AP Statistics Test
Describing Data

User Name: _____ **Instructor:** _____ **Date:** _____
(print clearly)

Directions:

- Neatly write your responses in the spaces provided. Use a blue or black pen. Don't write in the margins.
- Remember to complete the submission information on **every** page you turn in.

Free Response: 2 questions, 52 points, 30 minutes

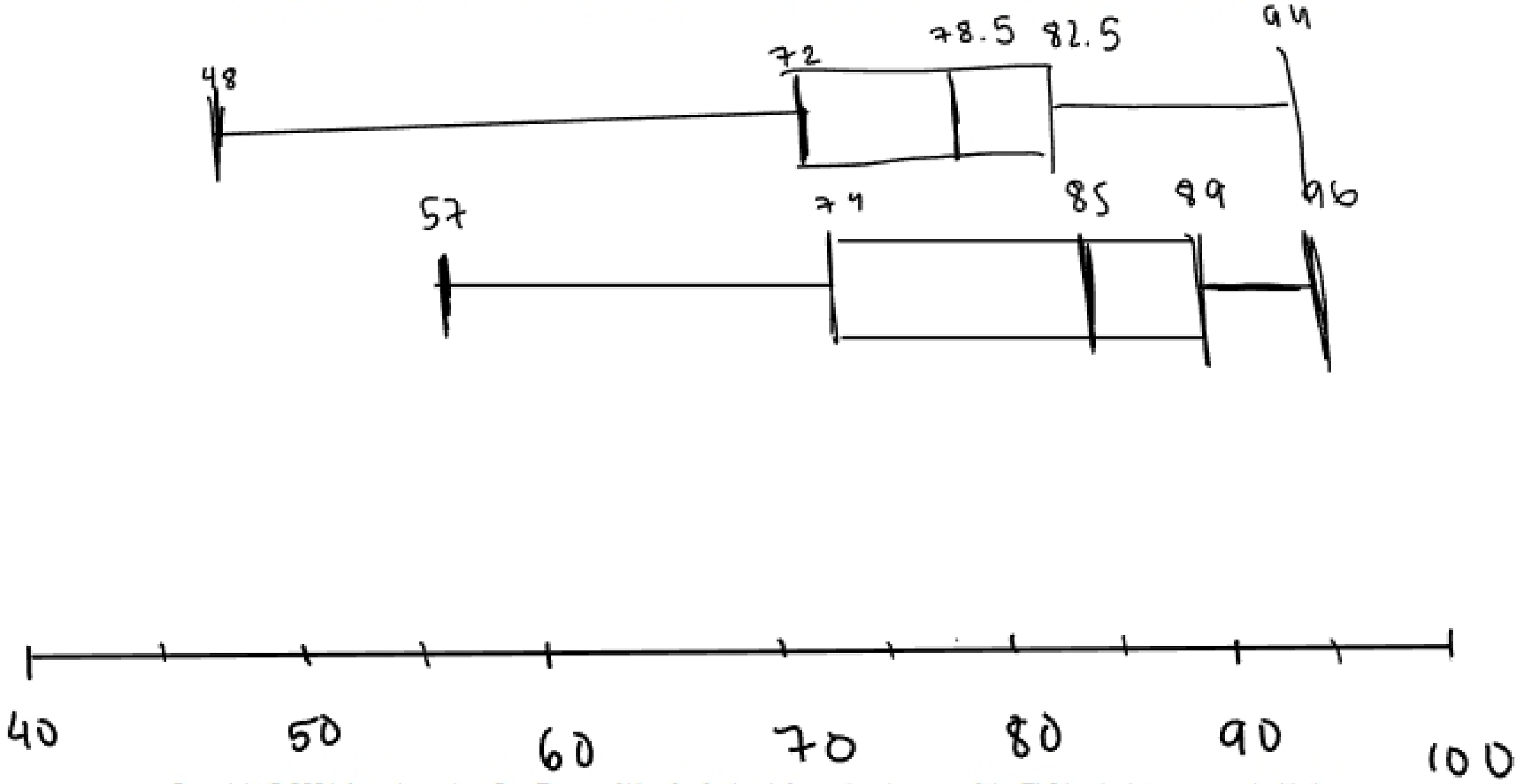
1. **Shown below are the summary statistics for test scores from two different AP Statistics classes — one taught first period and one taught third period.** no outliers

	Mean	Std. Dev.	n	Median	Q1	Q3	Min.	Max.
First-period test scores	75.7	11.93	20	78.5	72	82.5	48	94
Third-period test scores	81.8	9.6	29	85	74	89	57	96

outlier check: 1st: ← 56.25 & 98.25 → 2nd ← 51.5 & 111.5 →

NOTE: A student saw a paper that showed the two lowest grades for the first-period class, which were 48 and 68.

A. **Construct box plots for the two classes on the same horizontal axis with one box plot above the other box plot. Be sure to check for outliers.** (10 points)



User Name: _____ **Instructor:** _____ **Date:** _____
(print clearly)

B. Which class had the better overall performance? Justify your answer. (12 points)

Based off the created graphs, it can be seen that the second period class had the best test scores. There are low deviations in the second-period data, and there are higher means & medians in it as well. Since the standard deviation in the first period class is higher, it tells us that there is more scattered data from the mean. This shows how period 2 had higher consistent data.

C. The first-period class has a standard deviation of 11.93. Explain how this value summarizes variability in this class. (4 points)

The first-period class has a rather larger standard deviation of 11.93 (compared to second period of 9.6). This tells me that it is much more scattered and deviated from the mean. This means that first period had much more spread & variability in its test scores, depicting the variety of differing scores

AP Statistics Test
Describing Data

User Name: _____ **Instructor:** _____ **Date:** _____
(print clearly)

2. The stemplot below shows the weights of 30 male high school students from Small Rock High School.

9	6
10	2
11	5
12	1 3
13	3 5 6 6
14	0 5 7 7 8 8 9
15	1 3 4 6 6 8 9
16	0 2 2 3 4 6 7
17	

A. Using your knowledge of statistics, describe the distribution. Include statistics and discussion of shape, center, and spread. (20 points)

This data set has a tail that is extended towards the left, meaning the data is negatively skewed / left skewed. The best measure of central tendency is the median and spread is IQR. The 15th data point is 148 & the 16th data point is 149, so the median(center) is 148.5. The IQR (spread, is Q3 - Q1 = 159 - 136 = 23. This shows that the center is at 148.5 and the spread is 23. From the data, it can be concluded that maximum (more than 85%) of students weight range from 130 - 170 with center at approx 145. The data is biased because the data sets tail extends to the left

B. Would it be reasonable to use this information to generalize about the distribution of weights for the entire population of high school boys? Why or why not? (6 points)

no, because all the weights of only 30 male high school students from Small Rock High School is not normally distributed. This is only a small sample size from the high school, generalizing the distribution of weights for the entire population could be biased Biasness may occur in generalizing the distribution of weights for the entire population.

It is unreasonable to generalize that might