Stat 481
HW 4

1 (a) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$$\hat{\beta}_1 x = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} =$$

$$= \frac{1156.31}{579.71} = 1.99$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 36.63 - 1.99 \cdot 17.86 = 1.09$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{250}{14} = 17.86$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{512.9}{14} = 36.63$$

$$\hat{y} = 1.09 - 1.99x$$

(b) (i) If the baking temperature increases by 1 unit, the percent yield is expected to increase by the slope of the regression line, which is 1.99

ii' If the baking temperature increases by 10 units, the percent yield is expected to increase by 10 times the slope of the regression line, which is 19.9.

2 a $x = 119$ & $y = 2.1$

for the first observation:

$\hat{y} = -9.33 + 0.097 \cdot 119 = \boxed{2.213}$

↓

the fitted value

Residual:

$\hat{\epsilon}_i = y_1 - \hat{y}_1 = 2.1 - 2.213 = \boxed{-0.113}$

for second observation:

Fitted value

$\hat{y} = -9.33 + 0.097 \cdot 130 = \boxed{3.28}$

Residual:

$\hat{\epsilon}_1 = 3.2 - 3.28 = \boxed{-0.08}$

(b) We can find the rest of the residuals as we did previosly and then use Shapiro-Wilk Test:

```
> residuals <- c(-0.113, -0.08, 0.193, -0.11, -0.368, 0.329, -0.092, -0.095, -0.074, 0.414)
> shapiro.test(residuals)

	Shapiro-Wilk normality test

data:  residuals
W = 0.86532, p-value = 0.08812
```

Yes, residuals are normally distributed, because p-value is 0.08812 which is > than $\alpha = 0.05 \Rightarrow$ do not reject H_0.

(c) $\beta_1 \pm t(\alpha/2, n-2) \cdot SE(\beta_1)$

$$\beta_1 = \frac{S_{xy}}{S_{xx}} = \frac{26.57}{273.6} = 0.097$$

$$SE(\beta_1) = \sqrt{3.081 - 0.097 \cdot 26.54 / (8 / 273.6)} =$$

$n-2 =$
$= 10-2 = 8$

$= 0.015$

$t(\alpha/2, n-2) = t(0.025, 8) = 2.306$
by t-table

β_1 CI: $0.097 \pm 2.306 \cdot 0.015$

0.097 ± 0.035

$(0.062, 0.132)$

(3) (a) the x-values are
10, 10, 12, 15, 14.

(b) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$\hat{y} = -0.25 + 0.25$
because we have $\hat{\beta} = \begin{bmatrix} -0.25 \\ 0.25 \end{bmatrix}$

the first one is $\hat{\beta}_0$ and
second one is $\hat{\beta}_1$

(c) $var(\hat{\beta}_0) = 9 * 7.356 = \boxed{66.204}$
$var(\hat{\beta}_1) = 9 * 0.048 = \boxed{0.432}$

$Cov(\hat{\beta}_0, \hat{\beta}_1) = -9 * 0.587 = \boxed{-5.283}$

(4) (a) $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \epsilon_i$

$y_i = \beta_0 + \sum c_i y_i x_i + c_i$

$\bar{y} = \beta_0 \bar{x} + \sum c_i y_i x_i + \bar{\epsilon}$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$\hat{\beta}_0 = \bar{y} - \sum c_i y_i \bar{x}$

$\hat{\beta}_0 = \bar{y} - \sum c_i y_i \frac{1}{n} \sum x_i$

$\hat{\beta} = \bar{y} - (\sum (\frac{1}{n} - \bar{x} c_i) \sum y_i / s_{xx})$

$\hat{\beta}_0 = \sum k_i y_i$

where $k_i = \frac{1}{n} - \bar{x} c_i$

(b) $E(\hat{\beta}_0) = \Sigma(\sum k_i y_i)$

$E(\hat{\beta}_0) = \sum k_i E y_i$

$E(y_i) = \beta_0 + \beta_1 \sum k_i x_i$

$\sum k_i = 0$ & $\sum k_i x_i = 1$;

$E(\hat{\beta}_0) = \beta_0$