1. A company database contains the following information about each employee: age, date hired, sex (male or female), ethnic group (Asian, black, Hispanic, etc.), job category (clerical, management, technical, etc.), yearly salary. Which of the following lists of variables are *all* categorical?

(a) age, sex, ethnic group. (b) sex, ethnic group, job category.

(c) ethnic group, job category, yearly salary. (d) yearly salary, age.

(e) age, date hired.

2. Were the extinctions that occurred in the last ice age more frequent among species of animals with large body sizes? A researcher gathers data on the average body mass (in kilograms) of all species known to have existed at that time. These measurements are values of

(a) a categorical variable. (b) a quantitative variable.

(c) an invalid variable. (d) a margin of error.

3. In the situation of the previous question, what are the explanatory and response variables?

(a) There is no explanatory-response distinction in this situation.

(b) Explanatory: body mass of a species. Response: whether the species went extinct.

(c) Explanatory: the ice age. Response: whether a species went extinct.

(d) Explanatory: whether a species went extinct. Response: the body mass of the species.

(e) Explanatory: the ice age. Response: the body mass of a species.

4. An example of a categorical variable is

(a) the name of the college a student attends. (b) a student's weight in kilograms.

(c) a student's class rank, such as 25th out of 364. (d) a student's sex (male or female).

(e) Both (a) and (d).

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:

**The next five questions** are related to this situation.

5. The median return during this period is

(a) 5.5 (b) 20.07 (c) 23.0 (d) 25.8 (e) 28.6

6. The third quartile of these returns is

(a) 7.6 (b) 30.5 (c) 31.1 (d) 31.7 (e) 33.4

7. The mean return is

(a) 20.07 (b) 20.69 (c) 22.3 (d) 25.8 (e) 33.4

8. The standard deviation of the returns is

(a) 13.75 (b) 13.98 (c) 14.74 (d) 20.07 (e) 25.8

9. You have similar data on returns on common stocks for all years since 1945. To show clearly how returns have changed over time, your best choice of graph is

(a) a bar graph (b) a line graph (c) a pie chart

(d) a histogram (e) a scatterplot

10. 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.

Here are boxplots of the number of calories in 20 brands of beef hot dogs, 17 brands of meat

hot dogs, and 17 brands of poultry hot dogs.



11. The main advantage of boxplots over stemplots and histograms is

(a) boxplots make it easy to compare several distributions, as in this example

(b) boxplots show more detail about the shape of the distribution

(c) boxplots use the five-number summary, whereas stemplots and histograms use the mean and standard deviation

(d) boxplots show skewed distributions, whereas stemplots and histograms show only symmetric distributions

12. This plot shows that

(a) all poultry hot dogs have fewer calories than the median for beef and meat hot dogs

(b) about half of poultry hot dog brands have fewer calories than the median for beef and meat hot dogs

(c) hot dog type is not helpful in predicting calories, because some hot dogs of each type are high and some of each type are low

(d) most poultry hot dog brands have fewer calories than most beef and meat hot dogs, but a few poultry hot dogs have more calories than the median beef and meat hot dog

13. We see from the plot that the median number of calories in a beef hot dog is about

(a) 190 (b) 179 (c) 153 (d) 139 (e) 129

14. The box in each boxplot marks

(a) the full range covered by the data

(b) the range covered by the middle half of the data

(c) the range covered by the middle three-quarters of the data

(d) the span one standard deviation on each side of the mean

(e) the span two standard deviations on each side of the mean

15. The calorie counts for the 17 poultry brands are:

129 132 102 106 94 102 87 99 170 113 135 142 86 143 152 146 144

The median of these values is

(a) 129 (b) 132 (c) 130.5 (d) 121 (e) 170

16. The first quartile of the 17 poultry hot dog calorie counts is

(a) 99 (b) 102 (c) 100.5 (d) 143.5 (e) 143

17. To illustrate a talk you are giving, you want to make a graph to compare the percents of adults in several countries who have finished university. For example, this percent is 9% in France and 24% in the United States. You should make a

(a) bar graph (b) histogram (c) line graph (d) pie chart

18. 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

19. Jorge's score on Exam 1 in his statistics class was at the 64th percentile of the scores for all students. His score falls

(a) between the minimum and the first quartile (b) between the first quartile and the median

(c) between the median and the third quartile (d) between the third quartile and the maximum

20. Which of these statements about the standard deviation *s* is true?

(a) *s* is always 0 or positive.

(b) *s* should be used to measure spread only when the mean  is used to measure center.

(c) *s* is a number that has no units of measurement.

(d) Both (a) and (b), but not (c).

(e) All of (a), (b), and (c).

Here is a table of the undergraduate enrollment at a large state university, broken down by

class:



21. To make a correct graph of the distribution of students by class, you could use

(a) a bar graph. (b) a pie chart. (c) a histogram.

(d) all of (a), (b), and (c). (e) (a) or (b), but not (c).

22. A bar graph compares the size of the armed forces for China, Germany, Russia, and the

United States. To make the graph look nicer, the artist replaces each bar by a picture of a soldier as tall as the bar. This graph is misleading because

(a) it's the wrong kind of graph -- use a pie chart instead.

(b) it's the wrong kind of graph -- use a boxplot instead.

(c) China's armed forces are twice as large as those of the U.S., but the choice of vertical scale can change this to half as large or four times as large.

(d) China's armed forces are twice as large as those of the U.S., but the area of China's soldier picture is four times as large as the U.S. picture.

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

**The next six questions** use these data.

23. What is the median number of study hours?

(a) 2.5 (b) 4 (c) 4.5 (d) 5.6 (e) 56

24. What is the mean number of study hours?

(a) 2.5 (b) 4 (c) 4.5 (d) 5.6 (e) 56

25. What is the third quartile of the number of study hours?

(a) 1.5 (b) 2 (c) 5.5 (d) 6 (e) 6.5

26. What is the standard deviation of the number of study hours?

(a) 5.8 (b) 6.1 (c) 37.2 (d) 33.4 (e) 56

27. The data contain one high outlier (22 hours). Which of your results for the previous four questions would change if this were 7 hours instead of 22 hours? (You do *not* need to calculate new values.)

(a) All four would change.

(b) The mean, the third quartile, and the standard deviation would change.

(c) The mean and the standard deviation would change.

(d) Only the mean would change.

(e) None of the four would change.

28. Which of your results for the same four questions are measured in hours?

(a) All four are measured in hours.

(b) All except the standard deviation are measured in hours.

(c) Only the median and the mean are measured in hours.

(d) Only the median is measured in hours.

(e) None of the four are measured in hours.

29. An outlier will usually have a large effect on the

(a) first quartile (b) mean (c) third quartile (d) median (e) sample size

30. Fifty percent of the observations will be at or above the

(a) maximum (b) mean (c) median (d) third quartile

31. For a distribution that is skewed to the left, usually

(a) the mean will be larger than the median (b) the median will be larger than the mean

(c) the first quartile will be larger than the third quartile (d) the standard deviation will be negative

(e) the minimum will be larger than the maximum

32. For a distribution that is skewed to the right, usually

(a) the mean will be larger than the median (b) the median will be larger than the mean

(c) the first quartile will be larger than the third quartile (d) the standard deviation will be negative

(e) the minimum will be larger than the maximum

33. The five numbers in the five-number summary are

(a) the five smallest observations

(b) the five largest observations

(c) the mean, the median, the maximum, the standard deviation and the sample size

(d) *x*1, *x*2, *x*3, *x*4, and *x*5.

(e) the median, the quartiles, the minimum and the maximum

















