AP Statistics Ch 5 - Required Practice   Name: ___Solution___ Per: ____

A student studying the sleeping habits of seniors at his school asked 34 randomly-selected seniors how many hours of sleep they got the previous night. The data, rounded to the nearest half-hour, is given in the table below.

8	7.5	9	7.5	9	6	5	9	7.5	7	8	7
6.5	8.5	8	6.5	8.5	6	7	7.5	7	6	8.5	
7	8	7	7.5	7	6	7	8	7.5	6	7	

#1. Find the mean, standard deviation, and 5-number summary for these data.

1 Var-stats 4

$\bar{x} = 7.3$ hrs

$s = 0.977$ hrs

min	Q1	med	Q3	max
5	7	7.25	8	9

#2. Determine if there are any outliers (calculate both the lower and upper fences to show work).

$IQR = 8 - 7 = 1$

upper fence: $Q3 + 1.5(IQR) = 8 + 1.5(1) = 9.5$
no data above 9.5, so no high outliers,

lower fence: $Q1 - 1.5(IQR) = 7 - 1.5(1) = 5.5$
one value (5) is a low outlier, b/c it is below 5.5.

#3. Suppose 4 more values were added to the data, each exactly equal to the mean.
Would this have any impact on standard deviation? Explain, without using any calculations.

Yes, with more values closer to the mean, the average distance of data from the mean decreases (which is what standard deviation measures).
— or —
Yes: Although the numerator of the std.dev. fraction doesn't change, we are dividing by 4 more in the denominator, so std dev decreases.

#4. Create a set of five positive numbers (repeats allowed) that have a median of 10 and mean of 7.

1, 2, 10, 11, 11

- Start with a middle number at median (10)
- Mean = 7 suggests $\frac{x+x+x+x+x}{5} = 7$ so sum of numbers must be 35.
- Add pairs of numbers above and below median so numbers sum is 35.

Tempe and Alex wanted to know if the number of words students could recall from a list they studied was influenced by the kind of music they were listening to. They asked students to study a list of words for a fixed amount of time while listening to either classical music or rap. Then they counted how many words each student could recall from the list. Forty different students listened to each type of music. The results are shown in the boxplots below.
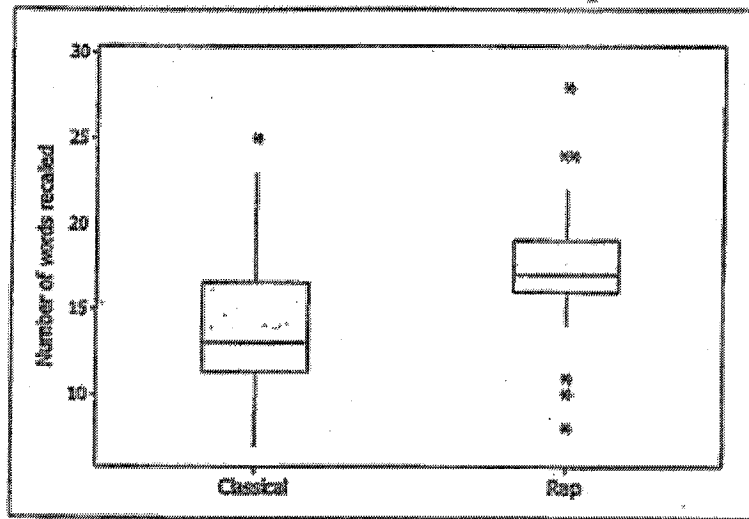
#5. Approximate the interquartile range for each set of data. Why is this the appropriate measure of spread to use for these two data sets?

Classical IQR = $16 - 11 = 5$

rap IQR = $18 - 16 = 2$

Both data sets contain outliers, so we should use IQR instead of standard deviation because IQR is

#6. Write two or three sentences comparing the word-recall performances of students listening to each type of music.

Both word recall distributions are slightly skewed right. Rap median of 17 vs 13 for classical suggests that students typically recall more words listening to rap than classical. There is more variation for classical (IQR 5) compared to rap (IQR 2). Classical has one high outlier, while rap distribution has 3 high and 3 low outliers (mainly due to the lower IQR).

1. A survey conducted in a college intro stats class during Autumn 2003 asked students about the number of credit hours they were taking that quarter. The number of credit hours for a random sample of 16 students is:

10	10	12	14	15	15	15	15
17	17	19	20	20	20	20	22

a. Sketch a histogram for these data.
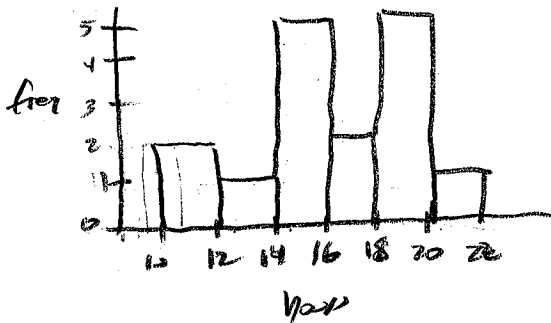
b. Find the mean and standard deviation for the number of credit hours.

$1-var \; stats$

$\bar{x} = 16.313 \; hr.$

$s = 3.7 \; hrs.$

c. Find the median and IQR for the number of credit hours.

| Q1	med	Q3
14.5	16	20

median = 16 hrs.

IQR = 20 - 14.5 = 5.5 hrs

d. Is it more appropriate to use the mean and standard deviation or the median and IQR to summarize these data? Explain.

Median and IQR because the data gaps mean data may not be unimodel

2. The five-number summary for midterm scores (number of points; the maximum possible score was 50 points) from an intro stats class is:

Min	Q1	Median	Q3	Max
16.5	32	39	43.5	48.5

a. Would you expect the mean midterm score of all students who took the midterm to be higher or lower than the median? Explain.
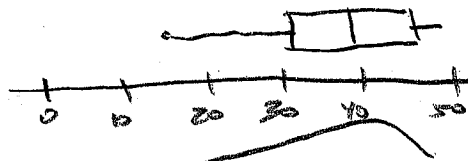
mean would be lower than median because the data is skewed left and skew pulls the mean away from the median toward the tail.

skewed left

b. Based on the five-number summary, are any of the midterm scores outliers? Explain.

IQR = 43.5 - 32 = 11.5

UF = 43.5 + 1.5(11.5) = 60.75

LF = 32 - 1.5(11.5) = 14.75

no high outliers

no low outliers

Suppose that the student who scored 48.5 on the midterm got her grade raised to 50. Indicate whether changing the midterm score for that student would make each of the following summary statistics increase, decrease, or stay about the same:

a. mean

increase

b. median

Same

c. range

increase

d. IQR

Same

e. standard deviation

increase

4. The side-by-side boxplots show the cumulative college GPAs for sophomores, juniors, and seniors taking an intro stats course in Autumn 2003.

a. Which class (sophomore, junior, or senior) had the lowest cumulative college GPA? What is the approximate value of that GPA?
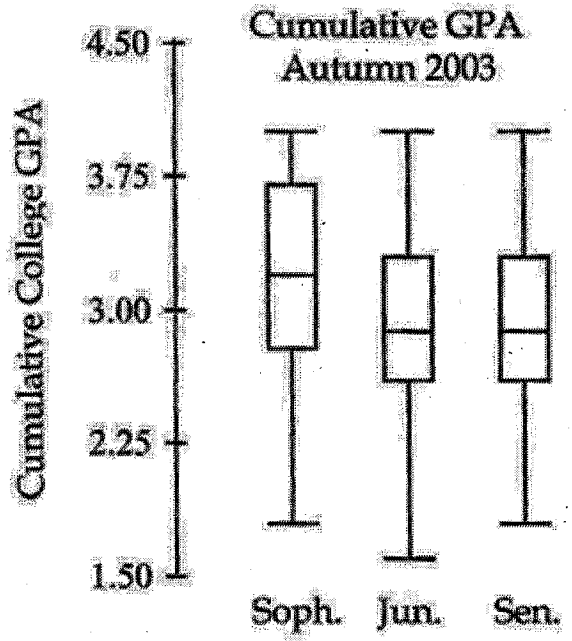
Juniors had a 1.6 GPA

b. Which class has the highest median GPA, and what is that GPA?

Sophomores had 3.2 median GPA.

c. Which class has the largest range for GPA, and what is it?

Juniors   (4.0 - 1.6 = 2.4 range)

d. Which class has the most symmetric set of GPAs? The most skewed set of GPAs?

Seniors had most symmetric GPA's.
Sophomores had most skewed (left)