Q6

Let $\bar{x}_j$ = arithmetic mean of first j numbers

$\overline{x_{j+1}}$ = " " " " $j+1$ "

$$\overline{x_{j+1}} = \frac{1}{j+1}\sum_{i=1}^{j+1} x_i = \frac{1}{j+1}\sum_{i=1}^{j}\left(x_i + x_{j+1}\right)$$

$$= \frac{j}{j+1}\bar{x}_j + \frac{x_{j+1}}{j+1} = \overline{x_j} + \frac{x_{j+1} - \overline{x_j}}{j+1}$$

Thus to compute $\overline{x_{j+1}}$ you need to know only $\overline{x_j}$ and of course the new data value x_{j+1}. You do not need to loop over the entire dataset to find $\overline{x_{j+1}}$.

For the standard deviation, we have

$$j s_{j+1}^2 = \sum_{i=1}^{j}\left(x_i - \overline{x_{j+1}}\right)^2 + \left(x_{j+1} - \overline{x_{j+1}}\right)^2$$

$$= \sum_{i=1}^{j}\left(x_i - \overline{x_j} - \frac{x_{j+1} - \overline{x_j}}{j+1}\right)^2 + \left(x_{j+1} - \overline{x_{j+1}}\right)^2$$

$$= \sum_{i=1}^{j}\left\{\underbrace{(x_i - \overline{x_j})^2}_{(j-1)\, s_j^2} + \left(\frac{x_{j+1} - \overline{x_j}}{j+1}\right)^2 - 2(x_i - \overline{x_j})\left(\frac{x_{j+1} - \overline{x_j}}{j+1}\right)\right\}$$

$$+ \left(x_{j+1} - \overline{x_{j+1}}\right)^2$$

$$\therefore\ j\, s_{j+1}^2 = (j-1)\, s_j^2 + j\left(\frac{x_{j+1} - \overline{x_j}}{j+1}\right)^2$$

$$- 2j\left(\underbrace{\sum_{i=1}^{j}(x_i - \overline{x_j})}_{\to\, 0}\right)\left(\frac{x_{j+1} - \overline{x_j}}{j+1}\right) + \left(x_{j+1} - \overline{x_{j+1}}\right)^2$$

$$= (j-1)\, s_j^2 + j\left(\frac{x_{j+1} - \overline{x_j}}{j+1}\right)^2 + \left(x_{j+1} - \overline{x_{j+1}}\right)^2$$

$$= (j-1)\, s_j^2 + j\left(\overline{x_{j+1}} - \overline{x_j}\right)^2 + \left(x_{j+1} - \overline{x_{j+1}}\right)^2$$

$\hookrightarrow$ see formula for mean

$$\therefore\ s_{j+1}^2 = \left(1 - \frac{1}{j}\right) s_j^2 + \left(\overline{x_{j+1}} - \overline{x_j}\right)^2 + \frac{\left(x_{j+1} - \overline{x_{j+1}}\right)^2}{j}$$

This is the formula to update the standard deviation.