Example in multiple linear regression

The percent survival of a certain type of animal semen after storage was measured at various combinations of concentrations of three materials used to increase chance of survival. The data are as follows:

y (% survival)	x_1 (weight %)	x_2 (weight %)	x_3 (weight %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Summary statistics:

$$\sum_{i=1}^{13} y_i = 377.5 \qquad \sum_{i=1}^{13} y_i^2 = 11,400.15 \qquad \sum_{i=1}^{13} x_{i1} = 59.43$$

$$\sum_{i=1}^{13} x_{i2} = 81.82 \qquad \sum_{i=1}^{13} x_{i3} = 115.40 \qquad \sum_{i=1}^{13} x_{i1}^2 = 394.7255$$

$$\sum_{i=1}^{13} x_{i2}^2 = 576.7264 \qquad \sum_{i=1}^{13} x_{i3}^2 = 1035.96 \qquad \sum_{i=1}^{13} x_{i1}y_i = 1877.567$$

$$\sum_{i=1}^{13} x_{i2}y_i = 2246.661 \qquad \sum_{i=1}^{13} x_{i3}y_i = 3337.78 \qquad \sum_{i=1}^{13} x_{i1}x_{i2} = 360.6621$$

$$\sum_{i=1}^{13} x_{i1}x_{i3} = 522.078 \qquad \sum_{i=1}^{13} x_{i2}x_{i3} = 728.31 \qquad n = 13$$

$$\begin{pmatrix} 13 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.078 \\ 81.82 & 360.6621 & 576.7264 & 728.31 \\ 115.40 & 522.078 & 728.31 & 1035.96 \end{pmatrix}^{-1} = \begin{pmatrix} 8.0648 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0866 \end{pmatrix}$$

Or

$$(\underset{\sim}{X}_c^T \underset{\sim}{X}_c)^{-1} = \begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 123.039 & -13.3812 & -5.4775 \\ 0 & -13.3812 & 61.7639 & 2.0002 \\ 0 & -5.4775 & 2.0002 & 11.5631 \end{pmatrix}^{-1} = \begin{pmatrix} 0.07692 & 0 & 0 & 0 \\ 0 & 0.0085 & 0.0017 & 0.0037 \\ 0 & 0.0017 & 0.0166 & -0.0021 \\ 0 & 0.0037 & -0.0021 & 0.0886 \end{pmatrix}$$

(a) Estimate the multiple linear regression model for the given data.

(b) Estimate σ^2.

(c) Test the hypothesis that $\beta_2 = -2.5$ at the 0.05 level of significance against the alternative that $\beta_2 > -2.5$.

(d) $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

(e) Test $\beta_3 = 0$ using the increase in the regression sum of squares.

(f) Construct 95% confidence limits for the mean response $\mu_{Y|x}$ when $x_1 = 3$, $x_2 = 8$ and $x_3 = 9$.

(g) Construct a 95% confidence interval for the predicted response when $x_1 = 3$, $x_2 = 8$ and $x_3 = 9$.

(h) Calculate the coefficient of multiple determination.