
AP[®] Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free-Response Question 1

- ☒ **Scoring Guidelines**
- ☒ **Student Samples**
- ☒ **Scoring Commentary**

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

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each part of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

	Model Solution	Scoring
A	<p>The distribution of gas mileage for the sample of cars manufactured in Country A has a lower center than the distribution of gas mileage for the sample of cars manufactured in Country B. The median gas mileage for the sample of cars manufactured in Country A (18 mpg) is less than the median gas mileage for the sample of cars manufactured in Country B (32 mpg).</p> <p>The range of the gas mileages for the sample of cars manufactured in Country A (24 mpg) is slightly greater than the range of the gas mileages for the sample of cars manufactured in Country B (22 mpg). However, the IQR of the gas mileages for the sample of cars manufactured in Country A (8 mpg) is less than the IQR of the gas mileages for the sample of the cars manufactured in Country B (12 mpg).</p> <p>The car manufactured in Country A with 38 mpg (the maximum of the sample of cars manufactured in Country A) is an outlier, while the distribution of gas mileage for the sample of cars manufactured in Country B has no outliers.</p>	<p>Essentially correct (E) if the response satisfies at least three of the following four components:</p> <ol style="list-style-type: none"> Directly compares the center for the two distributions Directly compares the spread (either IQR or range) for the two distributions Indicates that the gas mileage of one of the cars manufactured in Country A is an outlier Provides sufficient context, which includes the manufacturing countries (“Country A” and “Country B”) AND the dependent variable (“gas mileage” or “mpg”) <p>Partially correct (P) if the response satisfies only two of the four components required for E.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Scoring Notes:

- The response need not include specific numerical values to satisfy any given component or to score E on part A.
- A response that only uses the “means” in the comparison of center would not satisfy component 1.
- Any acceptable mention of shape in the response should be ignored because complete shape information cannot be determined from a boxplot. Acceptable mentions of shape include:
 - The shape of the distribution of the gas mileage for cars manufactured in Country A can be described as skewed, positively skewed, or right skewed.
 - The shape of the distribution of the gas mileage for cars manufactured in Country B can be described as skewed, negatively skewed, left skewed, or **approximately** symmetric.

- If the response describes the shape of either distribution as just “symmetric,” “normal,” “unimodal,” or an incorrect shape (e.g., “the distribution of gas mileages for Country A is left skewed” or “the distribution of gas mileages for Country B is right skewed”), then part A cannot be scored E.
 - A response that only lists values for center and/or spread and does not directly compare them does not satisfy components 1 and/or 2.
 - A response that just refers to “A” or “B” AND the dependent variable (“gas mileage” or “mpg”) may satisfy component 4.
-

Model Solution	Scoring
<p>B The mean of the distribution of gas mileage for the sample of cars manufactured in Country A is expected to be greater than 18 mpg, the median of the distribution. Because the distribution of gas mileage for the sample of cars manufactured in Country A has an outlier to the right (or is skewed to the right), the mean of the distribution (which is not resistant) is expected to be pulled above the median (which is resistant) toward the higher values of gas mileage.</p>	<p>Essentially correct (E) if the response satisfies the following three components:</p> <ol style="list-style-type: none"> 1. Indicates that the mean of the distribution of gas mileage for the sample of cars manufactured in Country A is expected to be greater than 18 mpg 2. Indicates 18 as the value of the median 3. Provides a reasonable justification for the relationship between the mean and the median based on a right-skewed distribution or the expectation that an outlier in the right tail will pull the mean above the median <p>Partially correct (P) if the response satisfies component 1 AND either component 2 or component 3.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Scoring Notes:

- A response that states the median is 18 mpg in either part A or part B satisfies component 2.
- A response that provides a weak justification (e.g., “because the distribution is skewed”) does not satisfy component 3. However, a response that provides a justification based on the distribution being right skewed may satisfy component 3.
- Component 3 may be satisfied with an appropriate numerical argument that shows that the mean must be greater than 18 mpg.
 - An example of an inappropriate numerical argument would be averaging portions of the five-number summary.

Model Solution	Scoring
<p>C i. The maximum value in the combined data is 40 mpg because 40 mpg is the maximum gas mileage for the sample of cars manufactured in Country B, and as shown in the boxplot, all the gas mileages for the sample of cars manufactured in Country A are less than 40 mpg. The minimum value in the combined data is 14 mpg, because 14 mpg is the minimum mpg for the sample of cars manufactured in Country A, and as shown in the boxplot, all the gas mileages for the sample of cars manufactured in Country B are greater than 14. Thus, the range of the combined data set is $40 - 14 = 26$ mpg.</p> <p>ii. In the combined data, there are 200 gas mileages. The median is a value where at least half, or 100, of the gas mileages in the combined data are less than or equal to the median value and at least half, or 100, of the gas mileages in the combined data are greater than or equal to the median value. From the boxplot for the sample of cars manufactured in Country A, the third quartile, Q_3, is 24 mpg indicating there are at least 75 gas mileages less than or equal to 24 mpg and at least 25 gas mileages greater than or equal to 24 mpg. From the boxplot for the sample of cars manufactured in Country B, the first quartile, Q_1, is 24 mpg indicating there are at least 25 gas mileages less than or equal to 24 mpg and at least 75 gas mileages greater than or equal to 24 mpg. Thus, in the combined data set, there are at least 100 gas mileages less than or equal to 24 mpg and at least 100 gas mileages greater than or equal to 24 mpg, which implies 24 is the value of the median of the combined data set.</p>	<p>Essentially correct (E) if the response satisfies the following four components:</p> <ol style="list-style-type: none"> 1. Correctly calculates 26 mpg as the range of the combined data set 2. Provides a justification for the range with an argument based on identifying 40 as the maximum of the combined data set and identifying 14 as the minimum of the combined data set 3. Calculates 24 mpg as a possible value of the median of the combined data set 4. Provides a justification for the median by indicating that at least half of the combined data values are at most 24 and at least half of the combined data values are at least 24 <p>Partially correct (P) if the response satisfies only two or three of the four components.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Scoring Notes:

- A response that solely indicates a range of values (e.g., “the gas mileages range from 14 mpg to 40 mpg”) does not satisfy component 1 but may satisfy component 2.
- A response that states the range as a value of 25 or 27 from misinterpreting the minimum value in the distribution of gas mileages for Country A as 13 or 15 may be considered a minor error and satisfies components 1 and 2.

- A response that gives the correct value for the median (24) and justifies the value by looking at the quartiles on one side of 24 (e.g., identifying that one-fourth of Country B cars and three-fourths of Country A cars have mpg less than or equal to 24) satisfies component 4.
 - A response that provides justification of the median based on a value near the 100th data value satisfies component 4.
 - A response that provides justification of the median based on counting the segments in the boxplots may satisfy component 4. The response is not required to specify equal sample sizes.
 - For example, one-half of the boxplot segments are no greater than 24 mpg (four of eight segments—three segments from Country A and one segment from Country B) and one-half of the boxplot segments are at least 24 mpg (four of eight segments—one segment from Country A and three segments from Country B), which indicates that 24 mpg is the median of the combined data set.
-

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

Question 1

(1)

Begin your response to QUESTION 1 on this page.

(A) The distribution of gas mileages is skewed right ~~for~~ ^{for} cars from country A, whereas it is ^{slightly} skewed left ~~for~~ ^{for} cars from country B. ~~Cars from country B are also skewed right.~~

The distribution of cars from country A has one evident ~~outlier~~ upper outlier at 30 mpg, whereas country B has no evident outliers.

Cars from country A have lower median gas mileages than that of country B, with a median of only 18 mpg compared to country B's median of 32 mpg.

Cars from country A ~~also vary~~ also vary less than country B, with an IQR of $24 - 16 = 8$ mpg, compared to B's $36 - 24 = 12$ mpg IQR.

(B) For the distribution of gas mileage for the sample of cars from country A, it would be expected for the mean to be greater than 18 mpg. ^{This is because} ~~Since~~ the distribution is skewed right and has an upper outlier. Since mean is not resistant to outliers ~~and~~, it will be ~~impacted~~ impacted by the upper outlier and therefore be greater than the ~~mean~~ median of 18 mpg.

(C) ~~next~~ ^{next} page

Question 1

Continue your response to QUESTION 1 on this page.

(1)

(C)

(i.)

$$\text{Range} = 40 - 14 = \underline{26} \text{ mpg}$$

The range of the new dataset is 26 mpg, since the max in the dataset from country B is 40 mpg and the min of the dataset from country A is 14 mpg. Their difference (and therefore the range) is 26 mpg.

(ii.) A possible median of the combined dataset is 24 mpg.

This is because 75% of country A's dataset is ≤ 24 mpg (Q_3 , 75 cars), and 25% of country B's dataset is ≤ 24 mpg (Q_1 , 25 cars). This adds to 100 cars, which ~~is~~ ^{is the} ~~the median~~ implies that 24 mpg is a possible median of the new dataset.

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

0135586

Question 1

Begin your response to QUESTION 1 on this page.

A. The gas mileage for cars manufactured in country A is slightly skewed right with an outlier at 38 miles per gallon. The median is 18 miles per gallon and has an Interquartile range of 8.

Meanwhile the gas mileage for cars manufactured in country B is approximately normal with no outliers and a higher median of 32 mpg and more variation with an IQR of 12.

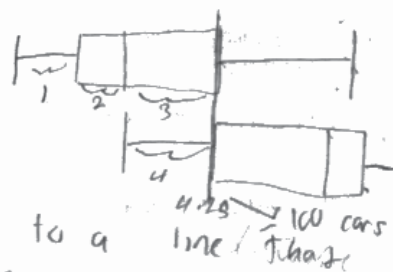
B. Because the data for country A is skewed to the right, we will expect the mean to be greater than the median of 18 mpg.

Ci. Range = max of both - min of both

max of both = 40 mpg

min of both = 14 mpg

$$\text{range} = 40 - 14 = 16$$



Cii. 24 mpg because each interval to a line that represents a quartile is roughly 28 cars. The median of the combined set would be at 100 cars so if you count the quartiles, adding the overlap so that they equal 100, you can get a good estimate of the median @ 24 mpg.

Question 1

Begin your response to QUESTION 1 on this page.

A) Country A is skewed right with an outlier at 38 and a lower median at 18. Country B is skewed left with no outliers and a median at 32. Country B has a much larger range of 22 by 14.

B) If the sample from country A is not biased, the mean should also be 18. If the population mean is 18, the sample mean should be equal.

Ci) Range:

Country B = 18 \rightarrow 40 Country A = 14 \rightarrow 30

Combined = 14 \rightarrow 40

$$40 - 14 = \boxed{26 \text{ mpg}}$$

Cii) $\bar{X}_A = 18$ $\bar{X}_B = 32$

$$\bar{X}_{A+B} = \frac{18 + 32}{2} = \boxed{25 \text{ mpg}}$$

Question 1

Note: Student samples are quoted verbatim and may contain spelling and grammatical errors.

Overview

NEW for 2025: The question overviews can be found in the *Chief Reader Report on Student Responses on AP Central*.

Sample: 1A

Score: 4

The response earned the following: part A – E, part B – E, and part C – E.

In part A the response correctly compares the measures of center (median), satisfying component 1; correctly compares the measures of spread (IQR), satisfying component 2; correctly identifies the outlier in the distribution of gas mileage for Country A, satisfying component 3; and contains sufficient context, satisfying component 4. This part of the response was scored essentially correct (E).

In part B the response correctly states “it would be expected for the mean to be greater than 18 mpg,” satisfying component 1; correctly indicates the median is equal to 18 mpg, satisfying component 2; and correctly justifies the mean is greater than the median using the skewed right and upper outlier justifications (although only one justification is needed), satisfying component 3. This part of the response was scored essentially correct (E).

In part C the response correctly calculates the range of the combined data as 26 mpg, satisfying component 1; provides justification for the range using the maximum and minimum of the combined data set, satisfying component 2; correctly indicates 24 mpg as a possible value of the median of the combined data set, satisfying component 3; and justifies the median by indicating that there are 25 cars in Country A and 75 cars in Country B, representing 100 cars at or below 24 mpg, satisfying component 4. This part of the response was scored essentially correct (E).

Sample: 1B

Score: 2

The response earned the following: part A – P, part B – E, and part C – P.

In part A the response correctly compares the measures of center (median), satisfying component 1; correctly compares the measures of spread (IQR), satisfying component 2; correctly identifies the outlier in the distribution of gas mileage for Country A, satisfying component 3; and contains sufficient context satisfying component 4. The response satisfies all four components for an (E), but the response refers to the distribution for Country B as “approximately normal,” which lowered the score. (See the fourth bullet in the scoring notes.) This part of the response was scored partially correct (P).

In part B the response correctly states the mean for Country A is greater than 18 mpg, satisfying component 1; correctly indicates the median is equal to 18 mpg, satisfying component 2; and correctly justifies the mean is greater than the median using “skewed to the right,” satisfying component 3. This part of the response was scored essentially correct (E).

In part C the response does not correctly calculate 26 mpg as the range, so it does not satisfy component 1. The response provides justification for the range using the maximum and minimum of the combined data set, satisfying component 2; correctly identifies the median of the combined data set as 24, satisfying component 3; and justifies the median by indicating that four of the eight segments are at or below 24, as illustrated in the drawing, represent 100 cars, which satisfies component 4. For satisfying three of the four components, this part of the response was scored partially correct (P).

Question 1 (continued)**Sample: 1C****Score: 1**

The response earned the following: part A – P, part B – I, and part C – P.

In part A the response correctly compares the measures of center by stating “lower median,” satisfying component 1. Although the response attempts to compare the ranges of the two distributions, it incorrectly identifies the range of Country A as 16 by erroneously omitting the outlier, so it does not satisfy component 2. The response correctly identifies the outlier in the distribution for Country A, satisfying component 3. The response does not contain sufficient context as it fails to include the dependent variable (mpg), so it does not satisfy component 4. Because only two of the four components are satisfied, this part of the response was scored partially correct (P).

In part B the response incorrectly states that the mean should be equal to 18, so it does not satisfy component 1. The response correctly identifies that the median is equal to 18 in part A, satisfying component 2. The response does not provide a reasonable justification for the relationship between the mean and median, so it does not satisfy component 3. Because this part of the response fails to satisfy component 1, it was scored incorrect (I).

In part C the response correctly calculates 26 mpg as the range, satisfying component 1. The response provides justification for the range using the maximum and minimum of the combined data set, satisfying component 2. The response does not correctly calculate 24 mpg as a possible value of the median, so it does not satisfy component 3. The response uses an incorrect justification based on an average of the medians of the two countries, so it does not satisfy component 4. For satisfying two of the four components, this part of the response was scored partially correct (P).