

EXAM QUESTIONS

The *Journal of the American Medical Association* recently published an article titled "The Spread of the Obesity Epidemic in the United States." Here is a plot of the percent of adults who are obese in each of the 45 states that participated in a large study. The data are in percents and tenths of a percent. For example, 19.5% of Indiana adults are obese.

The next eight questions concern this study of obesity.

12	7
13	8
14	047777
15	233479
16	068
17	0456899
18	25777
19	003558899
20	2777
21	3
22	09

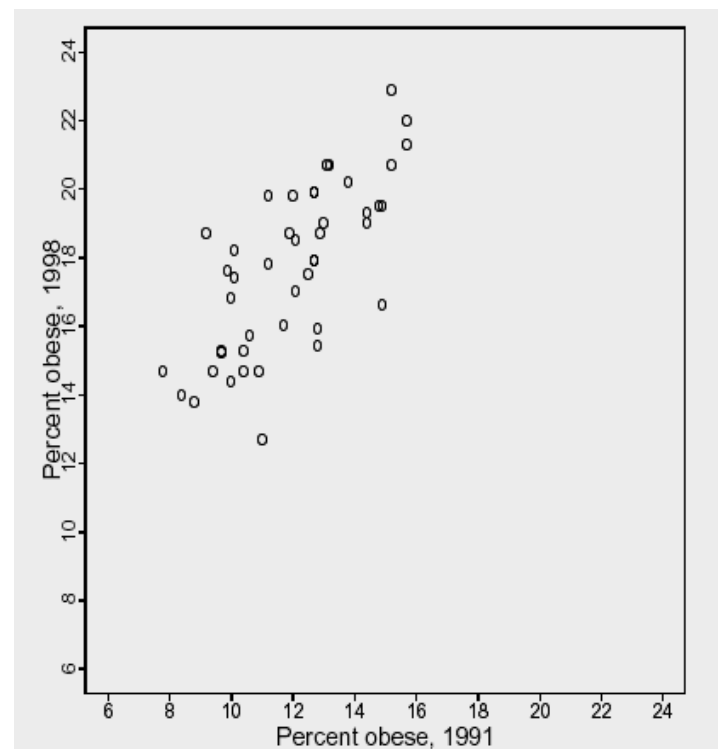
1. This type of graph is called a

- (a) boxplot
- (b) histogram
- (c) line graph
- (d) scatterplot
- (e) stemplot

2. The median percent of obese adults in these 45 states is:

- (a) 17.69%
- (b) 17.8%
- (c) 17.85%
- (d) 17.9%
- (e) 23%

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



3. This type of graph is called a

- (a) boxplot
- (b) histogram
- (c) line graph
- (d) scatterplot
- (e) stemplot

4. 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) $r = 1$

5. Arizona had the lowest percent obese in 1998, 12.7%. About what percent of Arizona adults were obese in 1991?

- (a) 7.8%
- (b) 9.0%
- (c) 11.0%
- (d) 12.7%
- (e) 14.7%

6. The least-squares regression line for predicting 1998 percent obese from 1991 percent obese is $y = 7.4 + 0.86x$. In 1991, 14.8% of Indiana adults were obese. Based on this information, what percent would you predict to be obese in 1998?

- (a) 5.3%
- (b) 7.5%
- (c) 12.7%
- (d) 19.5%
- (e) 20.1%

7. The 11 states in the Midwest have these percents of obese adults:

19.5 19.5 17.9 20.7 17.9 15.7 19.3 19.8 18.7 15.4 17.5

The mean and standard deviation of these values are about:

- (a) 18.35 and 18.7
- (b) 18.35 and 19.5
- (c) 18.7 and 1.68
- (d) 18.35 and 1.68
- (e) 18.35 and 2.00

8. Indiana has 19.5% obese adults. Where does Indiana stand in the distribution for the 11 Midwest states?

- (a) Below the median.
- (b) Exactly at the median.
- (c) Above the median but below the third quartile
- (d) Exactly at the third quartile.
- (e) Above the third quartile.

9. A table of random numbers is used to select 25 students from a large class to rate a CD by the group Wilco. The ratings that these students give are used to estimate the ratings that would be given if the entire class were asked to rate the CD. The average of the ratings of the 25 students in the sample is

- (a) a convenience sample
- (b) a census
- (c) an unbiased estimate of the class rating
- (d) a biased estimate of the class rating
- (e) the population

10. You read that drinking moderate amounts of alcohol may reduce the risk of heart disease. To be convinced this is true, you would like to have data from

- (a) a random sample of people with heart disease that asked about their drinking habits.
- (b) a random sample of people with different drinking habits, followed for several years to compare their future heart disease.
- (c) a comparison of heart disease rates in countries where people drink lots of wine with countries where people drink little.
- (d) a randomized comparative experiment in which some people drink moderately and others do not drink at all.
- (e) methods (b), (c), and (d) are equally effective.

12. You take an SRS of size 100 from the 48,300 students at Ohio State University. You then take an SRS of size 100 from the 8,500,000 adults in the state of Ohio. The margin of error in a 95% confidence statement for the Ohio State student sample is

- (a) the same as for the sample of Ohio adults because both are samples of size 100.
- (b) smaller than for the sample of Ohio adults because the population is much larger.
- (c) larger than for the sample of Ohio adults because the population is much larger.
- (d) either larger or smaller than for the sample of Ohio adults because it changes at random when we take a sample.
- (e) either larger or smaller than for the sample of Ohio adults, depending on what confounding variables are present.

A 2008 Gallup Poll asked 1010 randomly chosen adults to rate the "honesty and ethical standards" of 21 occupations. Nurses were at the top, with 84% of the poll respondents rating them "very high" or "high." Lobbyists were at the bottom with just 5% "very high" or "high" ratings. **The next two questions** concern this poll.

14. The population for this poll is

- (a) all adults.
- (b) the 1010 respondents.
- (c) people in 21 occupations.
- (d) nurses and lobbyists.
- (e) honesty and ethical standards.

16. When we take a census, we collect data from

- (a) a stratified random sample
- (b) every individual selected in a simple random sample
- (c) every individual in the population
- (d) a voluntary response sample
- (e) a convenience sample

18. Divide undergraduate students at State Tech into two groups: those who live in university housing and those who do not. Take a random sample from each group. This is an example of

- (a) multistage sampling
- (b) stratified random sampling
- (c) clustered sampling
- (d) simple random sampling
- (e) systematic random sampling

19. The correlation between two variables x and y is -0.6 . If we used a regression line to predict y using x , what percent of the variation in y would be explained?

- (a) 20%
- (b) 36%
- (c) -36%
- (d) 60%
- (e) -60%

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

- (a) maximum
- (b) standard deviation
- (c) mean
- (d) median
- (e) first quartile

23. Which correlation indicates a strong negative straight line relationship?

- (a) 0.5
- (b) -1.5
- (c) -0.5
- (d) -0.9
- (e) 0.9

A college football team wins 8 games and loses 3 during its season. The margins of victory in the 8 wins are (in points scored):

11 14 12 16 25 45 1 49

The next three questions are based on these data.

25. What were the mean and median margins of victory in these eight games?

- (a) mean = 20.5, median = 21.6
- (b) mean = 21.6, median = 20.5
- (c) mean = 15, median = 21.6
- (d) mean = 21.6, median = 15
- (e) Can't tell without knowing the results of the three losses.

26. The relationship between the mean and median margin of victory is typical of distributions which, like this one, are

- (a) nearly symmetric.
- (b) skewed to the left.
- (c) skewed to the right.
- (d) all positive.
- (e) football-shaped.

27. Suppose the "margins of victory" in the team's three losses were also reported, as negative numbers; for example, the "margin of victory" in a 42 to 17 was -25. What effect would including these three negative numbers, in addition to the eight positive numbers, have on the mean and median?

- (a) Both the mean and the median would be smaller.
- (b) The mean would be smaller, but the median would stay the same.
- (c) The median would be smaller, but the mean would stay the same.
- (d) Both the mean and the median would stay the same.
- (e) Can't tell without knowing the results of all three losses.

Here are the attendance figures for the lectures in a large class. **The next three questions** are based on these data.

Date	Attendance	Date	Attendance
08/26	398	10/16	306
08/28	389	10/21	303
09/02	359	10/24	285
09/04	373	10/28	290
09/09	370	10/31	298
09/11	351	11/04	278
09/16	341	11/11	275
09/18	317	11/13	270
09/23	321	11/18	269
09/30	318	11/21	255
10/02	309	11/25	211
10/07	317	12/02	262
10/09	265	12/04	259

29. To show the evolution of attendance during the semester, what type of graph should you draw?

- (a) boxplot
- (b) histogram
- (c) line graph
- (d) scatterplot
- (e) stemplot

31. In which of these cases would the confidence interval be wider than the one in the previous question?

- (a) If the confidence level were 90% instead of 95%.
- (b) If the sample size were 498 instead of 398.
- (c) If the entire student population were larger.
- (d) Both (a) and (b).
- (e) None of the above.

A news report says, "New research suggests that an active social life may increase resistance to respiratory infections. In a recent study, 300 healthy volunteers were given nasal sprays of cold virus, then quarantined for five days. The volunteers were asked about relationships with spouse, children, friends, neighbors, and coworkers. Those who had only one to three types of relationships were four times as likely to get a cold as those with six or more types. The most significant factor for resisting the cold virus seemed to be the diversity rather than the sheer quantity of social contacts." **The next two questions** concern this report.

32. How would you categorize this study?

- (a) An experiment with the nasal spray as the treatment.
- (b) An observational study with the various possible "social networks" as the treatments.
- (c) An experiment with the various possible "social networks" as the treatments.

- (d) A randomized, controlled clinical trial.
- (e) A voluntary response study.

33. Why do these results fall short of demonstrating that by broadening our social networks we can increase our resistance to respiratory infections? (Pick the best answer.)

- (a) Volunteers were used.
- (b) We're not told the age of the volunteers.
- (c) We're not told what day of the week they got the nasal spray.
- (d) Maybe the subjects lied about their relationships.
- (e) Some other variables, associated with the breadth of one's social network, may be the real causes of higher resistance to colds.

34. According to the FBI Uniform Crime Report, the robbery rate in the United States is 202 per 100,000 people. At that rate, how many robberies would there be in a state the size of Indiana (5.8 million people)?

- (a) about 1,200
- (b) about 12,000
- (c) about 120,000
- (d) about 1,200,000
- (e) impossible to say

35. The difference between an experiment and an observational study is

- (a) Observational studies don't have explanatory and response variables.
- (b) Experiments don't have placebos.
- (c) An experiment imposes treatments on subjects, while an observational study measures variables of interest without attempting to influence responses.
- (d) Experiments are double-blind, while observational studies are only single-blind.
- (e) All of the above.

36. An auto assembly plant reports these sales figures for the models it makes:

Vehicle	Nov 97	YTD-97	Nov 96	YTD-96
Subaru Legacy	8,570	83,994	8,615	85,256
Isuzu Rodeo	4,536	57,810	3,889	57,099
Honda Passport	1,569	20,567	2,259	25,982

By what percent did November Honda Passport sales drop from '96 to '97?

- (a) 69%
- (b) 44%
- (c) 31%
- (d) 21%
- (e) They didn't drop. They rose by 44%.

The next three questions concern a baseball team that had the following record for the season:

	Home Games	Away Games
Won	45	25
Lost	35	55

38. What percent of their games did the team win for the season?
(a) 28.125% (b) 31.25% (c) 43.75% (d) 56.25% (e) 64.29%
39. What percent of their home games did the team win?
(a) 28.125% (b) 31.25% (c) 43.75% (d) 56.25% (e) 64.29%
40. The two previous questions lead to this conclusion about the relationship between where the team played and how well it did:
(a) The team is more likely to win when it plays at home.
(b) The team is less likely to win when it plays at home.
(c) There is a positive association between where the team played and whether it won or lost.
(d) Both (a) and (c).
(e) Both (b) and (c).
42. You are planning an experiment to determine the effect of the brand of gasoline and the weight of a car on gas mileage measured in miles per gallon. You will use a single test car, adding weights so that its total weight is 3000, 3500, or 4000 pounds. The car will drive on a test track at each weight using each of Amoco, Marathon, and Speedway gasoline. The response variable in your experiment is
(a) the weight of the car.
(b) the brand of gasoline.
(c) Both (a) and (b) are response variables.
(d) gas mileage in miles per gallon.
(e) random.
44. In the gas mileage experiment of the previous two questions,
(a) gasoline brand is a categorical variable and weight is a quantitative variable.
(b) gasoline brand and weight are both categorical variables.
(c) gasoline brand and weight are both quantitative variables.
(d) gasoline brand is a quantitative variable and weight is a categorical variable.
(e) gasoline brand and car model are categorical variables and weight is a quantitative variable.
48. Scores on the Scholastic Assessment Test are reported on a scale that yields a normal distribution with mean 500 and standard deviation 100. The percent of scores above 500 on the SAT is
(a) 99.7% (b) 95% (c) 68% (d) 50% (e) 34%
49. Scores on the Scholastic Assessment Test are reported on a scale that yields a normal distribution with mean 500 and standard deviation 100. Julie scores 600 on the SAT. Her standard score is
(a) $z = -1$ (b) $z = 0$ (c) $z = 1$ (d) $z = 6$ (e) $z = 100$
50. In any normal distribution, the percent of observations falling between standard score $z = 0$ and standard score $z = 2$ is about
(a) 95% (b) 81.5% (c) 61% (d) 50% (e) 47.5%

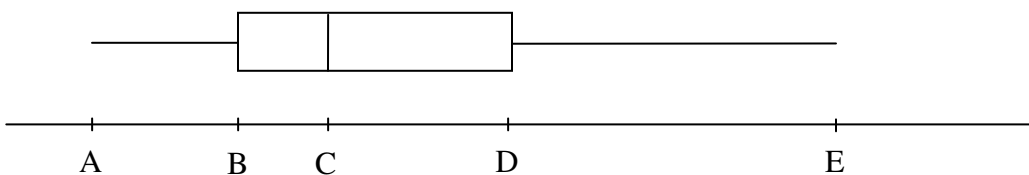
51. George has an average bowling score of 180 and bowls in a league where the average for all bowlers is 150 and the standard deviation is 20. Bill has an average bowling score of 190 and bowls in a league where the average is 160 and the standard deviation is 15. Who ranks higher in his own league, George or Bill ?

- (a) Bill, because his 190 is higher than George's 180.
- (b) Bill, because his standard score is higher than George's.
- (c) George, because his standard score is higher than Bill's.
- (d) Bill and George have the same rank in their leagues, because they are both 30 pins above the mean.
- (e) Bill and George have the same rank in their leagues, because their standard scores are equal.

52. Scores of adults on the Wechsler Adult Intelligence Scale (a common "IQ" test) follow a normal distribution. The middle 95% of scores on this test range from 70 to 130. What is the standard deviation of the test scores?

- (a) 30 points
- (b) 15 points
- (c) 10 points
- (d) 7.5 points
- (e) 5 points

53. 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

65. A sportswriter wants to know how strongly Lafayette residents support the local minor league baseball team, the Lafayette Leopards. She stands outside the stadium before a game and interviews the first 20 people who enter the stadium. The intended *population* for this survey is

- (a) all residents of Lafayette.
- (b) all Leopard fans.
- (c) all people attending the game the day the survey was conducted.
- (d) the 20 people who gave the sportswriter their opinion.
- (e) all American adults.

66. The *sample* for the survey in the previous question is

- (a) all residents of Lafayette.
- (b) all Leopard fans.
- (c) all people attending the game the day the survey was conducted.
- (d) the 20 people who gave the sportswriter their opinion.
- (e) the sportswriter.

67. The newspaper asks you to comment on their survey of local opinion, described in the previous two questions. You say

- (a) This is a simple random sample. It gives very accurate results.
- (b) This is a simple random sample. The results are not biased, but the sample is so small that variation will be high.

- (c) This is a census, because all fans had a chance to be asked. It gives very accurate results.
- (d) The sampling method is biased. It will almost certainly overestimate the level of support among all Lafayette residents.
- (e) The sampling method is biased. It will almost certainly underestimate the level of support among all Lafayette residents.

68. A *stratified random sample* is a sample in which

- (a) every member of the population has the same chance of being selected.
- (b) every member of the population has a known chance of being selected.
- (c) the population is first divided into groups of similar individuals, then a separate simple random sample is selected from each group and combined to form the full sample.
- (d) every possible sample of the same size has the same chance of being selected.
- (e) we choose the final sample in several stages, for example, first states, then counties in those states, then households in those counties.

71. Can pleasant aromas help people work more efficiently? Researchers did this study to find out. Twenty-two students worked a paper and pencil maze six times. On three tries, they wore a mask with floral scents. On the other three tries they wore a mask with no scent. The six tries were done in random order and each used a different maze. The researchers found that the subjects took less time to complete the maze when wearing the scented mask. This study is

- (a) an experiment.
- (b) an observational study.
- (c) a sample survey.
- (d) a census.
- (e) a time series.

The table below gives the age in weeks of 6 laboratory mice and the best time (in three trials) for each of the mice to finish a maze.

Mouse age (wks)	10	15	16	18	20	20
Finish time (sec)	38	30	32	24	23	20

The regression equation for Finish time (x) vs. Age (y) is $y = 55.9 - 1.7x$, and $r^2 = 0.918$

The next four questions concern these data.

73. What is the predicted finish time for a mouse who is 16 weeks old?

- (a) 32
- (b) 28.7
- (c) 22.9
- (d) 16
- (e) 1.5

74. What is the correlation coefficient for the relationship between finish time and age?

- (a) -0.958
- (b) -0.918
- (c) -0.843
- (d) 0.918
- (e) 0.958

75. Why would it be risky to use this equation to predict the finish time for a 25-week-old mouse?

- (a) Mice rarely live that long.
- (b) We would be using the regression equation well beyond the range of the data.
- (c) None of the mice in our experiment are 25 weeks old.
- (d) The sample is too small to make such predictions.
- (e) The correlation coefficient is too small to make such predictions.

76. The residual for the 20-week-old mouse who took 23 seconds to finish the maze is
(a) -1.1 (b) 1.1 (c) 3 (d) 16.8 (e) 21.9

84. 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.

86. Which of the values below is impossible for the descriptive measure in question?
(a) $r = 1.25$
(b) $\bar{x} = -0.2$
(c) $s = 3.4$
(d) Both (a) and (b).
(e) Both (a) and (c).

87. You measure both the calories and the amount of salt in each of 33 brands of hot dogs. The correlation between these variables is $r = 0.49$. This shows that
(a) hot dogs with more calories tend to have less salt.
(b) calories and salt in hot dogs are not related at all.
(c) the mean amount of salt is less than the mean number of calories.
(d) the mean amount of salt is greater than the mean number of calories.
(e) hot dogs with more salt tend to also have more calories.

88. The label on a package of hot dogs tells you how much salt each hot dog has. You want to use this information to predict how many calories the hot dog has. The correlation is $r = 0.49$. This says that
(a) the fact that hot dogs have different amounts of salt explains about 24% of the observed variation in their calorie counts.
(b) the fact that hot dogs have different amounts of salt explains about 49% of the observed variation in their calorie counts.
(c) the fact that hot dogs have different amounts of salt explains about 70% of the observed variation in their calorie counts.
(d) the fact that hot dogs have different amounts of salt explains about 98% of the observed variation in their calorie counts.
(e) unless we know the slope of the regression line, we cannot determine what percentage of the observed variation in calorie counts is explained by variation in salt.

89. 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 (*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.

90. You want to make a graph to display the distribution of the salaries of the 1700 professors at a large university. The best choice is

- (a) a boxplot. (b) a histogram. (c) a line graph. (d) a pie chart. (e) a stemplot.

A study of drug addicts in Amsterdam recorded how often each addict had recently injected drugs and whether or not the addict was infected with HIV, the virus that causes AIDS. Here is a two-way table of the numbers of addicts in each condition:

		HIV Yes	HIV No
Inject?	Daily	32	45
	Less than daily	20	18
	No	18	23

The next six questions concern the data in this table.

91. How many addicts did the study gather data from?

- (a) 156 (b) 86 (c) 77 (d) 70 (e) Can't tell from the table.

92. About what percent of addicts who inject daily are infected with HIV?

- (a) 20.5%. (b) 41.6%. (c) 44.9%. (d) 45.7%. (e) 49.3%.

93. From the data in the two-way table, about what is the value of the correlation between whether or not an addict is infected with HIV and how often he or she injects drugs?

- (a) Small but positive, say $r = 0.2$.
- (b) Quite large and positive, say $r = 0.8$.
- (c) Very small, close to $r = 0$.
- (d) Small but negative, say $r = -0.2$.
- (e) r makes no sense for these data.

97. Until the scale was changed in 1995, SAT scores were based on a scale set many years ago. For math scores, the mean under the old scale in the 1990s was about 470 and the standard deviation was about 110. What is the standard score of someone who scored 500 on the old SAT?

- (a) $z = 0.27$ (b) $z = -0.27$ (c) $z = 30$ (d) $z = -30$ (e) $z = 0$

98. The change in scales makes it hard to compare scores on the 1994 math SAT (mean 470, standard deviation 110) and the 1996 math SAT (mean 500, standard deviation 100). Jane took the SAT in 1994 and scored 500. Her sister Colleen took the SAT in 1996 and scored 520. Who did better on the exam, and how can you tell?

- (a) Colleen -- she scored 20 points higher than Jane.
- (b) Colleen -- her standard score is higher than Jane's.
- (c) Jane -- her standard score is higher than Colleen's.
- (d) Jane -- the standard deviation was bigger in 1994.
- (e) Can't tell from the information given.

02. The poll questioned 1025 women and 472 men. The design of the sample choose separate samples of men and women and planned to interview more women than men. This is a

- (a) systematic random sample.
- (b) completely randomized design.
- (c) stratified random sample.
- (d) simple random sample.
- (e) matched pairs design

144. You gather data on the number of hours of television watched per week and the grade point average of juniors majoring in the School of Liberal arts. You expect that TV watching will help explain grades. In a scatterplot of your data,

- (a) hours of TV should be on the horizontal axis.
- (b) grade index should be on the horizontal axis.
- (c) it makes no difference which is horizontal.
- (d) a scatterplot is not an appropriate type of graph for these data, use a dot plot.
- (e) a scatterplot is not an appropriate type of graph for these data, use a box plot.

145. The plot of the data in the preceding question shows that students who watch more TV tend to have lower grade indexes. You calculate the correlation r between hours of TV and grade index. A plausible value is

- (a) $r = -1.2$ (b) $r = -1$ (c) $r = -0.4$ (d) $r = 0$ (e) $r = 0.4$

146. There is a strong straight-line relationship between the outdoor temperature and the amount of energy used to heat a house. Lower temperatures require more energy to keep the house warm. Knowing this, a reasonable value for the correlation coefficient between temperature and home energy consumption is:

- (a) $r = 1$ (b) $r = 0.8$ (c) $r = 0$ (d) $r = -0.8$ (e) $r = -1$

149. Which of the following pairs of variables is most likely to show a negative correlation?
- (a) a person's income and her years of education.
 - (b) a car's top speed and its gas mileage (miles per gallon).
 - (c) a student's grade point average and his SAT score.
 - (d) a person's height and his or her income.
 - (e) a person's age and his or her shoe size.

Below is a table of persons aged 18 to 21 years by school enrollment status and sex, collected by the Current Population Survey. **The next four questions** refer to this table.

Number of Persons (thousands)		
	Female	Male
High school drop-outs	965	1122
Enrolled in high school	355	695
HS graduate, not in college	2865	2540
HS graduate, in college	3068	2659
Total	7253	7016

150. How many persons age 18 to 21 are there?
- (a) 14,269 (b) 1,426,900 (c) 5,727,000 (d) 14,269,000 (e) 12,181,000
151. The percent of men age 18 to 21 who are high school drop outs is
- (a) 13.3% (b) 15.5% (c) 16.0% (d) 7.9% (e) 53.8%
152. On the whole, these data suggest that
- (a) young men are getting less education than young women.
 - (b) young men are getting more education than young women.
 - (c) almost all high school graduates go on to college.
 - (d) high school graduates are mainly men.
 - (e) most college students are male.

A medical researcher collects health data on many women in each of several countries. One of the variables measured for each woman in the study is her weight in pounds. The following list gives a **modified version of the five-number summary** for the weights of women in each of several countries. The modification is that first and last numbers for each country are the **deciles** (that is, the 10th and 90th percentiles) instead of the maximum and minimum.. **The next three questions** refer to this information.

- Country A : 100, 110, 120, 160, 200
- Country B : 113, 135, 151, 185, 240
- Country C : 84, 96, 110, 124, 136
- Country D : 100, 143, 182, 191, 200
- Country E: 112, 120, 128, 140, 150

155. In one of these countries the weights of women are approximately normally distributed, Which country is it?

- (a) Country A (b) Country B (c) Country C (d) Country D (e) Country E

156. In one of the five countries, the mean weight of women is less than the median weight. Which country is it most likely to be?

- (a) Country A (b) Country B (c) Country C (d) Country D (e) Country E

157. About what fraction of Country A women weigh between 110 and 200 pounds?

- (a) 50% (b) 65% (c) 75% (d) 85% (e) 95%

Suppose that the distribution of MATH SAT scores from your state this year is normally distributed with mean 480 and standard deviation 100 for Males, and mean 440 and standard deviation 120 for Females. **The next three questions** make use of the above information.

160. If a person who scores 780 or higher on MATH SAT can be considered a genius, the proportion of geniuses among male SAT takers in this state is about

- (a) 30% (b) 15% (c) 3% (d) 1.5% (e) 0.15%

161. Mary took the MATH SAT and scored 680. She did better than about _____ % of female students taking the test.

- (a) 99.9 (b) 99 (c) 97.5 (d) 97 (e) 95

162. How well did Mary's score of 680 rate in terms of the scores of male students? Mary did better than about _____ % of male students taking the test.

- (a) 99.9 (b) 99 (c) 97.5 (d) 97 (e) 95

A recent survey by censusatschool.org.uk found that the mean time it takes students in Canada to travel to school is 20.3 minutes with a standard deviation of 15.4 minutes. The distribution is skewed to the right. **The next two questions** refer to this information.

164. Which of the following best describes the sampling distribution of means for samples of size 80 from this population?

- (a) Mean 20.3; Standard deviation 15.4; shape unknown
(b) Mean 20.3; Standard deviation 15.4; shape approximately normal
(c) Mean unknown; Standard deviation 15.4; shape unknown
(d) Mean 20.3; Standard deviation 1.72; shape unknown
(e) Mean 20.3; Standard deviation 1.72; shape approximately normal