A multiple regression analysis yields the following partial results.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>4</td>
<td>750</td>
</tr>
<tr>
<td>Residual Error</td>
<td>35</td>
<td>500</td>
</tr>
</tbody>
</table>

(a) What is the total sample size? 40

(b) How many independent variables are being considered? 4

(c) Compute and interpret the coefficient of determination.

\[ R^2 = \frac{SSR}{SST} = \frac{750}{750 + 500} = 0.6 \]

(d) Estimate the error variance.

\[ \hat{\sigma}^2 = \frac{SSE}{n-k-1} = \frac{500}{35} = 14.2857 \]

(e) Test the hypothesis that at least one of the regression coefficients is not equal to zero. Let \( \alpha = .05 \).

\[ H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \]
\[ H_1 : \text{At least one variable is significant} \]

\[ F = \frac{SSR/k}{SSE/(n-k-1)} = \frac{750/4}{500/(35)} = 13.125 \]

Reject \( H_0 \) if \( F_0 > F_{\alpha,1,35} \) where \( F_{0.05,4,35} \approx 2.65 \)

We reject \( H_0 \) and conclude that at least one variable is significant.
The following sample observations have been obtained by a chemical engineer investigating the relationship between the weight of final product and the volume of raw material:

<table>
<thead>
<tr>
<th>Product</th>
<th>Volume</th>
<th>Weight</th>
<th>(Volume)²</th>
<th>(Weight)²</th>
<th>(Volume)*(Weight)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>68</td>
<td>196</td>
<td>4624</td>
<td>952</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>105</td>
<td>529</td>
<td>11025</td>
<td>2415</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>40</td>
<td>81</td>
<td>1600</td>
<td>360</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>79</td>
<td>289</td>
<td>6241</td>
<td>1343</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>81</td>
<td>100</td>
<td>6561</td>
<td>810</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>95</td>
<td>484</td>
<td>2090</td>
<td>1246.4</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>31</td>
<td>25</td>
<td>961</td>
<td>155</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>72</td>
<td>144</td>
<td>5184</td>
<td>864</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>45</td>
<td>36</td>
<td>2025</td>
<td>270</td>
</tr>
<tr>
<td>10</td>
<td>16</td>
<td>93</td>
<td>256</td>
<td>8649</td>
<td>1488</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td><strong>134</strong></td>
<td><strong>709</strong></td>
<td><strong>2140</strong></td>
<td><strong>55895</strong></td>
<td><strong>10747</strong></td>
</tr>
</tbody>
</table>

\( Y \rightarrow \text{Weight of the final product} \)

\( X \rightarrow \text{Volume of the raw material} \)

\[ \sum X = 134, \sum Y = 709, \sum X^2 = 2140, \sum Y^2 = 55895 \text{ and } \sum XY = 10747 \]

(a) Calculate the least squares estimates of the simple linear regression equation for predicting weight of the final product based on volume of raw material.

\[
S_{xx} = \sum X^2 - \left( \frac{\sum X}{n} \right)^2 = 2140 - \left( \frac{134}{10} \right)^2 = 344.4
\]

\[
S_{yy} = \sum Y^2 - \left( \frac{\sum Y}{n} \right)^2 = 55895 - \left( \frac{709}{10} \right)^2 = 5626.9
\]

\[
S_{xy} = \sum XY - \left( \frac{\sum X}{n} \right) \left( \frac{\sum Y}{n} \right) = 10747 - \left( \frac{134}{10} \right) \left( \frac{709}{10} \right) = 1246.4
\]

\[
\beta_1 = \frac{S_{xy}}{S_{xx}} = \frac{1246.4}{344.4} = 3.619
\]

Hence the regression equation is \( Y = 22.405 + 3.619X \)

(b) Find the error in estimating the product weight if the materials volume is 16 gallons.

\[
x_0 = 16 \quad y_0 = 3.619 + 22.405(16) = 80.309
\]

\[
e_0 = y_0 - y_0 = 93 - 80.309 = 12.691
\]

(c) Using a 95% confidence level, estimate an observation of the product weight if the materials volume is 16 gallons.

\[
x_0 = 16 \quad y_0 = 3.619 + 22.405(16) = 80.309
\]

95% confidence interval for a product:

\[
\alpha = 0.05, \quad \alpha = 0.025, \quad \nu = n - 2 = 10 - 2 = 8
\]

\[
t_{\alpha/2, \nu} = t_{0.025,8} = 2.306
\]

\[
y_0 \pm t_{\alpha/2, \nu} \frac{\sigma}{\sqrt{n}} \left[ 1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] = 80.309 \pm 2.306 \sqrt{139.522 \left( 1 + \frac{1}{10} + \frac{(16 - 13.4)^2}{344.4} \right)}
\]

\[
[51.487, 109.131]
\]
(d) At 5% level of significance, test the hypothesis that the regression line passes through origin.

\[ H_0 : \beta_0 = 0 \; ; \quad H_1 : \beta_0 \neq 0 \]

**Test Statistic:**

\[
T = \frac{\hat{\beta}_0 - \beta_0}{\sigma \left[ \frac{1}{n} + \frac{\bar{X}^2}{S_{xx}} \right]} = \frac{22.405 - 0}{\sqrt{139.522 \left( \frac{1}{10} + \frac{13.4^2}{344.4} \right)}} = 2.406
\]

**Decision Rule:** Reject \( H_0 \) if \( |T_0| > T_{\alpha/2, n-2} \)

**Critical Value:** \( T_{\alpha/2, n-2} = T_{0.025, 8} = 2.306 \)

**Decision:** As \( |T_0| > 2.306 \) so we reject \( H_0 \).

**Conclusion:** The regression line does not pass through the origin.