AP® STATISTICS
2019 SCORING GUIDELINES

Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) describe a procedure for identifying a potential outlier, (2) apply the outlier identification procedure to data presented in a stemplot and (3) describe the distribution of data presented in a stemplot.

Solution

Part (a):

Outliers are identified as any observation outside of the interval bounded by $Q1 - 1.5(IQR)$ and $Q3 + 1.5(IQR)$. The IQR is $Q3 - Q1 = 390 - 180 = \$210$. The interval lower bound is $180 - 1.5(210) = -\$135$ and the interval upper bound is $390 + 1.5(210) = \$705$. Since \$810 is outside of this interval, it is an outlier. No other observation is outside the interval.

Part (b):

The distribution of the amount of money students spent on textbooks is unimodal and skewed to the right. The sample median is between \$280 and \$290. Money spent on textbooks ranges from about \$120 to \$810 for this sample of students, with fifty percent of the amounts between about \$180 and \$390. The largest observation is a possible outlier.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 1 (continued)

Scoring

Parts (a) and (b) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

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

1. Describes a reasonable procedure for identifying potential outliers.
2. Correctly applies the described procedure for detecting potential outliers to the data presented in the stemplot *AND* clearly identifies potential outliers.

Partially correct (P) if the response includes only one of the two components.

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

Notes:

- Responses that that satisfy component 1 include, but are not limited to, checking for observations that are more than $(1.5)(\text{IQR})$ above the upper quartile or more than $(1.5)(\text{IQR})$ below the lower quartile, checking for observations that are more than two (or three) standard deviations away from the center of the data (sample mean or sample median), looking for one or more large gaps between any extreme observation and the rest of the data.
- Using the values in the stemplot, the sample mean is \$308.82 and the sample standard deviation is \$155.71. The sample median is between \$280 and \$290. There are no observations more than two standard deviations below the sample mean or sample median. The largest observation is a potential outlier because it is more than two (and three) standard deviations above the sample mean (or sample median).
- If the response does not have a clearly described procedure but shows correct work for the IQR or standard deviation approach, credit is given for component 2.
- Responses that interpret values in the stemplot as purchase amounts that are rounded down, should not be penalized. For example, the second largest observation may be interpreted as a purchase amount between \$620 and \$629.
- If the values in the stemplot are interpreted as being rounded down, the largest observation is an outlier because it is more than two (and three) standard deviations above the sample mean (or sample median). The second largest observation is a potential outlier based on the two standard deviation criterion because $(\text{sample mean}) + 2(\text{standard deviations}) = \$620.24$ and the second largest observation is interpreted as a purchase amount between \$620 and \$629.
- If the procedure checks for a gap, it must identify the largest observation as a potential outlier and it must not identify any other observation as a potential outlier.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 1 (continued)

Part (b) is scored as follows:

Essentially correct (E) if the response includes reasonable comments on the following four components.

1. The shape of the distribution is skewed to the right.
2. The center of the distribution is around $300. (Any value between $280 and $320 is acceptable.)
3. The spread of the distribution.
4. The response includes context.

Partially correct (P) if the response includes only three of the four components.

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

Notes:

- Because part (a) addresses possible outliers, the response to part (b) is not required to address potential outliers.
- Responses that satisfy component 3 include, but are not limited to, referring to the range ($690), referring to the standard deviation (any value between $150 and $160 is acceptable), referring to the interquartile range (any value between $200 and $240 is acceptable), or stating that most or all purchase amounts are between two reasonable amounts (simply stating the values without linking them as endpoints of an interval would not receive credit).
- Responses that interpret values in the stemplot as purchase amounts that are rounded down, should not be penalized; for example, stating that the mean amount of money spent on textbooks is between $308.82 and $317.82.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 1 (continued)

4 **Complete Response**

Both parts essentially correct

3 **Substantial Response**

One part essentially correct and one part partially correct

2 **Developing Response**

One part essentially correct and one part incorrect

OR

Both parts partially correct

1 **Minimal Response**

One part partially correct and one part incorrect

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 2

Intent of Question

The primary goal of this question were to assess a student's ability to (1) construct and interpret a confidence interval for the slope of a regression line; and (2) determine if the confidence interval contradicts or supports a prior belief/claim.

Solution:

(a) The 95 percent confidence interval for the slope is computed as follows:

$$-2.158 \pm t_{18,0.975} \times (0.149) = -2.158 \pm (2.101) \times (0.149)$$
$$= (-2.471, -1.845)$$

We are 95% confident that the slope of the true regression line is between -2.471 and -1.845 thousands of dollars per mile. This implies that for each additional mile that a three-bedroom house is away from the city center, the selling price of the house is expected to decline between \$1,845 and \$2,471.

(b) Because the confidence interval contains -2, corresponding to a \$2,000 decrease, it is a plausible value for the slope of the regression line. Consequently, the data do not contradict the agent's belief that the selling prices of three bedroom houses decrease about \$2,000 for every one-mile increase in the distance of the house from the city center.

Scoring

This question is scored in three sections. Section 1 consists of computing the 95% confidence interval for the slope of the true regression line in part (a). Section 2 consists of the interpretation of the confidence interval in part (a). Section 3 consists of the response to part (b). Sections 1, 2, and 3 are each scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the confidence interval displays the form

(estimated slope) $\pm$ critical value $\times$ *SE*(estimated slope)

AND

correct values are used for the estimated slope (-2.158), the standard error for the estimated slope (0.149), and the 0.975 percentile of the *t*-distribution with 18 degrees of freedom (2.101)

AND

correct values are reported for the endpoints of the confidence interval.

Partially correct (P) if the response displays the form (estimated slope) $\pm$ critical value $\times$ *SE*(estimated slope), but uses an incorrect value for one of the following: the t-percentile, the estimate of the slope, or the standard error of the slope.

OR

if the response gives the correct endpoints of the confidence interval but does not provide sufficient work.

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

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 2 (continued)

Section 2 is scored as follows:

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

1. Uses a 95% level of confidence
2. Provides a correct statement about potential values for the population or expected slope of the regression line.
3. Uses the end points of the confidence interval.
4. Presents the interpretation in context.

Partially correct (P) if the response satisfies only three of the four components.

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

Section 3 is scored as follows:

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

1. Provides a correct conclusion about whether the confidence interval contradicts the agent's belief based on the results from part (a). If the correct confidence interval is reported in part (a), then the correct conclusion is that the confidence interval does not contradict (or supports) the agents belief, but it should not indicate that it proves the agent's belief.
2. Links the conclusion to the confidence interval based on where $-2{,}000$, a decrease of 2,000, or -2, is relative to the interval.

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

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

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 2 (continued)

4 **Complete Response**

Three sections essentially correct

3 **Substantial Response**

Two sections essentially correct and one section partially correct

2 **Developing Response**

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 **Minimal Response**

One section essentially correct

OR

No section essentially correct and two sections partially correct

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 3

Intent of Question

The primary goals of this question were to assess a student's ability to (1) extract information about counts from a histogram; (2) use histograms to compare distributions; (3) estimate a median from the combined information in two histograms.

Solution

Part (a):

(i) Three sites were vacated before 8:30 A.M. This is the sum of the counts represented by the two leftmost bars of the histogram of exit times for campsites without young children. No campers with small children vacated their campsites before 8:30 A.M.

(ii) Eight sites were vacated at 11:00 A.M. or later. This is the sum of the counts represented by the two rightmost bars on each histogram.

Part (b):

The distribution of exit times for campers without young children is skewed to the left while the distribution for campers with young children is roughly symmetric. Both exit time distributions appear to be unimodal. The distribution of exit times is more spread out for campers without young children; the largest possible range is around 210 minutes compared to a largest possible range of around 105 minutes for campers with young children. Campers without young children tend to leave the campground earlier than campers with young children; the median exit time for campers without young children is between 60 and 75 minutes after 9:00 A.M. which is less than the median exit time for campers with young children, which is between 90 and 105 minutes after 9:00 A.M.

Part (c)

There are a total of 60 exit times, so the median falls between the 30^{th} and 31^{st} exit times. Since the 30^{th} and 31^{st} exit times fall between 10:15 and 10:30, any time between 10:15 A.M. and 10:30 A.M. provides a reasonable estimate of the median exit time.

Scoring

Parts (a), (b) and (c) are each scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following two components:
1. Correctly states that 3 sites were vacated before 8:30 A.M. in part (i).
2. Correctly states that 8 sites were vacated at 11:00 A.M. or later in part (ii)

Partially correct (P) if the response contains only one of the two components.
OR
Correctly estimates the counts at each site separately, but does not combine them for a total count.

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

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 3 (continued)

Part (b) is scored as follows:

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

1. Correct comparison of the centers of the two distributions
2. Correct comparison of the spread of the two distributions
3. Correct comparison of the shapes of the two distributions
4. Includes context

Partially correct (P) if the response satisfies only two or three of the four components.

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

Part (c) is scored as follows:

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

1. Reports a specific time or range of times between 10:15 A.M. and 10:30 A.M. (between 75 and 90 minutes after 9:00 A.M)
2. Gives a reasonable justification.

Partially correct (P) if the response satisfies only one of the two components
OR
if the response correctly estimates median exit times for both sets of sites (between 10:00 A.M and 10:15 A.M., or 60-75 minutes, for campers without young children, and between 10:30 A.M and 10:45 A.M., or 90-105 minutes, for campers with young children)

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

Notes: A reasonable justification can be:

- Markings on the histogram(s) that illustrate a histogram of the combined exit times.
- A combination of the medians that properly weights the medians in a 2:1 ratio.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 3 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 4

Intent of Question

The primary goals of this question were to assess a student's ability to (1) identify conditions under which a randomized block design would be better than a completely randomized design for detecting a difference in mean responses for two treatments, and (2) describe how an experiment can be conducted as a randomized block design.

Solution

Part (a):

A randomized block experiment will be better for detecting a difference between the abilities of the new and old filters to reduce arsenic concentrations when (1) there is at least one feature of the wells that enable the researchers to create blocks such that each block consists of wells with similar arsenic concentrations and (2) the average arsenic concentration varies among blocks. This would occur, for example, when well water arsenic concentrations at all four wells near the burned forest are higher than the arsenic concentration at any well near the granite ledge.

Part (b):

To take advantage of the differences among wells, well water arsenic concentrations must be similar for wells with each block, but average arsenic concentrations should vary among blocks. A randomized block experiment with two blocks could be conducted by including wells 1, 2, 3, and 4, on the granite ledge side of the field in one block and including wells 5, 6, 7, and 8, on the burned forest side of the field in the other block.

Part (c)

Within each block, the new filter should be randomly assigned to the same number of wells as the old filter. For the blocks identified in part (b), this can be accomplished by writing well numbers 1, 2, 3, and 4 on four slips of paper, placing the four slips into a hat, and mixing them. New filters would be used for the wells on the first two slips drawn from the hat and old filters would be used for the other two wells. Similarly, well numbers 5, 6, 7, and 8 can be written on four slips of paper that are put into a hat and mixed. New filters would be used for the wells on the first two slips drawn from the hat and old filters would be used for the other two wells.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 4 (continued)

Scoring

Parts (a), (b) and (c) are each scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:
1. Well water arsenic concentrations are similar for all wells within the same block.
2. Average well water arsenic concentrations vary substantially from block to block.
3. Statements are presented in the context of arsenic levels and wells.

Partially correct (P) if the response includes only two of the three components.

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

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following two components. Blocks are formed so that
1. Each block contains wells in locations with similar features, consistent with the scenario identified in the response,
2. Each block contains the same number of wells (either two or four).

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

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

Notes:
- Suppose the response to part (a) indicates that arsenic concentrations might be similar for wells on the granite ledge side of the field, and arsenic concentrations might be similar for wells of the burned forest side of the field, but arsenic concentrations might be quite different on those two sides of the field. Then, the two components required for E are satisfied by including wells 1, 2, 3, 4, in one block and including wells 5, 6, 7 and 8 in the other block. The two components required for E are also satisfied with four blocks with wells 1 and 2 in one block, wells 3 and 4 in a second block, wells 5 and 6 in a third block, and wells 7 and 8 in a fourth block. For this response to part (a), component 1 is <u>not</u> satisfied with four blocks consisting of wells 1 and 5, 2 and 6, 3 and 7, 4 and 8.
- If the response to part (a) uses a scenario in which arsenic levels decrease as one moves from the left toward the right side of the field, for example, then components 1 and 2 are satisfied with four blocks consisting of wells 1 and 5, 2 and 6, 3 and 7, 4 and 8. For this response to part (a), however, component 1 is <u>not</u> satisfied with four blocks consisting of wells 1 and 2, 3 and 4, 5 and 6, 7 and 8.
- If no conditions are identified in the response to part (a) and no conditions are given in the response to part (b), use the granite ledge and burned forest sides of the field as the conditions for creating blocks in scoring the response to part (b).
- If additional or alternative conditions are identified in part (b), use those conditions in scoring the response to part (b).

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 4 (continued)

Part (c) is scored as follows:

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

1. An indication that types of filters are randomly assigned to wells.
2. Within each block, each type of filter is assigned to half of the wells.
3. Describes how to correctly implement the random assignment process.

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.

Notes:

- Sufficient detail must be provided to enable a knowledgeable statistics user to implement the randomization method. Some additional acceptable methods are:
  - Using a random number generator to select half of the wells in a block for which the newer filters will be used.
  - For blocks consisting of just two wells, tossing a coin.
- Each well must have a 50% chance of using a newer filter and a 50% chance of using an older filter, and half of the wells within the block must be assigned to each filter.
- Responses to part (c) for designs that do not have an even number of wells in each block are scored no higher than P.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 4 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 5

Intent of Question

The primary goals of this question were to assess a student's ability to (1) recognize a situation in which a binomial probability distribution may be applied; (2) discuss conditions for accurately using a normal approximation to a binomial probability; (3) compute appropriate binomial probabilities.

Solution

Part (a):

The probability that Sasha will experience at least three days with traffic delays in the next 21 days that she travels to work is

$$\begin{aligned} P(\text{Number of delays} \geq 3) &= 1 - [P(\text{no delays}) + P(\text{exactly one delay}) + P(\text{exactly 2 delays})] \\ &= 1 - \left[\binom{21}{0}(0.2)^0(0.8)^{21} + \binom{21}{1}(0.2)^1(0.8)^{20} + \binom{21}{2}(0.2)^2(0.8)^{19} \right] \\ &\approx 1 - [0.00922 + 0.04842 + 0.12106] \\ &\approx 0.82130. \end{aligned}$$

Part (b):

The probability that Sasha's first traffic delay will occur **after** the fifth day is the probability that no traffic delay occurs on any of the first five days:

$$P(\text{No delay during first 5 days}) = (0.8)^5 = 0.32768.$$

Part (c):

No. The sampling distribution of the sample proportion is right skewed. It may not be well approximated by a normal distribution because the expected number of traffic delays in 21 days, $np = (21)(0.20) = 4.2$, is less than 10.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 5 (continued)

Scoring

Parts (a), (b) and (c) are each scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

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

1. Clearly indicates a binomial distribution with $n = 21$ and $p = 0.2$.
2. Indicates the correct boundary value and direction of the event.
3. Reports the correct probability.

Partially correct (P) if the response satisfies component 1 and it does not satisfy one or both of the other two components

OR

the response does not satisfy component 1 and both of the other two components are satisfied.

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

Notes:

- The response B(21, 0.2) satisfies component 1.
- Component 1 and 2 are satisfied by displaying the correct formula for computing the binomial probability using the correct values for n and p, e.g.
  $$1 - \left[\binom{21}{0}(0.2)^0(0.8)^{21} + \binom{21}{1}(0.2)^1(0.8)^{20} + \binom{21}{2}(0.2)^2(0.8)^{19}\right]$$
  Only component 1 is satisfied if the correct binomial distribution is used in an incorrect probability formula, e.g.
  $$\binom{21}{3}(0.2)^3(0.8)^{18}.$$
- For component 2, the boundary value and direction may be described in words, e.g. *P*(at least three delays in the next 21 days).
- Component 2 may be satisfied by displaying a bar graph of a binomial distribution with the appropriate bars shaded.
- The response of $1 - binomcdf\,(n = 21, p = 0.2, \text{upper bound} = 2) \approx 0.8213$ is scored E since *n, p* and the boundary value are clearly identified.
  The response of $1 - binomcdf\,(n = 21, p = 0.2, 2) \approx 0.8213$ is scored P since *n, p* are clearly identified and the boundary value is not identified.
  The response of $1 - binomcdf\,(21, 0.2, 2) \approx 0.8213$ is scored I.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 5 (continued)

- A normal approximation to the binomial is not appropriate since $np = 21 \times 0.2 = 4.2 < 5$.

  A response using the normal approximation can score at most P. To score P, the response must include all of the following:
  - An indication that the probability calculated is a normal approximation for the binomial probability
  - A correct mean and standard deviation based on the binomial parameters
  - Clear indication of boundary and direction with a z-score or diagram
  - The probability computed correctly

  An example of a response, which meets these four criteria, is $P\left(Z \geq \frac{3 - np}{\sqrt{np(1-p)}}\right) = P\left(Z \geq \frac{3 - (21)(0.2)}{\sqrt{(21)(0.2)(0.8)}}\right) \approx 0.744$ and the binomial distribution is mentioned.

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following two components:
1. The correct probability
2. An appropriate justification.

Partially correct (P) if the response satisfies only one of the two components
OR
if the response correctly computes the probability that the first traffic delay is on day 6, $(0.8)(0.8)(0.8)(0.8)(0.8)(0.2) = 0.065536,$ with supporting work.

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

Notes:
- Component 2 is satisfied if the response uses a binomial distribution with $n = 5$ and $p = 0.2$ to compute the probability of no successes in $n = 5$ trials; *or u*ses direct probability rules to obtain $(0.8)(0.8)(0.8)(0.8)(0.8) = 0.32768.$
- Responses that find the probability of the first delay occurring on day 6 or more, with an incorrect upper bound for the sum, should be scored P.

  For example, $0.8^5(0.2) + 0.8^6(0.2) + ... + 0.8^{20}(0.2).$
- Responses that find the probability of traffic delays on each of the first five days, $(0.2)(0.2)(0.2)(0.2)(0.2) = 0.00032,$ should be scored P.
- Responses using a normal approximation to the binomial distribution should be scored I.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 5 (continued)

Part (c) is scored as follows:

Essentially correct (E) if the response contains the following three components:

1. Calculates $np = (21)(0.2) = 4.2$
2. Indicates that np is less than 10 (or less than 5).
3. Reaches the conclusion that the condition is not met.

Partially correct (P) if the response satisfies only two of the three components;
OR
if the response satisfies component 3 and indicates that the sample size is too small with respect to a commonly used standard (for example, $n < 30$), but does not establish a link to the expected count condition.

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

Note: A response that simply indicates that the sample size is too small for the distribution of the sample proportion to be approximately normal is scored I.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 5 (continued)

4 **Complete Response**

Three parts essentially correct

3 **Substantial Response**

Two parts essentially correct and one part partially correct

2 **Developing Response**

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 **Minimal Response**

One part essentially correct

OR

No parts essentially correct and two parts partially correct

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question were to assess a student's ability to (1) analyze results from a completely randomized experiment; (2) identify a situation in which it may be better to compare medians instead of means; (3) develop a test procedure for comparing medians; and (4) interpret results of the test.

Solution

Part (a):

Because the samples are small and the dot plots suggest potential outliers for both keyboards, it is not appropriate to perform a two-sample *t*-test for the difference between the population means.

Part (b):

It is more appropriate to compare population medians because the dot plots indicate that the distributions of data entry times have outliers to the left for both keyboards. When outliers are present, the median provides a better indicator of the typical values than the mean. For each keyboard, the mean of the population of data entry times will be pulled away from the typical values toward the extreme values in the left tail of the distribution.

Part (c):

(i) The completed table rankings for the combined set of 11 data entry times is shown below

Rank	1	2	3	4	5	6	7	8	9	10	11
Keyboard	J	K	J	J	J	J	K	K	K	K	K
Time	158	184	240	248	251	261	267	279	280	284	305

(ii) Sum of ranks for keyboard J: $SR_J = 1 + 3 + 4 + 5 + 6 = 19$

Sum of ranks for keyboard K: $SR_K = 2 + 7 + 8 + 9 + 10 + 11 = 47$

Part (d):

$$W = SR_{\text{J}} - \frac{n_{\text{J}}(n_{\text{J}} + 1)}{2} = 19 - \frac{(5)(5 + 1)}{2} = 4$$

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 6 (continued)

Part (e):

(i) To obtain $W = 0$, we must have $0 = SR_J - \frac{n_J(n_J + 1)}{2}$ which implies that $SR_J = \frac{n_J(n_J + 1)}{2} = \frac{(5)(5 + 1)}{2} = 15$.

(ii) All five ranks for keyboard J must be smaller than any of the ranks for keyboard K.

Rank	1	2	3	4	5	6	7	8	9	10	11
Keyboard	J	J	J	J	J	K	K	K	K	K	K

Part (f):

For the one-sided alternative that the median of the distribution of data entry times for keyboard J is smaller than the median of the distribution of data entry times for keyboard K, the *p*-value is the proportion of possible arrangements with W values of 4 or less. From the graph, the *p*-value is $\frac{12}{462} = 0.026$. Because the *p*-value is smaller than the $\alpha = 0.05$ significance level, there is sufficient evidence to reject the null hypothesis in favor of the alternative that the median of the distribution of data entry times for keyboard J is smaller than the median of the distribution of data entry times for keyboard K.

Scoring

This question is scored in four sections. Section 1 consists of parts (a) and (b), section 2 consists of parts (c) and (d), section 3 consists of part (e), and section 4 consists of part (f). Sections 1, 2, 3 and 4 are each scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

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

1. The response to part (a) correctly argues that a two-sample *t*-test is not appropriate because the sample sizes are small and the dot plots show potential outliers (or left skewed distributions).
2. The response to part (b) uses the outliers (skewness) in the entry time distributions to argue that the medians are better indicators of typical entry times than the means.

Partially correct (P) if the response includes only one of the two components.

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

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 6 (continued)

Section 2 is scored as follows:

Essentially correct (E) if the response includes the following three components:
1. The table of rankings in part (c-i) is completed correctly.
2. The sum of ranks in part (c-ii) is computed correctly for both keyboards, using the ranks reported in the table in part (c-i).
3. The value of W is calculated correctly in part (d), based on the value of SR_J in part (c-ii) and the correct value of n_J.

Partially correct (P) if the response includes only two of the three components.

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

Section 3 is scored as follows:

Essentially correct (E) if the response includes the following two components:
1. The response to part (e-i) correctly finds the value of SR_J when $W = 0$, using the value of n_J from the response to part (d).
2. The correct assignment of ranks is displayed in the response to part (e-ii) to give $W = 0$ or the value of SR_J shown in part (e-i).

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

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

Section 4 is scored as follows:

Essentially correct if the response in part (f) includes the following four components:
1. The critical region is identified as the possible values of W that are less than or equal to the value of W reported in part (d). This can be done by marking on the graph shown in the stem prior to part (e).
2. A correct *p*-value is reported for the indicated critical region.
3. A correct conclusion is reached based on a reasonable reported *p*-value.
4. The conclusion is presented in the context of comparing the medians of the data entry time distributions for keyboards J and K.

Partially correct if the response satisfies only two or three of the four components.

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

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

AP® STATISTICS
2019 SCORING GUIDELINES

Question 6 (continued)

Each essentially correct (E) section counts as 1 point, and each partially correct (P) section counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ 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.

© 2019 The College Board.
Visit the College Board on the web: collegeboard.org.

2019 AP Statistics
Question Descriptors and Performance Data

Question	Skill	Learning Objective	Topic	Key	% Correct
32	3.D	UNC-4.C	Constructing a Confidence Interval for a Population Proportion	E	74
33	4.B	UNC-4.AF	Confidence Intervals for the Slope of a Regression Model	B	32
34	4.E	DAT-3.D	Carrying Out a Test for the Difference of Two Population Proportions	E	43
35	4.C	VAR-7.L	Setting Up a Test for the Slope of a Regression Model	E	25
36	4.B	DAT-3.A	Interpreting P-Values	E	55
37	3.B	VAR-5.E	Combining Random Variables	D	28
38	3.D	UNC-4.K	Confidence Intervals for the Difference of Two Proportions	A	77
39	3.E	VAR-8.L	Carrying Out a Chi-Square Test for Homogeneity or Independence	C	40
40	4.B	UNC-4.S	Justifying a Claim About a Population Mean Based on a Confidence Interval	E	66

Free-Response Questions

Question	Skill	Learning Objective	Topic	Mean Score
1	2.A|4.B	UNC-1.K|UNC-1.H|UNC-1.M	1.7|1.6|1.8	2.25
2	1.D|3.D|4.B|4.D	UNC-4.AC|UNC-4.AF|UNC-4.AG|UNC-4.AH	9.2|9.3	1.2
3	2.A|2.D	UNC-1.N|UNC-1.M	1.9|1.8	2.19
4	1.B|1.C	VAR-3.D|VAR-3.B	3.6|3.5	1.47
5	3.A|3.C	UNC-3.B|VAR-4.E|UNC-3.L	4.10|4.6|5.5	0.97
6	2.A|2.B|4.B|4.C	VAR-7.H|UNC-1.M	7.8|1.8	1.93