Now we continue to address the issue of comparing means between two groups, but we’ll consider data collected with a different (and often better) study design.

**Example 19-1: Marriage Ages**

A student of mine wanted to test whether husbands tend to be older than their wives on average. He went to the county courthouse and took a sample of 24 couples who had applied for marriage licenses, recording the ages of the man and woman in each case (MarriageAges.mtw). Some summary statistics are:

<table>
<thead>
<tr>
<th></th>
<th>Sample size</th>
<th>Sample mean</th>
<th>Sample std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Husbands</td>
<td>24</td>
<td>35.71</td>
<td>14.56</td>
</tr>
<tr>
<td>Wives</td>
<td>24</td>
<td>33.83</td>
<td>13.56</td>
</tr>
</tbody>
</table>

a) Perform a two-sample \( t \)-test of whether these sample data provide evidence that the population mean age of husbands exceeds that of wives.

b) Use a two-sample \( t \)-procedure to estimate the difference in population mean ages with a 90% confidence interval.

This two-sample \( t \)-procedure is *not valid* here because the samples are *not independent*. The age of a wife is related to the age of a husband, because older people tend to marry older people and younger people tend to marry younger people.

c) Look at a scatterplot to see the relationship between the ages of a husband and a wife (Graph> Scatterplot). Do the husband and wife ages appear to be related? Explain.
The two-sample t-procedures applied in a) and b) are completely inappropriate here. This analysis only would have been appropriate if the student had gathered ages for one sample of 24 husbands and then independently gathered ages for a different sample of 24 wives.

These data are paired, because the observational units are couples, not independent individuals. The appropriate analysis is to calculate the differences in ages for each couple, and then apply paired t-procedures to those differences.

- **Paired t-test, -interval**
  - Null hypothesis $H_0$: $\mu_d = 0$, where $\mu_d$ represents the population mean of the differences.
  - Test statistic: $t = \frac{\bar{x}_d}{s_d / \sqrt{n_d}}$, where
    - $\bar{x}_d$ is the sample mean of the differences
    - $s_d$ is the sample standard deviation of the differences,
    - $n_d$ is the sample size of the differences
  - The $p$-value is based on the $t$-distribution with $(n-1)$ degrees of freedom
  - Confidence interval for $\mu_d$: $\bar{x}_d \pm t^* \frac{s_d}{\sqrt{n_d}}$
  - Technical conditions
    - Random sample (of differences) from the population
    - Large sample size ($n_d \geq 30$) or normal distribution (of differences)

Some more summary statistics, with differences calculated as husband’s age minus wife’s age:

<table>
<thead>
<tr>
<th>Differences</th>
<th>Sample size</th>
<th>Sample mean</th>
<th>Sample std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differences</td>
<td>24</td>
<td>1.875</td>
<td>4.812</td>
</tr>
</tbody>
</table>

d) Determine and interpret a 90% confidence interval for the population mean difference in ages between husbands and wives. Also check and comment on the technical conditions for applying this procedure.
e) Conduct a paired $t$-test of whether the sample data provide strong evidence that the population mean difference exceeds zero. Include all components of a significance test. Indicate the test decision at the .10 significance level, and summarize your conclusion.

f) Explain why the paired analysis produces such a different conclusion than the independent-samples analysis.

g) Was the student wise for gathering paired data rather than independent-samples data to investigate his research question? Explain.

The pairing is effective here because there is a lot of variation in the ages of people who apply for marriage licenses, but there is a strong correlation between the ages of the husband and wife within a couple. Thus, there is much less variation in the differences in ages than there is in the ages themselves.

Notice that a paired analysis follows from the data having been collected according to a matched-paired design in the first place.

h) Would it be appropriate to compare men’s and women’s body temperatures with a paired $t$-test? Explain.
Example 19-2: Melting Morsels
Suppose that you want to investigate whether there is a difference in the melting times of semi-sweet milk chocolate chips and peanut butter chips. You could take a group of volunteers, randomly assign half to take a chocolate chip and the other half to take a peanut butter chip, and then time how long it takes before the chip melts in their mouth.

a) Would this be an independent-samples or a matched-pairs design? Explain.

b) Describe how you could alter the data collection process to make this a paired design.

c) Explain why it’s a good idea to conduct this study with a matched-paired design.

d) Describe how (and why) a paired design to investigate this issue should make use of randomness.

Example 19-3: Waterproofing Boots
Suppose that you want to compare two methods of water-proofing boots and that you have recruited 100 subjects to participate in an experiment.

a) Suggest a better experimental design than randomly assigning 50 subjects to wear boots with each type of water-proofing.

b) What is the primary advantage of the matched-pairs design in this study?