5. A student measured the heights and the arm spans, rounded to the nearest inch, of each person in a random sample of 12 seniors at a high school. A scatterplot of arm span versus height for the 12 seniors is shown.
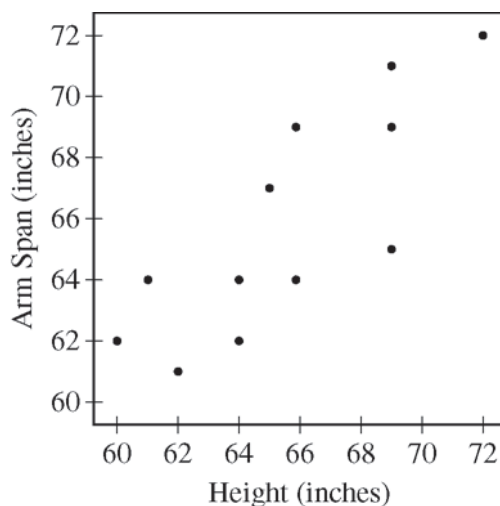
(a) Based on the scatterplot, describe the relationship between arm span and height for the sample of 12 seniors.

Let x represent height, in inches, and let y represent arm span, in inches. Two scatterplots of the same data are shown below. Graph 1 shows the data with the least squares regression line $\hat{y} = 11.74 + 0.8247x$, and graph 2 shows the data with the line $y = x$.

GRAPH 1

GRAPH 2

GO ON TO THE NEXT PAGE.
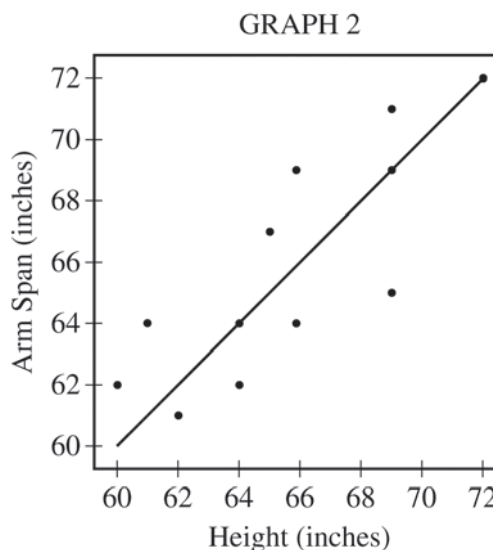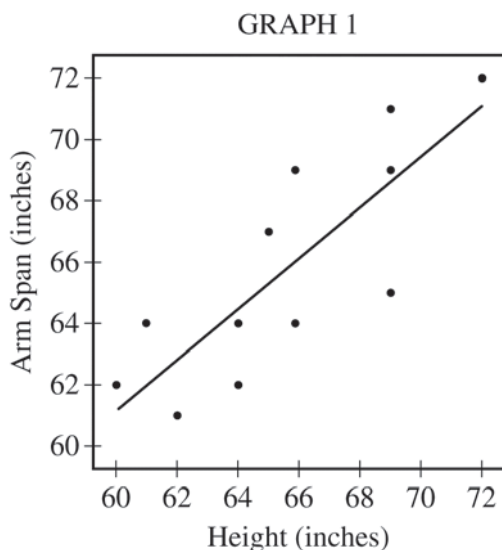
(b) The criteria described in the table below can be used to classify people into one of three body shape categories: square, tall rectangle, or short rectangle.

Square	Tall Rectangle	Short Rectangle
Arm span is equal to height.	Arm span is less than height.	Arm span is greater than height.

  (i) For which graph, 1 or 2, is the line helpful in classifying a student's body shape as square, tall rectangle, or short rectangle? Explain.

  (ii) Complete the table of classifications for the 12 seniors.

Classification	Square	Tall Rectangle	Short Rectangle
Frequency			

(c) Using the best model for prediction, calculate the predicted arm span for a senior with height 61 inches.

**STATISTICS**

SECTION II

Part B

Question 6

Spend about 25 minutes on this part of the exam.

Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Corn tortillas are made at a large facility that produces 100,000 tortillas per day on each of its two production lines. The distribution of the diameters of the tortillas produced on production line A is approximately normal with mean 5.9 inches, and the distribution of the diameters of the tortillas produced on production line B is approximately normal with mean 6.1 inches. The figure below shows the distributions of diameters for the two production lines.
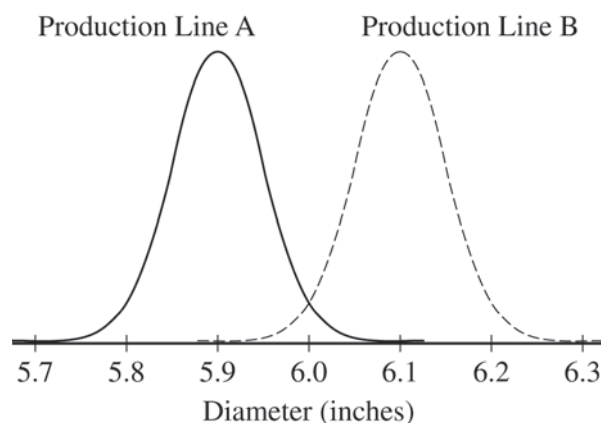
The tortillas produced at the factory are advertised as having a diameter of 6 inches. For the purpose of quality control, a sample of 200 tortillas is selected and the diameters are measured. From the sample of 200 tortillas, the manager of the facility wants to estimate the mean diameter, in inches, of the 200,000 tortillas produced on a given day. Two sampling methods have been proposed.

    Method 1: Take a random sample of 200 tortillas from the 200,000 tortillas produced on a given day. Measure the diameter of each selected tortilla.

    Method 2: Randomly select one of the two production lines on a given day. Take a random sample of 200 tortillas from the 100,000 tortillas produced by the selected production line. Measure the diameter of each selected tortilla.

(a) Will a sample obtained using Method 2 be representative of the population of all tortillas made that day, with respect to the diameters of the tortillas? Explain why or why not.

Question 5

Intent of Question

The primary goals of this question were to assess a student's ability to (1) use the information provided by a scatterplot to describe the relationship between two quantitative variables; (2) interpret and use the information given by lines displayed on a scatterplot; and (3) use a regression equation to estimate a predicted value of y for a given x value.

Solution

Part (a):

There is a moderately strong, positive, linear relationship between height and arm span so that taller students tend to have longer arm spans.

Part (b):

(i) The line in Graph 2 is the one that is helpful. For each student, the graph illustrates whether arm span is equal to height (points on the line), arm span is greater than height (points above the line), or arm span is less than height (points below the line).

(ii)

Classification	Square	Tall Rectangle	Short Rectangle
Frequency	3	4	5

Part (c):

The predicted arm span is $\hat{y} = 11.74 + 0.8247x = 11.74 + 0.8247(61) = 62.05$ inches.

Scoring

Parts (a), (b), and (c) 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 four components:
1. The relationship is approximately linear.
2. The relationship is positive.
3. There is a moderate to strong association (or relationship).
4. The response is given in context.

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

Incorrect (I) if the response includes at most one of the four components.

Notes:
- Correlation alone is not sufficient to indicate a linear relationship. However, linear correlation is sufficient to indicate a linear relationship.
- Reporting that the correlation coefficient is 0.81, the correlation coefficient is close to 1.0, or some other value or range of values for the correlation coefficient is not sufficient to satisfy component 3.

Part (b) is scored as follows:

Essentially correct (E) if the response contains the following four components:
1. The line in Graph 2 is selected.
2. A reasonable explanation for selecting the line in Graph 2 that links the body shapes to the regions defined by the line is given. The explanation should indicate that square shapes (arm span is equal to height) are represented by points on the line, tall rectangle shapes (arm span less than height) are represented by points below the line, and short rectangle shapes (arm span is greater than height) are represented by points above the line.
3. Correct counting of body types demonstrated by reporting correct frequencies (3, 4, and 5) or reporting correct proportions $\left(\frac{1}{4} = \frac{3}{12}, \frac{1}{3} = \frac{4}{12}, \text{and } \frac{5}{12} \right)$ for the square, tall rectangle, and short rectangle body types, respectively.
4. The correct frequencies (3, 4, and 5) are reported in the table.

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

Incorrect (I) if the response includes at most one of the four components.

Notes:
- Selecting the regression line on Graph 1 cannot satisfy either component 1 or component 2.
- To satisfy component 1, it is sufficient to refer to the $y = x$ line without explicitly mentioning Graph 2.
- Use of incorrect labels, such as regression line or least-squares regression line, in referring to the $y = x$ line in Graph 2 is an incorrect use of terminology that should be strongly discouraged, but it is ignored in this scoring rubric.
- The explanation for selecting Graph 2 is acceptable if it explicitly includes at least two of the following: (i) square body shapes (arm span equal to height) are represented by points *on* the line, (ii) tall rectangle body shapes (arm span less than height) are represented by points *below* the line, or (iii) short rectangle body shapes (arm span greater than height) are represented by points *above* the line.
- Incorrectly counting the points on the $y = x$ line in Graph 2 is considered a minor error. Frequencies (2, 4, and 5) are accepted for both component 3 and component 4. Reporting corresponding proportions $\left(\frac{2}{11}, \frac{4}{11}, \frac{5}{11} \right)$ or $\left(\frac{2}{12}, \frac{4}{12}, \frac{5}{12} \right)$ in the table satisfies component 3 but does not satisfy component 4.
- Frequencies reported in the table as (2, 5, 4) or (3, 5, 4), interchanging the counts for tall and short rectangle body shapes, satisfies component 3, but does not satisfy component 4. Reporting corresponding proportions, $\left(\frac{2}{11}, \frac{5}{11}, \frac{4}{11} \right)$ or $\left(\frac{3}{12}, \frac{5}{12}, \frac{4}{12} \right)$, in the table does not satisfy either component 3 or component 4.

Part (c) is scored as follows:

Essentially correct (E) if the response contains the following three components:
1. Correct formula for predicting arm span with 61 inserted for x.
2. Correct value for the predicted arm span.
3. Units for the predicted arm span given as inches.

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

Incorrect (I) if the response includes at most one of the three components.

Notes:
- Any value for the predicted arm span between 61.5 and 62.5 is acceptable; values outside the interval do not satisfy component 2. This allows for inaccuracy in obtaining a value for the predicted arm span directly from the line displayed on Graph 1 and for rounding in applying the prediction formula; but it excludes predictions based on the $y = x$ line and other unreasonable predictions.
- Reporting a prediction equation that is similar to the equation given in the stem, with 61 inserted for x, satisfies components 1 and 2 if the value of the predicted arm span is between 61.5 and 62.5 inches. Otherwise, neither of those two components is satisfied. This could occur if a student enters the data from the graph into a calculator to compute the least squares regression line, instead of using the formula provided in the stem.

4   Complete Response

> All three parts essentially correct

3   Substantial Response

> Two parts essentially correct and one part partially correct

2   Developing Response

> Two parts essentially correct and one part incorrect
>
> *OR*
>
> One part essentially correct and two parts partially correct
>
> *OR*
>
> One part essentially correct, one part partially correct, and one part incorrect
>
> *OR*
>
> Three parts partially correct

1   Minimal Response

> One part essentially correct and two parts incorrect
>
> *OR*
>
> Two parts partially correct and one part incorrect