# 4.2 AP HW (variation)

**4.5** a

| stem : tens | leaf : ones |
|---|---|
| 32 | 55 |
| 33 | 49 |
| 34 | |
| 35 | 6699 |
| 36 | 34469 |
| 37 | 03345 |
| 38 | 9 |
| 39 | 2347 |
| 40 | 23 |
| 41 | |
| 42 | 4 |



The stem and leaf display suggests that the mean and median will be fairly close to each other. Most values are between 356 and 375 and there are approximately equal amounts of values larger or smaller than these central values.

b

$$\bar{x} = \frac{9638}{26} = 370.69$$

$$\text{median} = \frac{369 + 370}{2} = 369.5$$

c The largest value, 424, could be increased by any arbitrary amount without affecting the sample median. The median is insensitive to outliers. However, if the largest value was decreased below the sample median, 369.5, then the value of the median would change.

Interquartile range: 18

b The IQR for inpatient cost-to-charge ratio in Example 4.9 is 14. There is more variability in cost-to-charge ratios for outpatient services.

4.25

| $x$ | $x-\bar{x}$ | $(x-\bar{x})^2$ |
|---|---|---|
| 244 | 51.4286 | 2644.9009 |
| 191 | -1.5714 | 2.4693 |
| 160 | -32.5714 | 1060.896 |
| 187 | -5.5714 | 31.0405 |
| 180 | -12.5714 | 158.0401 |
| 176 | -16.5714 | 274.6113 |
| 174 | -18.5714 | 344.8969 |
| 205 | 12.4286 | 154.4701 |
| 211 | 18.4286 | 339.6133 |
| 183 | -9.5714 | 91.6117 |
| 211 | 18.4286 | 339.6133 |
| 180 | -12.5714 | 158.0401 |
| 194 | 1.4286 | 2.0409 |
| 200 | 7.4286 | 55.1841 |
| 2696 | .0004 | 5657.4285 |

$\bar{x} = 192.5714$

$$s^2 = \frac{5657.4285}{13} = 435.1868$$

$s = 20.8611$

$s^2$ could be interpreted as the mean squared deviation from the average leg power at a high workload. This is 435.1868. The standard deviation, s, could be interpreted as the typical amount by which leg power deviates from the average leg power. This is 20.8611.

4.26 a $\bar{x} = \frac{532}{11} = 48.364$

b

| Observation | Deviation | (Deviation)$^2$ |
|---|---|---|
| 62 | 13.636 | 185.9405 |
| 23 | −25.364 | 643.3325 |
| 27 | −21.364 | 456.4205 |
| 56 | 7.636 | 58.3085 |
| 52 | 3.636 | 13.2205 |
| 34 | −14.364 | 206.3245 |
| 42 | −6.364 | 40.5005 |
| 40 | −8.364 | 69.9565 |
| 68 | 19.636 | 385.5725 |
| 45 | −3.364 | 11.3165 |
| 83 | 34.636 | 1199.6525 |
| | −0.004 | 3270.5455 |

$s^2 = (3270.5455)/10 = 327.05$ and $s = \sqrt{327.05} = 18.08$. $s^2$ could be interpreted as the mean squared deviation from the average distance when detection takes place. This is 327.05 squared centimeters. The standard deviation s could be interpreted as the typical amount by which a distance deviates from the average distance when detection first takes place. This is 18.08 centimeters.

**4.27** Subtracting 10 from each data point yields:

| x | $(x - \bar{x})$ | $(x - \bar{x})^2$ |
|---|---|---|
| 52 | 13.636 | 185.9405 |
| 13 | −25.364 | 643.3325 |
| 17 | −21.364 | 456.4205 |
| 46 | 7.636 | 58.3085 |
| 42 | 3.636 | 13.2205 |
| 24 | −14.364 | 206.3245 |
| 32 | −6.364 | 40.5005 |
| 30 | −8.364 | 69.9565 |
| 58 | 19.636 | 385.5725 |
| 35 | −3.364 | 11.3165 |
| 73 | 34.636 | 1199.652 |

$$\bar{x} = \frac{422}{11} = 38.364$$

These deviations from the mean are identical to the deviations from the mean for the original data set. This would result in a variance for the new data set that is equal to the variance of the original data set. In general, adding the same number to each observation has no effect on the variance or standard deviation.

**4.28** Multiplying each data point by 10 yields:

| x | $(x - \bar{x})$ | $(x - \bar{x})^2$ |
|---|---|---|
| 620 | 136.36364 | 18595.04132 |
| 230 | −253.63636 | 64331.40496 |
| 270 | −213.63636 | 45640.49587 |
| 560 | 76.36364 | 5831.40496 |
| 520 | 36.36364 | 1322.31405 |
| 340 | −143.63636 | 20631.40496 |
| 420 | −63.63636 | 4049.58678 |
| 400 | −83.63636 | 6995.04132 |
| 680 | 196.36364 | 38558.67769 |
| 450 | −33.63636 | 1131.40496 |
| 830 | 346.36364 | 119967.76860 |
| | 0.00004 | 327054.54545 |

$\bar{x} = 483.63636$

$$s^2 = \frac{(327054.54545)}{10} = 32705.4545$$

$$s = \sqrt{32705.4545} = 180.846$$

The standard deviation for the new data set is 10 times larger than the standard deviation for the original data set.