

AP Statistics Summer Packet

THE RESULTS ARE PRETTY CONCLUSIVE, IT SEEMS THAT 75.8% OF THE 65.2% OF GPs WHO BOTHERED TO VOTE WERE 29.3% HAPPY WITH 14.2% OF THE PROPOSALS...AND THE REST WEREN'T SURE!



© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com

- Part 1: Chapter 1 Multiple Choice Questions**
- Part 2: Chapter 4 Multiple Choice Questions**
- Part 3: 4 Previous AP free response questions**

- Packet Due Date: First Day of Classes
- Be prepared for a test on chapters 1 and 4 on the third day of class!
- This packet will count as a 50 point quiz grade, graded for both completeness and accuracy.

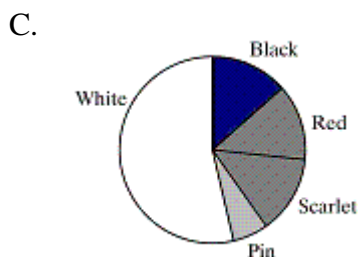
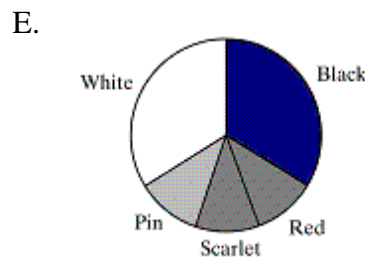
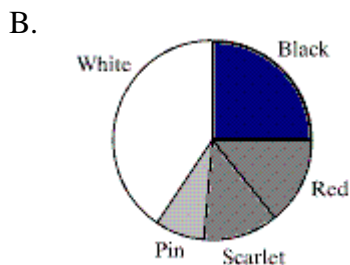
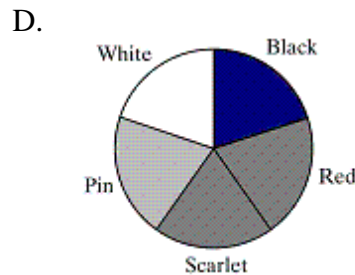
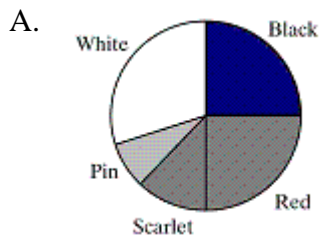
Part 1:

Chapter 1 Multiple Choice

- ___ 1. You measure the age, marital status and earned income of an SRS of 1463 women. The number and type of variables you have measured is
- 14563.
 - four; two categorical and two quantitative.
 - four; one categorical and three quantitative.
 - three; two categorical and one quantitative.
 - three; one categorical and two quantitative.
- ___ 2. A statistics teacher asks the 29 students in his statistics class how many minutes they spent on one homework assignment. The distribution of the variable “time on homework” is
- the difference between the longest time and the shortest time among the students’ responses
 - a description of what values the variable takes and how often it takes them.
 - the average distance between each value of the variable.
 - the average time the students spent on the assignment.
 - the number of students who were asked the questions—that is, 29.
- ___ 3. Deciduous forests in the Eastern United States often have many different species of oak trees. Below is a frequency distribution for five different species of oaks found in sample plots a certain forest.

Species of oak	Black	Red	Scarlet	Pin	White
Frequency	25	14	12	8	40

Which of the following pie charts describes the same distribution?



- _____ 4. X and Y are two categorical variables. The best way to determine if there is a relation between them is to
- construct parallel box plots of the X and Y values.
 - draw dot plots of the X and Y values.
 - make a two-way table of the X and Y values.
 - compare medians and interquartile ranges of the X and Y values.
 - compare means and standard deviations of the X and Y values.
- _____ 5. In a study of the link between high blood pressure and cardiovascular disease, a group of white males aged 35 to 64 was followed for 5 years. At the beginning of the study, each man had his blood pressure measured and it was classified as either "low" systolic blood pressure (less than 140 mm Hg) or "high" blood pressure (140 mm Hg or higher). The following table gives the number of men in each blood pressure category and the number of deaths from cardiovascular disease during the 5-year period.

Blood pressure	Deaths	Total
Low	10	2000
High	5	3500

Based on these data, which of the following statements is correct?

- These data are consistent with the idea that there is a link between high blood pressure and death from cardiovascular disease.
- The mortality rate (proportion of deaths) for men with high blood pressure is 5 times that of men with low blood pressure.
- These data probably understate the link between high blood pressure and death from cardiovascular disease, because men will tend to understate their true blood pressure.
- Although there were more deaths in the high blood pressure group, this is expected, because there were 1500 more men in that group.
- All of the above.

Scenario 1-1

A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.

	Male	Female
Democrat	300	600
Republican	500	300
Other	200	100

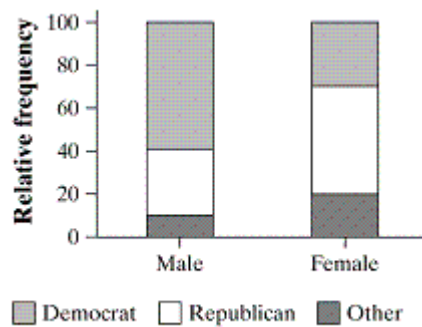
- _____ 6. Use Scenario 1-1. The proportion of males that are registered as Democrats is
- 300
 - 30
 - 0.33
 - 0.30
 - 0.15

7. Use Scenario 1-1 Your percentage from question number 12 is part of
- A. The marginal distribution of political party registration.
 - B. The marginal distribution of gender.
 - C. The conditional distribution of gender among Democrats.
 - D. The conditional distribution of political party registration among males.
 - E. The conditional distribution of males within gender.

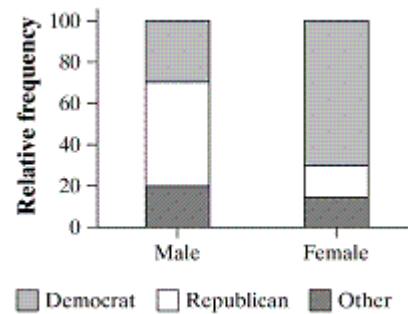
8. Use Scenario 1-1. The proportion of registered Democrats that are male is
- A. 300
 - B. 33
 - C. 0.33
 - D. 0.30
 - E. 0.15

9. Use Scenario 1-1. Which of the following graphs accurately represents the distribution for political party registration for each gender?

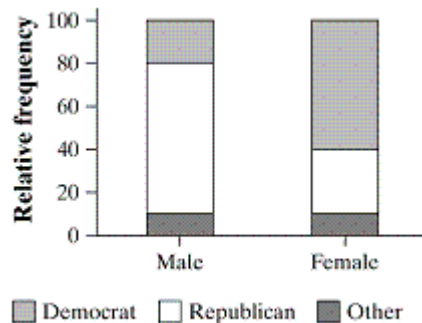
A.



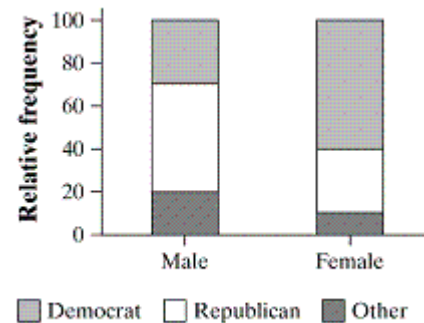
D.



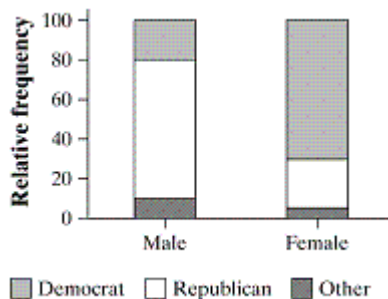
B.



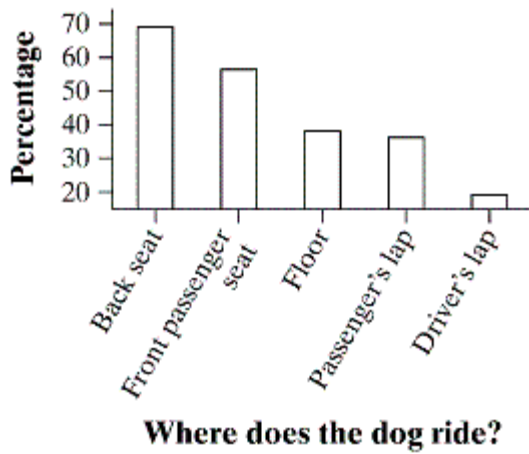
E.



C.

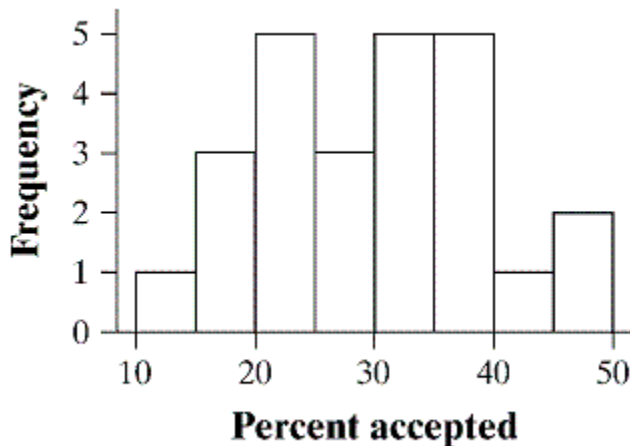


10. The bar graph below summarizes responses of dog owners to the question, “Where in the car do you let your dog ride?”



Which of the following statements is false?

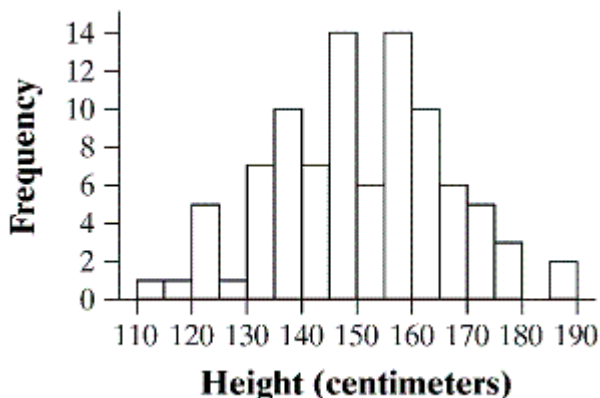
- A. Some owners let their pets ride in more than one place in the car.
 - B. A majority of owners allow their pets to ride in the front passenger seat.
 - C. The most common place dogs ride is in the back seat.
 - D. The vertical scale of this graph exaggerates the difference between the percentage who let their dogs ride in the driver's lap *versus* a passenger's lap.
 - E. These data could also be presented in a pie chart.
11. The following histogram represents the distribution of acceptance rates (percent accepted) among 25 business schools in 1997.



What percent of the schools have an acceptance rate of under 20%?

- A. 3%
- B. 4%
- C. 12%
- D. 16%
- E. 24%

- ___ 12. The histogram below shows the distribution of heights for 100 randomly selected school children in Great Britain.



Which of the following descriptions best fits this distribution?

- A. Roughly uniform, centered at about 150, range 110 to 190.
 - B. Roughly uniform, centered at about 150, range 80
 - C. Roughly symmetric, centered at about 150, range 110 to 190.
 - D. Roughly symmetric, centered at about 150, range 80.
 - E. Roughly symmetric, centered at about 150, range about 135 to 165.
- ___ 13. Which of the following statements is NOT true?
- A. In a symmetric distribution, the mean and the median are equal.
 - B. Fifty percent of the scores in a distribution are between the first and third quartiles.
 - C. In a symmetric distribution, the median is halfway between the first and third quartiles.
 - D. The median is always greater than the mean.
 - E. The range is the difference between the largest and the smallest observation in the data set.
- ___ 14. A consumer group surveyed the prices for a certain item in five different stores, and reported the average price as \$15. We visited four of the five stores, and found the prices to be \$10, \$15, \$15, and \$25. Assuming that the consumer group is correct, what is the price of the item at the store that we did not visit?
- A. \$5
 - B. \$10
 - C. \$15
 - D. \$20
 - E. \$25

___ 15. The ages of people in a college class are as follows:

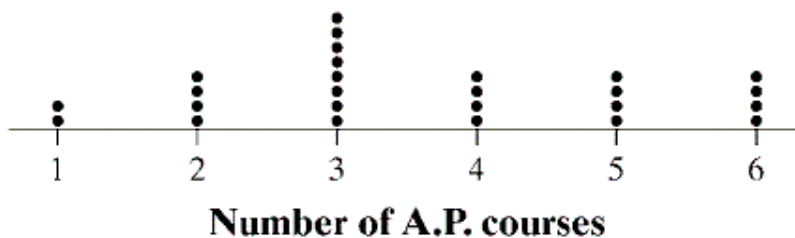
Age	18	19	20	21	22	23	24	25	32
Number of students	14	120	200	200	90	30	10	2	1

What is true about the median age?

- A. It must be 20.
 - B. It must be 20.5.
 - C. It could be any number between 19 and 21.
 - D. It must be 21.
 - E. It must be over 21.
- ___ 16. The median age of five elephants at a certain zoo is 30 years. One of the elephants, whose age is 50 years, is transferred to a different zoo. The median age of the remaining four elephants is
- A. 40 years.
 - B. 30 years.
 - C. 25 years.
 - D. less than 30 years.
 - E. Cannot be determined from the information given.
- ___ 17. A set of data has a mean that is much larger than the median. Which of the following statements is most consistent with this information?
- A. The distribution is symmetric.
 - B. The distribution is skewed left.
 - C. The distribution is skewed right.
 - D. The distribution is bimodal.
 - E. The data set probably has a few low outliers.

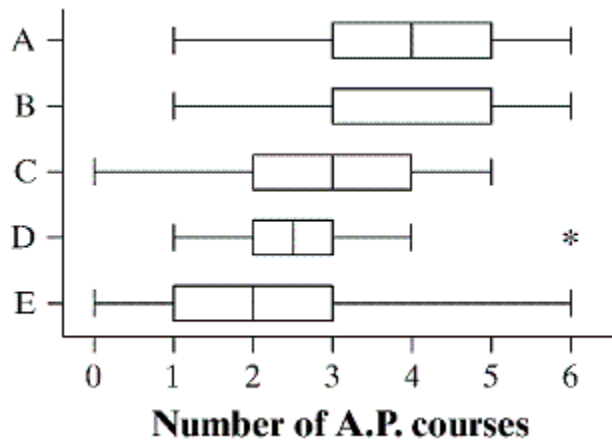
Scenario 1-4

Mr. Williams asked the 26 seniors in his statistics class how many A.P. courses they had taken during high school. Below is a dot plot summarizing the results of his survey.



- ___ 18. Use Scenario 1-4. The interquartile range for the number of A.P. Courses is
- A. 3 to 4
 - B. 2.5 to 5
 - C. 3 to 5
 - D. 2
 - E. 2.5

___ 19. Use Scenario 1-4. Which of the following is a correct box plot for these data?



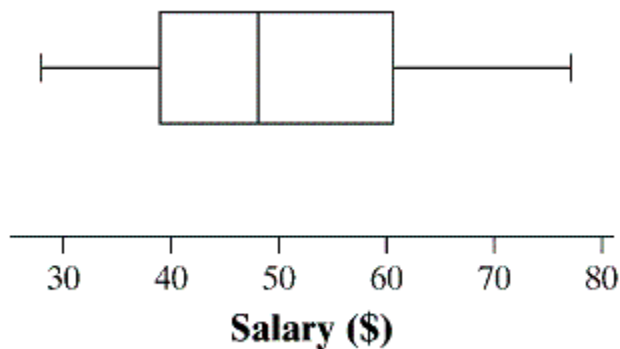
- A. A
- B. B
- C. C
- D. D
- E. E

___ 20. The mean age of four people in a room is 30 years. A new person whose age is 55 years enters the room. The mean age of the five people now in the room is

- A. 30.
- B. 35.
- C. 37.5.
- D. 40.
- E. Cannot be determined from the information given.

Scenario 1-5

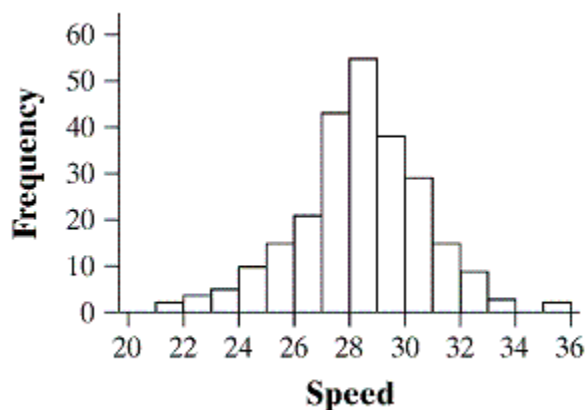
A sample was taken of the salaries of 20 employees of a large company. The following boxplot shows the salaries (in thousands of dollars) for this year.



___ 21. Use Scenario 1-5. Based on the boxplot, which of the following statements is true?

- A. The maximum salary is between \$60,000 and \$70,000.
- B. The minimum salary is \$20,000.
- C. The range of the middle half of the salaries is about \$20,000.
- D. The median salary is about \$40,000.
- E. 25% of the employees make more than \$70,000.

- ___ 22. Use Scenario 1-5. Based on the boxplot, the five-number summary is
- 28, 39, 48, 60.5, 77.
 - 28, 41, 48, 58, 77.
 - 28, 39, 51, 58, 77.
 - 28, 41, 51, 60.5, 77.
 - 26, 39, 48, 60.5, 81.
- ___ 23. There are three children in a room, ages three, four, and five. If a four-year-old child enters the room the
- mean age will stay the same but the variance will increase.
 - mean age will stay the same but the variance will decrease.
 - mean age and variance will stay the same.
 - mean age and variance will increase.
 - mean age and variance will decrease.
- ___ 24. The standard deviation of 16 peoples' weights (in pounds) is computed to be 5.4. The variance of these measurements is
- 2.24.
 - 29.16.
 - 52.34.
 - 256.
 - 21.6.
- ___ 25. The standard deviation of 16 peoples' weights (in pounds) is computed to be 5.4. The units for the variance of these measurements is
- pounds.
 - square root pounds.
 - pounds squared.
 - no units. Variance never has units.
 - percentiles.
- ___ 26. A policeman records the speeds of cars on a certain section of roadway with a radar gun. The histogram below shows the distribution of speeds for 251 cars.



Which of the following measures of center and spread would be the best ones to use when summarizing these data?

- Mean and interquartile range.
- Mean and standard deviation.
- Median and range.
- Median and standard deviation.
- Median and interquartile range.

- _____ 27. You want to use numerical summaries to describe a distribution that is strongly skewed to the left. Which combination of measure of center and spread would be the best ones to use?
- A. Mean and interquartile range.
 - B. Mean and standard deviation.
 - C. Median and range.
 - D. Median and standard deviation.
 - E. Median and interquartile range.

- _____ 28. A lobster fisherman is keeping track of the productivity of a set of traps he has placed in a favorite location. Below are the numbers of lobsters in these traps over the course of 12 different hauls.

0 3 3 3 4 5 5 6 7 7 12 14

According to the $1.5 \times \text{IQR}$ rule, which values in the above distribution are outliers?

- A. 0 only
 - B. 14 only
 - C. 12 and 14
 - D. 0 and 14
 - E. 0, 12, and 14
- _____ 29. The stemplot below shows the number of home runs hit in 2008 by members of the Philadelphia Phillies, who won major League Baseball's World Series that year. (Each of the 13 players who appeared in at least half the Phillies' games that year is included). Note that $4 | 8$ represents 48 home runs.

0	0 2 4
0	9 9 9
1	1 4 4
1	
2	4
2	
3	3 3
3	
4	
4	8

The five number summary for these data is:

- A. 0, 9, 1, 3, 8
 - B. 0, 9, 11, 33, 48
 - C. 0, 6.5, 11, 28.5, 48
 - D. 0, 6.5, 11, 28.5, 33
 - E. 0, 4, 11, 24, 48
- _____ 30. A study of the salaries of full professors at Upper Wabash Tech shows that the median salary for female professors is considerably less than the median male salary. Further investigation shows that the median salaries for male and female full professors are about the same in every department (English, physics, etc.) of the university. This apparent contradiction is an example of
- A. a fallacy.
 - B. Simpson's paradox.
 - C. concealed gender bias.
 - D. a conditional distribution.
 - E. negative association.

Part 2:

Chapter 4 Multiple Choice

Scenario 4-1

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.

- _____ 1. Use Scenario 4-1. 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.
- _____ 2. Use Scenario 4-1. The sample 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. The sportswriter.
- _____ 3. Use Scenario 4-1. The newspaper asks you to comment on their survey of local opinion. 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 too small to have high precision
 - C. This is a census, because all fans had a chance to be asked. It gives very accurate results.
 - D. This is a convenience sample. It will almost certainly overestimate the level of support among all Lafayette residents.
 - E. This is a convenience sample. It will almost certainly underestimate the level of support among all Lafayette residents.
- _____ 4. A study sponsored by American Express Co. and the French government tourist office found that old stereotypes about French unfriendliness were not true. The respondents were more than 1000 Americans who have visited France more than once for pleasure over the past two years. The results of this study are probably
- A. Very accurate, given the large sample size
 - B. Very inaccurate because the sample is only a small fraction of all Americans who have visited France.
 - C. Extremely variable, because people's opinions differ so greatly.
 - D. Biased, overstating the extent to which the old stereotypes were not true.
 - E. Biased, understating the extent to which the old stereotypes were not true.
- _____ 5. In order to assess the opinion of students at the University of Minnesota on campus snow removal, a reporter for the student newspaper interviews the first 12 students he meets who are willing to express their opinion. The method of sampling used is
- A. A census
 - B. A systematic sample
 - C. A voluntary sample
 - D. A convenience sample
 - E. A simple random sample

- _____ 6. A television station is interested in predicting whether voters in its viewing area are in favor of offshore drilling. It asks its viewers to phone in and indicate whether they support/are in favor of or are opposed to this practice. Of the 2241 viewers who phoned in, 1574 (70%) were opposed to offshore drilling. The viewers who phoned in are
- A. A voluntary response sample
 - B. A convenience sample
 - C. A probability sample
 - D. A population
 - E. A simple random sample
- _____ 7. A simple random sample of size n is defined to be
- A. A sample of size n chosen in such a way that every unit in the population has the same chance of being selected
 - B. A sample of size n chosen in such a way that every unit in the population has a known nonzero chance of being selected
 - C. A sample of size n chosen in such a way that every set of n units in the population has an equal chance to be the sample actually selected
 - D. A sample of size n chosen in such a way that each selection is made independent of every other selection
 - E. All of the above. They are essentially identical definitions
- _____ 8. Simple random sampling
- A. Reduces bias resulting from poorly worded questions
 - B. Offsets bias resulting from undercoverage and nonresponse
 - C. Reduces bias resulting from the behavior of the interviewer
 - D. Reduces variability
 - E. None of the above

Scenario 4-2

We wish to choose a simple random sample of size three from the following employees of a small company. To do this, we will use the numerical labels attached to the names below

- | | | |
|--------------|------------|-----------|
| 1. Bechhofer | 4. Kesten | 7. Taylor |
| 2. Brown | 5. Kiefer | 8. Wald |
| 3. Ito | 6. Spitzer | 9. Weiss |

We will also use the following list of random digits, reading the list from left to right, starting at the beginning of the list.

11793 20495 05907 11384 44982 20751 27498 12009 45287 71753 98236 66419 84533

- _____ 9. Use Scenario 4-2. The simple random sample is
- A. 117
 - B. Bechhofer, Bechhofer again, and Taylor
 - C. Bechhofer, Taylor, Weiss
 - D. Kesten, Kiefer, Taylor
 - E. Taylor, Weiss, Ito

- _____ 10. Use Scenario 4-2. Which of the following statements is true?
- A. If we use another list of random digits to select the sample, we would get the same results as that obtained with the list actually used.
 - B. If we use another list of random digits to select the sample, we would get a completely different sample than that obtained with the list actually used.
 - C. If we use another list of random digits to select the sample, we would get, at most, one name in common with that obtained with the list actually used.
 - D. If we use another list of random digits to select the sample, the result obtained with the list actually used would be just as likely to be selected as any other set of three names.
 - E. If we use another list of random digits to select the sample, the result obtained with the list actually used would be far less likely to be selected than any other set of three names.
- _____ 11. Use Scenario 4-2. Which of these statements about the table of random digits is true?
- A. Every row must have exactly the same number of 0's and 1's.
 - B. In the entire table, there are exactly the same number of 0's and 1's.
 - C. If you look at 100 consecutive pairs of digits anywhere in the table, exactly 1 pair is 00.
 - D. All of these are true.
 - E. None of these are true.
- _____ 12. A public opinion poll in Ohio wants to determine whether or not registered voters in the state approve of a measure to ban smoking in all public areas. They select a simple random sample of fifty registered voters from each county in the state and ask whether they approve or disapprove of the measure. This is an example of a
- A. Systematic random sample
 - B. Stratified random sample
 - C. Multistage sample
 - D. Simple random sample
 - E. Cluster sample
- _____ 13. A stratified random sample is appropriate when
- A. It is impractical to take a simple random sample because the population is too large.
 - B. The population can be easily subdivided into groups according to some categorical variable, and the variable you are measuring is quite different within the groups but very similar between groups.
 - C. The population can be easily subdivided into groups according to some categorical variable, and the variable you are measuring is very similar within the groups but quite different between groups.
 - D. You intend to take a sample of more than 100 individuals.
 - E. You want to avoid undercoverage of certain groups.
- _____ 14. In order to select a sample of undergraduate students in the US, I select a simple random sample of four states. From each of these states, I select a simple random sample of two colleges or universities. Finally, from each of these eight colleges or universities, I select a simple random sample of 20 undergraduates. My final sample consists of 160 undergrads. This is an example of
- A. Simple random sampling
 - B. Stratified random sampling
 - C. Multistage sampling
 - D. Convenience sampling
 - E. Cluster sampling

- _____ 15. A marine biologist wants to estimate the mean size of the barnacle *Semibalanus balanoides* on a stretch of rocky shoreline. To do so, he randomly selected twenty 10-cm square plots and measured the size of every barnacle in each plot. This is an example of
- A. Convenience sampling
 - B. Cluster sampling
 - C. Stratified random sampling
 - D. Simple random sampling
 - E. Multistage sampling
- _____ 16. A 1992 Roper poll found that 22% of Americans say that the Holocaust may not have happened. The actual question asked in the poll was “*Does it seem possible or impossible to you that the Nazi extermination of the Jews never happened?*” and 22% responded possible. The results of this poll cannot be trusted because
- A. Undercoverage is present. Obviously, those people who did not survive the Holocaust could not be in the poll
 - B. The question is worded in a confusing manner
 - C. We do not know who conducted the poll or who paid for the results
 - D. Nonresponse is present. Many people will refuse to participate, and those who do will be biased in their opinions.
 - E. The question is clearly biased in the direction of a “possible” answer.
- _____ 17. Frequently, telephone poll-takers call near dinner time—between 6 pm and 7 pm—because most people are at home then. This is an effort to avoid
- A. Voluntary response bias
 - B. Calling people after they have gone to bed
 - C. A convenience sample
 - D. Nonresponse
 - E. Response bias
- _____ 18. The Bradley effect is a theory proposed to explain observed discrepancies between voter opinion polls and election outcomes in some elections where a white candidate and a non-white candidate run against each other. The theory proposes that some voters tend to tell pollsters that they are undecided or likely to vote for a non-white candidate, and yet, on election day, vote for the white component. This is an example of
- A. Voluntary response bias
 - B. Bias resulting from question wording
 - C. Undercoverage
 - D. Nonresponse
 - E. Response bias

Scenario 4-3

In order to assess the effects of exercise on reducing cholesterol, a researcher took a random sample of fifty people from a local gym who exercised regularly and another random sample of fifty people from the surrounding community who did not exercise regularly. They all reported to a clinic to have their cholesterol measured. The subjects were unaware of the purpose of the study, and the technician measuring the cholesterol was not aware of whether or not subjects exercised regularly.

- ____ 19. Use Scenario 4-3. This is a(n)
- A. Observational study
 - B. Experiment, but not a double blind experiment
 - C. Double blind experiment
 - D. Matched pairs experiment
 - E. Block design
- ____ 20. Use Scenario 4-3. Which of the following best describes the inferences the researcher can make based on his results?
- A. He can make inferences about cause and effect, but not about the populations from which the samples were taken.
 - B. He can make inferences about the populations from which the samples were taken, but not about cause and effect.
 - C. He can make inferences about both cause and effect and the populations from which the samples were taken.
 - D. He cannot make inferences about either cause and effect or the populations from which the samples were taken.
 - E. There is not enough information to make judgments about the scope of inference.

Scenario 4-4

Does caffeine improve exam performance? Suppose all students in the 8:30 section of a course are given a “treatment” (two cups of coffee) and all students in the 9:30 section are not permitted to have any caffeine before a mid-term exam.

- ____ 21. Use Scenario 4-4. Unfortunately, any systematic difference between the two sections on the exam might be due to the fact that the 8:30 and 9:30 classes have different instructors. This is an example of
- A. Placebo effect
 - B. Bias
 - C. Confounding
 - D. Observational study
 - E. Stratification
- ____ 22. Use Scenario 4-4. The response variable in this study is
- A. Two cups of coffee
 - B. The time the class is held
 - C. Class attendance
 - D. Teacher’s performance
 - E. Exam performance

- _____ 23. Use Scenario 4-4. Instead of giving all students in the 8:30 section two cups of coffee, students in the 8:30 section are randomly assigned to a treatment group (two cups of coffee) or a control group (two cups of decaffeinated coffee). The coffee is so bad that students cannot tell whether they are in the treatment or the control group. As it turns out, students in both groups do better on the exam than students in the 9:30 section, who weren't given anything. This could be the result of
- A. The placebo effect
 - B. An observational study
 - C. Voluntary response
 - D. Sampling variability
 - E. All of the above
- _____ 24. Use Scenario 4-4. Suppose half of the 8:30- students are randomly allocated to the treatment group (two cups of coffee), the other half to the control group (two cups of decaf). In addition, half of the 9:30 students are randomly allocated to the treatment group, the other half to the control group. This is an example of a
- A. Voluntary response study
 - B. Stratified sampling procedure
 - C. Matched pairs design
 - D. Completely randomized design
 - E. Block design
- _____ 25. An experiment was conducted by some students to explore the nature of the relationship between a person's heart rate (measures in beats per minute) and the frequency at which that person stepped up and down on steps of various heights. Three rates of stepping and two different step heights were used. A subject performed the activity (stepping at one of the three stepping rates at one of the two possible heights) for three minutes. Heart rate was then measured at the end of this period. The variables "stepping rate" and "step height" are the
- A. Factors
 - B. Levels
 - C. Controls
 - D. Units
 - E. Response variables
- _____ 26. Medical researchers are excited about a new cancer treatment that destroys tumors by cutting off their blood supply. To date, the treatment has only been tried on mice, but in mice it has been nearly 100% effective in eradicating tumors and appears to have no side effects. As evidence of the effectiveness of the new treatment in treating cancer in humans, these studies
- A. Display a high degree of statistical significance and so with nearly 100% certainty will work in humans
 - B. Are convincing, assuming the results have been replicated in a large number of mice
 - C. Are convincing, assuming that proper randomization and control were used
 - D. Suffer from lack of realism
 - E. Suffer from placebo effect

- _____ 27. A lurking variable is
- A. A variable that is not among the variables studied but that affects the response variable
 - B. The true cause of a response
 - C. Any variable that produces a large residual
 - D. The true variable that is explained by the explanatory variable
 - E. Another response variable
- _____ 28. For one kindergarten class in his district, a researcher determines which children already can read simple words and which children cannot upon entering kindergarten. The children are followed until third grade, at which point they are tested to determine the grade level at which they are reading. Those children who were reading simple words on entering kindergarten are found to be reading at a higher level than those who could not read simple words on entering kindergarten. The researcher
- A. Can conclude that children should be taught to read in preschool, as there are clear benefits to reading early
 - B. Cannot conclude that being able to read before entering kindergarten is beneficial, as there may be confounding variables in this study
 - C. Needs to have taken a random sample of kindergarten students instead of one class to conclude a cause-and-effect relationship
 - D. Needs to check the reading level of the children's parents
 - E. Needs to retest in sixth grade or no conclusions can be reached
- _____ 29. The principle reason for the use of *random assignment* in designing experiments is that it
- A. Distinguishes a treatment effect from the effects of confounding variables
 - B. Allows double-blinding
 - C. Reduces sampling variability
 - D. Creates approximately equal groups for comparison
 - E. Examines the placebo effect
- _____ 30. The principle reason for the use of *controls* in designing experiments is that it
- A. Distinguishes a treatment effect from the effects of confounding variables
 - B. Allows double-blinding
 - C. Reduces sampling variability
 - D. Creates approximately equal groups for comparison
 - E. Eliminates the placebo effect
- _____ 31. The principle reason for *replication* in designing experiments is that it
- A. Distinguishes a treatment effect from the effects of confounding variables
 - B. Allows double-blinding
 - C. Reduces sampling variability
 - D. Creates approximately equal groups for comparison
 - E. Eliminates the placebo effect

- _____ 32. When controlled experiments are impractical or unethical, which of the following would be necessary to establish a cause-and-effect relation between two variables?
- A. Strong association between the variables
 - B. An association between the variables is observed in many different settings
 - C. The alleged cause is plausible
 - D. There is no obvious lurking variable that would affect the response variable
 - E. All of the above
- _____ 33. In an experiment, an observed effect so large it would rarely occur by chance is called
- A. An outlier
 - B. Influential
 - C. Statistically significant
 - D. Bias
 - E. Replication
- _____ 34. In comparative trials in medicine, the placebo effect and subconscious bias on the part of the physicians evaluating treatment outcomes can be avoided by using
- A. The double-blind technique
 - B. Randomized complete block designs
 - C. Response variables
 - D. Stratified random samples
 - E. All of the above
- _____ 35. Twelve people who suffer from chronic fatigue syndrome volunteer to take part in an experiment to see if shark fin extract will increase one's energy level. Eight of the volunteers are men, and four are women. Half of the volunteers are to be given shark fin extract twice a day, and the other half are to be given a placebo twice a day. We wish to make sure that four men and two women are assigned to each of the treatments, so we decide to use a block design with the men forming one block and the women the other. A block design is appropriate in this experiment if
- A. We want to be able to compare effects on energy level in men and women
 - B. We believe men and women will respond differently to treatments
 - C. Gender equity is an important legal consideration in this study
 - D. We want the conclusions to apply equally to men and women
 - E. All of the above
- _____ 36. An experiment compares the taste of a new spaghetti sauce with the taste of a commercially successful sauce readily available in grocery stores. Each of a number of tasters tastes both sauces (in random order) and says which tastes better. This is called a
- A. Simple random sample
 - B. Stratified random sample
 - C. Completely randomized design
 - D. Matched pairs design
 - E. Double-blind design

Chapter 1 “FRAPPY”

{Free Response AP Problem...Yay!}



The following problem is taken from an actual Advanced Placement Statistics Examination. Your task is to generate a complete, concise statistical response in 15 minutes (or less). You will be graded based on the AP rubric. After grading, keep this problem in your files for your AP Exam preparation.

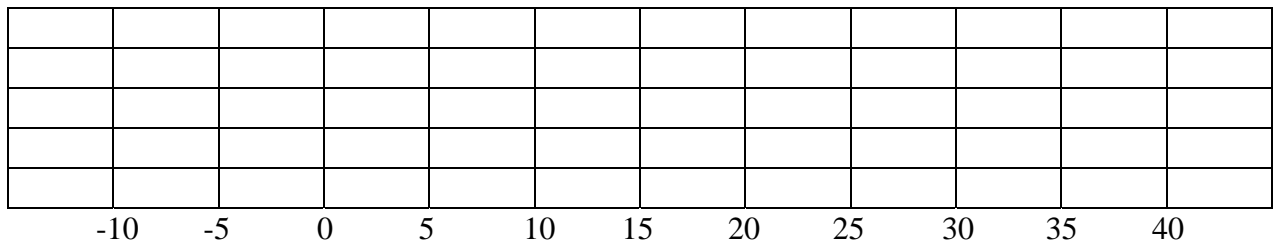
1. A consumer advocate conducted a test of two popular gasoline additives, A and B. There are claims that the use of either of these additives will increase gasoline mileage in cars. A random sample of 30 cars was selected. Each car was filled with gasoline and the cars were run under the same driving conditions until the gas tanks were empty. The distance traveled was recorded for each car.

Additive A was randomly assigned to 15 of the cars and additive B was randomly assigned to the other 15 cars. The tank of each car was filled with the gasoline and the assigned additive. The cars were again run under the same driving conditions until the tanks were empty. The distance traveled was recorded and the difference in the distance with the additive minus the distance without the additive for each car was calculated.

The following table summarizes the calculated differences. Note that negative values indicate less distance was traveled with the additive than without the additive.

Additive	Values below Q1	Q1	Median	Q3	Values above Q3
A	-10, -8, -2	1	3	4	5, 7, 9
B	-5, -3, -3	-2	1	25	35, 37, 40

A. On the grid below, display parallel boxplots (showing outliers, if any) of the differences of the two additives.



Show all of your work. Indicate clearly the methods you use.

B. Two ways that the effectiveness of the gasoline additive can be evaluated are by looking at either

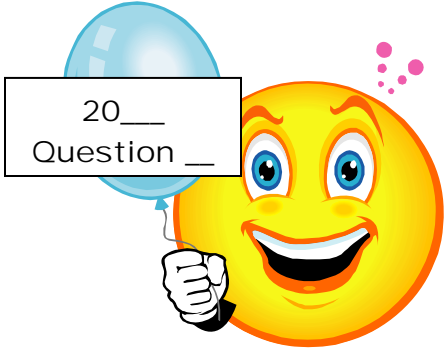
- the proportion of cars that have increased gas mileage when the additive is used in those cars.

OR

- the mean increase in gas mileage when the additive is used in those cars.

i. Which additive, A or B, would you recommend if the goal is to increase gas mileage in the highest proportion of the cars? Explain your choice.

ii. Which additive, A or B, would you recommend if the goal is to have the largest mean increase in gas mileage? Explain your choice.

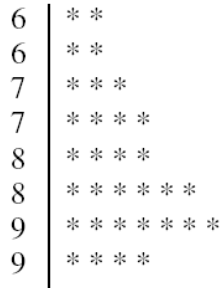


Chapter 1 “FRAPPY”

{Free Response AP Problem...Yay!}

The following problem is taken from an actual Advanced Placement Statistics Examination. Your task is to generate a complete, concise statistical response in 15 minutes (or less). You will be graded based on the AP rubric. After grading, keep this problem in your files for your AP Exam preparation.

1. The graph below displays the scores of 32 students on a recent exam. Scores on this exam ranged from 64 to 95 points.



- (a) Describe the shape of this distribution.
- (b) In order to motivate her students, the instructor of the class wants to report that, overall, the class’s performance on the exam was high. Which summary statistic, the mean or the median, should the instructor use to report that overall exam performance was high? Explain.
- (c) The midrange is defined as $\frac{\text{maximum} + \text{minimum}}{2}$. Compute this value using the data on the preceding page.

Is the midrange considered a measure of center or a measure of spread? Explain.



Chapter 4 “FRAPPY”

{Free Response AP Problem...Yay!}

The following problem is taken from an actual Advanced Placement Statistics Examination. Your task is to generate a complete, concise statistical response in 15 minutes (or less). You will be graded based on the AP rubric. After grading, keep this problem in your files for your AP Exam preparation.

3. An apartment building has nine floors and each floor has four apartments. The building owner wants to install new carpeting in eight apartments to see how well it wears before she decides whether to replace the carpet in the entire building.

The figure below shows the floors of apartments in the building with their apartment numbers. Only the nine apartments indicated with an asterisk (*) have children in the apartment.

11* 14	1st Floor	12 13	21 24	2nd Floor	22* 23*	31 34	3rd Floor	32 33	* = Children in the apartment
41 44	4th Floor	42 43	51* 54	5th Floor	52 53	61 64	6th Floor	62 63	
71 74*	7th Floor	72 73*	81 84*	8th Floor	82 83	91 94	9th Floor	92* 93*	

- a. For convenience, the apartment building owner wants to use a cluster sampling method, in which the floors are clusters, to select the eight apartments. Describe a process for randomly selecting eight different apartments using this method.
- b. An alternative sampling method would be to select a stratified random sample of eight apartments, where the strata are apartments with children and apartments with no children. A stratified random sample of size eight might include two randomly selected apartments with children and six randomly selected apartments with no children. In the context of this situation, give one statistical advantage of selecting such a stratified sample as opposed to a cluster sample of eight apartments using the floors as clusters.



Chapter 4 “FRAPPY”

{Free Response AP Problem...Yay!}

The following problem is taken from an actual Advanced Placement Statistics Examination. Your task is to generate a complete, concise statistical response in 15 minutes (or less). You will be graded based on the AP rubric. After grading, keep this problem in your files for your AP Exam preparation.

1. As part of its twenty-fifth reunion celebration, the class of 1988 (students who graduated in 1988) at a state university held a reception on campus. In an informal survey, the director of alumni development asked 50 of the attendees about their incomes. The director computed the mean income of the 50 attendees to be \$189,952. In a news release, the director announced, “The members of our class of 1988 enjoyed resounding success. Last year’s mean income of its members was \$189,952!”

(a) What would be a statistical advantage of using the median of the reported incomes, rather than the mean, as the estimate of the typical income?

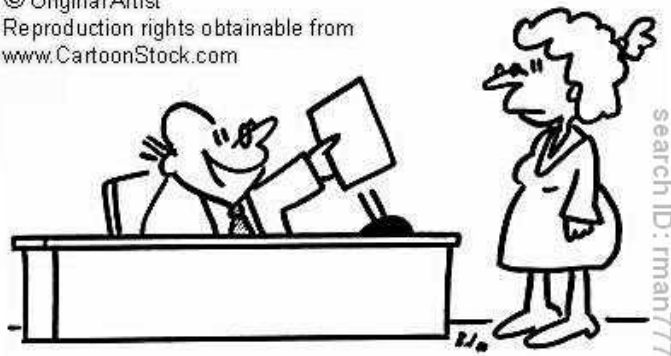
(b) The director felt the members who attended the reception may be different from the class as a whole. A more detailed survey of the class was planned to find a better estimate of the income as well as other facts about the alumni. The staff developed two methods based on the available funds to carry out the survey.

Method 1: Send out an e-mail to all 6,826 members of the class asking them to complete an online form. The staff estimates that at least 600 members will respond.

Method 2: Select a simple random sample of members of the class and contact the selected members directly by phone. Follow up to ensure that all responses are obtained. Because method 2 will require more time than method 1, the staff estimates that only 100 members of the class could be contacted using method 2.

Which of the two methods would you select for estimating the average yearly income of all 6,826 members of the class of 1988? Explain your reasoning by comparing the two methods and the effect of each method on the estimate.

© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com



"Take these statistics, Miss Findlay,
and make them warm and fuzzy."

Part 1: Chapter 1 Multiple Choice answers

- | | | |
|-----------|-----------|-----------|
| 1. _____ | 11. _____ | 21. _____ |
| 2. _____ | 12. _____ | 22. _____ |
| 3. _____ | 13. _____ | 23. _____ |
| 4. _____ | 14. _____ | 24. _____ |
| 5. _____ | 15. _____ | 25. _____ |
| 6. _____ | 16. _____ | 26. _____ |
| 7. _____ | 17. _____ | 27. _____ |
| 8. _____ | 18. _____ | 28. _____ |
| 9. _____ | 19. _____ | 29. _____ |
| 10. _____ | 20. _____ | 30. _____ |

Part 2: Chapter 4 Multiple Choice answers

- | | | |
|-----------|-----------|-----------|
| 1. _____ | 13. _____ | 25. _____ |
| 2. _____ | 14. _____ | 26. _____ |
| 3. _____ | 15. _____ | 27. _____ |
| 4. _____ | 16. _____ | 28. _____ |
| 5. _____ | 17. _____ | 29. _____ |
| 6. _____ | 18. _____ | 30. _____ |
| 7. _____ | 19. _____ | 31. _____ |
| 8. _____ | 20. _____ | 32. _____ |
| 9. _____ | 21. _____ | 33. _____ |
| 10. _____ | 22. _____ | 34. _____ |
| 11. _____ | 23. _____ | 35. _____ |
| 12. _____ | 24. _____ | 36. _____ |

Free Response - answers should be on the FRAPPYs