For problems 1 through 3, refer to the following data sets:

<p>| | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>s-1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2.2</td>
<td>2.5</td>
<td>2.9</td>
<td>3.3</td>
<td>3.5</td>
<td>3.5</td>
<td>3.6</td>
<td>3.6</td>
<td>3.7</td>
<td>3.9</td>
<td>3.9</td>
<td>4.2</td>
<td>4.4</td>
<td>4.6</td>
<td>4.9</td>
</tr>
<tr>
<td>s-2</td>
<td>2.6</td>
<td>3.6</td>
<td>4.5</td>
<td>5.2</td>
<td>5.9</td>
<td>6.3</td>
<td>6.7</td>
<td>6.8</td>
<td>7.1</td>
<td>7.6</td>
<td>7.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(\textit{the top row is just labels, they’re not actual data values})

1. Make a stemplot for each data set. (Try to keep your columns neat)

   - S-1
   - S-2

2. If you wanted to describe the center of these data sets, would it be better to use the mean or median? Why?

   - S – 1  Mean  Median  \text{ (circle one) Why?}
   - S – 2  Mean  Median  \text{ (circle one) Why?}

3. Using \textit{all} of the data from \textit{both} sets, create a histogram with 5 bars. Clearly label your histogram.

<table>
<thead>
<tr>
<th>Class</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. A teacher measured how long it took her students to take their Final Exams last semester and summarized the data with these boxplots:

Which class would have the larger standard deviation? ______________

The middle 50% of her MAT 114 students finished during what range of times? ______________________

Half of her 114 students were done with their Final before anyone finished their 238 Final. True / False

Would you consider the distribution for MAT 238 to be skewed right or skewed left? Right / Left

The mean of the MAT 114 times would be ______________________ the median (fill in the blank with: greater than, less than, or about the same as).

5. The number of beats per minute (bpm) was measured for several popular contemporary songs and the resulting data is summarized in the stem-and-leaf plot below.

State the five number summary for this data set.

Now, make a boxplot for this data set, using the axis provided below. (4 points)

6. Suppose my data set has 13 values in it. How many data values are larger than Q₃?
7. Calculate the standard deviation of the following data set.

23 27 31 44 52 54 62 67

<table>
<thead>
<tr>
<th>Data values</th>
<th>Subtract the mean</th>
<th>Square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. A university wishes to summarize certain information at the start of the academic year. Which type of graph discussed in class would be the most appropriate for the following:

a) To show the proportion of total student enrollment broken down by its six colleges. (Arts, Science, Business, Education, Social Science and Professional)

b) To track the research monies awarded to the 6 colleges over the past 5 years.
9. Suppose you want to use a paired sample to compare the mean TV viewing times of married men and married women. Use this information to answer the following questions:

What is the variable under consideration?
A. Mean TV viewing time.
B. Marital status.
C. TV viewing time.
D. None of the above.

What are the two populations under consideration?
A. Married men and married women.
B. Married people and single people.
C. Men and women.
D. None of the above.

10. Why is a sample often a better way to obtain information about a population than a census?
A. A census is more costly than taking a sample.
B. A census is frequently impractical.
C. A census is too time consuming.
D. All of the above.
E. None of the above.

11. In sampling, why is obtaining a representative sample important?
A. Obtaining a representative sample is not important.
B. Because representative samples guarantee that our results are reliable.
C. Because we are using the sample to draw conclusions about the entire population.
D. None of the above.

12. A summary measure that is computed from a sample to describe a characteristic about a population is called A:
A. Statistic
B. Parameter
C. Population
D. Both a parameter and population since they are the same.
A college professor decides to run for Congress in a district with 450,000 registered voters. In a survey she commissioned 58% of the 4,000 registered voters interviewed indicated they plan to vote for her.

Match each vocabulary word below with its value or description on the right:

---

--- population
--- sample
--- parameter
--- statistic
--- sample size

a) percentage of all registered voters who will vote for her
b) all registered voters in the entire district
c) voluntary response bias
d) 58%
e) the registered voters who were actually interviewed
f) bricks which were inspected
g) 4000
h) 450,000
i) 446,000