude to succeed in this class on a scale of 1 to 10" where 1 = Lowest Aptitude and 10 = Highest Aptitude. The instructor examined the data for men and women separately. Below is the distribution of this variable for the 30 women in the class.
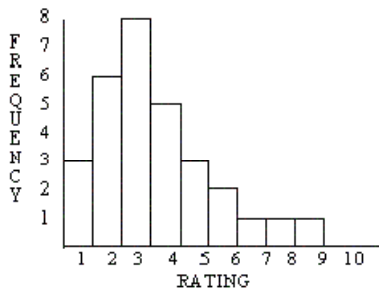
20. How should the instructor interpret the women's perceptions regarding their success in the class?
    a. A majority of women in the class do not feel that they will succeed in statistics although a few feel confident about succeeding.
    b. The women in the class see themselves as having lower aptitude for statistics than the men in the class.
    c. If you remove the three women with the highest ratings, then the result will show an approximately normal distribution.

21. Which of the following boxplots represents the same data set as the histogram shown above?
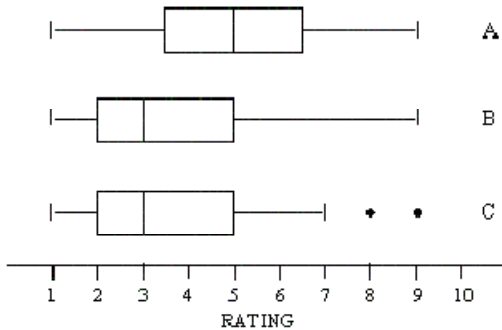

    a. Graph A.
    b. Graph B.
    c. Graph C.

22. The school committee of a small town wanted to determine the average number of children per household in their town. They divided the total number of children in the town by 50, the total number of households. Which of the following statements must be true if the average children per household is 2.2 children?
    a. Half the households in the town have more than 2 children.
    b. There are a total of 110 children in the town.
    c. The most common number of children in a household is 2.2.
    d. None of the above.

23. The distribution of the top 1% of individual incomes in the US is strongly skewed to the right. In 1997, the two measures of center for the top 1% of individual incomes were $330,000 and $675,000. Which number represents the mean income of the top 1% and which number represents the median income of the top 1%? Choose the best answer.
    a. $330,000 is the mean and $675,000 is the median.
    b. $330,000 is the median and $675,000 is the mean.
    c. Not enough information to tell which is which.

24 . For this graphical display of Quiz Scores, which estimates of the mean and median are most plausible?
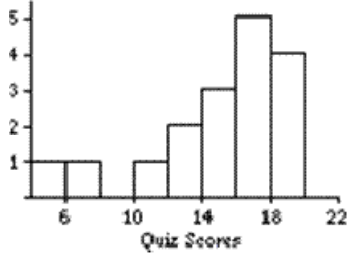
        a. median = 13.0 and mean = 12.0
        b. median = 14.0 and mean = 15.0
        c. median = 16.0 and mean = 14.3
        d. median = 16.5 and mean = 16.2

25. You give a test to 100 students and determine the median score. After grading the test, you realize that the 10 students with the highest scores did exceptionally well.
    You decide to award these 10 students a bonus of 5 more points. The median of the new score distribution will be _____ that of the original score distribution.
        a. lower than
        b. equal to
        c. higher than
        d. depending on skewness, higher or lower than

Items 26 and 27 refer to the following situation:
A college statistics class conducted a survey of how students spend their money. They gathered data from a large random sample of college students who estimated how much money they typically spent each week in different categories (e.g., food, entertainment, etc.). The following statistics were calculated for money spent weekly on food: mean = $31.52; median = $30.00; interquartile range = $34.00; standard deviation = $21.60; range = $132.50.

26. A student states that the median food cost tells you that a majority of students in this sample spend about $30 each week on food. How do you respond?
        a. Agree, the median is an average and that is what an average tells you.
        b. Agree, $30 is representative of the data.
        c. Disagree, a majority of students spend more than $30.
        d. Disagree, the median tells you only that 50% of the sample spent less than $30 and 50% of the sample spent more.

27. The class determined that a mistake had been made and a value entered as 138 should have been entered as 38. They recalculate all of the statistics. Which of the following would be true?
        a. The value of the median decreases, the value of the mean stays the same.
        b. The values of the median and mean both decrease.
        c. The value of the median stays the same, the value of the mean decreases.

28. A class of 30 introductory statistics students took a 15 item quiz, with each itemworth 1 point. The standard deviation for the resulting score distribution is 0. You know that:
        a. about half of the scores were above the mean.
   b. an arithmetic error must have been made.

c. everyone correctly answered the same number of items.

d. the mean, median, and mode must all be 0.

29. The 30 introductory statistics students took another quiz worth 30 points. On this quiz, the standard deviation of the scores of that quiz was 1 point. Which of the following gives the most suitable interpretation?

    a. all of the individual scores are one point apart

    b. the difference between the highest and lowest score is 1

    c. the difference between the upper and lower quartile is 1

    d. a typical score is within 1 point of the mean
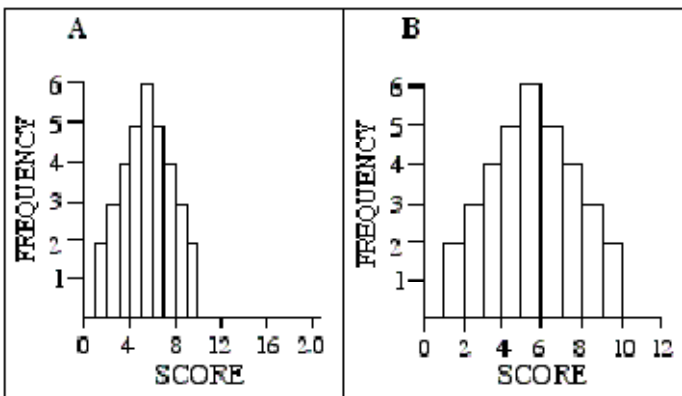
Items 30 and 31 refer to the following situation:

For each pair of graphs, determine which graph has the higher standard deviation (it is not necessary to do any calculations to answer these questions).

30.
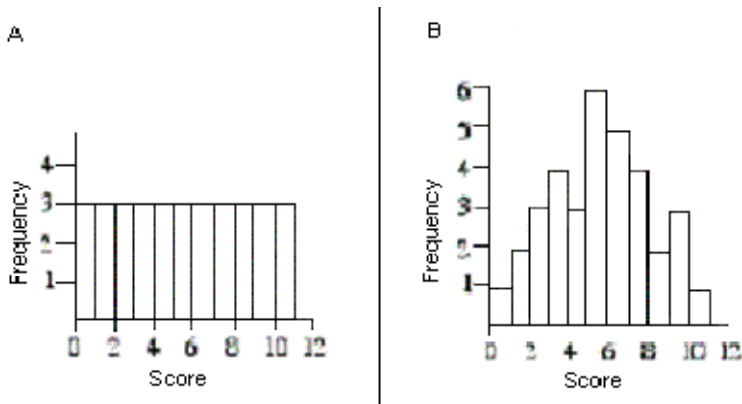
    a. A has a larger standard deviation than B.

    b. B has a larger standard deviation than A.

    c. Both graphs have the same standard deviation.

31.

    a. A has a larger standard deviation than B

    b. B has a larger standard deviation than A

    c. Both graphs have the same standard deviation

32. A teacher gives a 15 item science test. For each item, a student receives one point for a correct answer; 0 points for no answer; and loses one point for an incorrect answer. Total test scores could range from +15 points to -15 points. The teacher computes the standard deviation of the test scores for the class to be -2.30. What do we know?
    a. The standard deviation was calculated incorrectly.
    b. Most students received negative scores.
    c. Most students scored below the mean.
    d. None of the above.

33. Consider two populations in the same state. Both populations are the same size (22,000). Population 1 consists of all students at the State university. Population 2 consists of all residents in a small town. Consider the variable Age. Which population would most likely have the largest standard deviation?
    a. Population 1 would more likely have a higher standard deviation(SD) than Population 2.
    b. Population 2 would more likely have a higher standard deviation(SD) than Population 1.
    c. They would likely have the same standard deviation(SD) for age because they have the same population size.
    d. There is not enough information to tell.

Items 34 and 35 refer to the following situation:
For each list of test scores presented below (List A and List B), select the best estimate for the standard deviation. The mean for each list is 50. No calculations are required to answer these questions.
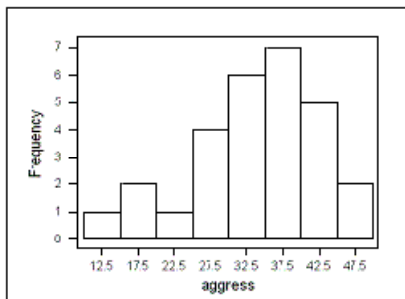34. LIST A: 49, 51, 49, 51, 49, 51, 49, 51, 49, 51
    a. 1
    b. 2
    c. 5

35. LIST B: 31, 36, 48, 50, 50, 53, 54, 56, 60, 62
    a. 1
    b. 3
    c. 8
    d. 20

36. A test to measure aggressive tendencies was given to a group of teenage boys who were members of a street gang. The test is scored from 10 to 60, with a high score indicating more aggression. The histogram represents the results for these 28 boys.
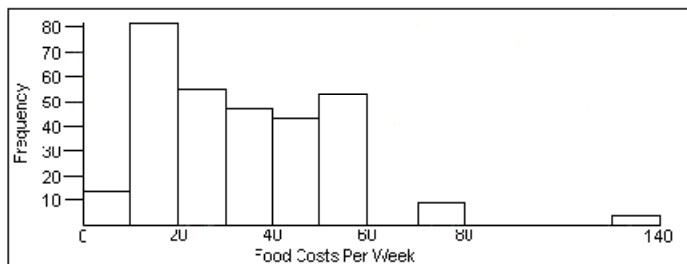

Which two measures would be most appropriate to describe center and spread for this distribution?
    a. Range and mean
    b. Mean and median
    c. Median and IQR
    d. Mean and standard deviation

Items 37 and 38 refer to the following situation:
This is a distribution of how much money was spent per week for a random sample of college students.

37. The range for this distribution is $132.50. Indicate your agreement or disagreement with the following statement: The range is not a useful summary of the variability of this data set.
    a. Agree, it is too vague.
    b. Agree, it is too easily influenced by outliers.
    c. Agree, it does not use information on the center of the data.
    d. Disagree, a range of $132.50 is a good measure of variability because students are apt to spend any amount of money between $0 and $132.50.

38. What is the best measure to use to summarize the variability of this data set?
    a. Range, because it tells you the overall spread of the data.
    b. Standard deviation, because it is based on all the information in the data set.
    c. Standard deviation, because it is the most commonly used measure of variability.
    d. Interquartile range, because it is resistant to outliers.

39. A random sample was taken to determine the left foot length of female bears based on measuring their tracks. The following statistics were calculated for this sample:
    Mean = 12.8 inches, median = 12.7 inches, standard deviation = 1.4 inches, interquartile range = 2 inches.
    The distribution is mound-shaped and symmetric.
    Based only on this information, choose the best estimates for the minimum and maximum values of the distribution.
    a. min = 11.4 and max = 14.2
    b. min = 10.7 and max = 14.7
    c. min = 8.6 and max = 17.0

Items 40 and 41 refer to the following situation:
The 1999 Consumer Reports new Car Buying Guide reported on the number of seconds required for a variety of cars to accelerate from 0 to 30 mph. The cars were also classified into six categories according to type. The following boxplots display the distribution of acceleration times for each type of car:
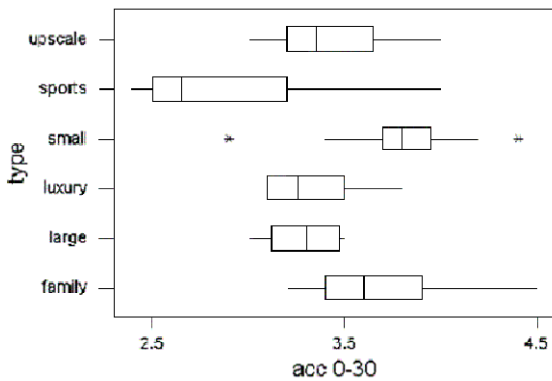
40. Which type of car has the smallest interquartile range for number of seconds to accelerate?
    a. Upscale.
    b. Sports.
    c. Small.
    d. Luxury.
    e. Large.
    f. Family.

41. If the outliers were removed from the dataset of Small cars, which of the following statistics would be least affected?
    a. Range.
    b. IQR.
    c. Standard Deviation.
    d. None of the above.

42. Scores on a standardized test are normally distributed with a mean of 100 and a standard deviation of 20. If these scores are converted to standard normal z scores, which of the following statements will be correct?
    a. Both the mean and median score will equal 0.
    b. The mean will equal 0, but the median cannot be determined.
    c. The mean of the standardized z-scores will equal 100.
    d. The mean of the standardized z-scores will equal 5.

43. A doctor collects a large set of heart rate measurements that approximately follow a normal distribution. He only reports 3 statistics, the mean = 110 beats per minute, the minimum o= 65 beats per minute, and the maximum = 155 beats per minute.
    Which of the following is most likely to be the standard deviation of the distribution?
    a. 5
    b. 15
    c. 35
    d. 90

**Items 44 and 45 refer to the following situation:**
A standardized measure of achievement motivation is normally distributed, with a mean of 35 and a standard deviation of 14. Higher scores correspond to more achievement motivation.

44. Gato scored 49 on this exam. Gato scored higher than what proportion of the population?
	a. 0.32
	b. 0.49
	c. 0.68
	d. 0.84

45. 2.5% of the students scored higher than Shamu. What was her achievement motivation score?
	a. 25
	b. 49
	c. 63
	d. 95

46. The distribution of heights of adult men is approximately normal with a mean of 69 inches and a standard deviation of 2 inches. Bob's height has a Z-score of -.5 when compared to all adult men. Which of the following is true?
	a. Bob is shorter than 69 inches tall.
	b. Bob's height is half of a standard deviation below the mean.
	c. Bob is 68 inches tall.
	d. All of the above.

47. A health insurance company is interested in the cholesterol levels for individuals ages 40 or older. A random sample of 100 individuals was chosen from the target population. The following information was obtained from the sample: average = 158 mg, median = 159 mg, s.d. = 20 mg. One individual has a cholesterol level at 175 mg. Based on only the summary statistics, is this an unusuallly high level of cholesterol for someone from this population? Why or why not?
	a. Yes, because 175 is 17 mg higher than the mean cholesterol level.
	b. Yes, because it is better to have a low cholesterol level.
	c. No, because 175 is less than one standard deviation above the mean.
	d. No, because it is not above 200 mg, the recommended maximum adult cholesterol level.

48. At a large high school, students took a series of advanced tests. Each test covered a different subject area, had a different number of items, and a different mean and standard deviation. One of the students, Marquita, wants to see on which tests she performed the best. Which of the following methods should she use?
	a. Compare her total scores on each test.
	b. Compare her percent correct for each test.
	c. Compare her standardized z score on each test.
	d. All of the above would lead to the same conclusion.

49. Chris is enrolled in a college algebra course and earned a score of 260 on a math achievement test that was given on the first day of class. The instructor looked at two distributions of scores, one is the distribution for all freshmen who took the test, and the other is a distribution for students enrolled in this algebra course. Both are approximately normally distributed and have the same mean, but the distribution for the algebra course has a smaller standard deviation. A z-score is calculated for Chris' test score in both distributions (all freshmen and all freshmen taking algebra). Given that Chris' score is well above the mean, which of the following would be true about these two z-scores?
	a. The z-score based on the distribution for the algebra students would be higher.
	b. The z-score based on the distribution for all freshmen would be higher.
	c. The two z-scores would be the same

50. Sam is interested in bird nest construction, and finds a correlation of .82 between the depth of a bird nest (in inches) and the width of the bird nest (in inches) at its widest point. Sue, a classmate of Sam, is also interested in looking at bird nest construction, and measures the same variables on the same bird nests that Sam does, except she does her measurements in centimeters, instead of inches. What should her correlation be?

    a. Sue's correlation should be 1, because it will match Sam's exactly.

    b. Sue's correlation would be 1.64(.82) = 1.3448, because you need to change the units from inches to centimeters and 1 inch = 1.64 centimeters.

    c. Sue's correlation would be .82, the same as Sam's.

51. The correlation between height and weight for a certain breed of plant is found to be 0.75. What percentage of the variability in plant weight is NOT explained by height?

    a. 1-.75 = .25 or 25%

    b. (.75)2 = .5625 or 56.25%

    c. 1-(.75)2 = .4375 or 43.75%

    d. (1-.75)2 = .0625 or 6.25%

52. A student was studying the relationship between how much money students spend on food and on entertainment per week. Based on a sample size of 270, he calculated a correlation coefficient (r) of .013 for these two variables. Which of the following is an appropriate interpretation?

    a. This low correlation of .013 indicates there is no relationship.

    b. There is no linear relationship but there may be a nonlinear relationship.

    c. This correlation indicates there is some type of linear relationship.

Items 53 to 54 refer to the following situation:
Consider the five scatterplots that are shown below:
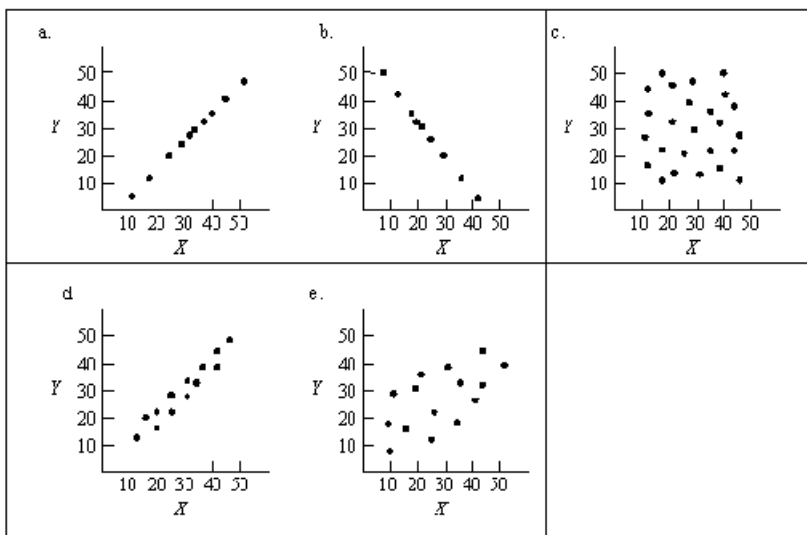
53. Select the scatterplot that shows a correlation of zero?

    a. a

    b. b

    c. c

    d. d

    e. e

54. Select the scatterplot that shows a correlation of about .60?
   a. a
      b. b
      c. c
      d. d
      e. e

55. Select the scatterplot that shows the strongest relationship between the X and Y variables?
      a. a
      b. b
      c. a and b
      d. a and d
      e. a, b, and d

Items 56 and 57 refer to the following situation:
A statistics instructor produces the following scatterplot and regression line to see if her students' exam scores can be predicted from their scores on a standard test of mathematical ability.
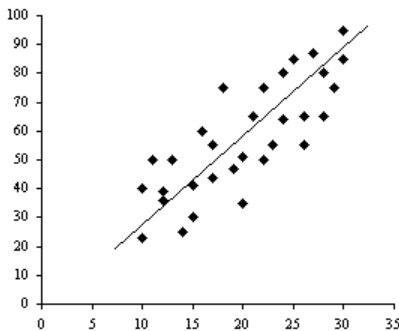
56. What do the numbers on the horizontal axis represent?
      a. Statistics exam scores
      b. The number of people earning each exam score
      c. The response variable
      d. Mathematics ability scores

57. What do the numbers on the vertical axis represent?
      a. Statistics exam scores
      b. The number of people earning each exam score
      c. Mathematics ability scores

58. A random sample of 25 Real Estate listings for houses in the Northeast section of a large city was selected from the city newspaper. A correlation coefficient of -.80 was found between the age of a house and its list price. Which of the following statements is the best interpretation of this correlation?
      a. Older houses tend to cost more money than newer houses.
      b. Newer houses tend to cost more money than older houses.
      c. Older houses are worth more because they were built with higher quality materials and labor.
      d. New houses cost more because supplies and labor are more expensive today.

Items 59 to 60 refer to the following situation:

Dr. Jones gave students in her class a pretest about statistical concepts. After teaching about hypotheses tests, she then gave them a posttest about statistical concepts. Dr. Jones is interested in determining if there is a relationship between pretest and posttest scores, so she constructed the following scatterplot and calculated the correlation coefficient.
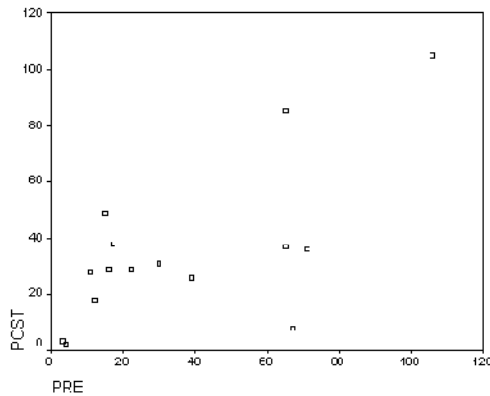

59. Which of the following is the best interpretation of the scatterplot?
    a. There is a moderate positive correlation between pretest and posttest scores.
    b. There is no correlation between pretest and posttest scores.
    c. All of the students' scores increased from pretest to posttest.

60. Locate the point that shows a pretest score of 107. This point, which represents John's scores, is actually incorrect. If John's scores are removed from the data set, how would the correlation coefficient be affected?
    a. The value of the correlation would decrease.
    b. The value of the correlation would increase.
    c. The value of the correlation would stay the same.

61. It turns out that John's pretest score was actually 5, and his posttest score was 100. If this correction is made to the data file and a new correlation coefficient is calculated, how would you expect this correlation to compare to the original correlation?
    a. The absolute value of the new correlation would be smaller than the absolute value of the original correlation.
    b. The absolute value of the new correlation would be larger than the absolute value of the original correlation.
    c. The absolute value of the new correlation would be the same as the absolute value of the original correlation.
    d. It is impossible to predict how the correlation would change.

62. A statistics student gathered data on a large numbers of cars of a particular model, from new cars to those that were up to 10 years old. Using the data on car ages (in years) and car prices (in US dollars) he found a linear relationship and produced the following regression model:
    Predicted Price = 5620 - 440 * Age
    A friend asked him to predict the price of a 5 year old model of this car, using his equation. Which of the following is the most correct response to provide?
    a. Plot a regression line, find 5 on the horizontal axis, and read off the corresponding value on the y axis.
    b. Substitute 5 in the equation and solve for "price".

c. Both of these methods are correct.

d. Neither of these methods is correct.

63. A statistics instructor wants to use the number of hours studied to predict exam scores in his class. He wants to use a linear regression model. Data from previous years shows that the average number of hours studying for a final exam in statistics is 8.5, with a standard deviation of 1.5, and the average exam score is 75, with a standard deviation of 15. The correlation is .76. Should the instructor use linear regression to predict exam scores for a student who studied 10 hours for the final?

a. Yes, there is a high correlation, so it is alright to use linear regression.

b. Yes, because linear regression is the statistical method used to make predictions when you have bivariate quantitative data.

c. Linear regression could be appropriate if the scatterplot shows a clear linear relationship.

d. No, because there is no way to prove that more hours of study causes higher exam scores.

64. Figure A represents the weights for a sample of 26 pebbles, each weighed to the nearest gram. Figure B represents the mean weights of a random sample of 3 pebbles each, with the mean weights rounded to the nearest gram. One value is circled in each distribution. Is there a difference between what is represented by the X circled in A and the X circled in B? Please select the best answer from the list below.
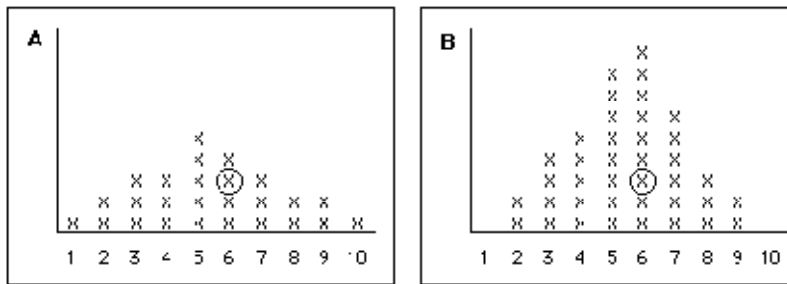

a. No, in both Figure A and Figure B, the X represents one pebble that weights 6 grams.

b. Yes, Figure A has a larger range of values than Figure B.

c. Yes, the X in Figure A is the weight for a single pebble, while the X in Figure B represents the average weight of 3 pebbles.

65. In a geology course, students were learning to use a balance scale to make accurate weighings of rock samples. One student plans to weigh a rock 20 times and then calculate the average of the 20 measurements to estimate her rock's true weight. A second student plans to weigh a rock 5 times and calculate the average of the 5 measurements to estimate his rock's true weight. Which student is more likely to come the closest to the true weight of the rock he or she is weighing?

a. The student who weighed the rock 20 times.

b. The student who weighed the rock 5 times.

c. Both averages would be equally close to the true weight.

66. Suppose half of all newborns are girls and half are boys. Hospital A, a large city hospital, records an average of 50 births a day. Hospital B, a small, rural hospital, records an average of 10 births a day. On a particular day, which hospital is less likely to record 80% or more female births?

a. Hospital A (with 50 births a day), because the more births you see, the closer the proportions will be to .5.

b. Hospital B (with 10 births a day), because with fewer births there will be less variability.

c. The two hospitals are equally likely to record such an event, because the probability of a boy does not depend on the number of births.

67. A random sample of 25 college statistics textbook prices is obtained and the mean price is computed. To determine the probability of finding a more extreme mean than the one obtained from this random sample, you would need to refer to:

a. the population distribution of all college statistics textbook prices.

b. the distribution of prices for this sample of college statistics textbooks.

c. the sampling distribution of textbook prices for all samples of 25 textbooks from this population.

68. Consider the distribution of average number of hours that college students spend sleeping each weeknight. This distribution is very skewed to the right, with a mean of 5 and a standard deviation of 1. A researcher plans to take a simple random sample of 18 college students. If we were to imagine that we could take all possible random samples of size 18 from the population of college students, the sampling distribution of average number of hours spent sleeping will have a shape that is

a. Exactly normal.

b. Less skewed than the population.

c. Just like the population (i.e., very skewed to the right).

d. It's impossible to predict the shape of the sampling distribution.

69. Imagine you have a huge jar of candies that are a generic version of M&Ms. We know that 40% of the candies in the jar are brown. Imagine that you create a sample by randomly pulling 20 candies out of the jar. If you repeated this 10 times to create 10 samples, each with 20 candies, about how many browns would you expect to find in each of the 10 samples?

a. Each sample would have exactly 8 brown candies.

b. Most of the samples would have 0 to 8 brown candies.

c. Most of the samples would have 8 to 20 brown candies.

d. Most of the samples would have 6 to 10 brown candies.

e. You are just as likely to get any count of brown candies between 0 and 20.

Items 70 and 71 refer to the following situation:
The distribution for a population of measurements is presented below.
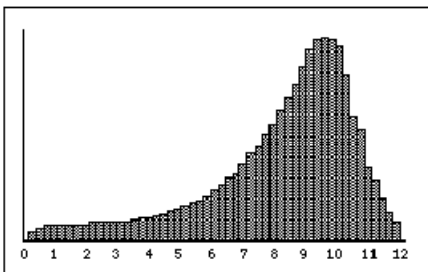
70. A sample of 10 randomly selected values will be taken from the population and the sample mean will be calculated. Which of the following intervals is MOST likely to include the sample mean?

a. 4 to 6

b. 7 to 9

c. 10 to 12

71. Another sample of 10 randomly selected values will be taken from the population and the sample mean will be calculated. Which of the following intervals is LEAST likely to include the sample mean?

      a. 0 to 3
      b. 4 to 7
      c. 8 to 11

Items 72 to 77 refer to the following situation:

A hypothetical distribution for a population of test scores is displayed below. The population has a mean of 60.4, a median of 62.8, and a standard deviation of 6.404. Each of the other four graphs labeled A to D represent possible distributions of sample means for random samples drawn from the population.
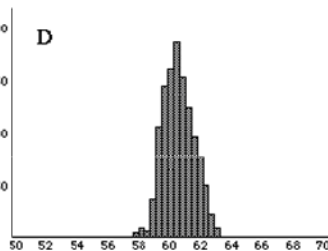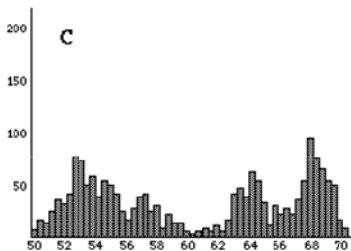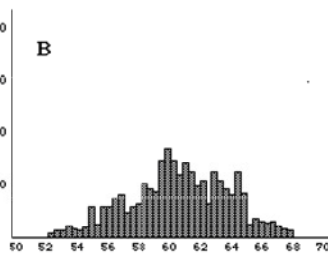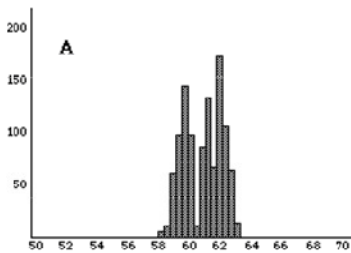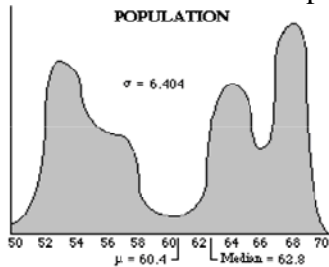
72. Which graph best represents a distribution of sample means for 1000 samples of size 4?

      a. A
      b. B
      c. C
      d. D

73. What do you expect for the shape of the sampling distribution (the distribution of sample means for all possible samples of size n = 4)?

      a. Shaped more like a normal distribution then like the population distribution.
      b. Shaped more like the population distribution then like a normal distribution.
      c. Shaped like neither the population or the normal distribution.

74. What do you expect for the variability (spread) of the sampling distribution?

      a. Same as the population.

b. Less variability than the population (a narrower distribution).

c. More variability than the population (a wider distribution).

75. Which graph best represents a distribution of sample means for 1000 samples of size 50?

   a. A

   b. B

   c. C

   d. D

76. What do you expect for the shape of the sampling distribution (the distribution of sample means for all possible samples of size n = 50)?

   a. Shaped more like a normal distribution.

   b. Shaped more like the population.

   c. Shaped like neither the population or the normal distribution.

77. What do you expect for the variability (spread) of the sampling distribution?

   a. Same as the population.

   b. Less variability than the population (a narrower distribution).

   c. More variability than the population (a wider distribution).

78. Two different samples will be taken from the same population of test scores where the population mean and standard deviation are unknown. The first sample will have 25 data values, and the second sample will have 64 data values. A 95% confidence interval will be constructed for each sample to estimate the population mean. Which confidence interval would you expect to have greater precision (a smaller width) for estimating the population mean?

   a. I expect the confidence interval based on the sample of 64 data values to be more precise.

   b. I expect both confidence intervals to have the same precision.

   c. I expect the confidence interval based on the sample of 25 data values to be more precise.

79. A 95% confidence interval is computed to estimate the mean household income for a city. Which of the following values will definitely be within the limits of this confidence interval?

   a. The population mean

   b. The sample mean

   c. The standard deviation of the sample mean

   d. None of the above

80. Each of the 110 students in a statistics class selects a different random sample of 35 Quiz scores from a population of 5000 scores they are given. Using their data, each student constructs a 90% confidence interval for μ the average Quiz score of the 5000 students. Which of the following conclusions is correct?

   a. About 10% of the sample means will not be included in the confidence intervals.

   b. About 90% of the confidence intervals will contain μ .

   c. It is probable that 90% of the confidence intervals will be identical.

   d. About 10% of the raw scores in the samples will not be found in these confidence intervals.

81. A 95% confidence interval for the mean reading achievement score for a population of third grade students is (43, 49). The margin of error of this interval is:

   a. 5

b. 3

c. 6

82. Justin and Hayley conducted a mission to a new planet, Planet X, to study arm length. They took a random sample of 100 Planet X residents and calculated a 95% confidence interval for the mean arm length. What does a 95% confidence interval for arm length tell us in this case? Select the best answer:

a. I am 95% confident that this interval includes the sample mean (x) arm length.

b. I am confident that most (95%) of all Planet X residents will have an arm length within this interval.

c. I am 95% confident that most Planet X residents will have arm lengths within this interval.

d. I am 95% confident that this interval includes the population mean arm length.

83. Suppose that a random sample of 41 state college students is asked to measure the length of their right foot in centimeters. A 95% confidence interval for the mean foot length for students at this university turns out to be (21.709, 25.091). If instead a 90% confidence interval was calculated, how would it differ from the 95% confidence interval?

a. The 90% confidence interval would be narrower.

b. The 90% confidence interval would be wider.

c. The 90% confidence interval would be the same as the 95% confidence interval.

84. A pollster took a random sample of 100 students from a large university and computed a confidence interval to estimate the percentage of students who were planning to vote in the upcoming election. The pollster felt that the confidence interval was too wide to provide a precise estimate of the population parameter. What could the pollster have done to produce a narrower confidence interval that would produce a more precise estimate of the percentage of all university students who plan to vote in the upcoming election?

a. Increase the sample size to 150.

b. Increase the confidence level to 99%.

c. Both a and b

d. None of the above

85. A newspaper article states with 95% confidence that 55% to 65% of all high school students in the United States claim that they could get a hand gun if they wanted one. This confidence interval is based on a poll of 2000 high school students in Detroit. How would you interpret the confidence interval from this newspaper article?

a. 95% of large urban cities in the United States have 55% to 65% high school students who could get a hand gun.

b. If we took many samples of high school students from different urban cities, 95% of the samples would have between 55% and 65% high school students who could get hand guns.

c. You cannot use this confidence interval to generalize to all teenagers in the United States because of the way the sample was taken.

d. We can be 95% confident that between 55% and 65% of all United States high school students could get a hand gun.

86. The Gallup poll (August 23, 2002) reported that 53% of Americans said they would favor sending American ground troops to the Persian Gulf area in an attempt to remove Hussein from power. The

poll also reported that the "margin of error" for this poll was 4%. What does the margin of error of 4% indicate?

    a. There is a 4% chance that the estimate of 53% is wrong.

    b. The percent of Americans who are in favor is probably higher than 53% and closer to 57%.

    c. The percent of Americans who are in favor is estimated to be between 49% and 57%.

87. Suppose two researchers want to estimate the proportion of American college students who favor abolishing the penny. They both want to have about the same margin of error to estimate this proportion. However, Researcher 1 wants to estimate with 99% confidence and Researcher 2 wants to estimate with 95% confidence. Which researcher would need more students for her study in order to obtain the desired margin of error?

    a. Researcher 1.

    b. Researcher 2.

    c. Both researchers would need the same number of subjects.

    d. It is impossible to obtain the same margin of error with the two different confidence levels.

88. The makers of Mini-Oats cereal have an automated packaging machine that is set to fill boxes with 24 ounces of cereal. At various times in the packaging process, a random sample of 100 boxes is taken to see if the machine is filling the boxes with an average of 24 ounces of cereal. Which of the following is a statement of the null hypothesis being tested?

    a. The machine is filling the boxes with the proper amount of cereal.

    b. The machine is not filling the boxes with the proper amount of cereal.

    c. The machine is not putting enough cereal in the boxes.

89. A research article gives a p-value of .001 in the analysis section. Which definition of a p-value is the most accurate?

    a. the probability that the observed outcome will occur again.

    b. the probability of observing an outcome as extreme or more extreme than the one observed if the null hypothesis is true.

    c. the value that an observed outcome must reach in order to be considered significant under the null hypothesis.

    d. the probability that the null hypothesis is true.

90. If a researcher was hoping to show that the results of an experiment were statistically significant they would prefer:

    a. a large p-value

    b. a small p-value

    c. p-values are not related to statistical significance

91. A researcher compares men and women on 100 different variables using a two-sample t-test. He sets the level of significance to .05 and then carries out 100 independent t-tests (one for each variable) on these data. If, for each test, the null hypothesis actually is true, about how many "statistically significant" results will be produced?

    a. 0

    b. 5

    c. 10

    d. none of the above

92. A newspaper article claims that the average age for people who receive food stamps is 40 years. You believe that the average age is less than that. You take a random sample of 100 people who receive food stamps, and find their average age to be 39.2 years. You find that this is significantly lower than the age of 40 stated in the article ($p < .05$). What would be an appropriate interpretation of this result?

    a. The statistically significant result indicates that the majority of people who receive food stamps is younger than 40.

    b. Although the result is statistically significant, the difference in age is not of practical importance.

    c. An error must have been made. This difference is too small to be statistically significant.

93. A newspaper article stated that the US Supreme Court received 812 letters from around the country on the subject of whether to ban cameras from the courtroom. Of these 812 letters, 800 expressed the opinion that cameras should be banned. A statistics student was going to use this sample information to conduct a test of significance of whether more than 95% of all American adults feel that cameras should be banned from the courtroom. What would you tell this student?

    a. This is a large enough sample to provide an accurate estimate of the American public's opinion on the issue.

    b. The necessary conditions for a test of significance are not satisfied, so no statistical test should be performed.

    c. With such a large number of people favoring the notion that cameras be banned, there is no need for a statistical test.

94. A researcher conducts an experiment on human memory and recruits 15 people to participate in her study. She performs the experiment and analyzes the results. She obtains a p-value of .17. Which of the following is a reasonable interpretation of her results?

    a. This proves that her experimental treatment has no effect on memory.

    b. There could be a treatment effect, but the sample size was too small to detect it.

    c. She should reject the null hypothesis.

    d. There is evidence of a small effect on memory by her experimental treatment.

95. It is reported that scores on a particular test of historical trivia given to high school students are approximately normally distributed with a mean of 85. Mrs. Rose believes that her 5 classes of high school seniors will score significantly better than the national average on this test. At the end of the semester, Mrs. Rose administers the historical trivia test to her students. The students score an average of 89 on this test. After conducting the appropriate statistical test, Mrs. Rose finds that the p-value is .0025. Which of the following is the best interpretation of the p-value?

    a. A p-value of .0025 provides strong evidence that Mrs. Rose's class outperformed high school students across the nation.

    b. A p-value of .0025 indicates that there is a very small chance that Mrs. Rose's class outperformed high school students across the nation.

    c. A p-value of .0025 provides evidence that Mrs. Rose is an exceptional teacher who was able to prepare her students well for this national test.

    d. None of the above.