1. (d) all scores are identical (they do not HAVE to equal 0 or 1)

2. The data points for subjects scoring above the mean on X would have to have a negative slope to them. The data points for subjects below the mean on X could be any pattern as long as the to whole set of data points forms a positive slope, as below.

![Graph showing data points forming a positive slope]

3. $SS_Y$ (all data points fall on a straight line, so the predicted $Y$ values are the same as the observed $Y$ values)

4. (a) .34; $(-.58)^2$
   (b) 5.77; $6.3+(-.10)(30-24.7)$
   (c) 0.98; standard error of the estimate $= 1.2 \sqrt{(1 - (-.58)^2)(\frac{83}{82})}$

5. (b) and (f) (A sample mean near the extreme end of the distribution of means implies that all of the scores in the sample are very large (or very small if the mean is at the lower end), and so variability is likely to be low. A sample mean near the mean of the distribution could be composed primarily of scores that are similar to the mean, yielding a small variance or of scores that are widely different, coming from each end of the distribution of raw scores, or anything in between. So for a sample whose mean is near the middle of the distribution of means, there is a very wide range of possible variances.)

6. (a) B (b) B (Sample B has more scores in it, so it is more likely that it will include any particular score. With a larger sample, the variability of sample means is less, so there is a greater chance that a sample mean will be close to the mean of the population.)

7. (a) In symbolic form: $H_0: \mu = 50$ $H_1: \mu > 50$
   In natural language: The null hypothesis is that the mean for children taught with the whole-word method is equal to 50; the alternative hypothesis is that the mean for children taught with the whole-word method is greater than 50.
   (b) The standard error of the mean is $\sigma_m = \frac{\sigma}{\sqrt{N}} = \frac{10}{\sqrt{28}} = 1.89$
so \( z = \frac{(52.4 - 50)/1.89 = 1.27; \) with a directional test and \( \alpha = .05, \) the critical \( z \) score for rejecting \( H_0 \) is 1.65. The observed \( z \) value is not extreme enough to allow rejection of \( H_0. \)

(c) This study provides no evidence that the whole-word method is better than the normal method.

(d) To reject \( H_0 \) with a directional test and \( \alpha = .05, \) it would be necessary to have a mean that produced a \( z \) score of at least 1.65. A \( z \) score of that value, in the context of the current example, would correspond to a raw-score mean of 

\[ M = \mu + z(\sigma_M) = 50 + 1.65(1.89) = 53.1. \]

8. (a) \( s_M = \frac{s}{\sqrt{N}} = \frac{8.4}{\sqrt{28}} = 1.59, \) so \( t = (52.4 - 50)/1.59 = 1.51; \) critical \( t \) with \( df = 27 \) is 1.703, so do not reject \( H_0. \)

(c) This study fails to provide evidence that the whole-word method is superior to the standard method.

9. With \( N = 25, \sigma_M = \frac{20}{\sqrt{25}} = 4.00; \) the critical \( z \) value for a directional test with \( \alpha = .05 \) is 1.65, so the minimum or least extreme value of \( M \) needed to reject the null hypothesis is \( M = 70 + 1.65(4) = 76.6. \)

With \( N = 50, \sigma_M = \frac{20}{\sqrt{60}} = 2.58; \) the minimum value of \( M \) needed to reject the null hypothesis is \( M = 70 + 1.65(2.58) = 74.3. \) The minimum value is less extreme with larger \( N \) because sample means are less variable with larger sample sizes and it is therefore less likely that a sample mean will deviate very much from \( \mu. \)

10. The sample standard deviation is different for these two researchers, but the sample sizes and sample means are the same for both researchers. The researcher with the smaller \( s \) will obtain the larger \( t \) ratio because the denominator of the \( t \) ratio will be smaller in that case (i.e., the standard error of the mean will be smaller).

11. First read in the data file with this command,

\[
> \text{data} = \text{read.table(file.choose(new=T),header=T)}
\]

then run the \( \text{t.test} \) function.

\[
> \text{t.test(data, alternative = "two.sided", mu = 35)}
\]

One Sample t-test

\[
data: \text{data} \\
t = -2.1031, \text{df} = 39, \text{p-value} = 0.04196 \\
\text{alternative hypothesis: true mean is not equal to 35} \\
\text{95 percent confidence interval:} \\
30.56382 \text{ 34.91359} \\
\text{sample estimates:} \\
\text{mean of x} \\
32.73871
From the R output you can see that the obtained mean is 32.74 (bottom line of the output) and the t ratio is −2.10. The resulting p value is .04, which means that under the null hypothesis, there would be a .04 probability of obtaining a sample mean as extreme or more extreme than 32.74. So the obtained mean is among the 5% least likely, or most extreme, sample means that could have been obtained. As a result the null hypothesis can be rejected. Because the obtained mean is less than the mean of the general population, one can conclude that those who have recently sustained a concussion have, on average, lower scores on the short-term memory test. Would it be safe to conclude that concussion caused this difference? No, because there could be other differences between people who do and who do not experience concussions and these other differences might explain the difference in short-term memory performance.

If this had been a directional test, the p value would have been half of what is reported here, namely, .021. This is because with a nondirectional test, the reported probability includes both the upper and lower tails of the distribution, with half of the probability in each tail. In this case, a t ratio of +2.1031 cuts off the upper .021 of the t distribution with 39 degrees of freedom, and a t ratio of −2.1031 cuts off the lower .021 of that t distribution.

12. To compute the related-samples t test, obtain a difference score for each subject, then compute the mean and standard deviation of the difference scores. After that, all that is required is to compute a one-sample t test on the difference scores.

<table>
<thead>
<tr>
<th>Subject</th>
<th>S</th>
<th>P</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>8</td>
<td>−2</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>

Working with the difference scores, we have $M_D = 2.50$ and $s_D = 2.777$.

\[
s_{M_D} = \frac{2.777}{\sqrt{8}} = 0.982 \quad t = \frac{2.50-0}{0.982} = 2.546
\]

With $N = 8$, there are 7 degrees of freedom for this t test. According to the t table, with $df = 7$, $\alpha = .05$, and a nondirectional test, the critical t ratio is ±2.365. The observed t ratio is more extreme than the critical t ratio, so the null hypothesis is rejected. The researcher can conclude that the words processed under a survival instruction were remembered better than words processed under a pleasantness instruction. Note, however, that these two sets of words might differ on a number of other characteristics as well which could explain the difference in memory performance.

13. After reading the data into a variable called data and loading the psych library, we can use the describe function to get the descriptive statistics.
You can see that the mean for the control condition is 4.23 and the mean for the alcohol condition is 4.80. Now we can test this difference to see whether it is significant.

```r
> t.test(data$A, data$C, paired = T)
```

Paired t-test

data:  data$A and data$C
t = 2.5381, df = 29, p-value = 0.01678
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.1100408 1.0232925
sample estimates:
mean of the differences
  0.5666667

This $t$ test produces a $t$ ratio of 2.538. The corresponding $p$ value is .01678, which is less than the standard $\alpha$ value of .05, so the null hypothesis can be rejected. Because the mean for the alcohol condition is higher than for the control condition, the researcher can conclude that consuming alcohol increased perceived facial attraction.

14. Using the `cor` function, we have the following for the data in question 13 (attract1.txt):

```r
> cor(data)
```

```
   A     C
A 1.0000000 0.5110622
C 0.5110622 1.0000000
```

For the `attract2.txt` data set, we have

```r
> describe(data2)
```

```
   vars n mean   sd median trimmed mad min max range skew kurtosis   se
A 2 30 4.80 1.30 5  4.83 1.48 2 7  5 -0.37  -0.24 0.24
C 1 30 4.23 1.17 4  4.21 1.48 2 7  5  0.06  -0.37 0.21
```

The means and standard deviations are exactly the same as for the `attract1.txt` data set. Indeed, all that happened was that the scores in the control condition were reassigned to subjects in a random way. The $t$ test produces this result:

```r
> t.test(data2$A, data2$C, paired = T)
```

Paired t-test

data:  data2$A and data2$C
t = 1.8306, df = 29, p-value = 0.07746
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
The \( t \) test for these data yields \( t = 1.831 \), with a \( p \) value of .077. In this case, the null hypothesis is not rejected, even though the means are the same as in question 13. The reason for the discrepancy can be found in the variability of the difference scores. That variability is much larger in this second case and is reflected in the much smaller correlation between scores in the two conditions for this new data set.

```r
> cor(data2)
A  C
A 1.0000000 0.05475667
C 0.05475667 1.0000000
```

In this case, the correlation is only .055, whereas for the original data set the correlation was .511.

Looking at the difference scores for the two data sets, for the original data set we have

```r
> diff=data$A-data$C
> sd(diff)
[1] 1.222866
```

and for the new data set we have

```r
> diff2 = data2$A-data2$C
> sd(diff2)
[1] 1.695498
```

The variability (standard deviation) of the difference scores is larger in the second case, and as a result the \( t \) ratio is smaller (and not significant).

With a directional test and an alternative hypothesis specifying that the ratings should be higher in the alcohol condition, the null hypothesis could be rejected if the observed difference between means goes in that direction (i.e., the sample mean for the alcohol condition is indeed larger than for the control condition) and if the probability from the nondirectional test computed by \( \mathbf{R} \) is no greater than .10. Remember that for a nondirectional test with \( \alpha = .10 \), each tail of the \( t \) distribution has .05 of the area beyond the critical \( t \) ratio. In the data from the \texttt{attract2.txt} file, the observed difference between means goes in the right direction (higher ratings in the alcohol condition) and the nondirectional \( p \) value is less than .10, so the null hypothesis could be rejected in this case if a directional test were adopted. But if the alternative hypothesis had specified that higher ratings should occur in the control condition, then these data would not allow rejection of the null hypothesis. Any time an effect goes in the direction opposite to that predicted by a directional alternative hypothesis, the null hypothesis cannot be rejected, no matter how small the obtained \( p \) value is.