Example: Wolf River

You’ve been given the responsibility to analyze the Wolf River data to test whether the true mean aldrin concentrations at the bottom and mid-depth might differ. That is, you want to set up a test of significance for the difference between two population means based on data from independent random samples. Use $\alpha = 0.10$.

### 9.3 How Large a Sample? The 15/40 rule for t-procedures.

- If your sample size is less than 15: Be **very careful**. Your data or transformed data must look like they came from a normal distribution—little skewness, no outliers.
- If your sample size is between 15 and 40: **Proceed with caution**.
  - Strongly skewed distributions should be transformed to a scale that makes them more nearly symmetric before using the t-procedures.
  - If you have gross outliers, a transformation may be in order. If you don’t transform or if the outliers remain even after a change of scale, do two versions of your test or interval, one with and one without the outliers.
  - Don’t rely on any conclusions that depend on whether you include the outliers.
- If your sample size is over 40: You’re in **good shape**. Your sample size is large enough that skewness will not reduce capture rates or alter significance levels enough to matter.
  - Still, if your sample shows strong skewness, it is worth asking whether a change of scale would make the usual summary statistics (especially the standard deviation) more meaningful. Even though outliers may not have much effect on capture rates or significance levels, you should still check by doing two versions of your t-procedure.
Examples of Transformations

<table>
<thead>
<tr>
<th>Species</th>
<th>Brain Weight (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African elephant</td>
<td>5712</td>
</tr>
<tr>
<td>African giant pouched rat</td>
<td>64</td>
</tr>
<tr>
<td>Arctic fox</td>
<td>44.5</td>
</tr>
<tr>
<td>Arctic ground squirrel</td>
<td>57</td>
</tr>
<tr>
<td>Baby elephant</td>
<td>4603</td>
</tr>
<tr>
<td>Baboon</td>
<td>178.5</td>
</tr>
<tr>
<td>Barracuda</td>
<td>3.83</td>
</tr>
<tr>
<td>Big brown bat</td>
<td>0.3</td>
</tr>
<tr>
<td>Blue whale</td>
<td>6800</td>
</tr>
<tr>
<td>Brown trout</td>
<td>0.17</td>
</tr>
<tr>
<td>Canary</td>
<td>0.85</td>
</tr>
<tr>
<td>Cat</td>
<td>25.6</td>
</tr>
<tr>
<td>Catfish</td>
<td>1.84</td>
</tr>
<tr>
<td>Chimpanzee</td>
<td>440</td>
</tr>
<tr>
<td>Chimchilla</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Significance Tests
Various Examples

1. A job placement director claims that the average starting salary for nurses is $24,000. A sample of 10 nurses’ salaries has a mean of $23,450 and a standard deviation of $400. Is there enough evidence to reject the director’s claim at $\alpha=0.05$?

2. A statistician read that at least 77% of the population oppose replacing $1 bills with $1 coins. To see if this claim is valid, the statistician selected a sample of 80 people and found that 55 were opposed to replacing the $1 bills. At $\alpha=0.01$, test the claim that at least 77% of the population are opposed to the change.

3. A researcher reports that the average salary of an assistant professor is more than $42,000. A sample of 30 assistant professors has a mean salary of $43,260 and a standard deviation of $5,230. At $\alpha=0.05$, test the claim that assistant professors earn more than $42,000 a year.

4. In a sample of 200 surgeons, 15% thought the government should control health care. In a sample of 200 general practitioners, 21% felt the same way. At $\alpha=0.01$, is there a difference in the proportions?

5. The average size of a farm in Indiana County, Pennsylvania, is 191 acres. The average size of a farm in Greene County, Pennsylvania, is 199 acres. Assume the data were obtained from two samples with standard deviations 38 and 12 acres, respectively, and sample sizes of 8 and 10, respectively. Can it be concluded at $\alpha=0.05$ that the average size of the farms in the two counties is different? Assume the populations are normally distributed.

6. A survey found that the average hotel room rate in New Orleans is $88.42 and the average room rate in Phoenix is $80.61. Assume that the data were obtained from two samples of 50 hotels each and that the standard deviations of these samples were $5.62$ and $4.83$ for New Orleans and Phoenix, respectively. At $\alpha=0.03$, can it be concluded that there is a significant difference in the rates?