9:56AM <u>Practice Problems for Exam 1</u>

1. 10, 6, 2, 7, 100

a) $\bar{x} = \frac{16+9+100}{5} = \frac{125}{5} = 25$

median: 2, 6, 7, 10, 100 median = 7

b) The median is better. The outlier, 100, distorts the mean.

c)

x	$x-\bar{x}$	$(x-\bar{x})^2$
2	$2-25=-23$	$(-23)^2 = 529$
6	$6-25=-19$	$(-19)^2 = 361$
7	$7-25=-18$	$(-18)^2 = 324$
10	$10-25=-15$	$(-15)^2 = 225$
100	$100-25=75$	$75^2 = 5625$

$7064 = \Sigma(x-\bar{x})^2$

$$s = \sqrt{\frac{7064}{5-1}} = 42.02$$

d) No. The standard deviation~~s~~ measures how spread out from the mean the dataset is. Adding k to each value doesn't change the spread of the data. Think of a dotplot, adding k to each data value will simply shift the dotplot to a new location, the hence the variation shape will remain the same. An ~~alg~~ algebraic argument would ~~be an~~ also work.

e) Yes, the lowest score is an outlier~~s~~ and increases the spread of the data about the mean. Hence, removing it will cause the st. dev. to <u>decrease</u>.

2. a) 250

b) $\approx 36/200$

c) $36 + 8 + 6 \approx 50$

2. d) $\bar{x} \pm 2s = 269 \pm 2(14.2) = 269 \pm 28.4$

$$(240.6,\ 297.4) \approx (240,\ 300)$$

e) $9+36+60+50+25+15 \doteq 195$ about $\frac{195}{200} \approx .975$

f) The data are approx bell-shaped so the Empirical Rule will give a better approx. of the % of data in the interval in part (d). Empirical Rules says about 95% of the data will be in the interval, which is close to the actual percent (97.5%).

3a)

	position	value
Q_1	$\frac{1}{4}(n+1)=\frac{1}{4}(28+1)=7.25$	$\frac{63+66}{2}=$ $\boxed{64.5=Q_1}$
median	$\frac{1}{2}(n+1)=\frac{1}{2}(29)=14.5$	$\frac{76+79}{2}=$ $\boxed{77.5=m}$
Q_3	$\frac{3}{4}(n+1)=\frac{3}{4}(29)=21.75$	$\frac{85+88}{2}=$ $\boxed{86.5=Q_3}$

b) lower fence $=Q_1-1.5(Q_3-Q_1)=64.5-1.5(86.5-64.5)$

$=64.5-1.5(22)$

$=31.5$

upper fence $=Q_3+1.5(Q_3-Q_1)$

$=86.5+1.5(22)$

$=119.5$

Since no scores are below 31.5 or above 119.5, there are <u>no</u> outliers.

6 a)

	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$
	6	18	$6-8.5=-2.5$	$18-55.75=-37.75$	$(-2.5)(-37.75)=94.375$
	7	47	$7-8.5=-1.5$	$47-55.75=-8.75$	13.125
	10	67	$10-8.5=1.5$	$67-55.75=11.25$	16.875
	11	91	$11-8.5=2.5$	$91-55.75=35.25$	88.125
Total	34	223	0	0	$212.5 = \Sigma(x-\bar{x})(y-\bar{y})$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{34}{4} = 8.5$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{223}{4} = 55.75$$

$$s_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n-1} = \frac{212.5}{4-1} = \frac{212.5}{3} = 70.8\bar{3}$$

b) positive since $s_{xy} > 0$

c) a Scatterplot

7 a) qualitative

b) quantitative - continuous

c) qualitative

d) qualitative

e) quantitative - discrete

f) quantitative - continuous