Jon is interested in whether or not implementing certain energy saving measures (e.g., turning off lights in rooms that aren’t being used) actually saves people money even with the increases in electricity rates. He asked a random sample of 200 people (a) whether or not they have made a conscious effort to conserve energy in the last few months (‘Yes’ or ‘No’) and (b) if their utility bills had gone up, down or stayed the same (‘Up’, ‘Same’ or ‘Down’). 120 people said they tried to save energy. Of those 120, 20 had bills go up and 70 had bills that stayed the same. Of those not conserving energy, 10 had bills that went down and 15 had bills that stayed the same.

a. Create the appropriate frequency distribution for Jon’s data.

b. Test the null hypothesis that conservation is independent of change in utility bill.

c. State the relation, if any, between the two variables.

Discrete Frequency Distributions

Exam grades

\[ \begin{array}{cccc}
 & (1) & (2) & (3) \\
 f_i & rf_i & rf_i' & rf_i'' \\
 A & 15 & .12 & .45 & .3 \\
 B & 98 & .78 & .40 & .6 \\
 C & 12 & .10 & .15 & .1 \\
 \hline
 & 125 & 1.0 & 1.0 & 1.0 \\
\end{array} \]

Location

Dispersion/Variability

[...some text...]

ABBACCBABBCBBCCBBBABAABBCBABBBCBA
BABCBABBCBABAABABBCBABBBCBBBABB
CBBBBBBCBABBABAABBBBABBBCBABBBCB
BBABABCBABCBBCB
### Continuous Frequency Distributions

#### Anxiety scores

<table>
<thead>
<tr>
<th>Interval</th>
<th>( f_i )</th>
<th>( rf_i )</th>
<th>( rcf_i )</th>
<th>Lower-Upper real limits</th>
<th>Mid-pt ( X_i )</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-11</td>
<td>14</td>
<td>.093</td>
<td>1.0</td>
<td>9.5-11.5</td>
<td>10.5</td>
<td>2</td>
</tr>
<tr>
<td>8-9</td>
<td>41</td>
<td>.272</td>
<td>.907</td>
<td>7.5-9.5</td>
<td>8.5</td>
<td>2</td>
</tr>
<tr>
<td>6-7</td>
<td>44</td>
<td>.291</td>
<td>.635</td>
<td>5.5-7.5</td>
<td>6.5</td>
<td>2</td>
</tr>
<tr>
<td>4-5</td>
<td>34</td>
<td>.225</td>
<td>.344</td>
<td>3.5-5.5</td>
<td>4.5</td>
<td>2</td>
</tr>
<tr>
<td>2-3</td>
<td>15</td>
<td>.099</td>
<td>.119</td>
<td>1.5-3.5</td>
<td>2.5</td>
<td>2</td>
</tr>
<tr>
<td>0-1</td>
<td>3</td>
<td>.020</td>
<td>.020</td>
<td>-5-1.5</td>
<td>0.5</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Location

Dispersion/Variability

10 9 2 7 2 8 3 1 5 2 9 1 2 6 7 3 5 4 6 2 4 6 1 5 8 3 5 7 4 8 2
5 7 1 1 2 5 8 5 7 3 8 6 3 5 7 5 3 7 5 4 5 3 7 6 4 6 8 2 5 7 4 5
8 4 6 8
Central Tendency

The Mode – most freq.
Pro: Can be used on any kind of variable; Easy
Con: No value information; Unreliable

S: 500 500 500 570 620 690 700 700 720 800
B: 500 500 520 550 600 600 600 700 720 800

The Mean (arithmetic) – average: $\bar{x} = \frac{\sum x_i}{n}$
Pro: Mathematical center of distribution
Con: Sensitive to outliers; can’t be used with many kinds of data

B: 500 500 550 600 600 650 700 700 750 800
S: 100 500 550 600 650 700 700 750 800

The Median – 50th percentile
Pro: Not sensitive to outliers; can be used with ordinal/interval
Con: No nice, mathematical properties

B: 500 550 560 600 650 710 780
S: 550 560 600 630 690 710 780

Quiz grades
Mean: 198
Median: 215

US Average Income
Mean: $50,000
Median: $30,000

Sociobiology
How many partners would you ideally desire in the next 30 years?

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>2.8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Men</td>
<td>64.3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Examples
IQ scores: 60 72 63 68 74 90 86 74 80
Error scores: 10 15 18 15 14 13 42 15 12 14 42
Grades: B D C A B F C B C D D