Chief Reader Report on Student Responses:
2023 AP® Statistics Free-Response Questions

• Number of Students Scored	242,929		
• Number of Readers	1,385		
• Score Distribution	Exam Score	N	%At
	5	36,661	15.09
	4	53,909	22.19
	3	55,196	22.72
	2	39,248	16.16
	1	57,915	23.84
• Global Mean	2.89		

The following comments on the 2023 free-response questions for AP® Statistics were written by the Chief Reader, Dr. Barb Barnet, Ph.D. They give an overview of each free-response question and of how students performed on the question, including typical student errors. General comments regarding the skills and content that students frequently have the most problems with are included. Some suggestions for improving student preparation in these areas are also provided. Teachers are encouraged to attend a College Board workshop to learn strategies for improving student performance in specific areas.

Question 1

Task: Exploring Data
Max Points: 4
Mean Score: 1.84

What were the responses to the question expected to demonstrate?

The primary goals of this question were to assess a student's ability to (1) use data presented on a histogram to describe a variable within the context of a study; (2) create a box plot when provided summary statistics for a variable; and (3) determine which of two data sets match a description, when provided graphs and characteristics, and then provide a justification based on a comparison of the characteristics.

This question primarily assesses skills in skill category 2: Data Analysis. Skills required for responding to this question include (2.A) Describe data presented numerically or graphically, (2.B) Construct numerical or graphical representations of distributions, and (2.D) Compare distributions or relative positions of points within a distribution.

This question covers content from Unit 1: Exploring One-Variable Data of the course framework in the AP Statistics Course and Exam Description. Refer to topics, 1.6, 1.8, and 1.9, and learning objectives UNC-1.H, UNC-1.L, UNC-1.M, and UNC-1.N.

How well did the responses address the course content related to this question? How well did the responses integrate the skills required on this question?

- In part (a) most responses provided context by referencing "dissolved oxygen" and referenced at least two characteristics of the distribution. The majority of responses discussed shape, center, and spread but some did not include a discussion of unusual features such as potential outliers or the gap that was present on the graph.
- In part (b) the majority of responses were able to correctly construct a boxplot using the five relevant values listed in the table. Some responses attempted to plot outliers on the boxplot and did not extend the lines to the minimum or maximum values, missing two of the five components.
- In part (c) most responses provided a correct answer to the question asked and most provided a justification based on the characteristics of the two distributions. However, some provided incomplete comparisons of the distributions, particularly when comparing the centers or spreads. Most were able to correctly identify that the distribution of dissolved oxygen concentrations in cold streams was skewed left and that the distribution of dissolved oxygen concentrations in warmer streams was skewed right.

What common student misconceptions or gaps in knowledge were seen in the responses to this question?

Common Misconceptions/Knowledge Gaps	*Responses that Demonstrate Understanding*
• In part (a) failing to identify all characteristics of the distribution.	• The distribution of dissolved oxygen in streams colder than 8°C is unimodal and skewed left with a median between 11 mg/l and 12 mg/l, an IQR of approximately 2 mg/l, and a potential outlier at 2 mg/l – 3 mg/l.

• In part (a) using mode as a measure of center.	• The median dissolved oxygen in streams colder than 8°C is between 11 mg/l and 12 mg/l.
• In part (a) calculating a value such as the mean using grouped data and then using definitive language such as "the mean is 11.5 mg/l."	• The mean is approximately 11.5 mg/l.
• In part (a) incorrectly using the term range to describe an interval and not a single value.	• The range of dissolved oxygen in streams colder than 8°C is at most 14 mg/l − 2 mg/l = 12 mg/l.
• In part (b) failing to extend the lines to the minimum and maximum in the boxplot when the instructions were clear that outliers were not to be included.	• Provides a graph that extends the lines to the minimum and maximum values listed in the table.
• In part (c) failing to provide an answer as to which streams are healthier.	• Streams with water temperature colder than 8°C are generally healthier for wildlife.
• In part (c) failing to directly compare the centers of the distributions or not providing the values for an appropriate measure of center.	• The center (median) of the distribution of dissolved oxygen in colder streams is 11 to 12 mg/l which is higher than the center (median) of 5.43 mg/l for warmer streams.
• In part (c) not comparing the shapes of the two distributions.	• The distribution of dissolved oxygen in colder streams is left skewed, and the distribution of dissolved oxygen in warmer streams is right skewed.

Based on your experience at the AP® Reading with student responses, what advice would you offer teachers to help them improve the student performance on the exam?

- Emphasize the importance of discussing all characteristics of a distribution, such as shape, center, spread and unusual features, and including corresponding values.
- Remind students that definitive language when discussing grouped data is not always appropriate. They should use approximate language, such as the mean of the distribution is approximately 11.5 mg/l.
- Have students practice drawing boxplots using summary tables of data.
- Remind students that they should answer the question being asked and then provide a justification for their answer.
- Reinforce with students that when they are comparing distributions, they should directly compare and include appropriate values of the statistics being compared. Simply listing the characteristics of the two distributions is not a direct comparison.
- Additionally, when describing a distribution or comparing the characteristics of distributions, responses should be in the context of the study that provided the data and include appropriate units.

What resources would you recommend to teachers to better prepare their students for the content and skill(s) required on this question?

- The *AP Statistics Course and Exam Description* (CED), effective Fall 2020, includes instructional resources for AP Statistics teachers to develop students' broader skills. Please see page 227 of the CED for examples of key questions and instructional strategies designed to develop skill 2.A, describe data presented numerically or graphically; skill 2.B, construct numerical or graphical representations of distributions; and skill 2.D (page 228), compare distributions or relative positions of points within a distribution.

- A table of representative instructional strategies, including definitions and explanations of each, is included on pages 213-223 of the CED. The strategy “Create Representations,” for example, may allow students to have a deeper understanding of the differences in quantitative graphs.
- AP Classroom provides four videos focused on the content and skills to answer this question.
    - The daily video 1 for topic 1.6 discusses how to properly describe the characteristics of quantitative data distributions [shape, center, variability (spread), as well as any unusual features such as outliers, gaps, clusters, or multiple peaks, and context] of a histogram (see UNC-1.H.1). The key takeaway from this video that is relevant to this question is the best ways to discuss the important characteristics when describing a distribution of quantitative data (part (a)).
    - The daily video 1 for topic 1.8 explores how to use the five-number summary to create a boxplot (see UNC-1.L.2). A key takeaway from this video discusses how to create a boxplot using the five-number summary, which was the task in part (b) of the question.
    - The daily video 1 for topic 1.9 demonstrates how to compare distributions of quantitative data (see UNC-1.N.1). A key takeaway from this video explores what parts are necessary to have a complete response in part (c) for comparing the distributions of quantitative data.
- AP Classroom also provides topic questions for the formative assessment of topics 1.6, 1.8, and 1.9, as well as access to the question bank, which is a searchable database of past AP Questions on this topic.
- The Online Teacher Community features many resources shared by other AP Statistics teachers. For example, to locate resources to give your students practice discussing distributions of histograms, try entering the keywords “histograms” in the search bar, then selecting the drop-down menu for “Resource Library.” When you filter for “Classroom-Ready Materials,” you may find worksheets, data sets, practice questions, and guided notes, among other resources.

Question 2

Task: Collecting Data
Max Points: 4
Mean Score: 1.45

What were the responses to the question expected to demonstrate?

The primary goals of the question were to assess a student's ability to (1) identify the experimental units, treatments, and response variable from a description of a completely randomized experimental design; (2) describe a correct procedure that could be used to randomly assign two treatments to the experimental units in a completely randomized design; and (3) describe the benefit of using random assignment in an experimental design.

This question primarily assesses skills in skill category 1: Selecting Statistical Methods and skill category 4: Statistical Argumentation. Skills required for responding to this question include (1.B) Identify key and relevant information to answer a question or solve a problem, (1.C) Describe an appropriate method for gathering and representing data, and (4.B) Interpret statistical calculations and findings to assign meaning or assess a claim.

This question covers content from Unit 3: Collecting Data of the course framework in the AP Statistics Course and Exam Description. Refer to topics 3.5 and 3.7 and learning objectives VAR-3.A, VAR-3.B, and VAR-3.E.

How well did the responses address the course content related to this question? How well did the responses integrate the skills required on this question?

- In part (a) most responses correctly identified the experimental units and the response variable. Many responses correctly identified the treatments as having two groups, concrete with fibers and concrete without fibers. However, some responses were imprecise, identifying the experimental units as houses or the treatments as using concrete with fibers or not. Some responses incorrectly identified the response variable as a binary outcome, such as whether or not the driveway cracked, or identified the response variable as the number (or amount) of cracks rather than the severity. Some students confused the response variable with the research question.
- In part (b) most responses described a random process appropriate for a completely randomized design. However, many responses left out components of implementing the random process so that either the process was not random or the random process did not result in two equal-sized groups.
- In part (c) most responses linked random assignment to a reduction of the impact of confounding variables and described the impact of confounding but did not link the confounding to a cause-and-effect conclusion. A few linked the random assignment of driveways to causation. Some of these incorrectly included strong deterministic language, such as stating the random assignment "proves the fibers cause less cracking."

What common student misconceptions or gaps in knowledge were seen in the responses to this question?

Common Misconceptions/Knowledge Gaps	*Responses that Demonstrate Understanding*
• In part (a) some responses did not precisely use language. For example, when identifying the treatments, "driveways that are paved with concrete with fibers and driveways that are not."	• The treatments were concrete with fibers and concrete without fibers.
• In part (a) when identifying the treatments, some responses only identified the group receiving the	• The treatments were concrete with fibers and concrete without fibers.

concrete with fibers, confusing the treatment variable with the treatment factor level.	
• In part (a) some responses confused the response variable (severity of cracking) with the research question by using statements such as “does concrete with fiber reduce the severity of cracking on a scale of 0 to 10?”	• The response variable is the severity of cracks on a scale of 0 to 10 after one year.
• In part (b) some responses did not sufficiently describe how to implement the random assignment of driveways to treatment groups. For example, some responses that used a hat approach did not mix the slips of paper or did not indicate that sampling was done without replacement. Some responses did not provide sufficient detail, writing statements such as “randomly assign 30 driveways to receive concrete with fibers.”	• Label each driveway uniquely from 1 to 60. Write the numbers 1 to 60 on equal-sized slips of paper and put them in a hat. Mix the slips. Randomly draw 30 unique slips of paper without replacement; the driveways corresponding to these numbers will receive concrete with fibers. The remaining 30 driveways will receive concrete without fibers.
• In part (b) when using a random number generator, some responses did not indicate the range of numbers to enter into the random number generator. For example, some responses state that the first 30 unique numbers from a random generator will be assigned to concrete with fibers.	• Label each driveway uniquely from 1 to 60. Using a random number generator, generate 30 unique integers from 1 to 60. Assign the driveways with these 30 numbers to receive concrete with fibers and the other 30 driveways to receive concrete without fibers.
• In part (b) some responses did not create random groups, stating, for example, “label the driveways from 1 to 60 and put the even numbered driveways in one group and the odd numbered driveways in another group. Then toss a coin to decide which group receives concrete with fibers and which receives concrete without fibers.” In this example, the driveways are not randomly assigned to the groups.	• Label each driveway uniquely from 1 to 60. Write the numbers 1 to 60 on equal-sized slips of paper and put them in a hat. Mix the slips. Randomly draw 30 unique slips of paper without replacement; the driveways corresponding to these numbers will receive concrete with fibers. The remaining 30 driveways will receive concrete without fibers.
• In part (c) some responses confused the benefit of random assignment with random selection by concluding that the benefit was the results can be generalized to the population.	• The benefit of random assignment is that the developer has evidence to conclude that adding fibers to concrete in paving driveways does cause a reduction in cracking.
• In part (c) some responses were using deterministic language linking random assignment to causation by using words such as “ensures the fibers caused less cracking,” “we conclude the fibers caused less cracking,” “determines the fibers caused less cracking,” or “proves the fibers caused less cracking.”	• Due to the random assignment, there is evidence the type of concrete caused the severity of cracks.

Based on your experience at the AP® Reading with student responses, what advice would you offer teachers to help them improve the student performance on the exam?

- Encourage students to provide answers in context to provide clear communication.
- Have students practice describing a random assignment. Have peers read the response to see if they could carry out the written process step-by-step and obtain the desired random assignment.
- When students are using statistical words such as "confounding" or "bias," it is important that they use the word correctly. A stronger response will explain what is meant by the statistical terminology in context of the problem.
- Discuss the difference between random assignment and random selection. With random assignment, a cause-and-effect statement can be made and with random selection from a population, the results can be generalized to that population.
- Discuss different statistical designs, such as the completely randomized design, the randomized block design, and the matched-pairs design. Remind the students to closely read the stem of the problem to see if a type of design is specified.

What resources would you recommend to teachers to better prepare their students for the content and skill(s) required on this question?

- The *AP Statistics Course and Exam Description* (CED), effective Fall 2020, includes instructional resources for AP Statistics teachers to develop students' broader skills. Please see page 225 of the CED for examples of key questions and instructional strategies designed to develop skill 1.B, identify key and relevant information to answer a question or solve a problem, and skill 1.C, describe an appropriate method for gathering and representing data as well as skill 4.B (page 232), interpret statistical calculations and findings to assign meaning or assess a claim.
- A table of representative instructional strategies, including definitions and explanations of each, is included on pages 213–223 of the CED. The strategy "Team Challenge," for example, may be helpful in developing students' abilities to identify the treatments, experimental units, and response variable for an experiment.
- AP Classroom provides three videos focused on the content and skills to answer this question.
  - The daily video 1 for topic 3.5 discusses the basic components of an experiment (see VAR-3.A.1, VAR-3.A.2, and VAR-3.A.3) for part (a).
  - The daily video 2 for topic 3.5 describes the randomization process (see VAR-3.B.1). Key takeaways of this video were especially relevant to this question: "A well-designed experiment should include comparisons between at least two groups, random assignment of treatments to experimental units, replication of treatments to multiple experimental units, and control of possible confounding factors." This video can help students with part (b).
  - The daily video 1 for topic 3.7 discusses how to make decisions based on well-designed experiments (see VAR-3.E.2). There are several key takeaways of this video that are especially relevant to part (c). The first takeaway is that random assignment allows us to conclude that very large observed changes are not merely by chance (statistically significant). Another takeaway of the video is that statistically significant differences between and among experimental treatment groups serve as evidence that the treatments caused the effect. The final takeaway is that random selection of experimental units allows for results to be generalized to the population of interest.
- AP Classroom also provides topic questions for formative assessment of topics 3.5 and 3.7, as well as access to the question bank, which is a searchable database of past AP Questions on this topic.
- The Online Teacher Community features many resources shared by other AP Statistics teachers. For example, to locate resources to give your students practice discussing causation, try entering the keyword "experiments" in the search bar, then selecting the drop-down menu for "Resource Library." When you filter for "Classroom-Ready Materials," you may find worksheets, data sets, practice questions, and guided notes, among other resources.

Question 3

Task: Probability and Sampling Distributions
Max Points: 4
Mean Score: 1.86

What were the responses to the question expected to demonstrate?

The primary goals of the question were to assess a student's ability to (1) calculate single and compound discrete proportions when provided the probability distribution; (2) calculate a conditional probability for a discrete distribution; (3) calculate and interpret the expected value for a discrete distribution; and (4) convert an expected value to a different unit of measurement.

This question primarily assesses skills in skill category 3: Using Probability and Simulation. Skills required for responding to this question include (3.A) Determine relative frequencies, proportion, or probabilities using simulation or calculations, (3.B) Determine parameters for probability distributions, and (3.C) Describe probability distributions.

This question covers content from Unit 4: Probability, Random Variables, and Probability Distributions of the course framework in the AP Statistics Course and Exam Description. Refer to topics 4.3, 4.5. 4.8, and 4.9 with the following learning objectives VAR-4.A, VAR-4.D, VAR-5.C, and VAR-5.F.

How well did the responses address the course content related to this question? How well did the responses integrate the skills required on this question?

- In part (a) most responses correctly calculated the value of the two proportions and included supporting work for both. In part (a-ii) some responses misinterpreted "at least $10" as "at most $10" therefore calculating the incorrect proportion.
- In part (b) most responses were not able to correctly calculate the value for the conditional probability. While many responses included the formula for conditional probability, $P(A \mid B) = \dfrac{P\ (A\ and\ B)}{P(B)}$, there was a common misapplication of the formula. Many students did not recognize that the quantity in the numerator could be found directly from the probability distribution table and instead incorrectly multiplied the probabilities of events A and B together.
- In part (c) most responses correctly determined the expected value with supporting work, but many either did not provide an interpretation of the expected value of the distribution or provided an incomplete interpretation. Of responses with interpretations, many did not identify the expected value as the average of the distribution, and most did not note that this value is obtained from repeating a selection process over a long period of time.
- In part (d) most responses correctly converted the expected value of the distribution from dollars to euros with supporting work. The most common error was to not include units in the final answer either in words or as a symbol.

What common student misconceptions or gaps in knowledge were seen in the responses to this question?

Common Misconceptions/Knowledge Gaps	*Responses that Demonstrate Understanding*
• In part (a-ii) some responses calculated the proportion of bath fizzies that contain "at most" $10 instead of the requested "at least" $10, $0.68 + 0.20 + 0.05 = 0.93$.	• $P(X \geq 10) = 0.05 + 0.05 + 0.01 + 0.01 = 0.12$

• A common issue was the use of incorrect or missing probability notation. While this notation was not required, it does provide clear communication of the findings.	• In part (a) $P(X = \$1) = 1 - 0.32 = 0.68$ and $P(X \geq \$10) = 0.05 + 0.05 + 0.01 + 0.01 = 0.12$.
• In part (b) many responses demonstrated a misunderstanding of the conditional probability formula and the meaning of the values provided in the probability distribution table. Responses multiplied the probability of "at least \$10" and the probability of exactly \$100 together to form the numerator (incorrectly assuming independence) and divided this product by the probability of "at least \$10": $P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$ $= \frac{0.12(0.01)}{0.12} = 0.01.$	• $P(X = \$100 \mid X \geq \$10) = \frac{0.01}{0.12} = 0.0833$.
• In part (b) some responses unnecessarily converted the probabilities to whole numbers to find the conditional probability, e.g., $0.12(100) = 12$ and $0.01(100) = 1$, so $\frac{1}{12} = 0.0833.$	• $\frac{0.01}{0.12} = 0.0833$
• In part (c) a few responses rounded the value of the expected value to a whole number. This demonstrated that they did not recognize that this value was an average. "The expected value is \$4.68 rounded to \$5."	• $E(X) = 1(0.68) + 5(0.2) + 10(0.05)$ $+20(0.05) + 50(0.01) + 100(0.01)$ $= \$4.68$
• In part (c) many responses did not include an interpretation of the expected value of the distribution that included both the concepts of repeating the selection over a long period of time and the concept that this value is the average of the random variable. Some responses incorrectly identified the random variable as the bath fizzy. "The expected value of the average fizzy equals 4.68."	• The expected value is the average cash prize that results from the long run of many, many trials of randomly selecting bath fizzies and determining the amount of cash in each.
• In part (d) a few responses included the correct expected value but did not include units, e.g., $4.68(0.89) = 4.1652$.	• $4.68(0.89) = 4.17$ euros

• In part (d) many responses did not recognize they could simply multiply the expected value from part (c) by 0.89 to find the value in euros. Instead, the response converted all dollar values in the table to euros and multiplied each by the associated probability. While this process is not incorrect, it did result in some arithmetic errors: $0.89(0.68) + 4.45(0.20) + \ldots + 89(0.01) = 4.1652$ euros.	• $4.68(0.89) = 4.17$ euros

Based on your experience at the AP® Reading with student responses, what advice would you offer teachers to help them improve the student performance on the exam?

- Throughout the course, have students identify random variables in the context of the problem or data set.
- Have students reflect on an answer to determine if the value is valid, such as noticing that obtaining a probability value greater than one is not a valid answer.
- Have students practice writing probability notation to ensure clear communication of findings.
- Write out the specific conditional probability statements in words to help students understand the difference between finding a probability given another event has or has not occurred.
- Have students interpret responses to numerical values, not only the results of statistical tests, and explain what these values represent in the context of the problem or data set.
- Have students look at a plot of the probability distribution to understand that the expected value is the center of the distribution so it does not have to be an integer value.
- Have students conduct a probability experiment, such as flipping a coin many times, to show that the expected value is the result of a long run process.

What resources would you recommend to teachers to better prepare their students for the content and skill(s) required on this question?

- The *AP Statistics Course and Exam Description* (CED), effective Fall 2020, includes instructional resources for AP Statistics teachers to develop students' broader skills. Please see pages 229–230 of the CED for examples of key questions and instructional strategies designed to develop skill 3.A, determine relative frequencies, proportions, or probabilities using simulations or calculations; skill 3.B, determine parameters for probability distributions; and skill 3.C, describe probability distributions.
- A table of representative instructional strategies, including definitions and explanations of each, is included on pages 213–223 of the CED. The strategy "Odd One Out" for example, may be helpful in developing students' abilities to identify mutually exclusive events and conditional probabilities.
- AP Classroom provides four videos focused on the content and skills to answer this question.
    - The daily video 1 for topic 4.3 introduces probability rules for events and their complements (See VAR-4.A.2 and VAR-4.A.4). The key takeaway is how students can use complements to answer questions like part (a-ii).
    - The daily video for topic 4.5 discusses how to calculate and interpret conditional probabilities (see VAR-4.D.1). The key takeaways from this video describe different ways to calculate conditional probabilities (part (b)).
    - The daily video for topic 4.8 demonstrates calculating the mean and standard deviation of discrete random variables (see VAR-5.C.2). The key takeaway shows students how to calculate and interpret the mean (expected value) for a discrete random variable like in part (c).
    - The daily video 1 for topic 4.9 discusses the effects of a linear transformation on the mean and standard deviation of a random variable (see VAR-5.F.1). The key takeaway explains to students how to perform a linear transformation for the expected value.

- AP Classroom also provides topic questions for formative assessment of topics 4.3, 4.5, 4.8, and 4.9, as well as access to the question bank, which is a searchable database of past AP Questions on this topic.
- The Online Teacher Community features many resources shared by other AP Statistics teachers. For example, to locate resources to give your students practice discussing causation, try entering the keywords “random variables” in the search bar, then selecting the drop-down menu for “Resource Library.” When you filter for “Classroom-Ready Materials,” you may find worksheets, data sets, practice questions, and guided notes, among other resources.

Question 4

Task: Inference
Max Points: 4
Mean Score: 0.88

What were the responses to the question expected to demonstrate?

The primary goals of the question were to assess a student's ability to (1) identify an appropriate procedure for conducting a hypothesis test for paired data; (2) identify the correct hypotheses for conducting a paired *t*-test of a mean difference; (3) check the conditions for the hypothesis test for paired data; (4) calculate the test statistic and *p*-value for a paired *t*-test of a mean difference; (5) compare the *p*-value to a significance level to make a decision regarding the hypotheses; and (6) determine an appropriate conclusion for a hypothesis test for paired data.

This question primarily assesses skills in skill category 1: Selecting Statistical Methods, skill category 3: Using Probability and Simulation, and skill category 4: Statistical Argumentation. Skills required for responding to this question include (1.E) Identify an appropriate inference method for significance tests, (1.F) Identify null and alternative hypotheses, (3.E) Calculate a test statistic and find a *p*-value, provided conditions for inference are met, (4.C) Verify that inference procedures apply in a given situation, and (4.E) Justify a claim using a decision based on significance tests.

This question covers content from Unit 7: Inference for Quantitative Data: Means of the course framework in the AP Statistics Course and Exam Description. Refer to topics 7.4 and 7.5 and learning objectives VAR-7.B, VAR-7.C, VAR-7.D, VAR-7.E, and DAT-3.F.

How well did the responses address the course content related to this question? How well did the responses integrate the skills required on this question?

- A substantial number of responses failed to discern the difference between the "mean difference" and the "difference of means." This led to many responses incorrectly naming the two-sample *t*-test rather than correctly naming the one-sample t-test for a population mean difference which should be used for a matched-pairs design.
- Most responses successfully stated the correct equality in the null hypothesis and the correct inequality in the alternative hypothesis.
- A substantial number of responses provided the formula for a test statistic different from the test that was named.
- Most responses recognized that conditions must be checked before conducting a hypothesis test; however, a substantial number of responses failed to properly check those conditions.
- Most responses successfully checked the independence condition; however, some responses simply stated "random stated" without indicating the random assignment of treatments.
- Many responses had difficulty utilizing information from the boxplot(s) to address the normality of the sampling distribution condition.
- Most responses reported values of the test statistic and *p*-value consistent with the named test, however some responses omitted the test statistic value and only reported the *p*-value.
- Most responses did a very nice job of providing a comparison of the *p*-value to alpha and stating a correct decision.
- Most responses stated a correct conclusion in the context of the study with many using language from the stem of the problem.
- Although not required, some responses included an interpretation of the *p*-value in the conclusion; however, most attempts were incomplete.
- The most common error when writing conclusions was omitting one or more parts of the context.

What common student misconceptions or gaps in knowledge were seen in the responses to this question?

Common Misconceptions/Knowledge Gaps	*Responses that Demonstrate Understanding*
• Not understanding the difference between the "mean difference" and the "difference of means."	• The inference method to use is the one-sample t-test for a population mean difference which should be used for a matched-pairs design.
• Incorrectly stating the hypotheses for a mean difference between values in matched pairs.	• $H_0: \mu_d = 0$ $H_a: \mu_d > 0$
• For the independence condition: not stating that the treatments were randomly assigned or listing this condition without verifying it using information provided in the description of the study, i.e., only stating "random-check."	• The week in which the patient received the treatment was randomly assigned; therefore, the independence condition is met.
• For the condition of normality of the sampling distribution: incorrectly using the boxplot to verify the condition or not using the boxplot at all.	• The boxplot of the sample differences shows an approximately symmetric distribution.
• Not comparing the p-value to α, the level of significance when stating the decision of the hypothesis test., e.g., "because the $p\text{-value} = 0.0028$ and $\alpha = 0.05$, " or "because the p-value is small."	• Because the $p\text{-value} \approx 0.0028$ is less than the significance level, $\alpha = 0.05$, reject the null hypothesis.
• Not stating the conclusion in context, consistent with and in terms of the alternative hypothesis using non-deterministic language.	• The data provide convincing statistical evidence that the omega-3 supplement will decrease the mean irritability score of all patients with the medical condition similar to the volunteers who participated in the study.

Based on your experience at the AP® Reading with student responses, what advice would you offer teachers to help them improve the student performance on the exam?

- Provide opportunities for students to practice writing skills from the beginning of the course.
  - TIP: Assign previously released AP problems as practice.
  - TIP: Teach students organizational strategies (e.g., state/plan/do/conclude).
- Emphasize the importance of checking conditions before conducting inference procedures.
- Encourage students to name inference procedures in words instead of by formula or by using a calculator name.
  - TIP: Have students develop examples of mean differences versus difference of means and explain to others why they are so named and what test would be appropriate.
- Assess students' ability to quickly decide on appropriate inference procedures.
  - TIP: Encourage students to define the population parameter of interest in context.

- TIP: Give students practice with inference questions based on designed experiments, e.g., matched pairs or completely randomized designs with two treatments, or observational studies with two groups, in which they are only asked to identify which design was used and what inference procedure is appropriate.
- Provide opportunities for students to describe various types of graphical displays of data and review which characteristics can be determined from the different types of graphical displays.
- Assess students' knowledge of standard statistical notation.
    - TIP: Assign a notation quiz.

What resources would you recommend to teachers to better prepare their students for the content and skill(s) required on this question?

- The *AP Statistics Course and Exam Description* (CED), effective Fall 2020, includes instructional resources for AP Statistics teachers to develop students' broader skills.
    - Please see pages 226, 230, and 232 of the CED for examples of key questions and instructional strategies designed to develop skill 1.E, identify an appropriate inference method for significance tests; 1.F, identify an appropriate inference method for significance tests; skill 3.E, calculate a test statistic and find a *p*-value, provided conditions for inference are met; skill 4.C, verify that inference procedures apply in a given situation; and skill 4.E, justify a claim using a decision based on significance tests.
- A table of representative instructional strategies, including definitions and explanations of each, is included on pages 213–223 of the CED. The strategy "Build the Model Solution," for example, may be helpful in developing students' abilities to use precise language when justifying a claim based on a significance test.
- AP Classroom provides four videos focused on the content and skills to answer this question.
    - The daily video 1 for topic 7.4 discusses how to state the null and alternative hypotheses for a significance test for a population mean (see VAR-7.C.1 and VAR-7.C.2). The key takeaways from this video that were relevant to this question are identifying the null and alternative hypotheses for a significance test for a population mean with unknown sigma, including the mean difference between values in matched pairs.
    - The daily video 2 for topic 7.4 explores how to identify the procedure and check the conditions for performing a significance test for a population mean (See VAR-7.B.1, VAR-7.B.2, and VAR-7.D.1). The key takeaway from this video helps students understand how to verify conditions for a significance test for a population mean with unknown sigma, including the mean difference between values in matched pairs.
    - The daily video 1 for topic 7.5 describes how to calculate a test statistic and *p*-value for a significance test of a population mean (see VAR-7.E.1). The key takeaway of the video shows students how to use the calculator and tables when calculating the *p*-value and test statistic.
    - The daily video 2 for topic 7.5 demonstrates how to interpret the *p*-value and state a conclusion for a significance test for a population mean (see DAT-3.F.1 and DAT-3.F.2). A key takeaway from this video explains to students how to explain the conclusion using the alternative hypothesis.
    - The daily video 3 for topic 7.5 demonstrates how to complete a significance test for matched pair (See VAR-7.C.1, VAR-7.C.2, VAR-7.B.1, VAR-7.B.2, VAR-7.D.1, VAR-7.E.1, DAT-3.F.1, and DAT-3.F.2). The key takeaway from the video is an additional understanding of how to determine the direction of the difference for the matched pair data.
- AP Classroom also provides topic questions for formative assessment of topics 7.4 and 7.5, as well as access to the question bank, which is a searchable database of past AP Questions on this topic.
- The Online Teacher Community features many resources shared by other AP Statistics teachers. For example, to locate resources to give your students practice verifying conditions, try entering the keywords "conditions" in the search bar, then selecting the drop-down menu for "Resource Library." When you filter for "Classroom-Ready Materials," you may find worksheets, data sets, practice questions, and guided notes, among other resources.

Question 5

Task: Multi-Focus
Max Points: 4
Mean Score: 1.93

What were the responses to the question expected to demonstrate?

The primary goals of the question were to assess a student's ability to (1) describe the relationship exhibited on a scatterplot between an independent variable and response variable in context; (2) calculate the predicted y-value for a specific observation, when provided a least-squares regression equation; (3) calculate the residual for an observation; (4) interpret the slope for a provided least-squares regression equation, in context; (5) determine the p-value for a two-sided test of slope, when the test statistic is provided; and (6) compare the p-value for a test of slope to a significance level, then reach a conclusion regarding the slope, in context.

This question primarily assesses skills in skill category 2: Data Analysis, skill category 3: Using Probability and Simulation, and skill category 4: Statistical Argumentation. Skills required for responding to this question include (2.A) Describe data presented numerically or graphically, (2.B) Construct numerical or graphical representations of distributions, (2.C) Calculate summary statistics, relative positions of points within a distribution, correlation, and predicted response, (3.E) Calculate a test statistic and find a p-value, provided conditions for inference are met, (4.B) Interpret statistical calculations and findings to assign meaning or assess a claim, and (4.E) Justify a claim using a decision based on significance tests.

This question covers content from Unit 2: Exploring Two-Variable Data and Unit 9: Inference for Quantitative Data: Slopes of the course framework in the AP Statistics Course and Exam Description. Refer to topics 2.4, 2.6, 2.7, 2.8, and 9.5, and learning objectives DAT-1.A, DAT-1.D, DAT-1.E, DAT-1.H, DAT-3.N, and VAR-7.M.

How well did the responses address the course content related to this question? How well did the responses integrate the skills required on this question?

- In part (a) most responses identified at least two of the characteristics for a scatterplot, but many did comment on all five characteristics. The majority included direction, but some missed strength, form, or unusual features. Most of the responses included the context of chest circumference and weight of tule elk.
- In part (b) most responses correctly calculated the correct predicted value with adequate work shown. Most responses also correctly calculated the residual with adequate work shown. However, a significant number of students reversed the subtraction for the residual, resulting in an incorrect negative value.
- In part (c) most responses were able to correctly identify the slope given a regression equation in the context of their interpretation. However, many responses did not describe the change represented by the slope as a *predicted* change to clearly distinguish it from an observed change in the response variable. Clearly interpreting the slope value with an *increase* in chest circumference resulting in an *increase* in weight, including the correct units, was overlooked on many responses.
- In part (d) many responses did not recognize the need to double the probability found in the table to correctly determine the p-value for the two-sided t-test. Most responses were able to make a correct decision based on a p-value and at least attempt a conclusion statement in context.

What common student misconceptions or gaps in knowledge were seen in the responses to this question?

Common Misconceptions/Knowledge Gaps	*Responses that Demonstrate Understanding*
• In part (a) failing to identify all characteristics of a scatterplot.	• The scatterplot reveals a strong, positive, roughly linear association between the chest circumference and weight of tule elk, with no unusual features.
• In part (b) reversing the observed and predicted values when calculating the residual.	• The predicted weight of a male tule elk with a chest circumference of 145.9 cm is $-350.3 + 3.7455(145.9) \approx 196.17$ kg. The residual for a male tule elk with a chest circumference of 145.9 cm with an actual weight of 204.3 kg is $204.3 - 196.17 \approx 8.13$ kg.
• In part (c) failing to include the correct units (cm and kg), failing to include the concept of a *predicted* value, and failing to specifically reference an *increase* of 1 cm in chest circumference.	• The value of the slope of the least-squares regression line is 3.7455. This value indicates that the predicted weight of a tule elk increases by 3.7455 kilograms for each additional centimeter of chest circumference.
• In part (d) determining a *p*-value from a given test statistic appeared challenging. The responses that did provide a *p*-value often did not take into consideration that the alternative hypothesis is two-tailed. When writing the conclusion statement, many responses used language that was too definitive, implying proof of the alternative hypothesis.	• The degrees of freedom for the test of slope are $30 - 2 = 28.$ The *t*-table shows that for 28 degrees of freedom, the *p*-value for a one-sided test would be 0.001. Because this is a two-sided test, the *p*-value is 0.002. Because the *p*-value is less than $\alpha = 0.05,$ reject the null hypothesis. There is sufficient statistical evidence that the population slope for the linear regression of weight vs. chest circumference for male tule elk is different from 4.5 kg/cm.

Based on your experience at the AP® Reading with student responses, what advice would you offer teachers to help them improve the student performance on the exam?

- Emphasize the importance of the characteristics of a scatterplot: direction of association, strength of association, form of association, and unusual features. Additionally, when writing about these characteristics, responses should be in the context of the study that provided the data.
- Emphasize the importance of adequately showing work and carefully computing predicted values and residual values.
- Discuss the importance of units in an interpretation of slope and that for each *increase* of one unit in the explanatory variable, the slope is the *predicted* change in the units of the response variable.
- Have students practice determining the *p*-value given the test statistic, data, and summary statistics for both one-tailed and two-tailed hypothesis tests. Emphasize the importance of writing the conclusion statement in the context of the alternative hypothesis using non-definitive language justified by a comparison of the *p*-value to an alpha value.

What resources would you recommend to teachers to better prepare their students for the content and skill(s) required on this question?

- The *AP Statistics Course and Exam Description* (CED), effective Fall 2020, includes instructional resources for AP Statistics teachers to develop students' broader skills.
  - Please see pages 227, 230, and 232 of the CED for examples of key questions and instructional strategies designed to develop skill 2.A, describe data presented numerically or graphically; skill 2.B, construct numerical or graphical representations of distributions; skill 2.C, calculate summary statistics, relative positions of points within a distribution, correlation, and predicted response; skill 3.E, calculate a test statistic and find a p-value, provided conditions for inference are met; skill 4.B, interpret statistical calculations and findings to assign meaning or assess a claim; and skill 4.E, justify a claim using a decision based on significance tests.
- A table of representative instructional strategies, including definitions and explanations of each, is included on pages 213–223 of the CED. The strategy "Team FRQ," for example, may be helpful for students to think of how the content is aligned. This FRQ had Units 2 and 9 linked together, and students could mimic this FRQ or create new ones using Unit 1 (Exploring One-Variable Data) and Units 6, 7, or 8 (Inference for Categorical and Quantitative Data).
- AP Classroom provides four videos focused on the content and skills to answer this question.
  - The daily video 2 for topic 2.4 demonstrates how to properly describe a scatterplot (see DAT-1.A.1 to DAT-1.A.6). The key takeaways from this video are how to identify and describe a scatterplot in part (a). If students are struggling to understand the construction of a scatterplot, they can refresh with daily video 1 for topic 2.4.
  - The daily video 1 for topic 2.6 introduces students to the components of a linear regression model and demonstrates making predictions using that model (see DAT-1.D.2). The key takeaway from this video helps students understand the differences in the linear equation from algebra and the linear regression model from statistics. This video will help students with part (b-i).
  - The daily video 1 for topic 2.7 explores calculating and interpreting residuals (See DAT-1.E.1). The key takeaways from this video are how to identify if a residual is positive or negative and how to interpret it in terms of model underprediction and model overprediction after the residual has been calculated using the linear regression model. This video will help students with part (b-ii).
  - The daily video 2 for topic 2.8 describes the precise interpretations of the slope and y-intercept of a linear regression model (see DAT-1.H.2). The key takeaways of the video show students what the slope represents on a scatterplot, where to find the slope in the linear regression model, and how to interpret it (part (c)).
  - The daily videos 1 and 2 for topic 9.5 illustrates how to get the p-value from the t-table and how to conclude the significance test using the p-value and level of significance (see VAR-7.M.2, DAT-3.N.1, and DAT-3.N.2). The key takeaways from these videos explain to students how to use the t-table for a two-tailed test and describe how to conclude the test based on the alternative hypothesis which was completed in part (d).
- AP Classroom also provides topic questions for formative assessment of topics 2.4, 2.6, 2.7, 2.8, and 9.5, as well as access to the question bank, which is a searchable database of past AP Questions on this topic.
- The Online Teacher Community features many resources shared by other AP Statistics teachers. For example, to locate resources to give your students practice verifying conditions, try entering the keywords "regression" in the search bar, then selecting the drop-down menu for "Resource Library." When you filter for "Classroom-Ready Materials," you may find worksheets, data sets, practice questions, and guided notes, among other resources.

Question 6

Task: Investigative Task
Max Points: 4
Mean Score: 1.12

What were the responses to the question expected to demonstrate?

The primary goals of the question were to assess a student's ability to (1) calculate the probability a normally distributed random variable is between two values; (2) calculate the probability a sample mean is greater than a value using the distribution of the sample mean; (3) interpret a probability to determine whether it implies the population mean is a different value than what was assumed; (4) describe the sampling distribution of the sample range, when provided a graph of the results of a simulation; (5) describe how the sampling distribution of the sample range changes as the population standard deviation increases, based on simulated graphs; (6) determine if a value for the sample range is unlikely based on the results of a simulation; and (7) determine whether a machine is working properly based on the previously calculated probability for a mean and the simulated sampling distribution for the sample range.

This question primarily assesses skills in skill category 2: Data Analysis, skill category 3: Using Probability and Simulation, and skill category 4: Statistical Argumentation. Skills required for responding to this question include (2.A) Describe data presented numerically or graphically, (2.D) Compare distributions or relative positions of points within a distribution, (3.A) Determine relative frequencies, proportions, or probabilities using simulation or calculations, and (4.B) Interpret statistical calculations and findings to assign meaning or assess a claim.

This question covers content from Unit 1: Exploring One-Variable Data, Unit 4: Probability, Random Variables, and Probability Distributions, and Unit 5: Sampling Distributions of the course framework in the AP Statistics Course and Exam Description. Refer to topics 1.8, 1.9, 4.7 and 5.2, and learning objectives UNC-1.M, UNC-1.N, VAR-5.B, and VAR-6.A.

How well did the responses address the course content related to this question? How well did the responses integrate the skills required on this question?

- In part (a) most responses correctly indicated the normal distribution along with correct parameters, boundaries, and the event of interest. Most responses also calculated a correct probability. However, many responses satisfied the normality, parameters, and boundaries with minimal communication by indicating only calculator syntax. Some responses used calculator syntax without labeling the lower and/or upper bounds, the mean, and/or the standard deviation, which was not sufficient communication.
- In part (b) many responses identified the normal distribution with correct parameters, showed sufficient work for the standard deviation of the sample mean, and calculated a correct probability. However, many responses failed to recognize the need to divide the population standard deviation by the square root of 2 to calculate the standard deviation of the sample mean. In addition, several responses squared the calculated probability to find the probability that both randomly selected necklaces exceeded 303 mg in gold, rather than finding the probability that the *mean* of two randomly selected necklaces exceeded 303 mg of gold. Most responses correctly stated that the probability calculated in part (b-i) did not indicate the mean amount of gold in the necklaces differed from 300 mg. Many students included correct justification based on the size of the probability in (b-i). Many students correctly indicated 303 mg was within one standard deviation of the population mean of 300 mg. However, many responses made an incorrect argument based on sample size, lack of normality, or inability to appeal to the Central Limit Theorem. Finally, several responses used a formal inference procedure despite being told not to in the prompt.
- In part (c) most responses correctly commented on the right skewness of the simulated sampling distribution. Many responses correctly commented on the center and spread of the distribution. However, several responses gave too wide a range for the center, for example, stating, "It is centered at 0 – 10 mg." In addition, several responses stated, "The range is about 0-20," when the range should be reported as a single number. In (c-ii) many responses commented on a change in the amount of skewness, which was considered extraneous. Many responses

also commented on the increase in the variability in the sampling distributions. However, many responses failed to comment on the increase of the center of the sampling distributions. In addition, several responses had vague statements, such as, "As the population standard deviation increases, the sampling distribution of the sample range increases," without specifying which characteristic of the sampling distribution was increasing.

- In part (d) most responses correctly stated the sample range of 10 mg was not unusual. Many responses referred to the graph and commented on the relative frequency at which a sample range of 10 mg or more occurred. However, many responses did not discuss the tail of the distribution and rather focused on the relative frequency of obtaining a sample range of exactly 10 mg. Several responses did not refer to the graph of the sampling distribution, and based conclusions on a response using 5 mg as the standard deviation, which is the population standard deviation, not the standard deviation of the sampling distribution of the sample range. Furthermore, several responses treated the sampling distribution as normally distributed to justify the response. In (d-ii) many responses correctly concluded there was not convincing evidence the machine was not working properly. However, many responses did not link this conclusion back to the probability calculated in part (b), or the work that was done in part (d-i).

What common student misconceptions or gaps in knowledge were seen in the responses to this question?

Common Misconceptions/Knowledge Gaps	*Responses that Demonstrate Understanding*
• Some responses did not adequately indicate a normal distribution with correct parameters in part (a).	• The amount of gold in a necklace is normally distributed with a mean of 300 mg and a standard deviation of 5 mg. OR • Let X be the amount of gold in a necklace. Then $X \sim N(300, 5)$.
• Some responses did not indicate the boundary or event of interest in part (a).	• The probability a necklace contains between 296 mg and 304 mg of gold is $P(296 < X < 304)$.
• In part (b-i) many responses failed to recognize the need to use a sampling distribution and did not divide the population standard deviation by the square root of 2 for the standard deviation.	• $\sigma_{\bar{X}} = \frac{5}{\sqrt{2}}$ OR • $\bar{X} \sim N\left(300, \frac{5}{\sqrt{2}}\right)$
• In part (b-i) some responses squared the probability calculated from the normal distribution to find the probability that *both* of two randomly selected necklaces exceeded 303 mg.	• $P(\bar{X} > 303) = P\left(Z > \frac{303 - 300}{5/\sqrt{2}}\right) = 0.198$

• In part (b-ii) many responses used incorrect justification for why a sample mean of 303 mg did not indicate the population mean amount of gold applied to necklaces was different from 300. Some incorrect justifications included:
  ○ Use of an inference procedure despite the fact the prompt said, “without using an inference procedure.”
  ○ Claiming the sample size was too small to make a claim or to use normality, or for the Central Limit Theorem to be applied.
	• No, observing a sample mean of 303 mg would not provide convincing evidence that the population mean amount of gold being applied by the machine differs from 300 mg because the probability of observing a sample mean that differs from 300 mg by 3 mg or more is about $2(0.198) = 0.396,$ which is fairly large (greater than 0.1).

• In part (c-i) many responses described the sampling distribution of the sample range using only one or two of the desired characteristics—shape, center, spread—rather than all three.	• The sampling distribution of the sample range when $\sigma = 5$ is skewed to the right. It is centered at roughly 6 mg and sample ranges vary from about 0 mg to about 25 mg.
• In part (c-i) several responses included too large a range of values in describing center, for example, stating, “The distribution is centered at about 0 to 10 mg.”	• The sampling distribution of the sample range is centered at roughly 6 mg.
• In part (c-i) several responses used the mode as a description of center, claiming the distribution was centered at 0 mg.	• The sampling distribution of the sample range is centered at roughly 6 mg.
• In part (c-i) several responses reported the range of the simulated values was 0 to 23, rather than reporting the range as a single number.	• The sample ranges vary from about 0 mg to about 23 mg. OR • The range of the sampling distribution of the sample range is roughly 23 mg.
• In part (c-i) several responses reported that the standard deviation of the sampling distribution was 5 mg because that was the standard deviation of the amount of gold applied in the population.	• The sample ranges vary from about 0 mg to about 25 mg. OR • The range of the sampling distribution of the sample range is roughly 25 mg.
• In part (c-ii) several responses failed to mention that the center of the sampling distribution increases as the population standard deviation increases.	• The median of the sampling distribution increases as the population standard deviation increases.

• In part (c-ii) several responses included a vague comment about the sampling distribution increasing, as opposed to commenting on a specific characteristic of the sampling distribution increasing.	• As the population standard deviation increases, the median and range of the sampling distribution of the sample range both increase.
• In part (c-ii) some responses indicated the maximum of the sampling distribution increases as the population standard deviation increases. However, responses failed to state the minimum stays at 0, failing to communicate how the variability changes.	• While the minimum of the sampling distribution remains at 0, the maximum increases as the population standard deviation increases.
• In (d-i) many responses did not use the given graph of the sampling distribution to make a decision about whether a sample range of 10 mg was unusual.	• Based on the graph of the sampling distribution with a standard deviation of 5 mg, a sample range of 10 mg was not unusual because roughly 17% of the simulated ranges were 10 mg or higher.
• In (d-i) several responses stated that a range of 10 mg was unusual because a value of exactly 10 mg occurred less than 5% of the time in the simulation, rather than considering the relative frequency of obtaining a sample range of 10 or greater.	• Based on the graph of the sampling distribution with a standard deviation of 5 mg, a sample range of 10 mg was not unusual because roughly 17% of the simulated ranges were 10 mg or higher.
• Several responses did not link the work done in part (b) to the decision about whether the machine was working properly in part (d-ii).	• In part (b), we found there is a probability of $2(0.198) = 0.396$ that we would get a sample mean from a sample of two necklaces that differed from 300 mg by 3 mg or more if the true mean amount of gold applied is 300 mg. Also, in (d-i), it was determined a sample range of 10 mg or more is not unusual. Therefore, there is not convincing evidence that the machine is not working properly.
• Several responses did not follow through with the work done in part (d-i) to make a correct decision and conclusion in part (d-ii).	• In part (b), we found there is a probability of $2(0.198) = 0.396$ that we would get a sample mean from a sample of two necklaces that differed from 300 mg by 3 mg or more if the true mean amount of gold applied is 300 mg. Also, in (d-i), it was determined a sample range of 10 mg or more is not unusual. Therefore, there is not convincing evidence that the machine is not working properly.

Based on your experience at the AP® Reading with student responses, what advice would you offer teachers to help them improve the student performance on the exam?

- Students need to understand the type of response required by different task verbs used in the question.
    - TIP**:** In the AP Statistics Course and Exam Description (CED), page 239 contains a list of "Task Verbs in Free-Response Questions." Make sure students are familiar with this list as well as the descriptions for what a response should include when each of these verbs is used.
- Students need to be encouraged to read questions carefully and pay attention to specific information they are given. For example, if the question reads, "Justify your answer without using an inference procedure," an inference procedure should not be used. Justification must come from some other calculation.
- Students should be encouraged to use proper statistical communication. For example, it is good practice to specify how a random variable is defined and indicate how a random variable is distributed in either words or with proper notation [e.g., $X \sim N(\mu, \sigma)$] with the values of the mean and standard deviation specified.
- Students need more practice working with small sample sizes. Students must also understand that if a random variable is normally distributed, then the sampling distribution of the sample mean will be normally distributed without regard to the sample size.
- When determining normal probabilities, students need more practice in identifying when a sampling distribution should be used as opposed to using the distribution of the original random variable.
- Students need better understanding that when they use the results of a simulation to make a conclusion, they should be looking at the tail of the simulated sampling distribution.

What resources would you recommend to teachers to better prepare their students for the content and skill(s) required on this question?

- The *AP Statistics Course and Exam Description* (CED), effective Fall 2020, includes instructional resources for AP Statistics teachers to develop students' broader skills.
    - Please see pages 227, 230, and 232 of the CED for examples of key questions and instructional strategies designed to develop skill 2.A, describe data presented numerically or graphically; skill 2.D, compare distributions or relative positions of points within a distribution; skill 3.A, determine relative frequencies, proportions, or probabilities using simulation or calculations; and skill 4.B, interpret statistical calculations and findings to assign meaning or assess a claim.
- A table of representative instructional strategies, including definitions and explanations of each, is included on pages 213–223 of the CED. The strategy "FRQ Partner Quiz," for example, may be helpful for students to think of the investigative task in different ways. This FRQ used a combination of distributions and quantitative graphs to investigate the question. Additionally, the use of process mapping (or flow charts) could help students keep track of the parts of an FRQ.
- AP Classroom provides four videos focused on the content and skills to answer this question.
    - The daily video 1 for topic 5.2 shows how to calculate probabilities for continuous random variables associated with a normal distribution and determine the interval associated with a given area in a normal distribution (see VAR-6.A). The key takeaways from this video show how to write the parts of a normal distribution, including the correct notation for direction, mean, standard deviation, lower and upper bound, and the distribution (parts (a and b-i)). If students are struggling to understand the parameters and probabilities for a sampling distribution of a sample mean, they can refresh with daily video 2 for topic 5.7.
    - The daily video 2 for topic 4.7 shows students how to interpret or describe a probability distribution (see VAR-5.B.1). The key takeaway from this video helps students understand how probability distribution provides information about the shape, center, and spread of a population and allows one to make conclusions about the population of interest which can help with parts (b-ii and d).
    - The daily video 1 for topic 1.8 explores how the shape of the graph influences the relative relationship of the mean and median (See UNC-1.M.1 and UNC-1.M.2). The key takeaway from this video describes ways to use summary statistics to justify claims (part (c)).

    - The daily video 1 for topic 1.9 compare distributions of quantitative data (See UNC-1.N.1). The key takeaways of the video highlight that students need to address the four important characteristics (shape, center, variability(spread), and unusual features) when comparing distributions of quantitative data (part (c)).
- AP Classroom also provides topic questions for formative assessment of topics 1.8, 1.9, 4.7, and 5.2, as well as access to the question bank, which is a searchable database of past AP Questions on this topic.
- The Online Teacher Community features many resources shared by other AP Statistics teachers. For example, to locate resources to give your students practice verifying conditions, try entering the keywords “distributions” in the search bar, then selecting the drop-down menu for “Resource Library.” When you filter for “Classroom-Ready Materials,” you may find worksheets, data sets, practice questions, and guided notes, among other resources.