Warm Up:
Janelle cut a board 30 feet long into 2 pieces. The ratio of the lengths of the 2 pieces is 2:3. What is the length, to the nearest foot, of the shorter piece?
F. 5  G. 6  H. 12  J. 15  K. 18

30/5 = 6 ft

Vocabulary:
**Explanatory Variable:** may help explain or influence changes in a *response* variable.
**Response Variable:** measures an *outcome* of a study.
**Scatterplots:** shows the relationship between two *quantitative* variables measured on the *same* individuals. The values of one variable appear on the *horizontal* axis and the values of the other variable appear on the *vertical* axis. Each individual in the data appears as a *point* in the graph.
**Positive Association:** when above average values of one tend to accompany above average values of the other and when below average values tend to *occur* together.
**Negative Association:** when above average values accompany below average values of the other.

3.1 Scatterplots and Correlation

Learning Goal: I will be able to identify explanatory and response variables, make scatterplots, describe scatterplots, find outliers of scatterplots, calculate correlation and explain how correlation is influenced by extreme observations.

**Scatterplots:**
Are THE graph to use when dealing with bivariate, quantitative data. (This means 2 lists that you can make ordered pairs out of!)

**Tips for drawing Scatterplots:**
1. Put explanatory variable on x-axis and response variable on y-axis
2. Label scale and axes, title (does *NOT* have to start at 0).
3. Plot ordered pairs (x, y)
How to examine a Scatterplot:

1. Describe direction (positive/negative slope), form shape (linear or curved), strength (strong: points close together, moderate: points a little more spread out, or weak: spread out)
2. Outliers, distant or not following pattern.

Example 1:
1. Measure your height (in inches) and then measure the length of one arm from shoulder to finger tip (in inches).
2. Record data on whiteboard
3. Make a scatterplot on graph paper. Make height be on the x-axis and arm length be on the y-axis.
4. Answer the following:
   a.) What is the explanatory and response variable?
   exp: height resp: arm length
   b.) Describe the overall pattern by Direction, Outliers, Form and Strength
      Positive, linear, moderately weak
   c.) Any outliers?
      No
   d.) Positive or negative association?
      Positive

Explanatory and Response

Variables:
Identify each explanatory and response variable. 
Variable A: Variable B:
weed growth amount of rain
winning percentage attendance at games
of a basketball team
amount of daily exercise
weight height
with the following instances in mind:
1.) Julie wants to know if she can predict a student’s weight from his or her height. Information about height is easier to obtain than information about weight!
2.) Jim wants to know if there is a relationship between height and weight.
Scatterplots on Calculator:

1.) Enter x values in L1 and y values in L2 (let's use the data for the next example)
2.) You must have Plots turned ON
3.) Hit GRAPH, your graph will not show up, go to ZOOM and 9:ZSTAT
4.) Take note of scale:

Example 2: Track and Field Day!
The table below shows data for 13 students in a statistics class. Each member of the class ran a 40-yard sprint and then did a long jump (with a running start). Make a scatterplot of the relationship between sprint time (in seconds) and long jump distance (in inches).

<table>
<thead>
<tr>
<th>Sprint Time (s)</th>
<th>5.41</th>
<th>5.05</th>
<th>9.49</th>
<th>8.09</th>
<th>7.01</th>
<th>7.17</th>
<th>6.83</th>
<th>6.75</th>
<th>8.01</th>
<th>5.68</th>
<th>5.78</th>
<th>6.31</th>
<th>6.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Jump Distance (ft)</td>
<td>171</td>
<td>154</td>
<td>48</td>
<td>151</td>
<td>90</td>
<td>65</td>
<td>78</td>
<td>71</td>
<td>130</td>
<td>173</td>
<td>143</td>
<td>141</td>
<td></td>
</tr>
</tbody>
</table>

Sketch the Scatterplot:

Interpret the Scatterplot:
(remember DOFS)

Correlation:
The strength and direction of DOFS can be described by finding correlation (r), also called the correlation coefficient.

Correlation is a number between -1 and 1
Direction is the sign of the correlation number
Strength is strong near ±1, weak close to 0

Look at the track example you just did......

The correlation is r = -0.75.

a.) Explain what this value means.
There is a somewhat strong negative linear relationship between sprint time and long jump distance.
b.) What effect would removing the student at (8.09, 151) have on the correlation?
Closer to -1 (Stronger)
c.) What effect would removing the student at (9.49, 48) have on the correlation?
Closer to 0 (Weaker)
**Guess the correlation:**

- \( r = -0.2 \)
- \( r = 0.4 \)
- \( r = -0.6 \)
- \( r = 0.7 \)
- \( r = -0.9 \)

**How is correlation calculated?**

It is the summation of the \( x \) and \( y \) z-scores multiplied together and divided by \( n-1 \).

\[
r = \frac{1}{N-1} \sum \left( \frac{X_i - \bar{X}}{s_x} \right) \left( \frac{Y_i - \bar{Y}}{s_y} \right)^2
\]

**Example 3:**

The lengths of two bones in five fossil specimens of the extinct beast Archaeopteryx are:

- Femur: \( 38, 56, 59, 64, 74 \)
- Humerus: \( 41, 63, 70, 72, 84 \)

a.) Make a scatterplot (on your calculator and then sketch). Do you think that all five specimens come from the same species?

b.) Find the correlation \( r \) step-by-step using the formula.

\[
\bar{x} = \frac{56.2}{5}, \quad s_x = 13.2 \quad \bar{y} = 70.0 \quad s_y = 15.9
\]
### Example 4: Correlation Blunders

Each of the following statements contains an error. Explain in each case what is wrong.

a.) There is high correlation between the gender of American workers and their income.

Only variable is quantitative

b.) We found a high correlation \( r = 1.09 \) between students' ratings of faculty teaching and ratings made by other faculty members.

\[-1 \leq r \leq 1\]

c.) The correlation between planting rate and yield of corn was found to be \( r = 0.23 \) bushel.

\( r \) does not have units

### Summary

How do you describe scatterplots?

DOFS

What do you have to remember when sketching scatterplots from your calculator?

Zoom, 9 window

---

**Coursework:**

pg 158-162 # 2, 6, 8, 9, 15, 16, 22, 26, 27-32 all

---

**Table for Example 3:**

<table>
<thead>
<tr>
<th>femur</th>
<th>Humerus</th>
<th>zfemur</th>
<th>zhumerus</th>
<th>product</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>41</td>
<td>-1.53</td>
<td>-1.57</td>
<td>2.41</td>
</tr>
<tr>
<td>56</td>
<td>63</td>
<td>-0.17</td>
<td>-0.19</td>
<td>0.03</td>
</tr>
<tr>
<td>59</td>
<td>70</td>
<td>0.06</td>
<td>0.25</td>
<td>0.02</td>
</tr>
<tr>
<td>64</td>
<td>72</td>
<td>0.44</td>
<td>0.36</td>
<td>0.17</td>
</tr>
<tr>
<td>74</td>
<td>84</td>
<td>1.20</td>
<td>1.13</td>
<td>1.310</td>
</tr>
</tbody>
</table>

\[
\frac{3.9a}{5-1} = \frac{3.9a}{4} \quad \text{Sum} = 3.99 \quad r = 0.998
\]