SOLUTIONS TO QUESTION SET #1 (except 1-11)

1-1. Exercise 1.2: descriptive statistics in (a), (c) and (e); inferential statistics in (b) and (d).

Exercise 1.34: (a) observational; (b) one thousand people in British Columbia who were called and asked to respond to several questions; (c) can't say, as no randomization process is described.

1-2. (a) $\bar{x} = 6.04$, $s = 1.95$, $s^2 = 3.80$

(b) $\bar{x} = -6.04$, $s = 1.95$, $s^2 = 3.80$ (Note the values in (b) are -1 times the values in (a); the mean changes by -1 times, while all measures of variation remain the same.)

(c)) $\bar{x} = 0.32$, $s = 7.02$, $s^2 = 49.28$

(d) $\bar{x} = 26.04$, $s = 1.95$, $s^2 = 3.80$ (Values in (d) are those in (a) + 20, note mean increases by 20 but all measures of variation are the same as for (a).)

(e) $\bar{x} = 19.80$, $s = 0.00$, $s^2 = 0.00$

(f) $\bar{x} = -19.80$, $s = 0.00$, $s^2 = 0.00$

(g) $\bar{x} = 0.00$, $s = 22.86$, $s^2 = 522.72$

(h) $\bar{x} = 9.90$, $s = 19.80$, $s^2 = 392.04$

1-3. When 10 is added to each observation, the mean increases by 10, but there's no change for any of the measures of variation.

1-4. When each observation is tripled, the mean will also triple: $\bar{x} = 3(6.04) = 18.12$. The standard deviation will also triple: $s = 3(1.9501) = 5.85$. The variance will increase by $(3)^2$ times: $s^2 = (3)^2(3.803) = 34.23$.

1-5. The data sets in both (g) and (h) have the largest range, 39.6 (= 19.8 − (−19.8)).

1-6. (a) All summations are from 1 to n (the sample size). In the proof you need to recall that $\bar{x} = (\sum x_i)/n$, and for a constant c, $\sum c = nc$ and $\sum cx_i = c\sum x_i$.

$$
\begin{aligned}
\sum(x_i - \bar{x})^2 &= \sum(x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
&= \sum x_i^2 - 2\,\bar{x}\sum x_i + n\bar{x}^2 \\
&= \sum x_i^2 - 2[(\sum x_i)/n]\sum x_i + n[(\sum x_i)/n]^2 \\
&= \sum x_i^2 - 2(\sum x_i)^2/n + (\sum x_i)^2/n \\
&= \sum x_i^2 - (\sum x_i)^2/n \\
&= \sum x_i^2 - (n\bar{x})^2/n \\
&= \sum x_i^2 - n\bar{x}^2.
\end{aligned}
$$

(b) $s^2 = [\sum x_i^2 - n\bar{x}^2]/(n - 1) = [9402 - 26(18.6)^2]/25 = 16.2816$, $s = 4.035$.

1-7. (a) On a graph, $y = a + bx$ is a straight line. (The transformations $y = \log(x)$ and $y = x^2$ are examples of nonlinear transformations; graphically these are not straight lines.)

(b) For measures of central tendency ("average"), the same linear transformation applies to the relationship between the sample statistics. That is,

(i) $\bar{y} = a + b\bar{x}$

(ii) $\text{median}(y) = a + b[\text{median}(x)]$

(iii) $\text{mode}(y) = a + b[\text{mode}(x)]$.

Measures of variability are not affected by the additive constant a, only the multiplicative constant b. The sample statistics for variability are related by:

(iv) $\text{range}(y) = b[\text{range}(x)]$

(v) $s_y^2 = b^2 s_x^2$

(vi) $s_y = b s_x$

(c) Note that $z = (x - \bar{x})/s_x$ is a linear transformation of x with $a = -\bar{x}/s_x$ and $b = 1/s_x$. Thus the mean of the z values is

$$\bar{z} = -\bar{x}/s_x + (1/s_x)(\bar{x}) = 0$$

and the standard deviation is

$$s_z = (1/s_x)s_x = 1.$$

1-8. The total number of marks in the class is $\sum x_i = n\bar{x} = 540(65.4) = 35{,}316$. The section with 290 students and a mean of 67.7 will contribute 290(67.7) = 19,633 marks to this total, so the other section with 250 students contributes 35,316 − 19,633 = 15,683 marks. The mean for the section with 250 students is thus 15,683/250 = 62.7 (rounded to one decimal place).

1-9. To obtain an average purchase price of \$48.00 the total price of the 3000 shares must be (3000)(\$48.00) = \$144,000. The first 1000 shares purchased constitute 1000(\$46.50) = \$46,500 of this total, so the next 2000 shares must have a total price of \$144,000 – \$46,500 = \$97,500, or an average of \$97,500/2000 = \$48.75 per share.

1-10. For a linear transformation of a variable, what you do to the mean is the same as what you did to the variable. Constants you add don't affect the variance, but a multiplicative constant changes the variance by the product of the square of the constant. That is, for a linear transformation $y = a + bx$, then $\bar{y} = a + b\bar{x}$, $s_y^2 = b^2 s_x^2$, and $s_y = b s_x$. So we obtain:

In °C: $\bar{x}_C = -32(5/9) + (5/9)(\bar{x}_F) = -32(5/9) + (5/9)(98.8) = 37.11$ (37.1 rounded to one decimal place).

$s_C^2 = (5/9)^2 s_F^2 = (5/9)^2(3.23) = 0.997$ (1.00 rounded to two decimal places).

In °K: $\bar{x}_K = 273 + \bar{x}_C = 273 + 37.1 = 310.1$, $s_K^2 = s_C^2 = 1.00$.

1-11. Left for Assessment #2.