From the 2023 Administration

AP® CollegeBoard

AP® Statistics

This exam may not be posted on school or personal websites, nor electronically redistributed for any reason. This Released Exam is provided by the College Board for AP Exam preparation. Teachers are permitted to download the materials and make copies to use with their students in a classroom setting only. To maintain the security of this exam, teachers should collect all materials after their administration and keep them in a secure location.

Further distribution of these materials outside of the secure College Board site disadvantages teachers who rely on uncirculated questions for classroom testing. Any additional distribution is in violation of the College Board's copyright policies and may result in the termination of Practice Exam access for your school as well as the removal of access to other online services such as the AP Teacher Community and Online Score Reports.

©2023The College Board. College Board, Advanced Placement, AP, AP Central, and the acorn logo are registered trademarks of the College Board. Visit the College Board on the web: collegeboard.org.

制作者微信：liuxue119118
由于资料中间有修改添加过，请加微信确认资料是否完整
以免影响你的学习
老师微信：liuxue119118（真题有修改过，请加微信确认是否完整，以免影响您的学习！）

STATISTICS

SECTION I

Time—1 hour and 30 minutes

40 Questions

Directions: Each of the questions or incomplete statements below is followed by five suggested answers or completions. Select the one that is best in each case and then fill in the corresponding circle on the answer sheet.

1. Students in an online course are each randomly assigned to receive either standard practice exercises or adaptive practice exercises. For the adaptive practice exercises, the next question asked is determined by whether the student got the previous question correct. The teacher of the course wants to determine whether there is a difference between the two practice exercise types by comparing the proportion of students who pass the course from each group. The teacher plans to test the null hypothesis $p_S = p_A$ that versus the alternative hypothesis $p_S \neq p_A$, where p_S represents the proportion of students who would pass the course using standard practice exercises and p_A represents the proportion of students who would pass the course using adaptive practice exercises.

   The teacher knows that the 99 percent confidence interval for the difference in proportion of students passing the course for the two practice exercise types (standard minus adaptive) is $(0.03, 0.24)$ and the 90 percent confidence interval for the difference in the proportion of students passing the course for the two practice exercise types (standard minus adaptive) is $(0.06, 0.21)$. Which of the following is the most reasonable conclusion for the teacher's hypothesis test at the $\alpha = 0.05$ level?

   (A) The p-value is less than 0.01, so there is convincing statistical evidence that the proportion of students who would pass the course for the two practice exercise types differs at the 5% level.

   (B) The p-value is between 0.01 and 0.10, so there is convincing statistical evidence that the proportion of students who would pass the course for the two practice exercise types differs at the 5% level.

   (C) The p-value is between 0.01 and 0.10, so the researchers cannot determine whether the proportion of students who would pass the course for the two practice exercise types differs at the 5% level.

   (D) The p-value is between 0.01 and 0.10, so there is <u>not</u> convincing statistical evidence that the proportion of students who would pass the course for the two practice exercise types differs at the 5% level.

   (E) The p-value is greater than 0.10, so there is <u>not</u> convincing statistical evidence that the proportion of students who would pass the course for the two practice exercise types differs at the 5% level.

2. An experiment was conducted to assess the effect of baking temperature on the density of bread. A large batch of white bread dough and a large batch of whole wheat bread dough were prepared. Each of the batches was formed into 18 loaves, resulting in a total of 36 loaves. Each of the white bread loaves was randomly assigned to one of three baking temperatures (low, medium, and high), so that 6 loaves were baked at the low temperature, 6 were baked at the medium temperature, and 6 were baked at the high temperature. A similar process of random assignment was carried out for the whole wheat bread loaves.

   Which of the following are the treatments in the experiment?

   (A) The 36 loaves of bread

   (B) The type of bread (white, whole wheat)

   (C) The density of the bread

   (D) The baking temperatures (low, medium, high)

   (E) The two batches of bread dough

3. Concerned about the length of high school basketball games, the governing body for high school sports in a certain state is investigating the possibility of shortening game length if the mean game length is greater than 90 minutes. A random sample of 25 games from the current season was selected, and the sample mean game length was 95 minutes with a sample standard deviation of 10 minutes.

   Assuming that the distribution of game length is approximately normal, which of the following is the appropriate probability statement for determining the p-value of the hypothesis test to investigate whether the mean game length is greater than 90 minutes?

   (A) $P\left(t > \dfrac{95-90}{\dfrac{10}{\sqrt{25}}}\right)$, where t follows a t-distribution with 24 degrees of freedom.

   (B) $P\left(t > \dfrac{90-95}{\dfrac{10}{\sqrt{25}}}\right)$, where t follows a t-distribution with 24 degrees of freedom.

   (C) $P\left(t > \dfrac{90-95}{\dfrac{10}{\sqrt{25}}}\right)$, where t follows a t-distribution with 25 degrees of freedom.

   (D) $P\left(z > \dfrac{90-95}{\dfrac{10}{\sqrt{25}}}\right)$, where z follows a standard normal distribution.

   (E) $P\left(z > \dfrac{95-90}{\dfrac{10}{\sqrt{25}}}\right)$, where z follows a standard normal distribution.

4. A chocolatier, a person who makes chocolate candy, at a candy company has developed a new dark chocolate candy and believes it will be less bitter than the current dark chocolate candy on the market. The chocolatier hired 40 professional taste testers to participate in an experiment. The chocolatier flipped a coin for each participant. If the coin landed on heads, the participant tasted the new dark chocolate candy first, then tasted the current dark chocolate candy. If the coin landed on tails, the participant tasted the current dark chocolate candy first, then tasted the new dark chocolate candy. The participants rated the bitterness of each dark chocolate candy on a scale of 1 to 10, with 1 being the least bitter and 10 being the most bitter. Using a paired t-test to test the belief, the chocolatier calculated a p-value of 0.274.

   Assuming the conditions for inference have been met, which of the following is the best conclusion for the chocolatier's hypothesis test?

   (A) Because the p-value is less than any reasonable level of significance, there is convincing statistical evidence to support the belief that the new dark chocolate candy is less bitter than the current dark chocolate candy on average for those similar to the participants in the study.

   (B) Because the p-value is less than any reasonable level of significance, there is <u>not</u> convincing statistical evidence to support the belief that the new dark chocolate candy is less bitter than the current dark chocolate candy on average for those similar to the participants in the study.

   (C) Because the p-value is greater than any reasonable level of significance, there is convincing statistical evidence to support the belief that the new dark chocolate candy is less bitter than the current dark chocolate candy on average for those similar to the participants in the study.

   (D) Because the p-value is greater than any reasonable level of significance, there is <u>not</u> convincing statistical evidence to support the belief that the new dark chocolate candy is less bitter than the current dark chocolate candy on average for those similar to the participants in the study.

   (E) Since the taste testers were not randomly selected, this was not a well-designed experiment and a conclusion about the bitterness of the new dark chocolate candy cannot be made.

5. A public health analyst is interested in investigating whether there is a linear relationship between body mass index (BMI) and fasting blood glucose (FBG) for a particular population of adults. The analyst measured the MBI and FBG for each adult in a sample of 227 volunteers. The computer output for the linear regression analysis predicting FBG from BMI is shown.

Predictor	Coef	SE Coef	T	P($>\|t\|$)
Constant	94.207	3.126	30.141	0.000
BMI	0.204	0.109	1.868	0.060
$r^2 = 0.015$				

   Assume the conditions for inference have been met. Based on the computer output, which of the following conclusions is appropriate for all adults in the population similar to those who volunteered for the study?

   (A) There is a significant linear relationship at the 0.01 level between FBG and BMI.

   (B) There is a significant linear relationship at the 0.05 level but not at the 0.01 level between FBG and BMI.

   (C) There is a significant linear relationship at the 0.10 level but not at the 0.05 level between FBG and BMI.

   (D) There is a significant linear relationship at the 0.25 level but not at the 0.10 level between FBG and BMI.

   (E) There is not a significant linear relationship at the 0.25 level between FBG and BMI.

6. The distribution of body temperature for healthy adults is approximately normal with mean 98.2°F and standard deviation 0.7°F. Based on the description of the distribution, which of the following statements CANNOT be supported?

   (A) More than 95% of healthy adults have a body temperature greater than 96.8°F.

   (B) The first quartile of the distribution of healthy adults' body temperature is less than 97.5°F.

   (C) Approximately 50% of healthy adults have a body temperature greater than 98.2°F.

   (D) Approximately 95% of healthy adults have a body temperature between 96.8°F and 99.6°F.

   (E) A body temperature of 100.3°F or more would occur in less than 2% of healthy adults.

7. A research scientist is conducting an experiment to determine whether a new chemical process creates less toxic byproduct compared to the current chemical process. The scientist took a random sample of products made with the current chemical process and calculated the sample mean amount of toxic byproduct created. The scientist then took a random sample of products made with the new chemical process and calculated the sample mean amount of toxic byproduct created. The difference in the sample means (current minus new) was 2.31 liters of toxic byproduct. A hypothesis test was conducted using the following hypotheses.

$$H_0: \mu_{\text{current}} - \mu_{\text{new}} = 0$$
$$H_a: \mu_{\text{current}} - \mu_{\text{new}} \neq 0$$

   Assuming the conditions for inference were met, the scientist calculated the p-value of the test to be 0.072. Which of the following statements is the best interpretation of the p-value?

   (A) The probability that the null hypothesis is true is 0.072.

   (B) The probability that the alternative hypothesis is true is 0.072.

   (C) The probability of observing a difference in means of 2.31 liters of toxic byproduct is 0.072.

   (D) If the null hypothesis is true, the probability of observing a difference in means of at least 2.31 liters of toxic byproduct is 0.072.

   (E) If the null hypothesis is true, the probability of observing a difference in means of at most 2.31 liters of toxic byproduct is 0.072.

8. Consider a population consisting of four children with ages, in years, of 2, 5, 6, and 11. The population mean age is 6 years, and the population variance is 10.5 square years.

   For simple random samples of size 2 taken with replacement, what are the mean and standard deviation of the sampling distribution of the sample mean age, in years?

   (A) Mean 3 and standard deviation 5.25

   (B) Mean 3 and standard deviation 2.29

   (C) Mean 6 and standard deviation 4.24

   (D) Mean 6 and standard deviation 2.29

   (E) Mean 6 and standard deviation 1.87

9. For a chi-square test for independence with 6 degrees of freedom and a significance level of 0.05, the value of the test statistic was $\chi^2 = 10.497$. The associated p-value was 0.1052.

   Assuming all necessary conditions for the chi-square test were met, what type of error is possible?

   (A) A Type I error is possible, because the null hypothesis was rejected.

   (B) A Type I error is possible, because the null hypothesis was not rejected.

   (C) No error is possible, because the correct decision was made.

   (D) A Type II error is possible, because the null hypothesis was rejected.

   (E) A Type II error is possible, because the null hypothesis was not rejected.

10. A prize game at a circus has a spinner with 12 sections, where each section is equally likely to be selected when the spinner is spun. There are 7 blue sections and 5 orange sections. Camille has spun the spinner 4 times and the pattern of outcomes was orange, blue, orange, blue. Camille will spin the spinner two more times.

    Which of the following statements is true?

    (A) On Camille's fifth spin of the spinner, the probability that the spinner will land in a blue-colored section is equal to the probability that the spinner will land in an orange-colored section.

    (B) On Camille's fifth spin of the spinner, the probability that the spinner will land in a blue-colored section is less than the probability that the spinner will land in an orange-colored section.

    (C) On Camille's fifth and sixth spins of the spinner, the probability that the spinner will land in a blue-colored section for both spins is equal to the probability that the spinner will land in an orange-colored section for both spins.

    (D) On Camille's fifth and sixth spins of the spinner, the probability that the spinner will land in a blue-colored section for both spins is greater than the probability that the spinner will land in an orange-colored section for both spins.

    (E) On Camille's fifth and sixth spins of the spinner, the probability that the spinner will land in a blue-colored section for both spins is less than the probability that the spinner will land in an orange-colored section for both spins.

11. Suppose a new regulation requires businesses to pay all employees a minimum wage of \$15.00 per hour. A small coffee shop currently has 9 employees—5 employees whose wages are less than \$15.00 per hour and 4 employees whose wages are more than \$15.00 per hour. To meet the new regulation, the coffee shop plans to increase the wages of the 5 employees whose wages are less than \$15.00 per hour to \$15.00 per hour. The other 4 employees will not receive a wage increase at this time.

    Which of the following must be true about the median wage of the 9 employees after the wage increase occurs?

    (A) The median wage will be greater than \$15.00 per hour.

    (B) The median wage will increase but will be less than \$15.00 per hour.

    (C) The median wage will be \$15.00 per hour.

    (D) The median wage will not change.

    (E) The median wage cannot be determined from the information provided.

12. A technology company has 1,200 employees. The director of human resources at the company selects 250 of the employees at random, without replacement, and asks them to fill out an anonymous survey asking whether they plan to attend the company's holiday party. The proportion of employees in the sample who plan to attend the party is 0.816. Based on this proportion, the director calculates a 95 percent confidence interval for the proportion of all employees who plan to attend the holiday party as $0.816 \pm 1.96\sqrt{\frac{(0.816)(1-0.816)}{250}}$.

    Which of the following is true concerning whether it is appropriate to use the confidence interval to estimate the proportion of all employees who plan to attend the holiday party?

    (A) The confidence interval is appropriate to use because the employees were randomly selected.

    (B) The confidence interval is not appropriate to use because one of the values $n\hat{p}$ and $n(1-\hat{p})$ is less than than 10.

    (C) The confidence interval is not appropriate to use because 250 employees is more than 10% of all employees at the company.

    (D) The confidence interval is appropriate to use because $n\hat{p}$ and $n(1-\hat{p})$ are greater than 10.

    (E) The confidence interval is appropriate to use because $n > 30$.

13. A total of 13 distinct groups, but with each group similar to the other groups, participated in a craft fair. To investigate how satisfied the participants were with the fair, the fair director conducted a survey. The director randomly selected 5 of the groups participating in the fair and surveyed all members of the selected groups.

    Which statement best describes the sampling method used by the director?

    (A) It was convenience sampling, because participants from each group were easy to contact.

    (B) It was simple random sampling, because the participants were selected at random.

    (C) It was stratified random sampling, because the director ensured that members from each group were selected.

    (D) It was cluster sampling, because the sample consisted of all members of the 5 randomly selected groups.

    (E) It was systematic random sampling, because the director selected exactly 5 groups.

14. A grocery store received three shipments of bags of cranberries: one from farm A, one from farm B, and one from farm C. The distributions of the weight of a bag are unimodal and symmetric for all three shipments. The number of bags and the mean and standard deviation, in ounces (), for each shipment are shown in the table.

Farm	Number of Bags	Mean (oz)	Standard Deviation (oz)
A	100	16.0	0.5
B	100	16.0	0.1
C	100	16.0	0.1

    Which of the following statements is true?

    (A) The shipment from farm A is most likely to have the greatest number of bags that weigh less than 16 oz.

    (B) The shipment from farm B is most likely to have the greatest number of bags that weigh less than 16 oz.

    (C) The shipment from farm C is most likely to have the greatest number of bags that weigh less than 16 oz.

    (D) The shipments from farms B and C are the only shipments that are likely to have an approximately equal number of bags that weigh less than 16 oz.

    (E) The shipments from farms A, B, and C are all likely to have an approximately equal number of bags that weigh less than 16 oz.

15. A certain computer chip company has a machine that makes a large number of computer chips. Of the computer chips produced on this machine, 3% are defective. Miriam, the quality control inspector, will randomly select computer chips from the machine to inspect.

    What is the probability that the first defective computer chip Miriam finds is the sixth computer chip she inspects from the machine?

    (A) $(0.03)^5(0.97)$
    (B) $1-(0.03)^6$
    (C) $(0.97)^5(0.03)$
    (D) $(0.97)^6$
    (E) $1-(0.97)^6$

16. A county manager selected a random sample of 200 residents from a suburban area of the county and a separate random sample of 200 residents from an urban area of the same county. The residents in the samples were asked whether they thought the county should create dog parks where residents can take their dogs to play. From those who were sampled, 41 percent of the suburban residents and 35 percent of the urban residents thought the county should create dog parks.

    Let p_S represent the population proportion of suburban residents who think the county should create dog parks, and let p_U represent the population proportion of urban residents who think the county should create dog parks. The county manager will test the following hypotheses.

    $$\text{H}_0\text{: } p_S - p_U = 0$$
    $$\text{H}_\text{a}\text{: } p_S - p_U \neq 0$$

    Assuming the conditions for inference were met, which of the following is closest to the p-value of the county manager's test?

    (A) 0.7778
    (B) 0.4444
    (C) 0.2222
    (D) 0.2164
    (E) 0.1082

17. A fair coin will be tossed three times. The set of possible outcomes may be described by the ordered triples

    $$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

    , where H represents heads and T represents tails.

    What is the probability that all three tosses result in heads, given that the three tosses result in more heads than tails?

    (A) $\frac{1}{8}$
    (B) $\frac{1}{4}$
    (C) $\frac{3}{8}$
    (D) $\frac{1}{2}$
    (E) $\frac{3}{4}$

18. A researcher wants to examine the effect of 2 different strengths—low and high—of fertilizer on the growth of corn. A field will be divided into 54 equally sized plots of land, and the 2 fertilizer strengths will be randomly assigned so that 27 plots will be fertilized with low strength and 27 plots with high strength. The same type and amount of corn will be planted in each plot at the same time. The yield (amount of corn produced) will be measured for each plot.

    Which of the following would provide the best evidence that the strength of fertilizer had an effect on the yield?

    (A) The variation in yields within the fertilizer strength groups is greater than the variation in the average yields between the fertilizer strength groups.

    (B) The variation in yields within the fertilizer strength groups is less than the variation in the average yields between the fertilizer strength groups.

    (C) The variation in yields is about the same for all plots.

    (D) The variation in yields of plots receiving the low-strength fertilizer is greater than the variation in yields of plots receiving the high-strength fertilizer.

    (E) The variation in yields of plots receiving the low-strength fertilizer is less than the variation in yields of plots receiving the high-strength fertilizer.

19. According to the United States Social Security Administration (SSA), for children born years ago, the following table gives the percentage of children with a first name beginning with the letters A–C, D–J, K–M, and N–Z.

Percentages from the SSA

Beginning letter of first name	A–C	D–J	K–M	N–Z
Percent of names recorded by the SSA beginning with that letter	25.65%	26.30%	25.10%	22.95%

Daphne, a senior at a large high school, will investigate whether the distribution of the beginning letter of first names of students at her high school who were born 18 years ago is consistent with the percentages from the Social Security Administration. She took a random sample of 80 students from her school who were born 18 years ago. Daphne recorded the number of students with a first name beginning with the letters A–C, D–J, K–M, and N–Z. The results are shown in the following table.

Numbers from Daphne's Sample

Beginning letter of first name	A–C	D–J	K–M	N–Z
Number of names in Daphne's sample beginning with that letter	17	25	15	23

Assuming the conditions for inference are met, what is the appropriate chi-square test and calculation of the test statistic for her investigation?

(A) A χ^2 goodness-of-fit test; $\chi^2 = \dfrac{(17-25.05)^2}{25.05} + \dfrac{(25-26.30)^2}{26.30} + \dfrac{(15-25.10)^2}{25.15} + \dfrac{(23-22.95)^2}{22.95}$

(B) A χ^2 goodness-of-fit test; $\chi^2 = \dfrac{(17-20)^2}{20} + \dfrac{(25-20)^2}{20} + \dfrac{(15-20)^2}{20} + \dfrac{(23-20)^2}{20}$

(C) A χ^2 goodness-of-fit test; $\chi^2 = \dfrac{(17-20.52)^2}{20.52} + \dfrac{(25-21.04)^2}{21.04} + \dfrac{(15-20.08)^2}{20.08} + \dfrac{(23-18.36)^2}{18.36}$

(D) A χ^2 test for independence; $\chi^2 = \dfrac{(17-25.05)^2}{25.05} + \dfrac{(25-26.30)^2}{26.30} + \dfrac{(15-25.10)^2}{25.15} + \dfrac{(23-22.95)^2}{22.95}$

(E) A χ^2 test for independence; $\chi^2 = \dfrac{(17-20)^2}{20} + \dfrac{(25-20)^2}{20} + \dfrac{(15-20)^2}{20} + \dfrac{(23-20)^2}{20}$

20. To determine whether an herbal supplement helps relieve knee pain, researchers will select 50 patients with knee pain to participate in a study. The participants will be randomly assigned to one of two groups, with 25 participants in each group. The first group will receive the herbal supplement daily for one week, and the second group will receive a placebo daily for one week. Patients will continue to take their current pain medication during the week. At the end of the week, the patients will record their pain score on a scale of 1 to 10, where 1 represents no pain and 10 represents extreme pain.

    Assuming the conditions for inference are met, which of the following would be the most appropriate statistical method to estimate the difference in mean pain score between the herbal supplement and the placebo for all patients with knee pain who are similar to the participants in the study?

    (A) A one-sample z-interval for estimating a population proportion

    (B) A one-sample t-interval for estimating a population mean

    (C) A matched-pairs t-interval for estimating a population mean difference

    (D) A two-sample z-interval for estimating a difference between population proportions

    (E) A two-sample t-interval for estimating a difference between population means

21. It is known that 75 percent of all households in the United States have intermittent Internet connectivity problems.

    For 450 random samples of independent households in the United States, what are the mean, μ, and standard deviation, σ, of the distribution of the number of households that have intermittent Internet connectivity problems?

    (A) $\mu = 450(0.25)$, $\sigma = \sqrt{450(0.25)(0.75)}$

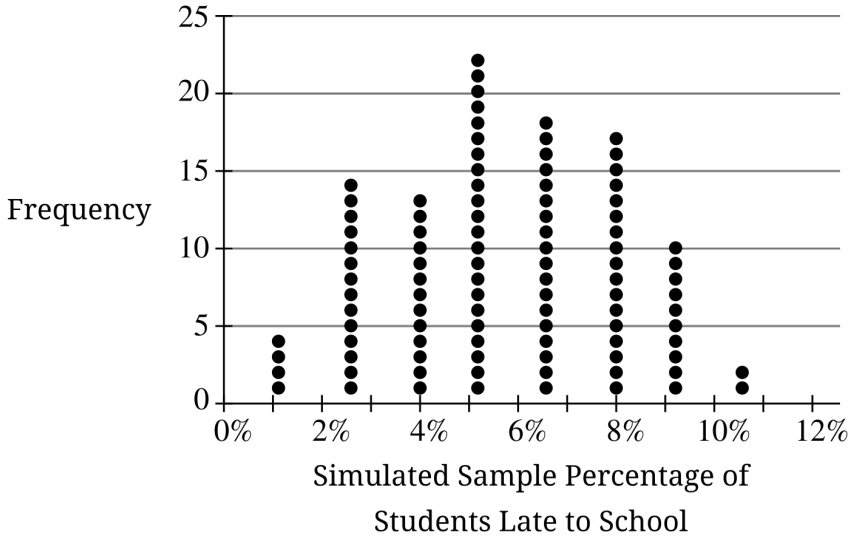
    (B) $\mu = 450(0.75)$, $\sigma = \sqrt{450(0.75)(0.25)}$

    (C) $\mu = 450(0.25)(0.75)$, $\sigma = \sqrt{\dfrac{(0.25)(0.75)}{450}}$

    (D) $\mu = 450(0.25)(0.75)$, $\sigma = \sqrt{450(0.25)(0.75)}$

    (E) $\mu = 0.75$, $\sigma = \sqrt{\dfrac{(0.75)(0.25)}{450}}$

22. Attendance records at a large high school show that 6 percent of students are late to school on a typical day. The students in an AP Statistics class believe the rate is higher than 6 percent. The students took a simple random sample of 75 students from the school and asked, "Were you late to school today?" In their sample, 6 out of the 75 students, or 8 percent of the sample, said they were late to school that day. The AP Statistics students conducted a simulation with 100 samples of 75 students from a population where 6 percent of students are late. The results of the simulation are shown in the dotplot.



Based on the simulation and the results of the AP Statistics class's sample, is there convincing statistical evidence that more than 6 percent of all students at the school are late to school on a typical day?

(A) Yes, because 8% of students in their sample said they were late to school that day.

(B) Yes, because about 47% of the simulated sample percentages of students late to school are 6% or more.

(C) No, because about 47% of the simulated sample percentages of students late to school are 6% or more.

(D) Yes, because about 29% of the simulated sample percentages of students late to school are 8% or more.

(E) No, because about 29% of the simulated sample percentages of students late to school are 8% or more.

23. The distribution of long jump distance in a statewide high school league is approximately normal with a mean of 18.5 feet and a standard deviation of 3.2 feet. Two students, Joshua and Mateo, from this league competed in the long jump. Joshua's distance was 0.52 standard deviation below the mean of the distribution. Mateo's distance was at the 48th percentile of the distribution.

At approximately what percentile was the distance jumped by Joshua, and how did his distance compare to Mateo's distance?

(A) Joshua's distance was at the 30th percentile of the distribution and was less than Mateo's distance.

(B) Joshua's distance was at the 30th percentile of the distribution and was greater than Mateo's distance.

(C) Joshua's distance was at the 48th percentile of the distribution and was less than Mateo's distance.

(D) Joshua's distance was at the 48th percentile of the distribution and was greater than Mateo's distance.

(E) Joshua's distance was at the 70th percentile of the distribution and was greater than Mateo's distance.

24. A large manufacturing company has tens of thousands of employees, and 30 percent of the employees have completed a management training course. Suppose a simple random sample of 50 employees is selected.

    Which of the following is closest to the probability that more than half of the sample will be employees who have completed the management training course?

    (A) 0.001
    (B) 0.065
    (C) 0.170
    (D) 0.300
    (E) 0.999

25. The recreation committee of a certain town is interested in creating a new nature trail for walkers, runners, and bikers. A random sample of 100 town residents were asked whether they are in favor of creating the new trail. Of those sampled, 78 said yes.

    What is the population of interest for which the committee can generalize its conclusions?

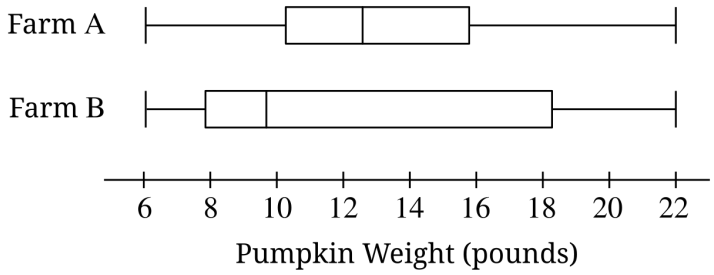
    (A) The 100 town residents who were asked about whether they are in favor of creating the new trail
    (B) The 78 town residents who responded that they are in favor of creating the new trail
    (C) All town residents who are walkers, runners, or bikers
    (D) All town residents who are in favor of the new trail
    (E) All town residents

26. Carol, a reporter in a large city, took a random sample of 300 households and found that the 90 percent confidence interval for the proportion of households that own a pet is $(0.63, 0.71)$. Jason, the director of a local animal shelter, wants to verify this result for the proportion of households that own a pet and takes a different random sample of 200 households and finds that 67 percent own a pet.

    Assume the conditions for inference have been met. Jason will calculate a 90 percent confidence interval for the proportion of households that own a pet using his sample data. How would the margin of error from Jason's confidence interval compare to the margin of error from Carol's confidence interval?

    (A) Jason's margin of error would be the same as Carol's margin of error because the confidence level did not change.
    (B) Jason's margin of error would be the same as Carol's margin of error because the point estimate did not change.
    (C) Jason's margin of error would be less than Carol's margin of error because Jason used a smaller sample size.
    (D) Jason's margin of error would be greater than Carol's margin of error because Jason used a smaller sample size.
    (E) Jason's margin of error would be greater than Carol's margin of error because Jason's point estimate is greater.

27. A botanist is studying the weights, measured in pounds, of pumpkins grown at two different farms. The boxplots show the distribution of pumpkin weight for two samples of equal size, one from farm A and the other from farm B.



Based on the boxplots, which of the following is a correct conclusion about the distribution of pumpkin weight for the samples from the two farms?

(A) The interquartile ranges of the pumpkin weights are the same for farm A and farm B.

(B) The range of the pumpkin weights for farm B is greater than the range for farm A.

(C) There are more pumpkin weights that are greater than 18 pounds from farm A than there are from farm B.

(D) There are more pumpkin weights that are greater than 12 pounds from farm B than there are from farm A.

(E) There are more pumpkin weights that are less than 10 pounds from farm B than there are from farm A.

28. Voters in a certain large county will have an opportunity to vote for or against a new environmental rule that would incentivize the return of certain used batteries to a county-certified recycling facility. The results of a previous poll found that 41 percent of registered voters in that county plan to vote for the environmental rule. Juwan, an engineer working at the recycling facility, randomly selected 15 registered voters and asked whether they plan to vote for the rule.

Assuming that 41 percent represents the actual percent of registered voters in the county who plan to vote for the environmental rule, which of the following is closest to the probability that at most 2 of the registered voters Juwan selected plan to vote for the environmental rule?

(A) 0.0042

(B) 0.0185

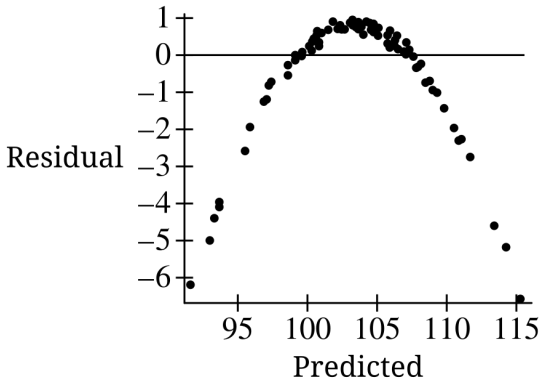
(C) 0.0227

(D) 0.9773

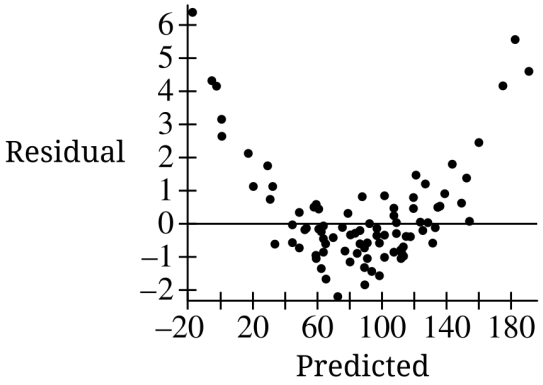
(E) 0.9958

29. A data scientist had different 5 sets of data. For each data set the scientist fit a least-squares regression line, then graphed the residuals versus the predicted values for each. Which residual plot indicates that the least-squares regression model was the most appropriate model to fit the data?

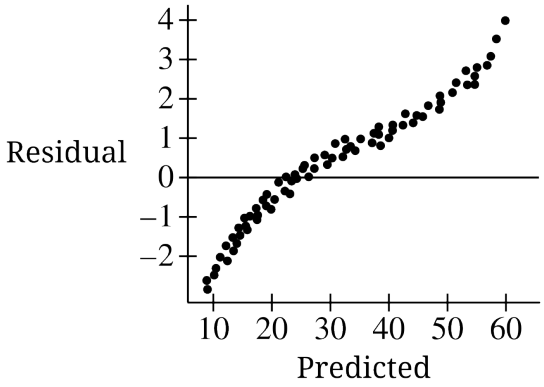
(A)



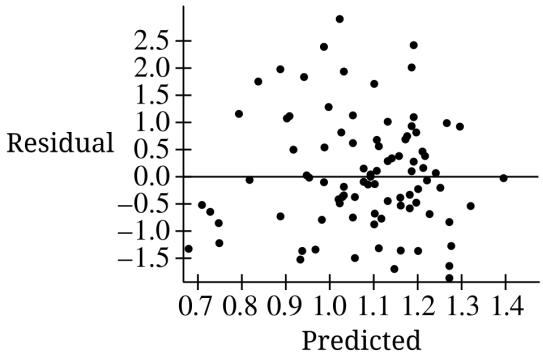
(B)



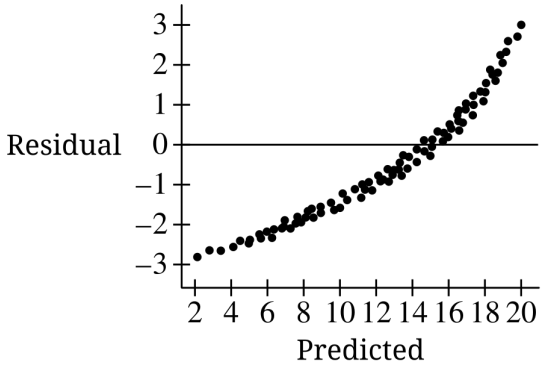
(C)



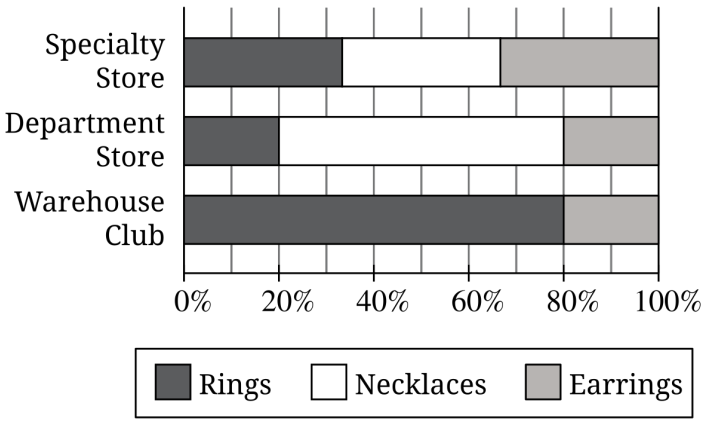
(D)



(E)



30. A jewelry manufacturer tracks sales of three types of jewelry (rings, necklaces, and earrings) across each of three different retail outlet categories (specialty store, department store, and warehouse club). The segmented bar chart shows the distribution for each retail outlet for the three jewelry types based on sales data for the past week.



Based on the chart, which statement must be true?

(A) The percent of earring sales at department stores was less than the percent of necklace sales at specialty stores.

(B) For each retail outlet category, less than 30% of sales were earrings.

(C) The number of earring sales for department stores was the same as the number of earring sales for warehouse clubs.

(D) The greatest number of ring sales came from warehouse clubs.

(E) The number of ring sales for specialty stores was greater than the number of ring sales for warehouse clubs.

31. A one-sample z-test for a population proportion was conducted to test $H_0 : p = 0.35$ versus $H_a : p \neq 0.35$.

Assuming the conditions for inference were met, which combination of sample proportion and sample size would result in the smallest p-value of the test?

(A) $\hat{p} = 0.35$ and $n = 200$

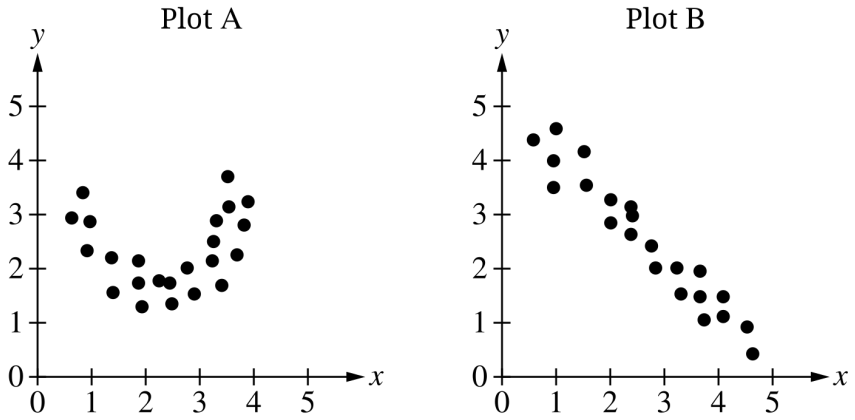
(B) $\hat{p} = 0.40$ and $n = 100$

(C) $\hat{p} = 0.40$ and $n = 60$

(D) $\hat{p} = 0.25$ and $n = 100$

(E) $\hat{p} = 0.25$ and $n = 60$

32. Two scatterplots are shown.



Based on the scatterplots, which statement CANNOT be true?

(A) Plot A indicates a nonlinear relationship between the variables.

(B) Plot B indicates a linear relationship between the variables.

(C) The correlation coefficient for the data in plot B is close to 1.

(D) The correlation coefficient for the data in plot A is close to 0.

(E) The r^2 value of the least-squares regression line for plot B is greater than the r^2 value of the least-squares regression line for plot A.

33. Researchers studying the relationship between screen time and developmental scores of children in Canada selected a random sample of 2,441 children. For each child, the researchers recorded the child's screen time, in hours per day, when they were 2 years old (x) and their scores on a developmental test, in points, when they were 5 years old (y). Higher scores on the developmental test are better. The 95 percent confidence interval for the slope of the least-squares regression line for predicting scores on the developmental test based on screen time was $(-0.13, -0.02)$.

The conditions for inference were met, and the researchers concluded that increased screen time is associated with lower scores on the developmental test. Which of the following provides the best reason that supports the researchers' conclusion?

(A) This was a 95% confidence interval, which is a high level of confidence.

(B) The sample size was large.

(C) The researchers took years to complete the study.

(D) The confidence interval for the slope of the regression line does not contain 0.

(E) The confidence interval for the slope of the regression line consists entirely of negative numbers.

34. To investigate whether more than half the adults in a country are fearful of riding in a self-driving car, researchers surveyed a random sample of 1,600 adults in the country. Of those in the sample, 1,280 adults responded that they would be fearful of riding in a self-driving car.

Which of the following represents the value of the test statistic for the appropriate hypothesis test for the investigation?

(A) $\dfrac{0.8-0.5}{\sqrt{\dfrac{(0.2)(0.8)}{1600}}}$

(B) $\dfrac{0.8-0.5}{\sqrt{\dfrac{(0.5)(0.5)}{1600}}}$

(C) $\dfrac{0.8-0.5}{\sqrt{\dfrac{(0.2)(0.8)}{1600}+\dfrac{(0.5)(0.5)}{1600}}}$

(D) $\dfrac{0.5-0.8}{\sqrt{\dfrac{(0.2)(0.8)}{1600}}}$

(E) $\dfrac{0.5-0.8}{\sqrt{\dfrac{(0.5)(0.5)}{1600}}}$

35. Researchers will conduct an experiment to compare the effectiveness of two brands of shampoo—Brand X and Brand Y. They will use a matched-pairs design in which each participant will use one brand for one week and then the other brand for one week. For each participant, the researchers will randomly assign which shampoo, Brand X or Brand Y, will be used first, and 15 milliliters of shampoo will be used for each hair wash. Because the temperature of the water is known to have an effect, a water temperature less than 100°F will be used for the first week and a water temperature greater than 100°F will be used for the second week. At the end of each week, the effectiveness of the shampoo being used will be measured.

Which of the following would improve this matched-pairs experiment?

(A) Have each participant use both brands, but have all participants use Brand X first, then switch to Brand Y

(B) Have each participant use only one of the shampoo brands, with brand randomly assigned for each participant

(C) Have each participant use only one of the shampoo brands, with brand determined by the subject's preference

(D) Let each participant decide for themselves how much shampoo to use

(E) Let each participant decide for themselves what water temperature to use, but each participant must keep the water temperature the same for both brands

36. Gwyneth measured the weights, in pounds, of ten striped bass taken from a local river. She calculated a mean weight of 13 pounds and a range of 5 pounds. For a report, Gwyneth converted the measurements from pounds to kilograms. (Note: There is approximately 0.454 kilogram in 1 pound.)

    Which of the following is closest to the mean and range of the converted weights?

    (A) The mean weight is 28.63 kilograms, and the range is 5 kilograms.

    (B) The mean weight is 28.63 kilograms, and the range is 2.27 kilograms.

    (C) The mean weight is 5.90 kilograms, and the range is 11.01 kilograms.

    (D) The mean weight is 5.90 kilograms, and the range is 5 kilograms.

    (E) The mean weight is 5.90 kilograms, and the range is 2.27 kilograms.

37. Karmen and Sophia are on a swim team and race in different events at a swim meet. The probability that Karmen will win her race at the swim meet is 0.1. The probability that Sophia will win her race at the swim meet is 0.3.

    If the events of Karmen winning her race and Sophia winning her race are independent, what is the probability that either Karmen or Sophia will win her race at the swim meet?

    (A) 0.03

    (B) 0.27

    (C) 0.30

    (D) 0.37

    (E) 0.40

38. As part of a survey, 118 adults with tattoos were asked whether they got their first tattoo before reaching the age of 25 and whether their first tattoo was done by a professional tattoo artist. The resulting data are summarized in the following table.

	Before the age of 25	**25 or older**	**Total**
Professional tattoo artist	10	87	97
Not a professional tattoo artist	9	12	21
Total	19	99	118

    Which of the following provides evidence that age at first tattoo and whether the first tattoo was done by a professional tattoo artist are <u>not</u> independent for the group of 118 adults with tattoos?

    (A) $\frac{10}{97} \neq \frac{87}{97}$

    (B) $\frac{10}{97} \neq \frac{19}{118}$

    (C) $\frac{10}{19} \neq \frac{10}{97}$

    (D) $\frac{19}{118} \neq \frac{99}{118}$

    (E) $\frac{97}{118} \neq \frac{21}{118}$

39. Daniela designed an experiment to investigate whether listening to music or burning a scented candle affects her performance while playing her favorite video game. Each day for one month, she randomly chose whether to listen to classical, rock, or no music while playing her favorite video game. Each day, she also randomly chose whether to burn a candle with an orange scent, burn a candle with a lavender scent, or not burn a candle. At the end of each day, she recorded the mean score for 10 plays of her favorite video game.

    How many explanatory variables are in the experiment?

    (A) One

    (B) Two

    (C) Three

    (D) Six

    (E) Nine

40. A real estate agent indicates that the mean selling price of all single-family homes in Washington, D.C., is $800,000. A family believes the mean selling price for all single-family homes that are currently for sale is less than the amount indicated by the real estate agent and will select a random sample of single-family homes that are currently for sale in Washington, D.C. Let μ represent the mean selling price, in dollars, of all single-family homes that are currently for sale in Washington, D.C.

    Assuming that the conditions for inference are met, which of the following hypotheses would be appropriate to test the family's belief?

    (A) $H_0 : \mu = 800,000$ versus $H_a : \mu > 800,000$

    (B) $H_0 : \mu = 800,000$ versus $H_a : \mu < 800,000$

    (C) $H_0 : \mu < 800,000$ versus $H_a : \mu = 800,000$

    (D) $H_0 : \bar{x} = 800,000$ versus $H_a : \bar{x} > 800,000$

    (E) $H_0 : \bar{x} = 800,000$ versus $H_a : \bar{x} < 800,000$

END OF SECTION I

STATISTICS

SECTION II

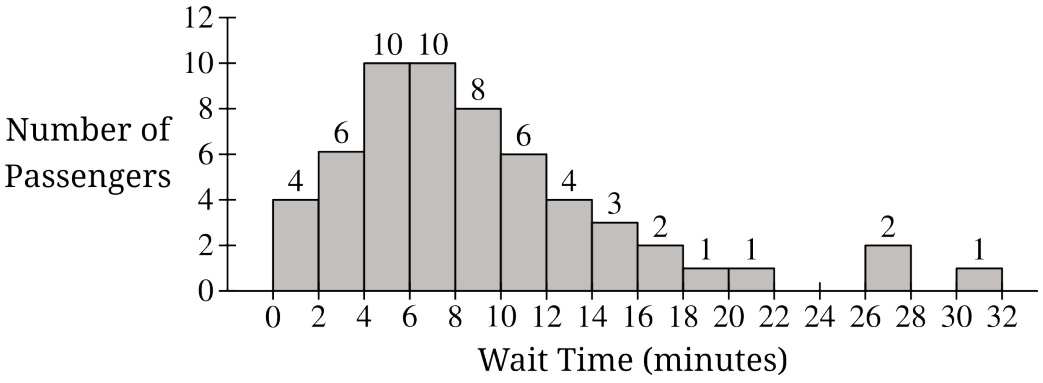
Time—1 hour and 30 minutes

6 Questions

Part A

Suggested Time — 1 hour and 5 minutes

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.



1. A large city is considering increasing the number of buses running on a certain route. To determine whether there is a need for more buses, a random sample of 58 people at bus stops along the route is selected. The amount of time a person waits until a bus arrives is recorded to the nearest tenth of a minute for each person. The histogram shows the distribution of wait time. The first bar includes values from 0 to 1.9 minutes, the second bar includes values from 2 to 3.9 minutes, and so on.

   (a) Based on the distribution of wait time shown in the histogram, what would be a statistical advantage of using the sample median rather than the sample mean as the estimate of the typical wait time?

   (b) Using the histogram, describe an appropriate procedure for estimating the sample median wait time. Use the procedure to provide an estimate for the sample median wait time.

   (c) Based on the histogram, the smallest possible interquartile range (IQR) for the wait times is 4.1 minutes and the largest possible IQR for the wait times is 7.9 minutes. Which of these two IQR values would identify more wait times as potential outliers? Justify your response.

2. Orthodontists use Temporary Anchoring Devices (TADs) to keep teeth in place. TADs can be made of stainless steel or titanium. A certain amount of damage is expected during use of the TADs, but too much damage to the TADs could result in a failure to achieve the desired results. An experiment was performed to compare the mean amount of damage to TADs made of stainless steel to the mean amount of damage to TADs made of titanium. Researchers recruited 386 volunteer orthodontic patients requiring TADs to participate in the experiment. Each volunteer had a TAD made of stainless steel placed on one side of the mouth and a TAD made of titanium placed on the other side of the mouth. For each volunteer, a flip of a fair coin determined which side of the mouth, left (Heads) or right (Tails), received the TAD made of stainless steel. The other side of the mouth received the TAD made of titanium. After six months, an orthodontist recorded the amount of damage observed on each TAD on a scale of 1 to 10, with 1 representing the least damage and 10 representing the most damage.

   (a) Identify the experimental units and treatments in the context of this experiment.
      - Experimental units
      - Treatments

   (b) Consider the design of this experiment.
      i. Identify the experimental design.
      ii. Characteristics of each volunteer, such as age, diet, and dental hygiene, influence the amount of damage to a TAD. Explain how the design of this experiment controls for the effects of these characteristics.

   (c) What would it mean for this experiment to be single-blind?

   (d) Suppose the experiment showed no significant difference in the mean amount of damage between the two types of TADs. Can these results be generalized to all orthodontic patients requiring TADs? Justify your response.

3. Jacob is taking a quiz that consists of a list of five quotations from history and a list of names of the five historical figures who said them. The quiz directions are to match each quotation on the first list with one of the names on the second list, matching each name to exactly one quotation.

   (a) Consider just the first two quotations. If Jacob randomly guesses two different names from the list of five names for the first two quotations, what is the probability that he correctly matches both?

   Now consider all five quotations and suppose that Jacob randomly guesses the matching name for each quotation, matching each name to exactly one quotation. Let the random variable X represent the number of correct matches out of the five quotations. The table shows the probability distribution of X.

Number of correct matches	0	1	2	3	4	5
Probability of x, $P(x)$	$\frac{44}{120}$	$\frac{45}{120}$	$\frac{20}{120}$	$\frac{10}{120}$	$\frac{0}{120}$	$\frac{1}{120}$

   (b) Calculate the conditional probability that Jacob will match exactly three quotations correctly, given that he matches at least two of the five quotations correctly. Show your work.

   (c) The expected value of the probability distribution of X is 1. What is the interpretation of this expected value in context?

   (d) Does the random variable X have a binomial distribution? Explain why or why not.

4. A large airport has parking spaces spread among 4 types of parking lots. The cost to park a car per day differs from lot to lot depending on the time it takes to get to the airport terminal. The parking lots are described as follows.

   - Premium parking is less than a 5-minute walk to the airport terminal. Premium parking spaces account for 15 percent of the total number of airport parking spaces.
   - General parking is a 5- to 10-minute walk to the airport terminal. General parking spaces account for 29.5 percent of the total number of airport parking spaces.
   - Economy parking is a 10- to 15-minute walk to the airport terminal. Economy parking spaces account for 37 percent of the total number of airport parking spaces.
   - Off-site parking uses a shuttle that operates every 15 minutes to transport people from the off-site parking lot to the airport terminal. Off-site parking spaces account for 18.5 percent of the total number of airport parking spaces.

   Tens of thousands of drivers park in the airport's parking lots each year. A random sample of 343 drivers who park in the airport's parking lots is selected during a given year, and each driver is asked where they would choose to park if all parking lot options were available to them. Airport officials counted the number of responses for each type of parking lot, and the results are summarized in the table.

Type of Parking Lot	**Number of Responses**
Premium	63
General	125
Economy	111
Off-site	44
Total	343

   Do the data in the table provide convincing statistical evidence that at least one of the proportions of all drivers who park in the airport's parking lots who would choose each type of parking lot is different from the proportion of the total number of airport parking spaces accounted for by that type of parking lot? Complete the appropriate inference procedure to support your answer using a significance level of $\alpha = 0.05$.

5. An important economic indicator is the labor participation rate. The labor participation rate is the proportion of the population of people 16 years and older that are working or actively looking for work. Dr. Ogden, an economist, believes this proportion for a certain city increased from January 2022 to December 2022. A random sample of people 16 years and older from that city was taken in January 2022 and an independent random sample of people 16 years and older from that city was taken in December 2022. The sample outcomes are shown in the table.

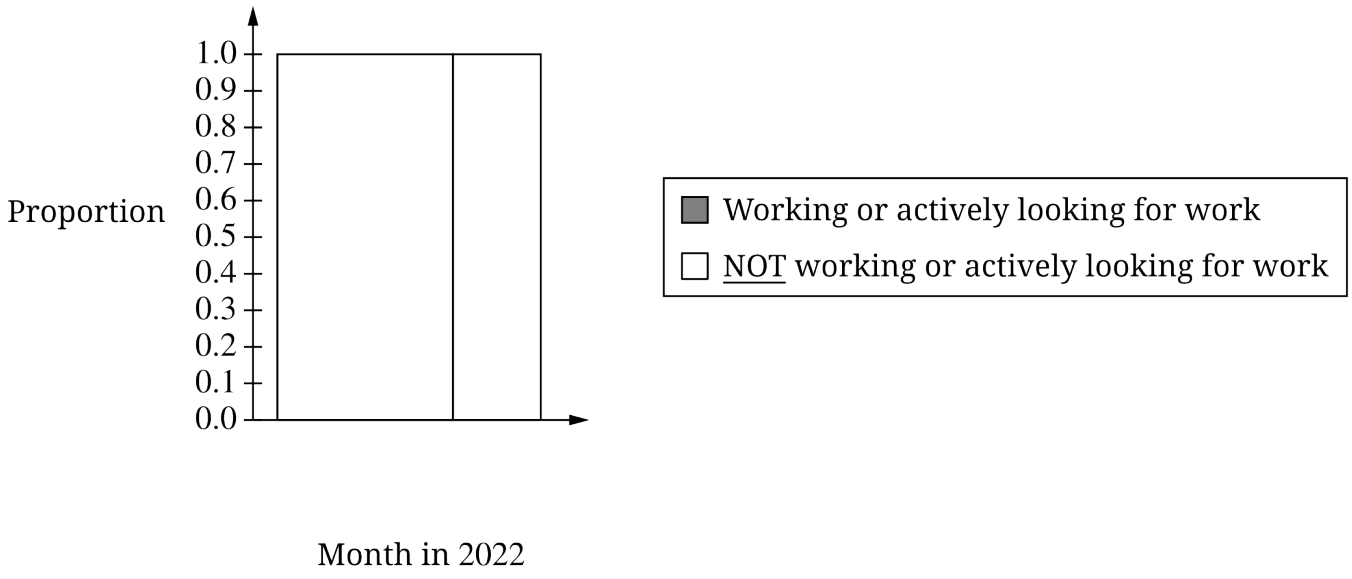
	January 2022	December 2022
Working or actively looking for work	55	35
NOT working or actively looking for work	50	15

(a) Consider the data in the table.

i. What proportion of people who were sampled in January 2022 were working or actively looking for work?

ii. What proportion of people who were sampled in December 2022 were working or actively looking for work?

(b) Complete the mosaic plot for these data.



Recall that Dr. Ogden believes the true proportion of 16 people years and older from the city that were working or actively looking for work increased from January 2022 to December 2022. To test this belief, Dr. Ogden will conduct a two-sample z-test for a difference in population proportions. Assume all conditions for inference have been met.

(c) Let p_J represent the true proportion of people 16 years and older from that city that were working or actively looking for work in January 2022 and p_D represent the true proportion of people 16 years and older from that city that were working or actively looking for work in December 2022.

i. What are the appropriate hypotheses for Dr. Ogden's test?

ii. The p-value of this test is $p = 0.019$. What conclusion should Dr. Ogden make at the $\alpha = 0.05$ level of significance?

STATISTICS

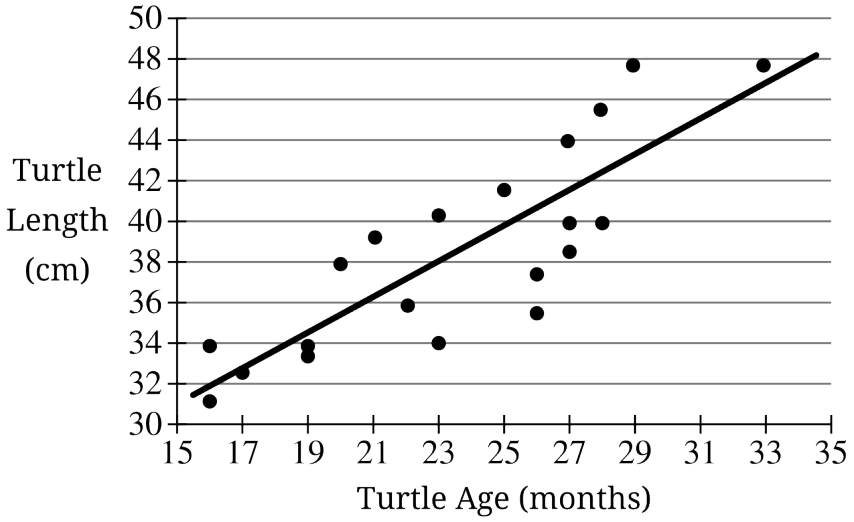
SECTION II, Part B

Suggested Time—25 30 minutes

1 Question

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. A zoo has 20 young turtles of a certain species. The zookeeper is interested in the relationship between the age in months and the length in centimeters (cm) of the 20 turtles. A scatterplot and the least-squares regression line relating age and length of the 20 turtles are shown.



The least-squares regression equation is predicted length $= 18.258 + 0.861$(age), and $r^2 = 0.701$.

(a) Calculate and interpret the correlation coefficient, r, between turtle age and length.

(b) There are 2 turtles that are 28 months old. One is female and one is male. The female turtle has a residual of -2.27 cm, and the male turtle has a residual of 3.13 cm.

i. On the scatterplot, circle the point that corresponds to the 28-month-old female turtle.

ii. Interpret the residual for the 28-month-old male turtle in context.

The 20 turtles make up the zoo's population for this species of turtle. There are 8 male turtles and 12 female turtles. The zookeeper wants to determine the strength of the relationship between sex of the turtle and length of the turtle. To do this, the zookeeper will use the point-biserial correlation. The point-biserial correlation is a measure of the strength of the relationship between a quantitative variable (e.g., length of the turtle) and a categorical variable with two categories (e.g., sex of the turtle). The value of the point-biserial correlation is between -1 and 1, and like the correlation coefficient, values closer to -1 or 1 indicate a stronger correlation between the two variables. A formula for the point-biserial correlation is $r_{pb} = \dfrac{(m_1 - m_2)\sqrt{p(1-p)}}{SD}$.

In this context, m_1 = the mean length of the male turtles, m_2 = the mean length of the female turtles, p = the proportion of the 20 turtles that are male, $1-p$ = the proportion of the 20 turtles that are female, and SD = the standard deviation of the lengths of the 20 turtles.

Numerical summaries for the turtle lengths are given in the following table.

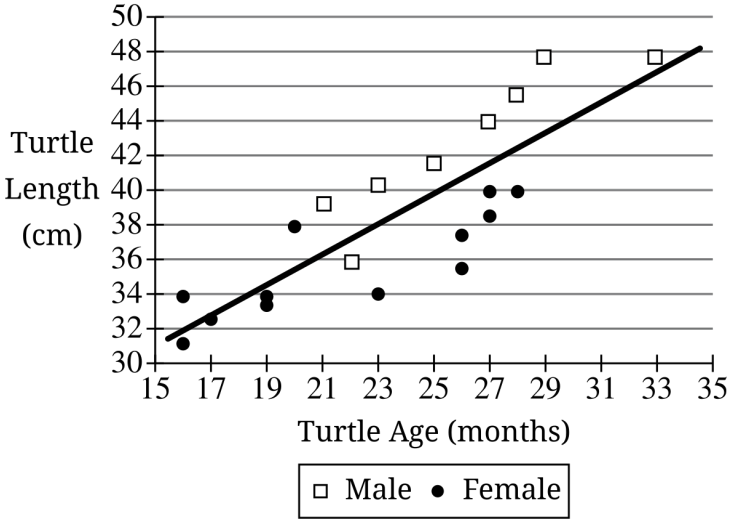
	Male	Female
Mean length (cm)	42.78	35.88
Number of turtles	8	12
Standard deviation of length of the 20 turtles (cm)	4.79	

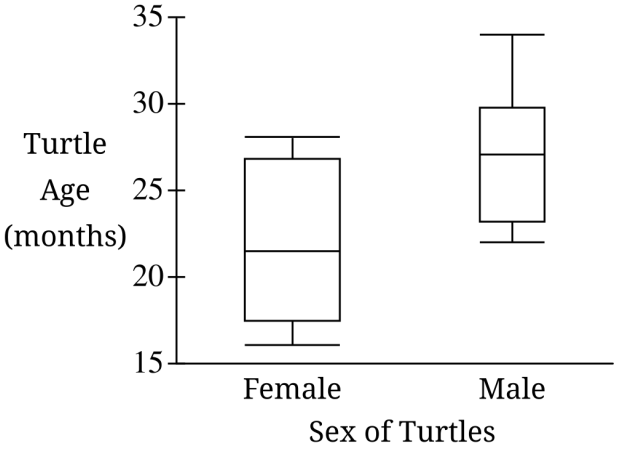
(c) Consider the strength of the relationship between sex and length.

i. Calculate the point-biserial correlation, r_{pb}, between length and sex of these 20 turtles.

ii. The zoo also has a population of 40 tortoises, of which 16 are male and 24 are female. The lengths of male and female tortoises tend to be more similar than the lengths of male and female turtles, but the overall variability in the lengths of tortoises is similar to that of the turtles. Based on the point-biserial correlations, would the strength of the relationship between length and sex for the tortoises be weaker or stronger than that of the turtles? Justify your answer.

Consider the original scatterplot of length versus age with the points identified by the sex of the turtle and summaries of the relationship between sex and age of the 20 turtles.





	Male	**Female**
Mean age (months)	27	22
Number of turtles	8	12
Standard deviation of age of the 20 turtles (months)	4.86	
Point-biserial correlation between sex and age	0.50	

(d) Is age a confounding variable in the relationship between length and sex of the 20 turtles? Justify your answer based on the analysis conducted and the information provided.

END OF EXAM