Begin your response to **QUESTION 1** on this page.

STATISTICS

SECTION II

Total Time—1 hour and 30 minutes

6 Questions

SECTION II, Part A

Suggested Time—1 hour and 5 minutes

5 Questions

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 exercise center has several thousand members from age 18 to 55 years and several thousand members age 56 and older. The manager of the center is considering offering online fitness classes. The manager is investigating whether members' opinions of taking online fitness classes differ by age. The manager selected a random sample of 170 exercise center members ages 18 to 55 years and a second random sample of 230 exercise center members ages 56 years and older. Each sampled member was asked whether they would be interested in taking online fitness classes.

   The manager found that 51 of the 170 sampled members ages 18 to 55 years and that 79 of the 230 sampled members ages 56 years and older said they would be interested in taking online fitness classes.

   At a significance level of $\alpha = 0.05$, do the data provide convincing statistical evidence of a difference in the proportion of all exercise center members ages 18 to 55 years who would be interested in taking online fitness classes and the proportion of all exercise center members ages 56 years and older who would be interested in taking online fitness classes? Complete the appropriate inference procedure to justify your response.

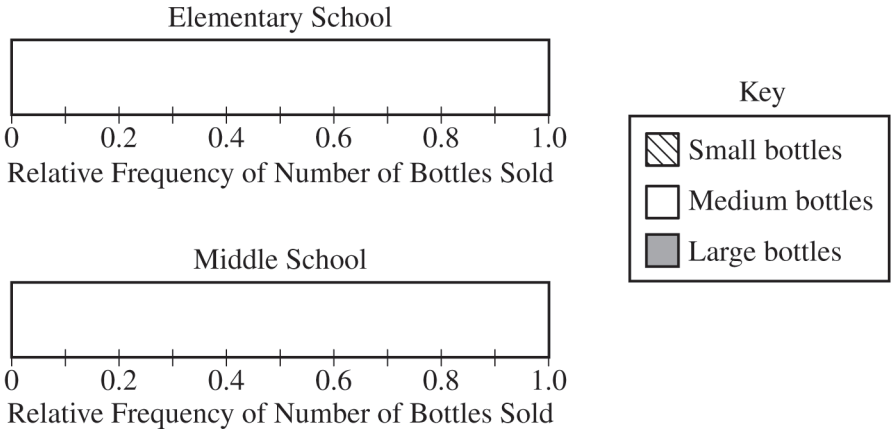
GO ON TO THE NEXT PAGE.

Begin your response to **QUESTION 2** on this page.

2. A local elementary school decided to sell bottles printed with the school district's logo as a fund-raiser. The students in the elementary school were asked to sell bottles in three different sizes (small, medium, and large). The relative frequencies of the number of bottles sold for each size by the elementary school were 0.5 for small bottles, 0.3 for medium bottles, and 0.2 for large bottles.

   A local middle school also decided to sell bottles as a fund-raiser, using the same three sizes (small, medium, and large). The middle school students sold three times the number of bottles that the elementary school students sold. For the middle school students, the proportion of bottles sold was equal for all three sizes.

   (a) Complete the segmented bar graphs representing the relative frequencies of the number of bottles sold for each size by students at each school.



   (b) An administrator at the elementary school concluded that the elementary school students sold more small bottles than the middle school students did. Is the elementary school administrator's conclusion correct? Explain your response.

GO ON TO THE NEXT PAGE.

Question 1: Focus on Inference — 4 points

General Scoring Notes

- This question is scored in three sections. Each section is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The first section includes statements of the null and alternative hypotheses and identification of the appropriate hypothesis test. The second section includes verifying the conditions for the test identified in the first section and calculating the value of the test statistic and the corresponding *p*-value. The third section includes the conclusion for the test identified in the first section. The response is then categorized based on the scores assigned to each section and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each section of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

	Model Solution	Scoring
Section 1	Let p_{younger} represent the proportion of all exercise center members from 18 to 55 years of age who would be interested in taking online fitness classes, and p_{older} represent the proportion of all exercise center members 56 years or older who would be interested in taking online fitness classes. The null hypothesis is $H_0: p_{\text{younger}} = p_{\text{older}}$ and the alternative hypothesis is $H_a: p_{\text{younger}} \neq p_{\text{older}}$. An appropriate inference procedure is a two-sample *z*-test for a difference of population proportions.	**Essentially correct (E)** if the response satisfies the following three components: 1. Identifies a two-sample *z*-test for a difference of population proportions by name (e.g., "two-proportion *z*-test" or "two-sample *z*-test") or by formula 2. States a correct null hypothesis of equal proportions *AND* a correct two-sided alternative hypothesis of unequal proportions 3. Provides sufficient context by referencing the two groups (18 to 55 years of age and 56 years or older) *AND* the populations (all exercise center members) **Partially correct (P)** if the response satisfies two of the three components required for E. **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:

- If the response identifies the correct test by name but also states an incorrect formula, then component 1 is not satisfied.
- If the response identifies the test using the correct formula but equates it with a *t* instead of a *z*, then component 1 is not satisfied.
- Neither the names of the groups nor the concept of *population* is required to satisfy component 2.

- A response that states the hypotheses in words (e.g., the null hypothesis is that the proportions are equal, and the alternative hypothesis is that the proportions are not equal) may satisfy component 2. If the hypotheses also include the group names (younger and older members) and reference to the populations, then both components 2 and 3 may be satisfied.
- Group or population aspects of component 3 may be satisfied anywhere in the response, provided there are no incorrect references to groups or population elsewhere in the response (e.g., using $\hat{p}$ in hypotheses, referring to samples in conclusion, or describing groups as "those who are interested in online exercise classes" and "those who are not interested in online exercise classes").
- If the response clearly refers to *sample* proportions instead of *population* proportions using words or symbols (e.g., $\hat{p}_o$ and $\hat{p}_y$), then component 2 may be satisfied, but component 3 is not satisfied unless the symbols used are defined as *population* proportions.
- A response may satisfy the population aspect of component 3 by
  - Referring to the population by using words such as: "population," "all," or "true" when defining the parameter.
  - Using notation such as p_1, p_2, p_y, p_o, π_y, π_o, pi_y, pi_o, when defining the hypothesis statements.

    Note that subscripts y and o would be sufficient to also satisfy the group aspect of component 3, whereas subscripts of 1 and 2 are not sufficient to satisfy the group aspect of component 3 without further clarification.

Confidence Interval Approach:

- If a two-sample *z*-interval for a difference in population proportions is identified correctly by name (e.g., "two-proportion *z*-interval" or "two-sample *z*-interval") or by formula, then component 1 may be satisfied.
- If the response uses individual one-sample *z*-intervals for the two proportions, which is not a correct approach, then component 1 is not satisfied.
- For a response using a confidence interval approach, component 2 is satisfied if the response clearly indicates that the confidence interval is used to assess the correct two-sided alternative to the null hypothesis of equal proportions.
- If a response uses a two-sample *z*-interval for a population proportion, it must include references to group names and populations to satisfy component 3. For example, component 3 is satisfied if the response indicates that it is a confidence interval for the difference in population proportions of younger and older exercise center members.

Chi-Square Test for Homogeneity Approach:

- If the response uses a chi-square test approach, identifying the procedure name as a "chi-square test for homogeneity," component 1 may be satisfied.
- If the response identifies the procedure as a "chi-square test for independence" or just a "chi-square test," then component 1 is not satisfied.

	Model Solution	Scoring
Section 2	The independent observations condition for performing the two-sample z-test for a difference in population proportions is satisfied because the data were obtained from a random sample of 170 exercise center members ages 18 to 55 years and a second random sample of 230 exercise center members ages 56 years and older. The 10% condition must be met by both samples because sampling of exercise center members is done without replacement. There are more than $10(170) = 1{,}700$ adults from 18 to 55 years of age who are members of the exercise center and more than $10(230) = 2{,}300$ adults ages 56 years and older who are members of the exercise center. The value of the sample proportions are $\hat{p}_{\text{younger}} = \frac{51}{170} = 0.3$ and $\hat{p}_{older} = \frac{79}{230} \approx 0.3435.$ The combined proportion is $\hat{p}_{\text{c}} = \frac{170(0.3) + 230(0.3435)}{170 + 230} \approx 0.325.$ The sample size is large enough to support an assumption that the sampling distribution of $\hat{p}_{\text{younger}} - \hat{p}_{\text{older}}$ is approximately normal because $170(0.325) = 55.25$, $(170)(1 - 0.325) = 114.75$, $230(0.325) = 74.75$, and $(230)(1 - 0.325) = 155.25$ are all at least 10. The value of the test statistic is $z = \frac{0.3 - 0.3435}{\sqrt{0.3250(1 - 0.3250)}\sqrt{\frac{1}{170} + \frac{1}{230}}}$ $z \approx -0.918.$ The corresponding p-value is $2 * P(z < -0.918) \approx 0.359.$	**Essentially correct (E)** if the response satisfies the following four components: 1. Checks the independence condition by referring to the two random samples *AND* indicates that there are more than $10(170) = 1{,}700$ exercise center members from 18 to 55 years of age and more than $10(230) = 2{,}300$ exercise center members ages 56 years and older 2. Checks that the sample sizes are large enough by verifying that $170(0.325) = 55.25$, $(170)(1 - 0.325) = 114.75$, $230(0.325) = 74.75$, and $(230)(1 - 0.325) = 155.25$ are all at least 5 (or 10) 3. Correctly reports the value of the z-statistic 4. Correctly reports the p-value, consistent with the reported test statistic and stated alternative hypothesis **Partially correct (P)** if the response satisfies only two or three of the four components required for E. **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:

- In order to satisfy the reference to the random selection of exercise center members in component 1, it is minimally acceptable to state "random samples—check" or "SRSs—check." However, component 1 is not satisfied if the response implies that random *assignment* was used or only states "random—check."

- If the response states that the random sample of members in each age group is less than 10 percent of total membership in that group (with sample sizes provided), component 1 may be satisfied. For example, "$170 < 10\%$ of all members ages 18 to 55, and $230 < 10\%$ of all members ages 56 and older."
- In order to satisfy component 2, the response must include values of the observed successes and failures, or values for the expected successes and failures, or formulas for the expected number of successes and failures with values inserted *AND* the response must make a comparison of the four values with some standard criterion, such as 5 or 10. If expressions such as $170(0.325)$ and $(170)(1 - 0.325)$ are used, simplification is not required.
- Examples of acceptable quantities (comparisons must still be made):
  - $55.25,\ 114.75,\ 74.75,$ and 155.25
  - $(170)(0.325),\ (170)(1 - 0.325),\ 230(0.325),$ and $(230)(1 - 0.325)$
  - 51, 119, 79, and 151 (observed counts)
  - $(170)(0.3),\ (170)(0.7),\ (230)(0.3435),$ and $(230)(0.6565)$
- Unless values of relevant parameters are explicitly identified in the response, the following quantities are unacceptable:
  - $170p_1,\ 170(1 - p_1),\ n_1p_1,\ n_1(1 - p_1),\ 230p_2,\ 230(1 - p_2),\ n_2p_2,\ n_2(1 - p_2)$
  - $170\hat{p}_1,\ 170(1 - \hat{p}_1),\ n_1\hat{p}_1,\ n_1(1 - \hat{p}_1),\ 230\hat{p}_2,\ 230(1 - \hat{p}_2),\ n_2\hat{p}_2,\ n_2(1 - \hat{p}_2)$
- A response that reports the correct value for the *z*-statistic but contains errors in supporting work may still satisfy component 3.
- A response that reports a value for the *z*-statistic using the incorrect estimate of the standard error of the difference $\hat{p}_{younger} - \hat{p}_{older}$ based on individual estimates for each group may still satisfy components 3 and 4, in which case $z = -0.9237$ and p-value $= 0.3557$.
- A response that uses Table A to determine the *p*-value from the rounded $z = -0.92$ should report p-value $= 0.3576$.
- A response that inputs correct values into the *z*-statistic formula but computes an incorrect value for the *z*-statistic, may satisfy component 3.
- If the response compares the value of the test statistic to a critical value instead of computing a *p*-value, then a comparison consistent with the stated alternative hypothesis satisfies component 4.
- If the response omits the hypotheses, the correct two-sided alternative hypothesis is assumed when scoring component 4.
- If an incorrect alternative hypothesis is stated, then the *p*-value must be consistent with the stated alternative hypothesis to satisfy component 4.
- If the response satisfies component 4, any supporting work for the *p*-value may be treated as extraneous.

Confidence Interval Approach:

- If the stated alternative hypothesis is correct or no alternative hypothesis is provided:
  - If a two-sided 95 percent confidence interval for $p_{\text{younger}} - p_{\text{older}}$ is correctly calculated as $(-0.1357,\ 0.0488)$, then component 3 is satisfied.
  - If the two-sided confidence interval is correctly interpreted based on whether zero is in the interval, then component 4 is satisfied.

- If the stated alternative hypothesis is incorrect (one-sided), the confidence interval approach must be consistent with the stated alternative to satisfy components 3 and 4:
  - An interval consistent with the stated alternative will satisfy component 3. A lower one-sided 95 percent confidence interval for $p_{\text{younger}} - p_{\text{older}}$ is $(-1, 0.3395)$ or an upper one-sided 95 percent confidence interval for $p_{\text{younger}} - p_{\text{older}}$ is $(-0.1209, 1)$.
  - If only the upper end of the lower one-sided confidence interval is used to reach a conclusion or only the lower end of the upper one-sided confidence interval is used to reach a conclusion, then component 4 is satisfied.
- A response that provides an interval for $p_{\text{younger}} - p_{\text{older}}$ should be scored the same way as the interval for $p_{\text{older}} - p_{\text{younger}}$.

Chi-Square Test for Homogeneity Approach:

- Component 2 is satisfied if the response verifies that all four expected counts are at least 5 (or 10), so that the test statistic has an approximate chi-square distribution.
- Component 3 is satisfied if the response correctly reports the chi-square statistic (no work is required).