Solution to Tutorial 8

1. (a)

   Hard hat: $E(Y) = \beta_0 + \beta_2 + \beta_1 X_1$

   Bump hat: $E(Y) = \beta_0 + \beta_3 + \beta_1 X_1$

   None: $E(Y) = \beta_0 + \beta_1 X_1$

   (b)

   (1) $H_0 : \beta_3 = 0; \quad H_a : \beta_3 \neq 0;$

   (1) $H_0 : \beta_3 = \beta_2; \quad H_a : \beta_3 \neq \beta_2;$

2. (1) β_3 means the difference in the intercepts between M2 and M4

   (2) $\beta_4 - \beta_3$ is the difference in the intercepts between M2 and M3

   (3) when all the other factors are fixed, the expected increment of Y as X_1 increase by one unit.

   (4) $\beta_7 = 0$ means there is no difference on the effect of X_1 on the response for M3 and M4.

   (5) $\beta_5 - \beta_6$ is difference on the effect of X_1 on the response for M1 and M2.

3. See **R-code**

4. See **R-code**

5. a.

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ \vdots & \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ \vdots & \\ 1 & 0 \end{pmatrix} \quad \text{thus} \quad X'X = \begin{pmatrix} n & n_1 \\ n_1 & n_1 \end{pmatrix}$$

We have

$$(X'X)^{-1} = \begin{pmatrix} \frac{1}{n-n_1} & \frac{-1}{n-n_1} \\ \frac{-1}{n-n_1} & \frac{n}{(n-n_1)n_1} \end{pmatrix}$$

b.

$$X'Y = \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \\ \sum_{i=1}^{n_1} Y_i \end{pmatrix}$$

Thus

$$b = \begin{pmatrix} \bar{Y}_d \\ \\ \frac{n}{n-n_1}(-\bar{Y} + \bar{Y}_c) \end{pmatrix}$$

where $\bar{Y}_c, \bar{Y}_d$ is means for incorporated firm and non-incorporated firm respectively. We have

$$\hat{Y} = \bar{Y}_d \qquad \text{if non-incorporated firm}$$

and

$$\hat{Y} = \bar{Y}_c \qquad \text{if incorporated firm}$$

c.

$$SSE = \sum_{i=1}^{n_1}(Y_i - \bar{Y}_c)^2 + \sum_{i=n_1+1}^{n}(Y_i - \bar{Y}_d)^2$$

and

$$SSR = n_1(\bar{Y}_c - \bar{Y})^2 + (n - n_1)(\bar{Y}_d - \bar{Y})^2$$

6. a. The fitted model is

Y	$=$	-207.5	$+$	$0.0005515X_1$	$+$	$0.1070X_2$	$+$	$149.0D_1$
(SE)		(70.28)		(0.0002835)		(0.01325)		(86.83)
					$+$	$145.5D_2$	$+$	$191.2D_3$
						(85.15)		(80.03)

SSE = 139093455,

b. The CIs for β_3 and β_4 are respectively

$$149.0 \pm 1.64 * 86.83 = [6.5988, 291.4012]$$

and

$$145.5 \pm 1.64 * 85.15 = [5.8540, 285.1460]$$

They have overlap (and include the other in the CI), thus, there is no difference in the regional effect

Another approach is by considering to models

$$Full: \qquad Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 D_1 + \beta_4 D_4 + +\beta_5 D_5 + \varepsilon$$

and

$$Reduced: \qquad Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (D_1 + D_4) + \beta_5 D_5 + \varepsilon$$

c. H_0 : there is no geographic effects The reduced model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

For teh reduced model $SSE = 140967081$. we have

$F = (140967081 - 139093455)/3/(139093455/434) = 1.948699 < F(0.99, 3, 434) = 2.62$

There is no geographic effects

See **R-code**