The first four questions on the quiz refer to the following scenario: A study measured the average weekly expenditure on alcoholic beverages (Y) and the average weekly expenditure on tobacco (X) for 11 regions of Great Britain. (Both are measured in the British unit of currency, pounds.) The data, some of the output from MINITAB from the Stat > Regression > Regression command for fitting a line to the data, and a scatter plot of the data are given on the next page.

1. Use the least squares line to predict the expenditure on alcoholic beverages for a region with a tobacco expenditure equal to 4 pounds.

   (a) 5.20  
   (b) 5.56  
   (c) 6.50  
   (d) 4.50

   **Solution.** From the MINITAB output we see that the least squares line is

   \[ \text{Alcohol} = 4.35 + 0.302 \text{ Tobacco}. \]

   Plugging 4 in for Tobacco gives the prediction,

   \[ 4.35 + (0.302)(4) = 5.558. \]

   So (b) is correct.

2. Compute the residual for the region with tobacco expenditure of 3.76 pounds.

   (a) .64  
   (b) 4.35  
   (c) .05  
   (d) .302

   **Solution.** The actual alcohol expenditure for the region is 6.13 pounds. The predicted alcohol expenditure for the region is

   \[ 4.35 + (0.302)(3.76) \approx 5.49. \]

   So the residual is \( 6.13 - 5.49 = 0.64 \).

3. One point on the scatter plot is by itself in the lower right hand corner of the plot. If this point were deleted from the data set, and a new least squares line were computed, the slope of the new line would be

   (a) the same as the slope of the line computed from the whole data set. 
   (b) smaller than the slope of the line computed from the whole data set. 
   (c) larger than the slope of the line computed from the whole data set.

   **Solution.** The correct answer is (c).

4. A researcher computes the correlation between tobacco and alcohol expenditures and says “These data show that there is not a strong relationship between the variables, because \( r \) is only 0.224.” Which of the following statements is a reasonable criticism of his position?

   (a) If the point in the lower right hand corner of the plot were removed, the value of \( r \) would be much higher than 0.224. We would see evidence of a strong relationship between \( X \) and \( Y \) if we excluded that data point.
   (b) Since all the \( Y \) values are between 4 and 7, the total variability in \( Y \) is small. This is what causes the small value of \( r \), not the lack of a relationship between the variables.
(c) The small value of \( r \) is caused by the scale of measurement. If the measurement were made in pence rather than pounds, \( r \) would be much higher.

(d) Both (a) and (c) are true.

**Solution.** Only (a) is true. Remember that correlation doesn’t change when units change, so (c) isn’t true.

5. On Mars, women always marry men who are exactly 2 years younger than they are. A sociologist collects data on husband’s age \( X \) and wife’s age \( Y \) for 40 Martian couples, and computes the correlation \( r \) between \( X \) and \( Y \). The value of \( r \) is

(a) \( r = -1 \)

(b) \( r = 0 \)

(c) \( r = 1 \)

(d) \( r = -2 \)

**Solution.** The correlation is \( r = 1 \), because all the points will lie on a line with positive slope. To see this, choose a few ages for husbands, say \( x = 35, x = 36, \) and \( x = 37 \). Compute the corresponding ages for wives, which will be \( y = 37, y = 38, \) and \( y = 39 \). Plot these values.

**THE DATA**

<table>
<thead>
<tr>
<th>Alcohol</th>
<th>6.47</th>
<th>6.13</th>
<th>6.19</th>
<th>4.89</th>
<th>5.63</th>
<th>4.52</th>
<th>5.89</th>
<th>4.79</th>
<th>5.27</th>
<th>6.08</th>
<th>4.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobacco</td>
<td>4.03</td>
<td>3.76</td>
<td>3.77</td>
<td>3.34</td>
<td>3.47</td>
<td>2.92</td>
<td>3.20</td>
<td>2.71</td>
<td>3.53</td>
<td>4.51</td>
<td>4.56</td>
</tr>
</tbody>
</table>

**THE MINITAB OUTPUT**

Regression Analysis: Alcohol versus Tobacco

The regression equation is

Alcohol = 4.35 + 0.302 Tobacco

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>4.35</td>
<td>1.607</td>
<td>2.71</td>
<td>0.024</td>
</tr>
<tr>
<td>Tobacco</td>
<td>0.3019</td>
<td>0.4388</td>
<td>0.69</td>
<td>0.509</td>
</tr>
</tbody>
</table>

\( S = 0.8196 \) \quad \text{R-Sq} = 5.0\% \quad \text{R-Sq(adj)} = 0.0\%

**THE SCATTER PLOT**
A scatter plot showing the relationship between Alcohol and Tobacco consumption levels.