12.12 $n=14$. $\sum x_i = 517$ $\sum y_i = 346$. $\sum x_i y_i = 25825$

$\sum x_i^2 = 39095$ $\sum y_i^2 = 17454$

a) $$\hat{\beta}_1 = \frac{\sum x_i y_i - \sum x_i \cdot \sum y_i / n}{\sum x_i^2 - (\sum x_i)^2/n} = \frac{25825 - 517 \cdot 346/14}{39095 - 517^2/14} = 0.6523$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \frac{346}{14} - 0.6523 \cdot \frac{517}{14} = 0.6261$$

$$\hat{y} = 0.6261 + 0.6523x$$

b) $x = 35$. $\hat{y} = 0.6261 + 0.6523 \cdot 35 = 23.4566$

residual $= y - \hat{y} = 21 - 23.4566 = -2.4566$

c) $$SSE = \sum(y_i - \hat{y}_i)^2 = \sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i$$

$$= 17454 - 0.6261 \cdot 346 - 0.6523 \cdot 25825$$

$$= 391.9610$$

$$\sigma = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{391.9610}{12}} = 5.7152$$

d) $$SST = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 17454 - \frac{346^2}{14} = 8902.8571$$

$$r^2 = 1 - \frac{SSE}{SST} = 1 - \frac{391.9610}{8902.8571} = 0.9560.$$

Thus 95.60% of observed variation in removal can be explained by the regression model.

e) $n=12$. $\sum x_i = 272$ $\sum y_i = 181$. $\sum x_i y_i = 5320$.

$\sum x_i^2 = 8322$ $\sum y_i^2 = 3729$

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \sum x_i \cdot \sum y_i / n}{\sum x_i^2 - (\sum x_i)^2 / n} = \frac{5320 - 272 \cdot 181/12}{8322 - 272^2/12} = 0.5645$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \frac{181}{12} - 0.5645 \cdot \frac{272}{12} = 2.288$$

$$\hat{y} = 2.288 + 0.5645x$$

$SSE = 311.79$

$$SST = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 3729 - \frac{181^2}{12} = 998.9167$$

$$r^2 = 1 - \frac{SSE}{SST} = 1 - \frac{311.79}{998.9167} = \underline{0.6879}.$$

The new regression line has a smaller slope, intercepts y-axis at a larger value. r^2 decreases to 68.79%. from 95%.

12.19 $n = 14$, $\Sigma x_i = 3300$ $\Sigma y_i = 5010$ $\Sigma x_i y_i = 1413500$

$\Sigma x_i^2 = 913750$ $\Sigma y_i^2 = 2207100$

a)

$$\hat{\beta}_1 = \frac{\Sigma x_i y_i - \Sigma x_i \cdot \Sigma y_i / n}{\Sigma x_i^2 - (\Sigma x_i)^2 / n} = 1.7114$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \frac{5010}{14} - 1.7114 \cdot \frac{3300}{14} = -45.5519$$

$$\underline{\hat{y} = -45.5519 + 1.7114x}$$

b) $x = 225$. $\mu_{225} = -45.5519 + 1.7114 \cdot 225 = \underline{339.5133}$

c) $1.7114 \cdot (-50) = \underline{-85.57}$

d) No, because the range of x data used to calculate the regression model is from 100 to 400. 500 is outside of this range. Extrapolation will not be accurate.

12.38 $H_0: \beta_1 = 0$, $t = \dfrac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$ model utility test

a) $H_a: \beta_1 \neq 0$

$\hat{\beta}_1 = 1.7114.$

$S_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 913750 - (3300)^2/14 = 135892.86$

$SSE = \sum y_i^2 - \hat{\beta}_0 \cdot \sum y_i - \hat{\beta}_1 \sum x_i y_i = 2207100 - (-45.5519) \cdot 5010 - 1.7114 \cdot 1413500$

$= 16251$

$s = \sqrt{\dfrac{SSE}{n-2}} = \sqrt{\dfrac{16251}{12}} = 36.8003.$

$s_{\hat{\beta}_1} = \dfrac{s}{\sqrt{S_{xx}}} = \dfrac{36.8003}{\sqrt{135892.85}} = 0.0998$

$t = \dfrac{1.7114}{0.0998} = 17.1438$

P-value $= 2 \cdot P(T > 17.1438) \approx 0.$

At $\alpha = 0.01$. $t_{0.005,12} = 3.0545$

$17.1438 > 3.0545$, Reject H_0.

Thus, there is a useful relationship between the two rates.

b). CI for 99% confidence level is.

$\hat{\beta}_1 \pm t_{\alpha/2, n-2} \cdot s_{\hat{\beta}_1} = 1.7114 \pm 3.0545 \cdot 0.0998$

$= (1.4065, 2.0164)$

With a 10 MBtu/hr ft^2 increase in liberation rate,

the interval is $(14.065, 20.164)$.

12.52 $n=9$. $\sum x_i = 24$ $\sum y_i = 312.5$ $\sum x_i y_i = 902.25$

$\sum x_i^2 = 70.50$ $\sum y_i^2 = 11626.75$

$\hat{\beta}_0 = 6.448718$, $\hat{\beta}_1 = 10.602564$.

a). $H_0: \beta_1 = 0$

$H_a: \beta_1 \neq 0$

$S_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 70.50 - 24^2/9 = 6.5$

$SSE = \sum y_i^2 - \hat{\beta}_0 \cdot \sum y_i - \hat{\beta}_1 \sum x_i y_i = 11626.75 - 6.448718 \cdot 312.5 - 10.602564 \cdot 902.25$

$= 45.3623$

$s = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{45.3623}{7}} = 2.5456$.

$s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}} = \frac{2.5456}{\sqrt{6.5}} = 0.9985$.

$t = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}} = \frac{10.602564}{0.9985} = 10.6186$.

for $\alpha = 0.01$. $t_{0.005,7} = 3.4995$, $10.6186 > 3.4995$. Reject H_0

Thus the relationship is useful specified by the regression model.

b). CI of $\hat{\beta}_1$ for $\alpha = 0.05$.

$\hat{\beta}_1 \pm t_{\alpha/2, n-2} \cdot s_{\hat{\beta}_1}$ $t_{0.025,7} = 2.3646$

$= 10.602564 \pm 2.3646 \cdot 0.9985$

$= (8.2415, 12.9636)$

c) $s_Y = s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}} = 2.5456 \cdot \sqrt{\frac{1}{9} + \frac{(3 - 24/9)^2}{6.5}} = 0.9115$

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x^* = 6.448718 + 10.602564 \cdot 3 = 38.2564$

CI of $\mu_{Y \cdot 3}$ at $\alpha = 0.05$ is

$38.2564 \pm 2.3646 \cdot 0.9115 = (36.1011, 40.4117)$

Sin $x^* - \bar{x}$ is small, it's precisely estimated.

d) PI of $\mu_{Y \cdot 3}$ at $\alpha = 0.05$

$$\hat{y} \pm t_{\alpha/2, n-2} \sqrt{s^2 + s_{\hat{Y}}^2} = 38.2564 \pm 2.3646 \cdot \sqrt{2.5456^2 + 0.9115^2}$$

$$= (31.8629, 44.6500)$$

It is likely to be accurate

e) It would be smaller, as 2.5 is closer to $\bar{x}$ than 3.

f) No, as 6 is not in the range of data use for calculating the regression model.

12.59 (-e). $n=18$, $\sum x_i = 1980$, $\sum y_i = 47.92$ $\sum x_i y_i = 5530.92$

$\sum x_i^2 = 251970$ $\sum y_i^2 = 130.6074$

a). $S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 251970 - \frac{1980^2}{18} = 40720$

$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 130.6074 - \frac{47.92^2}{18} = 3.0337$

$S_{xy} = \sum x_i y_i - \frac{\sum x_i \cdot \sum y_i}{n} = 5530.92 - \frac{1980 \cdot 47.92}{18} = 339.5867$

$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{339.5867}{\sqrt{40720 \cdot 3.0337}} = \underline{0.9662}$

b) $r>0$, thus x & y has a strong positive linear relationship.

~~The~~ sample with a bigger shear force tends to have a larger percent fiber dry weight.

c) r is independent of units ~~of~~ x & y.

d) $r^2 = 0.9662^2 = \underline{0.9335}$