

Question 2: Focus on Exploring Data

4 points

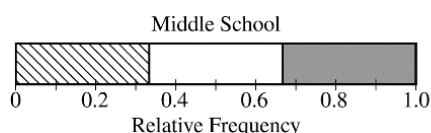
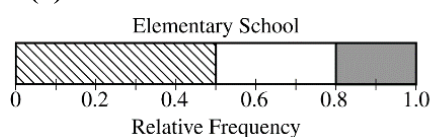
General Scoring Notes

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part 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 part of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

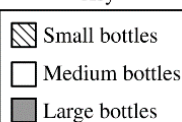
Model Solution

Scoring

(a)



Key



Essentially correct (E) if the response satisfies the following two components:

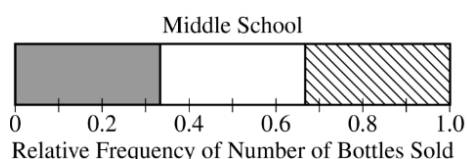
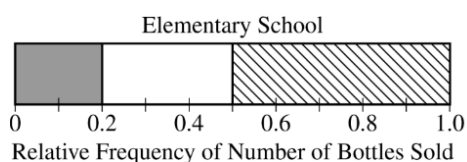
- The elementary school's segmented bar graph is correctly partitioned according to the given proportions
- The middle school's segmented bar graph is correctly partitioned to create three equal areas

Partially correct (P) if the response only satisfies one of the two components required for E.

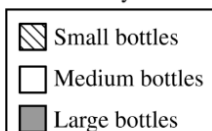
Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes:

- Responses that do not use the key given may still satisfy both components as long as a new key is provided that defines small, medium, and large bottles.
- Responses that do not use the key given and do not define a new key may receive credit for component 2 as long as the bar is vertically partitioned to create three equal areas, but do not satisfy component 1.
- Responses need to be within ± 0.05 of the actual cutoff to be acceptable for demonstrating equal areas.
- Responses do not need to be drawn in the order from smallest bottles to largest bottles as shown in the model solution. For example, the graph could be drawn from largest bottles to smallest bottles.



Key



Model Solution	Scoring
(b) No, the segment for small bottles for the elementary school is wider than the segment for small bottles for the middle school; however, the middle school students sold three times as many bottles as the elementary school students. So, if the elementary school sold x number of bottles, the middle school sold $3x$ number of bottles. For example, if the elementary students sold 100 bottles total, then they sold $0.5(100)$ or 50 small bottles. However, because the middle school students sold three times the total number of bottles as the elementary students, they would have sold 300 bottles total and $(0.\overline{3})(300) = 100$ small bottles. Because $100 > 50$, the middle school sold more small bottles, and the elementary school's administrator is not correct.	Essentially correct (E) if the response satisfies the following three components: 1. States that the elementary school's administrator is incorrect 2. Provides correct mathematical support verifying the middle school sold more small bottles than the elementary school, consistent with the response to component 1 3. Includes context Partially correct (P) if the response satisfies only two of the three components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes:

- A response that shows mathematical support by showing the inequalities $0.5x < 0.\overline{3}(3x)$, $0.5x < x$, $0.5(1) < 0.\overline{3}(3)$, or $(0.5)\left(\frac{1}{4}\right) < (0.\overline{3})\left(\frac{3}{4}\right)$ may satisfy component 2.
- In order to satisfy component 3, a response must include “elementary,” “middle,” and “bottles.”
- Part (b) should be scored consistent with the response to part (a).

Model Solution	Scoring
(c) (i) The mosaic plot shows that the proportion of large bottles sold by High School A was 0.7 and the proportion of large bottles sold by High School B was 0.6. High School A sold a greater proportion of large bottles because $0.7 > 0.6$. (ii) The number of bottles sold is represented by the area of the shaded region. The area of the rectangle representing large bottles sold by High School B is clearly larger than the area of the rectangle representing large bottles sold by High School A. Therefore, High School B sold more large bottles than High School A.	Essentially correct (E) if the response satisfies the following four components: 1. In part (c-i) the response indicates that High School A sold a greater proportion of large bottles 2. In part (c-i) the response bases reasoning on the height of the rectangles or the relative frequencies representing large bottles 3. In part (c-ii) the response indicates that High School B sold a greater number of large bottles 4. In part (c-ii) the response indicates that the area of the rectangle representing large bottles sold by High School B is larger than the area of the rectangle representing large bottles sold by High School A Partially correct (P) if the response satisfies 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:

- If a response refers to “proportions” rather than relative frequencies in part (c-i), the values 0.7 and 0.6 must be provided to satisfy component 2.
- If the response bases reasoning on the heights of rectangles in part (c-i), the response must clearly state that the rectangles representing large bottles are referenced.
- If the response bases reasoning on the areas of rectangles in part (c-ii), the response must clearly state that the areas representing large bottles are referenced.
- A response that provides reasonable estimates of the large bottle areas may satisfy component 4.

Scoring for Question 2	Score
Complete Response Three parts essentially correct	4
Substantial Response Two parts essentially correct and one part partially correct	3
Developing Response Two parts essentially correct and no part partially correct <i>OR</i> One part essentially correct and one or two parts partially correct <i>OR</i> Three parts partially correct	2
Minimal Response One part essentially correct and no part partially correct <i>OR</i> No part essentially correct and two parts partially correct	1

Question 5: Multi-Focus

4 points

General Scoring Notes

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part 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
(a) $P(11+ \text{ months} \cap \text{ majority regular cards})$ $= \frac{71 + 76 + 112}{500}$ $= \frac{259}{500}$ $= 0.518$	Essentially correct (E) if the response satisfies the following two components: - Provides the correct probability - Shows work for the correct probability Partially correct (P) if the response satisfies only one of the two components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes:

- An arithmetic or transcription error in a response can be ignored if correct work is shown.
- A correct fraction, decimal, or percentage may satisfy component 1.
- To satisfy component 2 work must demonstrate how the numerator of 259 was computed. This may be accomplished by showing the addition or clearly indicating the appropriate summands on the table.
- A response that satisfies components 1 and 2 for the probability that the collector has been collecting baseball cards for 11 or more months and has a majority of *rare* baseball cards may be scored P.
- A specific probability statement is not required, but if correctly given should be considered a positive in holistic scoring.

Model Solution	Scoring
(b) $P(\text{majority regular cards} \mid \text{fewer than 6 months})$ $= \frac{P(\text{majority regular cards} \cap \text{fewer than 6 months})}{P(\text{fewer than 6 months})}$ $= \frac{80}{500}$ $= \frac{80}{91}$ $= 0.879$	Essentially correct (E) if the response satisfies the following two components: 1. Provides the correct probability 2. Shows work for the correct probability Partially correct (P) if the response satisfies only one of the two components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes:

- An arithmetic or transcription error in a response can be ignored if correct work is shown.
- A response of $\frac{80}{91}$ satisfies both components 1 and 2.
- A response that satisfies components 1 and 2 for the probability that given a randomly selected collector has been collecting baseball cards for fewer than six months, they have a majority of *rare* baseball cards may be scored P.
- A specific probability statement is not required, but if correctly given should be considered a positive in holistic scoring.

Model Solution	Scoring
(c) (i) Michelle should conduct a chi-square test for independence between months collecting baseball cards and majority card status for all baseball card collectors at the convention. (ii) H_0: There is not an association between months collecting baseball cards and majority card status for all baseball card collectors at the convention. H_a: There is an association between months collecting baseball cards and majority card status for all baseball card collectors at the convention. <i>OR</i> H_0: Months collecting cards and majority card status are independent for all baseball card collectors at the convention. H_a: Months collecting cards and majority card status are not independent for all baseball card collectors at the convention.	Essentially correct (E) if the response satisfies the following three components: 1. In part (c-i) the response identifies a chi-square test for independence by name 2. In part (c-ii) the response states the correct null hypothesis to imply there is not an association (relationship) <i>AND</i> states the correct alternative hypothesis to imply there is an association (relationship) 3. In part (c-ii) the response provides sufficient context for at least one of the hypotheses by including reference to both “months collecting” and “majority card status” Partially correct (P) if the response satisfies only two of the three components required for E <i>OR</i> if the response identifies a “chi-square test for homogeneity” in part (c-i) by name or formula <i>AND</i> components 2 and 3 are consistent with the chi-square test for homogeneity. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes

- A response that states merely “chi-square” or “chi-square test,” without specifying independence, does not satisfy component 1. However, a response that states “chi-square test of association” does satisfy component 1.
- Any discussion of the degrees of freedom for the test should be ignored in scoring.
- If the hypotheses do not explicitly state the population of interest, it should be assumed that the population is all baseball card collectors at the convention. However, if the response clearly references the sample, component 2 is not satisfied.

Model Solution	Scoring
(d) Because the p-value of 0.0075 is less than any reasonable α level such as 0.05 or 0.10, the null hypothesis should be rejected. The data provide convincing statistical evidence that there is an association between months collecting baseball cards and majority card status for all baseball card collectors at the convention.	Essentially correct (E) if the response satisfies the following two components: 1. Provides correct comparison of the p-value to alpha (p-value is less than alpha) <i>AND</i> provides a correct decision about the null and/or alternative hypothesis 2. States a conclusion, in context, consistent with and in terms of the stated alternative hypothesis in part (c) using nondefinitive language Partially correct (P) if the response satisfies only one of the two components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes

- Because no alpha value is explicitly given, a response may satisfy component 1 by stating that the p -value is small *AND* providing a correct decision about the null and/or alternative hypothesis.
- An explicit decision about the null hypothesis is not required to satisfy component 1.
- If an explicit decision is stated and the conclusion is inconsistent with the decision, component 1 is not satisfied.
- The decision part of component 1 may be satisfied by implying the decision within the conclusion statement (sufficient evidence/insufficient evidence for the alternative hypothesis).
- To satisfy the context in component 2 the response must include discussion of both “months spent collecting baseball cards” and “which type of card is the majority in the collection.”
- If the response omits hypotheses in part (c), assume the correct alternative hypothesis is provided when scoring component 2.
- If the response states incorrect hypotheses in part (c), components 1 and 2 may be satisfied by either referring to the stated alternative hypothesis or referring to the inference question.
- Examples of nondefinitive language in component 2 include “evidence to accept the alternative,” “there is evidence for the alternative,” “there is not sufficient evidence for the alternative,” and “Michelle should conclude that her belief is correct.”
- Examples of definitive language in component 2 include “proves the null,” “proves the alternative,” “accepts the alternative,” “there is not evidence for the alternative,” “no evidence for the alternative,” “therefore the alternative is true,” and “there is an association.”
- If components 1 and/or 2 are satisfied and the response provides an incorrect interpretation of the p -value, the score is lowered from E to P or P to I.

Scoring for Question 5

Each essentially correct (E) part counts as 1 point, and each partially correct (P) part counts as $\frac{1}{2}$ point.

Score**Complete Response****4****Substantial Response****3****Developing Response****2****Minimal Response****1**

If a response is between two scores (for example, $2\frac{1}{2}$ points), use a holistic approach to decide whether to score up or down, depending on the strength of the response and quality of the communication.

Question 1: Focus on Exploring Data**4 points****General Scoring Notes**

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part 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 part 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
(a) The scatterplot reveals a strong, positive, roughly linear association between the mass and length of bullfrogs. There are no points that seriously deviate from the straight-line pattern of the points in the plot.	Essentially correct (E) if the response provides a description that includes at least three of components 1-4 and component 5: 1. Direction of association (positive or increasing) 2. Strength of association (strong) 3. Form of association (linear or approximately linear) 4. Unusual features (no points with large discrepancies from the pattern (straight line) exhibited by most of the points on the plot) 5. Context (association between length and mass of bullfrogs) Partially correct (P) if the response satisfies only one or two components out of components 1-4 and component 5 OR if the response satisfies at least three out of components 1-4 but does not satisfy component 5. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes:

- To satisfy component 4, it is sufficient to simply indicate that there are no unusual features.
- To satisfy component 5, it is minimally sufficient for the response to refer to the association or relationship between mass and length without explicitly mentioning bullfrogs.
- The strength of the response in part (a) may be considered if holistic scoring is needed.

Model Solution	Scoring
(b) The value of the slope of the least-squares regression line is 6.086. This value indicates that the predicted mass of a bullfrog increases by 6.086 grams for each additional millimeter of length.	Essentially correct (E) if the response satisfies the following three components: 1. Identifies the value of the slope as 6.086 2. Provides an interpretation that references an increase of a number of grams of mass for each one-millimeter increase in length 3. Indicates that the slope represents a change in a prediction using non-deterministic language such as “predicted,” “estimated,” “expected,” or “average” Partially correct (P) if the response satisfies only two of the three components. Incorrect (I) if the response does not meet the criteria for E or P.

Additional Notes:

- The value of the slope, 6.086, may be rounded to 6.09 or 6.1, but not to 6, to satisfy the numerical requirement in component 1.
- A response that only contains 6.086 in the interpretation satisfies component 1.
- A calculation of slope may satisfy component 1, provided that two points from the line are used in the calculation.
- Units of measurements must be correctly specified for both mass and length to satisfy component 2.
- It is not required to refer specifically to the “least-squares regression line.”

Model Solution	Scoring
(c) The coefficient of determination is $r^2 \approx 0.819$. This value indicates that 81.9% of the variation in bullfrog mass can be explained by variation in bullfrog length as described by the least-squares line.	Essentially correct (E) if the response provides a correct interpretation of r^2 in context. Partially correct (P) if the response provides a generic interpretation (no context) OR if the response provides a reasonable but incorrect interpretation of r^2 in context. Incorrect (I) if the response does not satisfy the criteria for E or P.

Additional Notes:

- Correct interpretations of r^2 include the concept that part of the variation in the response (dependent or y) variable is explained by the linear relationship with the explanatory (independent or x) variable. The response can take any of several equivalent forms, such as:
 - The proportion of the total variability in the dependent (response) variable y that is explained by the independent (explanatory) variable x .
 - The proportion of variation in y that is accounted for by the linear model.
 - The proportionate reduction of the total variation of the y -values that is associated with the use of the independent variable x .
 - The proportionate reduction in the sum of the squares of vertical deviations obtained by using the least-squares line instead of the sample mean to predict values of y .
- Correct interpretation of r^2 must explicitly relate to the dependent variable. Mention of the data, predicted values, or no mention of the dependent variable are incorrect interpretations. Common incorrect interpretations include:
 - The percent (or proportion or part of the total) variability in the predicted y -values that is explained by the linear relationship between y and x .
 - The percent (or proportion or part of the total) variability in the data that is explained by the linear relationship between y and x .
 - The percent (or proportion or part of the total) variability that is explained by the linear relationship between y and x .
 - The percent (or proportion or part of the total) variability in y that is on average explained by the linear relationship between y and x .
- A reasonable but incorrect interpretation of r^2 with context might include the following responses:
 - 81.9% of the variation in mass and length can be accounted for by the least-squares regression line.
 - 81.9% of the variability in predicted mass is accounted for by the length.
- For context, the response variable (y) must be identified as mass, and the explanatory variable (x) must be identified as length.
- An interpretation of the correlation between mass and length, $r = \sqrt{0.819} = 0.905$, is not considered a reasonable interpretation of r^2 .
- The value of the percentage (81.9%) or proportion (0.819) of variation does not need to be specified, but if an incorrect value is specified, the score is lowered by one level, from E to P or from P to I.
- The strength of the response in part (c) may be considered if holistic scoring is needed.

Model Solution	Scoring
(d) (i) The largest residual in absolute value belongs to the bullfrog with length 162 millimeters and mass 356 grams. (ii) The least-squares regression line overestimates the mass of the bullfrog with length 162 millimeters. Plot 2 shows that the point for the bullfrog with length 162 millimeters is below the least-squares regression line.	Essentially correct (E) if the response satisfies the following two components: 1. The response to part (d-i) identifies the correct bullfrog (length between 160 and 165 millimeters, mass between 350 and 375 grams) 2. The response to part (d-ii) explicitly indicates whether the linear model overestimates or underestimates mass for the bullfrog identified in part (d-i) and provides a correct justification based on a comparison of the identified observation to the least-squares regression line Partially correct (P) if the response satisfies only one of the two components. Incorrect (I) if the response does not satisfy the criteria for E or P.

Additional Notes:

- The comparison of the observation to the regression line in the response to part (d-ii) is satisfied if the response does one of the following:
 - Correctly indicates if the observation is below (above) the least-squares regression line in Plot 2.
 - Notes that observed mass is smaller (larger) than the mass predicted by the least-squares regression line.
 - Marks the observation selected in part (d-i) on Plot 2 with an indication of the vertical distance from the least-squares regression line.
 - Notes the correct sign of the residual.
- Numerical values are not required in the response to part (d-ii). If a numerical value is given for the predicted mass, however, it must be reasonable. A numerical value for the predicted mass could be computed with the formula given in the stem, e.g., $-546 + (6.086)(162) = 439.9$ grams, for a bullfrog of length 162 millimeters, or a value can be read from the line shown in Plot 2. Any value between 425 and 450 should be considered a reasonable value. Showing work is not required.
- The word overestimate with the calculated predicted value of mass is enough to satisfy component 2.
- If the wrong observation is identified in part (d-i), the response to part (d) may be scored P if the response to part (d-ii) correctly compares that observation to the least-squares regression line and states the correct conclusion about overestimating or underestimating mass with justification.
- It is not required to refer specifically to the “least-squares regression line.”

Scoring for Question 1

Each essentially correct (E) part counts as 1 point, and each partially correct (P) part counts as $\frac{1}{2}$ point.

Score**Complete Response****4****Substantial Response****3****Developing Response****2****Minimal Response****1**

If a response is between two scores (for example, 2 $\frac{1}{2}$ points), use a holistic approach to decide whether to score up or down, depending on the strength of the response and quality of the communication.