Assignment 4

Q1

1. $a = R^2$

$\bar{R}^2 = 0.694$

$0.694 = 1 - \frac{\hat{\sigma}^2(n-1)}{SST}$

$\hat{\sigma} = \sqrt{MSRes}$
$= \sqrt{16.2765}$
$\hat{\sigma} = 4.0337$

$0.694 - 1 = \frac{-\hat{\sigma}^2(n-1)}{SST}$

$-0.306 = \frac{-(4.0337)^2(15-1)}{SST}$

$-0.306 = \frac{-227.7947}{SST}$

$\frac{SST}{-227.7947} = \frac{1}{-0.306}$

$SST = 744.4271$

$\therefore R^2 = 1 - \frac{SSRes}{SST} = 1 - \frac{195.25264}{744.4271}$

$a = R^2 = 0.7377$

$b = Adj.\ R^2$

$\bar{R}^2 = 1 - \frac{\hat{\sigma}^2(n-1)}{SST}$

$R^2 = 1 - \frac{SSRes}{SST}$

$0.3486 = 1 - \frac{484.43832}{SST}$

$-0.6514 = \frac{-484.43832}{SST}$

$\frac{SST}{-484.43832} = \frac{1}{-0.6514}$

$SST = 744.4555$

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}^2(n-1)}{SST}$$

$$\hat{\sigma} = \sqrt{MSRes} = \sqrt{37.30295} = 6.1076$$

$$\bar{R}^2 = 1 - \frac{6.1076^2(14)}{744.455}$$

$$b = \bar{R}^2 = 0.29849$$

$c = C_{(k)}$

$$C_{13} = \frac{SSRes}{\hat{\sigma}^2_p} - (n-2K)$$

$K = p+1$

$$C_{13} = \frac{20.36709}{1.94213} - (15-2\times5)$$

$$= \frac{20.36709}{1.94213} - 5$$

$$c = 5.4870$$

d. MSRes

$$MSRes = \frac{SST - SSRes}{n-k}$$

$$R^2 = 1 - \frac{SSRes}{SST} = 1 - \frac{19.63927}{SST} = 0.9736$$

$$0.9736 - 1 = -\frac{19.63927}{SST}$$

$$\frac{SST}{-19.63927} = \frac{1}{-0.0264}$$

$$SST = 743.9117$$

$$MSRes = \frac{743.9117 - 19.63927}{15-5}$$

$$d = MSRes = 72.4272$$

e) $C = SSRes$

$$C = \frac{SSRes}{\sigma^2 p} - (n - 2u)$$

$$1.5851 = \frac{SSRes}{1.94213} - 5$$

$$1.5851 + 5 = \frac{SSRes}{1.94213}$$

$$SSRes = 6.5851 \times 1.94213$$

$$SSRes = 12.7891$$

$SST = SSR + SSRes$

2. SST:

$$R^2_{is} = 1 - \frac{SSRes}{SST}$$

$$0.9734 = 1 - \frac{19.42134}{SST}$$

$$0.9734 - 1 = -\frac{19.42134}{SST}$$

$$-0.0266 = -\frac{19.42134}{SST}$$

$$\frac{SST}{-19.42134} = \frac{1}{-0.0266}$$

$$SST = 744.1126$$

$n - p - 1 = 15 - 4 - 1$

∴ The total variation of the model, $SST = 744.1126$.

3.

Source	DF	Sum of square	Mean square	F-value
model	4	724.6914	181.1729	93.2854
Error	10	19.42134	1.942134	
Total	14	744.1126	53.1509	

4. The 5th and 6th models are the best since they have the largest adjusted R^2 values of 0.9678. This means that both models explain a larger percentage of the variation in the independent variable.

5. The 5th model is the best model based on $\hat{\sigma}^2$ since it is the minimum MSRes.

6. Reduced model: $y = \beta_0 + \beta_2 X_2$

$\alpha = 0.05$

$H_0: \beta_1 = \beta_3 = \beta_4 = 0$
H_a: At least one of $\beta_1, \beta_3, \beta_4$ does not equal zero

T.S. $F(X_1, X_3, X_4 \mid X_2) = \dfrac{(R^2_{full} - R^2_{reduced})/p-q}{(1 - R^2_{full})/n-k} = \dfrac{0.7381 - 0.6944/3}{(1 - 0.7381)/15 - 5}$

$TS = -2.9439$

Rejection point: $F_\alpha(p-q, n-k)$
$= F_{0.05}(3, 10)$
$= 3.708265$

$F(\text{model 1}) < F_\alpha(3, 10)$ therefore we do not reject H_0. Thus, it is not the case that at least one of $\beta_1, \beta_3, \beta_4$ does not equal zero, and the test is not significant.