1. Your college wants to gather student opinion about parking on campus. It isn’t practical to contact all students.

(a) Give an example of a way to choose a sample of students that is poor practice because it depends on voluntary response.

online survey

(b) Give another example of a bad way to choose a sample that doesn’t use voluntary response.

ask people going into the library

2. The two histograms show the distribution of word lengths in Shakespeare’s plays (left), and articles in *Popular Science* magazine (right).

(a) What should the units on the y-axis be?

per letter

(b) Do these distributions follow the Normal Curve? Explain briefly.

too - long right tail

(c) The mean word length for Shakespeare was 4.1 letters per word. Was the median larger or smaller than this? Explain briefly.

median will be smaller.

long right tail biases the mean to larger values.  

[TURN OVER]
(d) The data on word lengths for the *Popular Science* articles is given in the table below.

<table>
<thead>
<tr>
<th>Length</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>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent</td>
<td>3.6</td>
<td>14.8</td>
<td>18.7</td>
<td>16.0</td>
<td>12.5</td>
<td>8.2</td>
<td>8.1</td>
<td>4.4</td>
<td>3.6</td>
<td>2.1</td>
<td>0.9</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

What is the median number of letters per word?

3. A certain professor drives from Palo Alto to Santa Cruz each week. He records the time it takes. So far, the times recorded are 57, 55, 60, 58 and 54 minutes.

(a) What is the mean time?

\[ \frac{57 + 55 + 60 + 58 + 54}{5} = 56.8 \]

(b) What is the median time?

\[ \frac{54 \quad 55 \quad 57 \quad 58 \quad 60}{5} \text{ minutes} \]

(c) The numerical value of the standard deviation of the drive times is approximately 2.1. In what units is this standard deviation?

(d) Assuming the distribution of all journey times follows the Normal Curve, with mean as computed in a) and SD as in c) above, how long are the longest 2.5% of all journeys?

\[ \text{largest 2.5\% are } > 2 \text{ SD above mean} \]

\[ 56.8 + 2 \times 2.1 = 56.8 + 4.2 = 61 \text{ minutes} \]

(e) The professor's longest drive time was 60 minutes. Convert this to standard units.

\[ \frac{60 - 56.8}{2.1} = 1.5 \]

(f) Sketch a normal curve in standard units, and show the region corresponding to drive times longer than 60 minutes.

(g) Do you expect the distribution of all journey times to follow the normal curve? Explain briefly.