Name: _____ Period: _____ Date: _____

AP STATISTICS
Test 1 Review_**Free Response**

1. The table shows whether students in an introductory statistics class like dogs and/or cats.

		Like Dogs		
		Yes	No	Total
Like Cats	Yes	194	21	215
	No	110	10	120
	Total	304	31	335

a. What is the marginal distribution (in %) of "liking dogs"? $\frac{304}{335}$ 90.7% like dogs & $\frac{31}{335}$ (9.3%) don't like dogs

b. What is the conditional distribution (in %) of "liking dogs" for students who like cats? $\frac{194}{215}$ 90.2% like dogs & $\frac{21}{215}$ 9.8% don't.

c. What kind of display(s) would you use to examine the association between "liking dogs" and "liking cats"? (Just name a graph.) segmented bar graph

d. Do "liking dogs" and "liking cats" appear to be independent? Give statistical evidence to support your conclusion.
They maybe independent as the overall percent of those who like dogs (90.7%) is very close to those who like dogs given they like cats 90.2%.
Can also do "Like cats"
The overall percent of those who like cats (64.2%/215/335) is almost similar to those who like cats given they like dogs (63.8%/194/304).

2. The *World Almanac and Book of Facts 2004* reported the percent of people not covered by health insurance in the 50 states and Washington, D.C., for the year 2002. Computer output gives these summaries for the percent of people not covered by health insurance:

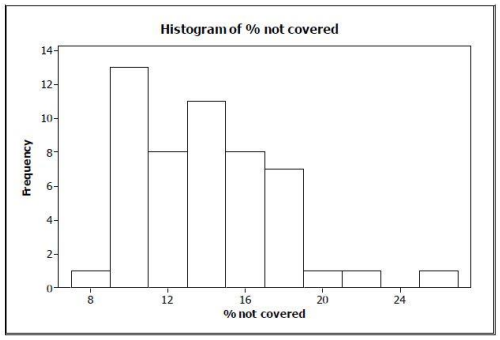
Min	Q1	Median	Q3	Max	Mean	SD
7.9	10.8	13.4	16.7	25.8	13.9	3.6

IQR = Q3 − Q1
= 16.7 − 10.8
= 5.9

10.8 − 1.5(5.9) = 1.95 ✓
16.7 + 1.5(5.9) = 25.55 ✓

a. Were any of the states outliers? Explain how you made your decision.
There is a high outlier as the maximum data point of 25.8 is higher than $Q_3 + 1.5$ IQR or 25.55%.
There is no low outlier because the minimum data point at 7.9% is well within $Q_1 - 1.5$ IQR or 1.95%.

b. A histogram of the data is as follows:



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

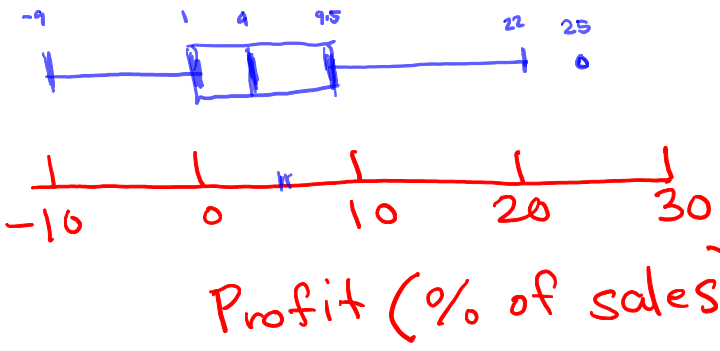
The distribution is skewed to the right, so it is appropriate to use the median and IQR to describe the data.

3. Here is a stem-and-leaf display showing profits as a percent of sales for 29 of the *Forbes* 500 largest U.S. corporations. The stems are split; each stem represents a span of 5%, from a loss of 9% to a profit of 25%.

-0 | 99 -0 | 1234 0 | 111123444 0 | 5555679 1 | 00113 1 | 2 | 2 2 | 5 Profits (% of sales) (-0|3 means a loss of 3%)	a) Find the 5-number summary. b) Draw a boxplot for these data. c) Find the mean and standard deviation. d) Describe the distribution of profits for these corporations.

a) The 5-number summary of the profits as percent of sales is:
-9, 1, 4, 9.5, 25

b)



c) The mean profit (% of sales) is 4.72%
• the standard deviation is 7.55%

d) The distribution of profit (% of sales) is unimodal and roughly symmetric. The center is around 4% (median). The middle 50% of the companies have profits between 1% to 9.5%. One company have high profit 22% while there is one high outlier with 25% profit.