Problem 1

1) $\hat{\beta}_1 = \dfrac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$\therefore Y = \hat{\beta}_0 + \hat{\beta}_1 X = 77.86 + 11.8X$

2) $H_0: \beta_1 = 0$; $H_A: \beta_1 \neq 0$

$SST = \sum(y_i-\bar{y})^2$; $SSR = \sum(\hat{y}_i-\bar{y})^2$; $SSE = \sum(y_i-\hat{y}_i)^2$

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ for $i=1,\dots,n$

$\hat{\sigma} = \sqrt{SSE/(n-2)}$

$s.e.(\hat{\beta}_1) = \hat{\sigma} / \sqrt{\sum(x_i-\bar{x})^2}$

$t = \dfrac{\hat{\beta}_1 - 0}{s.e.(\hat{\beta}_1)} \Longrightarrow$ p-value $= 0.00329 < \overbrace{.05}^{\text{default}} = \alpha$

so we reject null hypothesis and conclude there is a significant relationship

3) $R^2 = \dfrac{SSR}{SST} = 0.389$

4) $\hat{\beta}_1 \pm t_{(n-2,\ .05/2)} \cdot s.e.(\hat{\beta}_1) = [4.479, 19.122]$

5) $x_0 = 1.05$, $\hat{\mu}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$

$s.e.(\hat{\mu}_0) = \hat{\sigma}\sqrt{\dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$

$\hat{\mu}_0 \pm t_{(n-2,\ .05/2)} \cdot s.e.(\hat{\mu}_0) = [88.30, 92.20]$

6) $cor(y,x) = \dfrac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sqrt{\sum(y_i-\bar{y})^2 \sum(x_i-\bar{x})^2}} = 0.624 = r$

7) $t_1 = \dfrac{cor(Y,X)\sqrt{n-2}}{\sqrt{1-cor(Y,X)^2}} = 3.386$

$\therefore$ p-value $= 0.00329 < \alpha = .05$ by default, so we reject the null and conclude there is significant correlation

Problem 2

let $\bar{y} = \frac{1}{n}\sum y_i$ and $\hat{\beta}_1 = \sum_i c_i y_i$ for $c_i = \frac{x_i - \bar{x}}{S_{xx}}$

$\therefore \operatorname{cov}(\bar{y}, \hat{\beta}_1) = \operatorname{cov}\left(\sum \frac{y_i}{n}, \sum c_i y_i\right) = \sum \frac{c_i}{n} \cdot \operatorname{var}(y_i) + \sum_{i \neq j} \frac{c_j}{n} \operatorname{cov}(y_i, y_j)$

but $i \neq j \rightarrow \operatorname{cov}(y_i, y_j) = 0$ and $\operatorname{var}(y_i) = \sigma^2$

$\therefore \operatorname{cov}(\bar{y}, \hat{\beta}_1) = \frac{\sigma^2}{n}\sum c_i = \frac{\sigma^2}{n}\sum \frac{(x_i - \bar{x})}{S_{xx}} = 0$

Problem 3

data) let $a_{n\times 1} = (x_1 - \bar{x}, \dots, x_n - \bar{x})^T$ and $b_{n\times 1} = (y_1 - \bar{y}, \dots, y_n - \bar{y})^T$

$\therefore\ r = \frac{a^T b}{\|a\|_2 \cdot \|b\|_2} = \cos(\text{angle between } a, b) \in [-1, 1]$

SLR context) $SSE, SST, SSR \geq 0$ and $SST = SSE + SSR$ so $0 \leq \frac{SSR}{SST} = r^2 = 1 - \frac{SSE}{SST} \leq 1$

Problem 4

$$\begin{aligned}
\operatorname{var}(\hat{y}_h) &= \operatorname{var}(\hat{\beta}_0 + \hat{\beta}_1 x_h) = \operatorname{var}(\bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x_h) \\
&= \operatorname{var}(\bar{y} + \hat{\beta}_1 (x_h - \bar{x})) \\
&= \operatorname{var}(\bar{y}) + (x_h - \bar{x})^2 \operatorname{var}(\hat{\beta}_1) + 2(x_h - \bar{x}) \underbrace{\operatorname{cov}(\bar{y}, \hat{\beta}_1)}_{0 \text{ from problem 2}} \\
&= \frac{\sigma^2}{n} + \frac{\sigma^2 (x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2} = \sigma^2 \left[\frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]
\end{aligned}$$

Problem 5

a) $\operatorname{var}(y) = \frac{SST}{n-1} = 0.00301$, $\operatorname{cor}(x, y) = \sqrt{\frac{SSR}{SST}} = 0.630 = r$

and $r \geq 0$ since $\hat{\beta}_1 \geq 0$

b) $\hat{\beta}_0 + \hat{\beta}_1 x_0 = 0.499 = \hat{\mu}_0$

c) $s.e.(\hat{\mu}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$

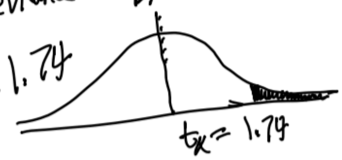
$\hat{\mu}_0 \pm t_{(n-2, .05/2)} \cdot s.e.(\hat{\mu}_0) = [0.465, 0.533]$

d) $s.e.(\hat{\beta}_1) = \hat{\sigma} / \sqrt{\operatorname{var}(x) \cdot (n-1)}$

$\hat{B}_1 \pm t_{(n-2,\ .05/2)} \cdot s.e.(\hat{B}_1) = [0.258,\ 1.054]$

e) $\hat{B}_1 = 0.656 < 1$ so we cannot reject the null hypothesis that $B_1 = 0$ since we have no evidence $B_1 > 1$

explicitly, $t = \dfrac{0.656 - 1}{s.e.(\hat{B}_1)} < 0 < 1.74$



so we cannot reject the null hypothesis

f) $r^2 = cor(Y, X)^2 = cor(X, Y)^2$