

**AP<sup>®</sup> STATISTICS**  
**2007 SCORING GUIDELINES (Form B)**

**Question 1**

**Intent of Question**

The three primary goals of this question are to assess a student's ability to: (1) construct a stemplot from a given data set; (2) describe the important features of the plot; and (3) discuss how a single measure of centrality fails to convey important features of the plot.

**Solution**

**Part (a):**

0| 89  
1| 26878993640  
2|  
3| 3856  
4| 143                      Legend: 1| 2 represents 12 questions answered correctly

*OR*, with ordered leaves (not required)

0| 89  
1| 02346678899  
2|  
3| 3568  
4| 134                      Legend: 1| 2 represents 12 questions answered correctly

*OR*, with repeated stems (leaves may be ordered or not)

0H 89  
1L 0234  
1H 6678899  
2 L  
2H  
3L 3  
3H 568                      Legend: 3H 6 represents 36 questions answered correctly  
4L 134                      4L 1 represents 41 questions answered correctly

**Part (b):**

The most striking feature of the plot is that the scores cluster into two groups, one concentrated in the mid-teens and the other in the high 30s (or one with relatively low scores on the exam and one with relatively high scores). There are no scores in the 20s.

**Part (c):**

A measure of center might fall between the two groups (as does the mean of 22.95 here) where there is no data and would not provide an accurate picture of student performance on the exam. It would not indicate that students tended to score either very well or very poorly on the exam.

**AP<sup>®</sup> STATISTICS**  
**2007 SCORING GUIDELINES (Form B)**

**Question 1 (continued)**

**Scoring**

This question is scored in four sections: section 1 is part (a), and sections 2 to 4 consist of elements of parts (b) and (c).

**Section 1** is scored as either essentially correct (E) or incorrect (I).

Section 1 is essentially correct (E) if in part (a) the student gives a correctly constructed stemplot. Any other type of plot is incorrect (I).

NOTE: One or two misplaced or omitted leaves can still be considered essentially correct as long as the important features of the display are not altered.

Parts (b) and (c) are scored together in three sections, each of which is scored as essentially correct (E), partially correct (P), or incorrect (I).

**Section 2** is essentially correct (E) if in either part (b) or (c) the student clearly notices:

1. that there are two groups;
2. that there is a gap in the middle of the distribution;
3. the relative or specific positions of the two groups,

*OR*

the location of the gap,

*OR*

a general measure of location (such as mean, median, or the fact that most scores fall between 10 and 19). (Median = 18, mean = 22.95)

Section 2 is partially correct (P) if the student notes two out of the three.

**Section 3** is essentially correct (E) if in part (b) or part (c) the solution is given in the context of the problem and is communicated well.

Section 3 is partially correct (P) if the student mentions the context (for instance, using the word “scores”), but communication of the context is weak.

Section 3 is incorrect (I) if the context is not mentioned at all.

**Section 4** is essentially correct (E) if in part (c) a valid reason is given for why a measure of center is not sufficient for data of this type (with the two groups and a gap). If, for instance, the reasoning would apply equally well to other shapes, it is not sufficient.

Section 4 can be at most partially correct (P) if a student does not recognize the groups or gap. It is partially correct if the student compares the mean and median and cites outliers or skewness as the reason why a measure of center is not sufficient, or if a general reason is given for why a measure of center is not sufficient. (For instance, the student may say that center alone without some measure of spread is never sufficient.)

**AP<sup>®</sup> STATISTICS**  
**2007 SCORING GUIDELINES (Form B)**

**Question 1 (continued)**

**4 Complete Response**

All four sections essentially correct

**3 Substantial Response**

Three sections essentially correct and no sections partially correct

*OR*

Two sections essentially correct and two sections partially correct

**2 Developing Response**

Two sections essentially correct and no sections partially correct

*OR*

One section essentially correct and two sections partially correct

Note: A score cannot exceed 2 if (1) the student fails to notice either of the two distinct groups of scores or the gap between the groups, and (2) the response to part (c) mentions neither the two groups nor the gap.

**1 Minimal Response**

One section essentially correct and no sections partially correct

*OR*

No sections essentially correct and two sections partially correct

**If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.**

42  
18

**STATISTICS**  
**SECTION II**  
**Part A**  
**Questions 1-5**

Spend about 65 minutes on this part of the exam.

Percent of Section II grade—75

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

1. The Better Business Council of a large city has concluded that students in the city's schools are not learning enough about economics to function in the modern world. These findings were based on test results from a random sample of 20 twelfth-grade students who completed a 46-question multiple-choice test on basic economic concepts. The data set below shows the number of questions that each of the 20 students in the sample answered correctly.

12 16 18 17 18 33 41 44 38 35  
19 36 19 13 45 8 16 14 10 9

- (a) Display these data in a stemplot.

```

0 | 8, 9
1 | 0, 2, 3, 4, 6, 6, 7, 8, 8, 8, 9
2 |
3 | 3, 5, 6, 8
4 | 1, 3, 4
  
```

- (b) Use your stemplot from part (a) to describe the main features of this score distribution.

The lapse of scores in the 20's suggest that the distribution is bi-modal. In addition, the distribution seems skew right (although a graph would more clearly show this.)

- (c) Why would it be misleading to report only a measure of center for this score distribution?

This is because of the distributions bimodal nature. Clearly there are two distinct groups of students (possibly one that has taken a class in economics and one that hasn't) so reporting two measures of center would be more informative. If you were only looking at a measure of center and had no sd, you could easily guess that many student's scores fell in the 20's which would be false

Unauthorized copying or reuse of  
any part of this page is illegal.

GO ON TO THE NEXT PAGE.

## STATISTICS

## SECTION II

## Part A

## Questions 1-5

Spend about 65 minutes on this part of the exam.

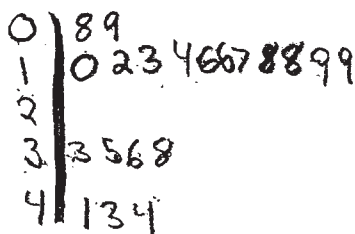
Percent of Section II grade—75

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

1. The Better Business Council of a large city has concluded that students in the city's schools are not learning enough about economics to function in the modern world. These findings were based on test results from a random sample of 20 twelfth-grade students who completed a 46-question multiple-choice test on basic economic concepts. The data set below shows the number of questions that each of the 20 students in the sample answered correctly.

12 16 18 17 18 33 41 44 38 35  
19 36 19 13 43 8 16 14 10 9

- (a) Display these data in a stemplot.



- (b) Use your stemplot from part (a) to describe the main features of this score distribution.

This score distribution is skewed to the right with a center at approximately 18 with scores ranging from 8-44, most of the scores tend to be in the teens. There are no scores in the 20's.

- (c) Why would it be misleading to report only a measure of center for this score distribution?

Because there are no scores within the middle values of the range (20's) and it is not normally distributed.

## STATISTICS

## SECTION II

## Part A

## Questions 1-5

Spend about 65 minutes on this part of the exam.

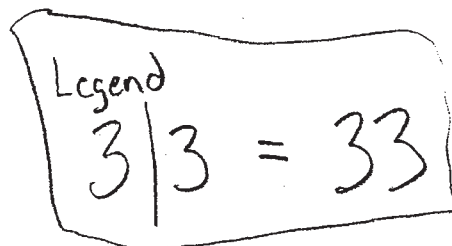
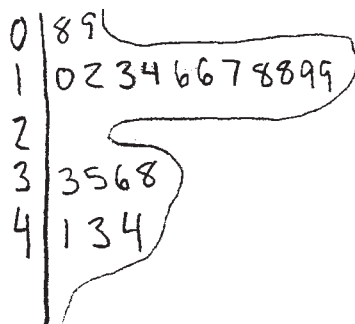
Percent of Section II grade—75

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

1. The Better Business Council of a large city has concluded that students in the city's schools are not learning enough about economics to function in the modern world. These findings were based on test results from a random sample of 20 twelfth-grade students who completed a 46-question multiple-choice test on basic economic concepts. The data set below shows the number of questions that each of the 20 students in the sample answered correctly.

~~12~~ ~~16~~ ~~18~~ ~~17~~ ~~18~~ ~~33~~ ~~41~~ ~~44~~ ~~38~~ ~~35~~  
~~19~~ ~~36~~ ~~19~~ ~~13~~ ~~43~~ ~~8~~ ~~16~~ ~~14~~ ~~10~~ ~~9~~

- (a) Display these data in a stemplot.



- (b) Use your stemplot from part (a) to describe the main features of this score distribution.

$$\sigma = 12.06 \quad \bar{x} = 22.95$$

$$\text{Range} = (8, 44) \quad Q_1 = 13.5$$

$$\text{Median} = 18$$

$$Q_3 = 35.5$$

$$IQR = [13.5, 35.5]$$

Large clump of # in 10's  
and little cluster in 30's  
+ 40's

- (c) Why would it be misleading to report only a measure of center for this score distribution?

The center is at about 18, however the larger numbers in the 30's and 40's raise the average considerably.

**AP<sup>®</sup> STATISTICS**  
**2007 SCORING COMMENTARY (Form B)**

**Question 1**

**Sample: 1A**  
**Score: 4**

This is a complete response that displays an appropriate stemplot and recognizes the main features of the data. The stem display presented in part (a) is correct except for a value of 18 that was incorrectly substituted for a value of 19. This was considered to be a minor error that did not detract from the overall score. The response to part (b) recognizes the gap between reports of 19 and 33 correctly answered questions on the test, and two distinct groups of scores are also noted in part (c). The discussion in the context of student scores is most evident in the response to part (c). The last sentence clearly explains how a single measure of center could mislead students to believe that there were exam scores in the gap where no scores were reported.

**Sample: 1B**  
**Score: 3**

This is a substantial response that displays an appropriate stemplot and recognizes the lack of tests with between 20 and 30 correctly answered questions, but it does not adequately explain why a single measure of center would be misleading. An appropriate stem display is presented in part (a). The response to part (b) recognizes the gap and describes its location by reporting that there are no scores in the 20s. By also giving the range of the test scores and noting that many scores are in the teens, the response recognizes two groups of scores without explicitly describing both groups. This discussion is presented in a test-score context. Although the response to part (c) refers to the lack of middle values and non-normality of the test scores, it fails to clearly explain why a single measure of center would be misleading.

**Sample: 1C**  
**Score: 2**

This is a developing response that displays an appropriate stemplot and recognizes two clumps of data, but it fails to explain why reporting only a single measure of center would be misleading. An appropriate stem display is presented in part (a). The response to part (b) reports a variety of summary statistics that are not useful in describing the two distinct groups feature of the recorded exam scores, but it does report the locations of two clumps of data in a way that implies recognition of a gap in the reported test scores. The response to part (c) fails to clearly explain why a single measure of center would be misleading. None of the discussion in parts (b) and (c) is clearly presented in the context of test scores or the number of questions answered correctly on a test.