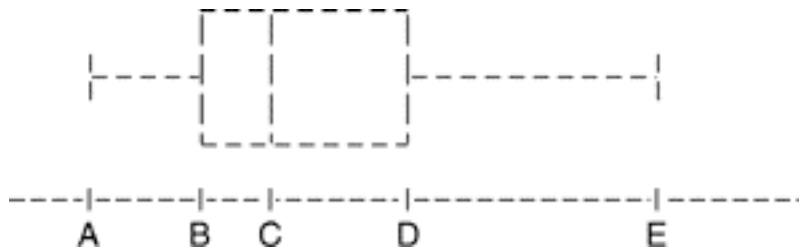


Use the following for 7-10. The length of pregnancy isn't always the same. In pigs, the length of pregnancies varies according to a normal distribution with mean 114 days and standard deviation 5 days.

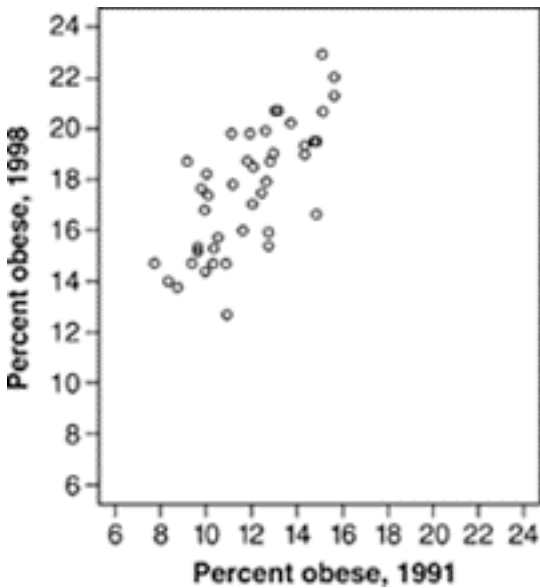
7. What range covers the middle 95% of pig pregnancies?
A) 109 to 119 days C) 99 to 129 days
B) 104 to 124 days D) 94 to 134 days
8. What percent of pig pregnancies are longer than 114 days?
A) 16% B) 34% C) 50% D) 84%
9. What percent of pig pregnancies are longer than 109 days?
A) 16% B) 34% C) 50% D) 84%
10. The median length of a pig pregnancy is
A) 119 days.
B) 114 days.
C) 109 days.
D) between 109 and 119 days, but can't be more specific.
E) greater than 114 days, but can't be more specific.
11. A set of measurements has this boxplot:



- Which point on this boxplot is the first quartile of the distribution?
A) A B) B C) C D) D E) E
12. If the least squares regression line for predicting y from x is $y = 500 - 20x$, what is the predicted value of y when $x = 10$?
A) 300 B) 500 C) 200 D) 700 E) 20

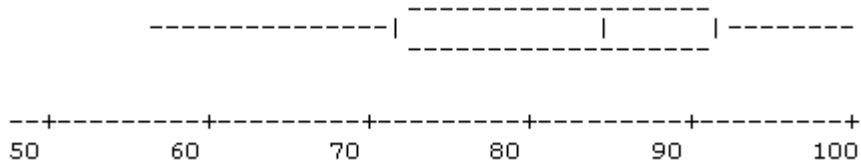
13. There is a close relationship between the correlation r and the slope b of the least-squares regression line. In particular, it is true that
- A) r and b always have the same sign, which shows whether the variables are positively or negatively associated.
 - B) r and b both always take values between -1 and 1.
 - C) the slope b is always at least as large as the correlation r .
 - D) the slope b is always equal to r^2 , the square of the correlation.
 - E) both (A) and (B) are true.
14. For categorical variables (like a person's occupation or sex) it only makes sense to speak of _____. It makes no sense to speak of the _____ or the _____ sex. The sequence of words to correctly complete this statement is
- A) quartiles, mean, median
 - B) median, mean, standard deviation
 - C) standard deviation, mean, median
 - D) counts or percents, mean, median
15. You are told that your score on an exam is at the 85 percentile of the distribution of scores. This means that
- A) your score was equal to or lower than approximately 85% of the people who took this exam.
 - B) your score was equal to or higher than approximately 85% of the people who took this exam.
 - C) you answered 85% of the questions correctly.
 - D) your score was the same as 85% of the people who took this exam.
 - E) you are 85% confident in your score is significant.
16. You are given the following data: 4, 6, 9, 3, 999. You calculate the mean and standard deviation (*don't* do it!). Then Professor Stats tells you to throw out the outlier 999 and redo your calculation without this value. In your revised calculation, the mean is _____, and the standard deviation is _____ than in the original calculation. The sequence of words to correctly complete this sentence is:
- A) smaller, smaller
 - B) smaller, larger
 - C) larger, smaller
 - D) larger, larger
 - E) Either (A) or (B) could be true.
17. Scores on the American College Testing (ACT) college entrance exam follow the normal distribution with mean 18 and standard deviation 6. Wayne's standard score on the ACT was -0.7 . What was Wayne's actual ACT score?
- A) 4.2
 - B) -4.2
 - C) 13.8
 - D) 22.2
 - E) 9.6

Below is a graph of the percent of adults in each state who were obese in 1991 and the percent who were obese in 1998:



18. This type of graph is called a
 A) boxplot. B) histogram. C) line graph. D) scatterplot. E) stemplot.
19. Which of these is a reasonable value of the correlation r for the data in this graph?
 A) $r = 0$ B) $r = 0.3$ C) $r = 0.7$ D) $r = 0.95$ E) $r = 1$

Here is a boxplot of some exam scores.



20. Approximately what percent of the scores are between 72 and 92?
 A) 50% B) 60% C) 70% D) 80% E) can't say
21. From the above boxplot, a student whose score is such that 60% of those taking the exam scored less than this student must have scored
 A) below 72. D) above 92.
 B) between 72 and 85. E) We can't say.
 C) between 85 and 92.

22. The distribution of heights of adult men is approximately normal. A man whose standardized height is 0.3 is at the 61.79th percentile. What percent of all men are taller than he is?
- A) 61.79%
 - B) 38.21%
 - C) 0.3%
 - D) We can't tell from the given information.
 - E) We can't tell without the normal table from the text.

23. One of the following 12 scores

84 76 92 92 88 96 68 80 92 88 76 96

was omitted from the stemplot below:

```

6 | 8
7 | 66
8 | 0488
9 | 2266

```

The missing number is

- A) 76.
 - B) 88.
 - C) 92.
 - D) 96.
 - E) 22266.
24. Here is a set of data: 2, 8, 10, 4, 6. The standard deviation is
- A) 6
 - B) 3.16
 - C) 10.
 - D) 2.
 - E) 40

Here are the number of hours that each of a group of students studied for this exam:

2 4 22 6 1 4 1 5 7 4

25. What is the mean number of study hours?
- A) 2.5
 - B) 4
 - C) 4.5
 - D) 5.6
 - E) 56
26. You calculate that the standard deviation of a set of observations is 0. This tells you that
- A) you made an arithmetic mistake.
 - B) all the observations have the same value.
 - C) there is no straight-line association.
 - D) the mean must also be 0.
 - E) Both (B) and (D) are true.
27. To display the percent of adult women who were in the work force in each of the years 1975 to 2000, a good choice of a graph would be a
- A) boxplot.
 - B) histogram.
 - C) line graph.
 - D) pie chart.
 - E) stemplot.

Use for 28-31. A study gathers data on the outside temperature during the winter, in degrees Fahrenheit, and the amount of natural gas a household consumes, in cubic feet per day. Call the temperature x and gas consumption y . The house is heated with gas, so x helps explain y . The least-squares regression line for predicting y from x is $y = 1344 - 19x$.

28. On a day when the temperature is 20°F , the regression line predicts that gas used will be about
- A) 1,724 cubic feet
 - B) 1,383 cubic feet
 - C) 1,325 cubic feet
 - D) 964 cubic feet
 - E) None of these.
29. We can see from the equation of the line that
- A) as the temperature x goes up, gas used y goes up, because the slope 1,344 is positive.
 - B) as the temperature x goes up, gas used y goes up, because the slope 19 is positive.
 - C) as the temperature x goes up, gas used y goes down because the slope 1,344 is bigger than 19.
 - D) as the temperature x goes up, gas used y goes down, because the slope -19 is negative.
30. When the temperature goes up 1 degree, what happens to the gas usage predicted by the regression line?
- A) It goes up 1 cubic foot.
 - B) It goes down 1 cubic foot.
 - C) It goes up 19 cubic feet.
 - D) It goes down 19 cubic feet.
 - E) Can't tell without seeing the data.
31. The correlation between temperature x and gas usage y is $r = -0.7$. Which of the following would *not* change r ?
- A) Measuring temperature in degrees Celsius instead of degrees Fahrenheit.
 - B) Removing two outliers from the data used to calculate r .
 - C) Measuring gas usage in hundreds of cubic feet, so that all values of y are divided by 100.
 - D) Both (A) and (C).
 - E) All of (A), (B), and (C).
32. Which correlation indicates a strong positive straight line relationship?
- A) 0.4
 - B) -0.75
 - C) 1.5
 - D) 0.0
 - E) 0.99

33. A well-drawn histogram should have
- A) bars all the same width.
 - B) no space between bars (unless a class has no observations).
 - C) a clearly marked vertical scale.
 - D) all of these.
34. The United Nations also has data on the percent of adult females who are illiterate in each of these 142 countries. The correlation between male illiteracy rate and female illiteracy rate is $r = 0.945$. This tells us that
- A) countries with high male illiteracy tend to also have high female illiteracy, and the relationship is very strong.
 - B) countries with high male illiteracy tend to also have high female illiteracy, but the two are only weakly related.
 - C) countries with high male illiteracy tend to have low female illiteracy, and the relationship is very strong.
 - D) countries with high male illiteracy tend to have low female illiteracy, but the two are only weakly related.
 - E) there is very little relationship between the illiteracy rates for males and females.
35. The risk of an investment is measured by the variability of the changes in its value over a fixed period, such as a year. More variation from year to year means more risk. The government's Securities and Exchange Commission wants to require mutual funds to tell investors how risky they are. A news article (the *New York Times*, April 2, 1995) says that some people think that "the proposed risk descriptions, especially one that goes by the daunting name standard deviation" are hard to understand. Explain to a friend what the standard deviation means, using the fact that the changes in a mutual fund's value over many years have a roughly normal distribution.
- A) The standard deviation is the distance between the first and third quartiles, so it spans half the yearly changes in the fund's value.
 - B) The standard deviation is the largest change we ever expect to see in a year.
 - C) The yearly change in the fund's value will be greater than the standard deviation half the time and less than the standard deviation half the time.
 - D) Start with the average (mean) change in the fund's value over many years; the actual change will be within one standard deviation of that average in about 68% of all years.
 - E) Start with the average (mean) change in the fund's value over many years; the actual change will be within one standard deviation of that average in about 95% of all years.

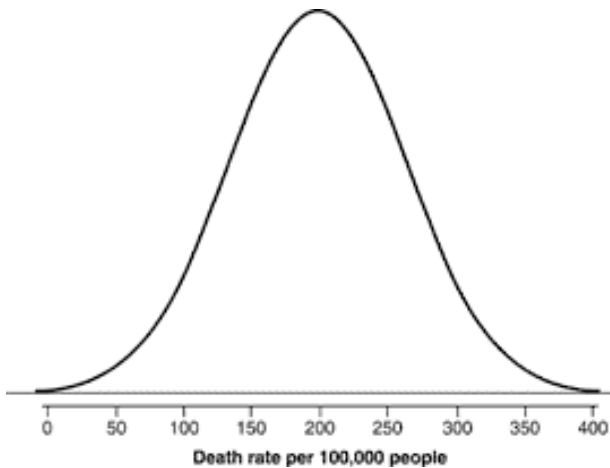
36. A common abuse of bar graphs is
- A) to use bars of equal width, so that we can't see differences among the bars.
 - B) to replace bars by pictures and increase both height and width as the variable plotted increases.
 - C) to use bars that don't touch each other.
 - D) to draw the bars vertically rather than horizontally.
 - E) to stretch or squeeze the scale at the base of the bars, so that our eyes are misled.

The stock market did well during the 1990s. Here are the percent total returns (change in price plus dividends paid) for the Standard & Poor's 500 stock index:

Year	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Return	31.7	-3.1	30.5	7.6	10.1	1.3	37.6	23.0	33.4	28.6

37. The median return during this period is
- A) 5.5.
 - B) 20.07.
 - C) 23.0.
 - D) 25.8.
 - E) 28.6.
38. The third quartile of these returns is
- A) 7.6.
 - B) 30.5.
 - C) 31.1.
 - D) 31.7.
 - E) 33.4.
39. You can roughly locate the mean of a density curve by eye because it is
- A) the point at which the curve would balance if made of solid material.
 - B) the point that divides the area under the curve into two equal parts.
 - C) the point at which the curve reaches its peak.
 - D) the point where the curvature changes direction.
40. The correlation between two variables is -0.8 . We can conclude
- A) one causes the other.
 - B) there is a strong positive association between the two variables.
 - C) there is a strong negative association between the two variables.
 - D) all of the relationship between the two variables can be explained by a straight line.
 - E) there are no outliers.

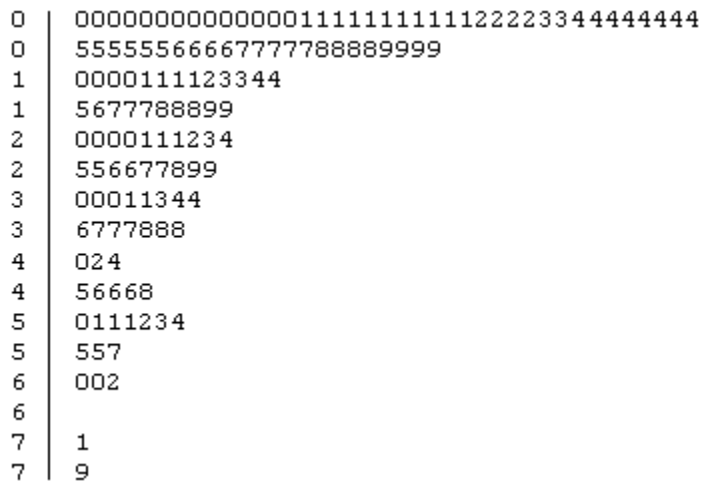
Record the death rate from heart disease per 100,000 people in a group of developed countries. The distribution is roughly described by this normal curve:



41. From this normal curve, we see that the mean heart disease death rate per 100,000 people is about
A) 60. B) 120. C) 190. D) 250. E) 400.
42. From the normal curve, we see that the standard deviation of the heart disease rate per 100,000 people is closest to
A) 25. B) 65. C) 100. D) 200. E) 400.
43. The least-squares regression line for predicting blood alcohol content from number of beers is $y = -0.013 + 0.018x$. The slope 0.018 of this line tells us that
A) the correlation between number of beers and BAC is 0.018.
B) on the average, BAC increases by 0.018 for each additional beer a student drinks.
C) a student who drinks no beer will still have a BAC of 0.018.
D) the average BAC of all the students in the study was 0.018.
44. A high correlation between two variables does not always mean that changes in one cause changes in the other. The best way to get good evidence that cause-and-effect is present is to
A) select a simple random sample from the population of interest.
B) arrange the data in a two-way table.
C) carry out a randomized comparative experiment.
D) make a scatterplot and look for a strong association.
E) make a histogram and look for outliers.

45. A study showed that students who study more hours tend to do better on statistics exams. In fact, number of hours studied explained 81% of the variation in exam scores among the students who participated in the study. What is the correlation between hours studied and exam score?
 A) $r = 0.9$ B) $r = 0.81$ C) $r = 0.656$ D) $r = -0.656$ E) $r = -0.9$

Here is a stemplot of the percent of adult males who are illiterate in 142 countries, according to the United Nations. For example, the highest illiteracy rate was 79%, in the African country Niger.



46. To locate the median of 142 observations, you would count up to what position after arranging the data in order from smallest to largest?
 A) Half-way between the 12th and 13th in the ordered list.
 B) 50th in the ordered list.
 C) 71st in the ordered list.
 D) 72nd in the ordered list.
 E) Half-way between the 71st and 72nd in the ordered list.
47. The mean of this distribution (*don't* try to find it) is certainly
 A) very close to the median.
 B) clearly less than the median.
 C) clearly greater than the median.
 D) can't say because the mean is random.

48. Ten members of a fraternity take a statistics course. Here are their scores on the first exam in the course:

61 74 47 60 62 63 65 79 55 45

To present the distribution of scores in a graph, you might choose
A) a line graph. B) a bar graph. C) a pie chart. D) a stemplot.

A study of home heating costs collects data on the size of houses and the monthly cost to heat the houses with natural gas. Here are the data.

Size of House	Heating Cost
1200 sq ft	\$50
2300 sq ft	\$125
1800 sq ft	\$90
2000 sq ft	\$105

49. Just by looking at the data (*don't* do a calculation) you can see that the correlation between house size and heating cost is
- A) close to zero.
 - B) clearly positive.
 - C) clearly negative.
 - D) not close to zero, but could be either positive or negative.
 - E) makes no sense for these data.
50. According to the student newspaper, the mean salary of male full professors in the School of Management is \$117,302. The median of these salaries
- A) would be lower, because salary distributions are skewed to the left.
 - B) would be lower, because salary distributions are skewed to the right.
 - C) would be higher, because salary distributions are skewed to the left.
 - D) would be higher, because salary distributions are skewed to the right.