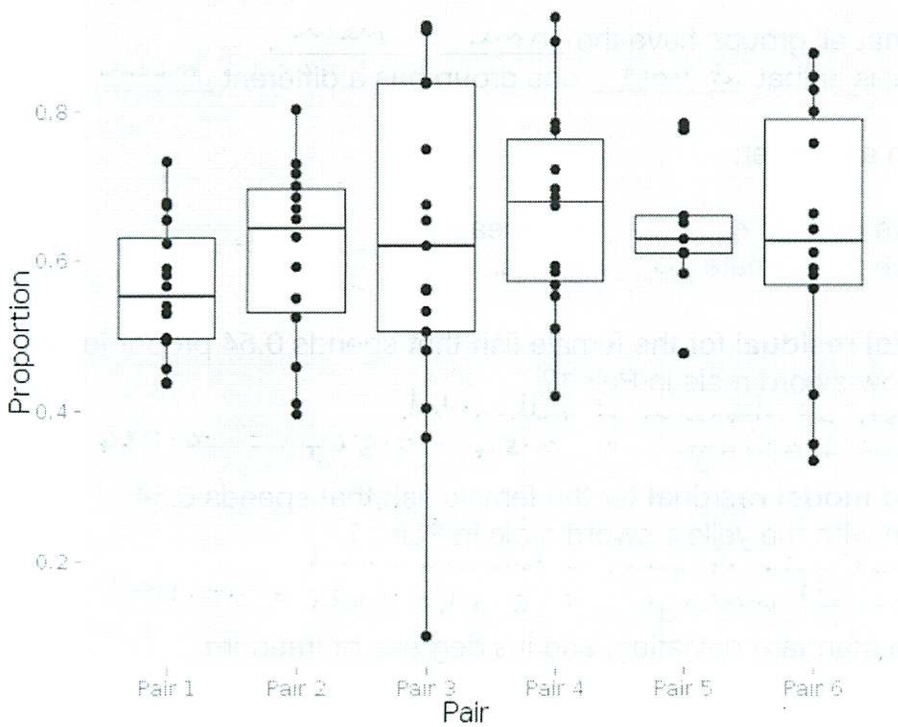
One way ANOVA worked example.
Pre-existing Preferences of Fish Case Study 6.2 in Sleuth



Summary Statistics:

	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Overall
Average	0.564	0.609	0.624	0.670	0.642	0.633	0.621
SD	0.090	0.125	0.223	0.143	0.094	0.177	0.154
Sample Size	16	14	17	14	9	14	84

We are going to perform a one-way analysis of variance to answer the question:
Does the pair of males make a difference in the preference of the females for the males with yellow swordtails?

1. Read the study description in Sleuth 6.1.2. Are causal inferences possible? Are population inferences possible?

The fish were not radomly selected from a larger population so population inferences are not possible.

The grouping variable here are the pairs, but the causal inference of interest is whether the yellow tail makes the male more attractive. To make this inference the male to have the yellow tail must be selected at random from the pair, the description doesn't actually say this is the case ...

2. Fill in the blanks:

The null hypothesis is that all groups have the same mean.
The alternative hypothesis is that at least one group has a different mean.

3. Fill in the blanks with a number:

In the reduced model we have to estimate 1 mean/s.
In the full model we have to estimate 6 mean/s.

4. What is the **full model residual** for the female fish that spends 0.54 proportion of the time with the yellow sword male in Pair 1?

= 0.54 − estimate of mean in full model
= 0.54 − pair 1 average = 0.54 − 0.564 = −0.024

5. What is the **reduced model residual** for the female fish that spends 0.54 proportion of the time with the yellow sword male in Pair 1?

= 0.54 − estimate of mean in reduced model
= 0.54 − overall average = 0.54 − 0.621 = −0.081

6. Calculate the pooled standard deviation, and it's degrees of freedom.

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + \dots (n_I-1)s_I^2}{(n_1-1)+(n_2-1)+\dots(n_I-1)}}$$

$$= \sqrt{\frac{15 \times 0.092^2 + 13(0.125)^2 + 16(0.223)^2 + 13(0.143)^2 + 8(0.094^2) + 13(0.177)}{84-6}}$$

$= 0.1546$

d.f. = 84 − 6
= 78

6. Find the within group sum of squares and the total sum of squares using the shortcuts from homework 6.

Within group sum of squares = s_p^2 x (degrees of freedom for s_p)
= $0.1546^2 \times 78$
= 1.864

Total sum of squares = (Overall SD)^2 x (n - 1)
= $(0.154)^2 \times 83$
= 1.968

8. Fill in the ANOVA table.

	Sum of squares	d.f	Mean Square	F-statistic	p-value
Between Groups	0.104	5	0.0208	0.870	0.51
Within Groups	1.864	78	0.0239		
Total	1.968	83			

If the null hypothesis is true, the F-statistic will be distributed like a F -distribution with 5 and 78 degrees of freedom.

The p-value can be found in R with: `1 - pf( 0.870, 5, 78)`

9. **Statistical summary.** Fill in the blanks:

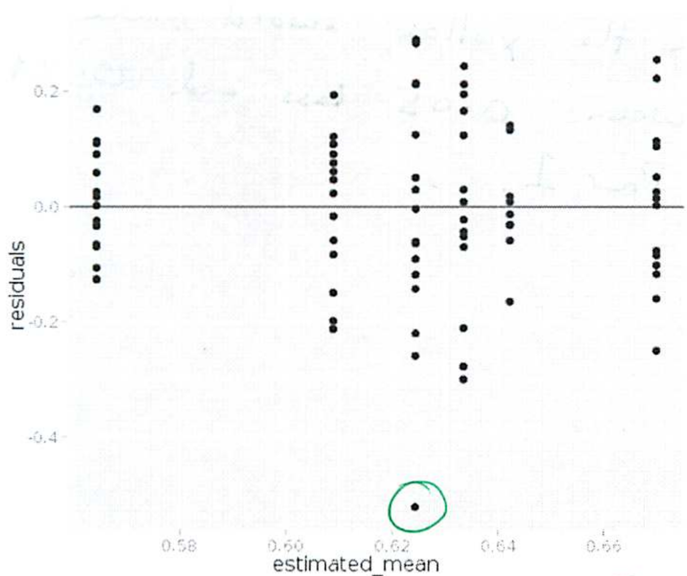
There is no evidence that the mean proportion of time spent with yellow sword male in different pairs are not equal (F-test, p-value = 0.51).

10. What are the three assumptions of the one-way ANOVA?

1. ~~Normality of popln.~~ Normality of pop.
2. Equal pop. SDs
3. Independence between & within groups

Which can you check by examining the plot below?

1 & 2



Some evidence of unequal SDs?

outlier?

11. Imagine we want to compare the male fish of Pair 1 to those of Pair 5.

a. The estimated difference in mean proportion of time spent with yellow sword male in Pair 1 and Pair 5 is: $\bar{Y}_5 - \bar{Y}_1 = 0.642 - 0.564$

$= 0.078$

b. The standard error on the difference estimated in a. is:

$$\mathrm{SE}_{\overline{Y}_2 - \overline{Y}_1} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= 0.1546 \times \sqrt{\frac{1}{16} + \frac{1}{9}}$$

$$= 0.0644$$

c. A 95% confidence interval for the difference in a. is:

$$\bar{Y}_5 - \bar{Y}_1 \pm qt(0.975, 78)\ \mathrm{SE}_{\bar{Y}_5 - \bar{Y}_1}$$

$$= 0.078 \pm 1.99 \times 0.0644$$

$$= (-0.050,\ 0.206)$$

(Hint: qt(0.975, 78) = 1.99)

With 95% confidence, the mean proportion of time spent with the yellow sword male in Pair 5 is between 0.05 less and 0.21 more than in Pair 1.