Today we consider data collected on a *quantitative* variable.

**Example 3-1: Matching Game**

Consider the following seven variables:

- (A) jersey numbers of Cal Poly football players
- (B) prices of properties on the Monopoly game board
- (C) blood pressure measurements for a sample of healthy adults
- (D) annual snowfall amounts for a sample of cities taken from around the U.S.
- (E) weights of rowers on the 2004 U.S. men’s Olympic team
- (F) point values of letters in the board game Scrabble
- (G) quiz percentages for a class of statistics students (quizzes were quite straight-forward for most students)

The following *dotplots* display the distributions of these variables, but they are not in the same order as the list above. Moreover, the scales are intentionally left off the axes! For each dotplot below, identify the variable that you believe it displays. Also provide a brief explanation of your reasoning in each case.

(1)

![Dotplot 1]

Variable:   
Explanation:

(2)

![Dotplot 2]

Variable:   
Explanation:

(3)

![Dotplot 3]
In any graph of data, look for an overall pattern and striking deviations from that pattern. Three important components of the overall pattern are shape, center, and variability. An outlier is an individual observation that falls outside the overall pattern of the graph. A distribution is symmetric if its left and right sides are roughly mirror images. It is skewed to the right if it extends much farther toward larger values, and it is skewed to the left if it extends much farther toward smaller values.
Which of the above graphs are skewed to the right? Which are skewed to the left? Which are symmetric?

**Example 3-2: Tallest Peaks**
The following histogram displays the distribution of altitudes of the tallest peaks (in feet) for each of the 50 states:

a) Sixteen states have a tallest peak between what two values?

b) Describe the shape of this distribution.

c) Between what two values (of highest peak in feet) does the middle value/state fall?

d) How many outliers are apparent in this graph? What is its value (approximately)? Guess which state is the outlier.

The following histograms separate these states by which are east and which are west of the Mississippi River:
Example 3-3: British Rulers’ Reigns
The following are the lengths of reign (in years) of British rulers since 1066:

<table>
<thead>
<tr>
<th>Ruler</th>
<th>Reign Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>William I</td>
<td>21</td>
</tr>
<tr>
<td>William II</td>
<td>13</td>
</tr>
<tr>
<td>Henry I</td>
<td>35</td>
</tr>
<tr>
<td>Stephen</td>
<td>19</td>
</tr>
<tr>
<td>Henry II</td>
<td>35</td>
</tr>
<tr>
<td>Richard I</td>
<td>10</td>
</tr>
<tr>
<td>John</td>
<td>17</td>
</tr>
<tr>
<td>Henry III</td>
<td>56</td>
</tr>
<tr>
<td>Edward I</td>
<td>35</td>
</tr>
<tr>
<td>Edward II</td>
<td>20</td>
</tr>
<tr>
<td>Edward III</td>
<td>0</td>
</tr>
<tr>
<td>Richard II</td>
<td>13</td>
</tr>
<tr>
<td>Mary I</td>
<td>5</td>
</tr>
<tr>
<td>Mary II</td>
<td>24</td>
</tr>
<tr>
<td>James I</td>
<td>22</td>
</tr>
<tr>
<td>James II</td>
<td>25</td>
</tr>
<tr>
<td>William III</td>
<td>2</td>
</tr>
<tr>
<td>George I</td>
<td>6</td>
</tr>
<tr>
<td>George II</td>
<td>33</td>
</tr>
<tr>
<td>George III</td>
<td>44</td>
</tr>
<tr>
<td>George IV</td>
<td>33</td>
</tr>
<tr>
<td>George V</td>
<td>25</td>
</tr>
<tr>
<td>George VIII</td>
<td>6</td>
</tr>
<tr>
<td>George VI</td>
<td>15</td>
</tr>
</tbody>
</table>

a) Who ruled the longest? How long was his/her reign?

b) What is the shortest reign? Who ruled it? What do you think this value really means?

c) Who is the current monarch? Why do you think her length of reign not included here?

Consider the following stemplot (also called a stem-and-leaf plot) of these reign lengths:

```
0 0123566799
1 0023333579
2 01224455
3 355589
4 4
5 069
6 3
```

d) How many rulers reigned for 13 years?

e) How many and what proportion of rulers reigned for more than 40 years?

f) Determine a value such that about half of these 40 rulers have reigned longer and half have reigned less long.
g) Describe this distribution (shape, center, variability). Relate your comments to the context.

**Example 3-4: Rowers’ Weights**
The following histogram displays the distribution of weights of the 28 individuals on the 2004 U.S. men’s Olympic rowing team:

![Histogram of rowers' weights](image)

a) How many and what proportion of these individuals weighed between 145 and 165 pounds?

b) Would you say that the distribution of these weights is skewed to the left, skewed to the right, or symmetric? Explain.

c) Identify the outlier in this graph—how much did he weigh? Can you explain why it makes sense for a rowing team to have one member who weighs substantially less than the others?

d) This graph also reveals a feature that the other graphs did not. Identify this feature, and offer an explanation for it in the context of an Olympic rowing team.
Example 3-5: February Temperatures
Data on daily high temperatures in February were recorded for three different locations. The data can be found in the Minitab worksheet FebTemps.

a) Use Minitab to produce dotplots of the temperature data for the three locations on the same scale (Graph> Dotplot> Multiple Ys).

b) Write a paragraph comparing and contrasting the distributions of February high temperatures across the three locations.

c) Based on the dotplots, which location tends to have the highest February temperatures? Which has the lowest?

Highest:  
Lowest:

d) Based on the dotplots, which location tends to have the most variability in February temperatures? Which has the least variability?

Most:  
Least: