Notes on Scatter Plots & Correlation

A scatter plot is often used to present bivariate quantitative data. Each variable is represented on an axis and the axes are labeled accordingly.

Make a scatter plot for this data:

<table>
<thead>
<tr>
<th>Number of Canned Beverages Sold</th>
<th>18</th>
<th>15</th>
<th>19</th>
<th>8</th>
<th>10</th>
<th>13</th>
<th>9</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Cans Recycled</td>
<td>8</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

A scatter plot displays data as points on a grid using the associated numbers as coordinates or ordered pairs \((x, y)\). The way the points are arranged by themselves in a scatter plot may or may not suggest a relationship between the two variables. For instance, by reading the graph below, do you think there is a relationship between the hours spent studying and exam grades?

If \(y\) tends to increase as \(x\) increases, then the data has **positive** correlation.

If \(y\) tends to decrease as \(x\) increases, then the data has **negative** correlation.

If \(x\) and \(y\) seem to have no relationship, or the data shows no pattern then the data has **NO** correlation.
A correlation coefficient, denoted by $r$, is a number from $-1$ to $1$ that measures how well a line fits a set of data pairs $(x, y)$. If $r$ is near 1, the points lie close to a line with a positive slope. If $r$ is near $-1$, the points lie close to a line with a negative slope. If $r$ is near 0, the points do not lie close to any line.

**Find the correlation coefficient in the calculator.**

a. 

<table>
<thead>
<tr>
<th>Beverage Can Recycling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Canned Beverages Sold</strong></td>
</tr>
<tr>
<td><strong>Number of Cans Recycled</strong></td>
</tr>
</tbody>
</table>

b. 

<table>
<thead>
<tr>
<th><strong>Hours (x)</strong></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Miles (y)</strong></td>
<td>35</td>
<td>29</td>
<td>26</td>
<td>20</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

**Practice Problems:**

For each scatter plot, tell whether the data has a positive correlation, a negative correlation, or no correlation. Then, tell whether the correlation is closest to $-1$, $-0.5$, $0$, $0.5$, or $1$.

3. Positive, negative, or no correlation?

a. Amount of exercise and percent of body fat ________________

b. A person's age and the number of medical conditions they have ________________

c. Temperature and number of ice cream cones sold ________________

d. The number of students at Lassiter and the number of dogs in Atlanta ________________

e. Age of a tadpole and the length of its tail ________________