Question 4-1

[Answer] When a measurement of a particular explanatory variable is in thousands kgs, which should be in millions of kgs in fact, then it is the same that all the observations on the explanatory variable are multiplied by $\lambda(= 1000)$. When $\hat{\beta}$ in multiple regression are estimated, we can see that the corresponding regression coefficient is multiplied by $1/\lambda$, therefore there is no effect on the estimate $\hat{y}$. We can see this from the linear regression $y = \beta_1 + \beta_2 x + \epsilon$:

$$\beta_1 = \overline{y} - \frac{\overline{x}\sum_t (x_t - \overline{x})y_t}{\sum_t (x_t - \overline{x})^2} \tag{1}$$

$$\beta_2 = \frac{\sum_t (x_t - \overline{x})y_t}{\sum_t (x_t - \overline{x})^2} \tag{2}$$

If you replace x_t with λx_t, β_2 will be:

$$\beta_2 = \frac{\sum_t (\lambda x_t - \lambda\overline{x})y_t}{\sum_t (\lambda x_t - \lambda\overline{x})^2} = \frac{1}{\lambda}\frac{\sum_t (x_t - \overline{x})y_t}{\sum_t (x_t - \overline{x})^2} \tag{3}$$

Therefore, λ will be canceled out during estimation and we can get the estimate $\hat{y}$ without the influence from the wrong scale.

Question 4-2

[Answer] Likewise 4-1, support that we replace x_t with $x_t + \lambda$:

$$\beta_2 = \frac{\sum_t (x_t + \lambda - \overline{x} - \lambda)y_t}{\sum_t (x_t + \lambda - \overline{x} - \lambda)^2} = \frac{\sum_t (x_t - \overline{x})y_t}{\sum_t (x_t - \overline{x})^2} \tag{4}$$

$$\beta_1 = \overline{y} - \frac{(\overline{x} + \lambda)\sum_t (x_t + \lambda - \overline{x} - \lambda)y_t}{\sum_t (x_t + \lambda - \overline{x} - \lambda)^2} = \overline{y} - \frac{\overline{x}\sum_t (x_t - \overline{x})y_t}{\sum_t (x_t - \overline{x})^2} - \frac{\lambda\sum_t (x_t - \overline{x})y_t}{\sum_t (x_t - \overline{x})^2} = \overline{y} - \overline{x}\hat{\beta}_2 - \lambda\hat{\beta}_2 \tag{5}$$

When we replace x_t with $x_t + \lambda$, there is no change in the corresponding regression coefficient. However, the constant term in linear regression is changed like in the equation above. When we think about a regression in logarithmic form $Y = AP^{\alpha}L^{\beta}M^{\gamma}$, we can get the following linear regression form by putting logarithmic function in each side of the equation:

$$ln(Y) = ln(A) + \alpha ln(P) + \beta ln(L) + \gamma ln(M) \tag{6}$$

As an example, if we multiple λ in P for showing that unit is not important, the above equation will be:

$$ln(Y) = ln(A) + \alpha ln(\lambda P) + \beta ln(L) + \gamma ln(M) \tag{7}$$

$$ln(Y) = ln(A) + \alpha ln(\lambda) + \alpha ln(P) + \beta ln(L) + \gamma ln(M) \tag{8}$$

This is the same case where a constant $\alpha ln(\lambda)$ is added in a particular explanatory variable in multiple linear regression. There is no change in its corresponding coefficient (α) in linear regression as we can see from 4.

Question 5-1

[Answer] $RSS = \sum (y_i - \hat{y}_i)^2$. Therefore, I think that if cubic regression is used, then its RSS is much less than the linear regression model for the given 100 observations as the difference between observed point (y) and its estimate ($\hat{y}$) is smaller due to higher order equation. But I do not think that it means the linear regression is inferior to cubic regression. It is as we can have the true value from linear regression when an estimate ($\hat{y}$) is done from unobserved data (x). The problem supposes that the true relationship is linear, so the cubic regression has overfitting, which is harmful for regression to produce more errors.

Question 5-2

[Answer] The problem supposes that the true relationship between X and Y is not linear, but we don't know how far it is from linear. This means that after the researcher fitted her model she can find a tendency of monotonic increase in Y while X is increased except several data points around the curvatures of cubic equations. As quadratic regression (the second lowest order function except linear regression) cannot show such monotonic increase, the cubic regression seems the most appropriate regression with the lowest order with the minimum calculation cost. Its RSS may be also smaller than the linear regression.