I'd like to work with a dataset I collected which contains big brown bat pup weights and forearm lengths, collected every two days from postnatal day two to 36. These data will be used as a measure of pup development. The dataset contains males (n=6) and females (n=6).

Hypothesis
I hypothesize that there is a difference between male and female pups, both in weight and forearm length. The null hypothesis would be that there is no difference between males and females, in weight or forearm length.

Statistical Test
As we've discussed in lecture, I don't want to just rely on a p-value to test this hypothesis. Instead, I want to consider effect sizes and confidence intervals (CI).

To test the hypothesis, I could take a non-inferiority approach. To do so, I would need to decide on the margin of non-inferiority (or the negative and positive "cut offs" as discussed in lecture). This would be based on prior work and also my prior knowledge. For example, I know that wild adult big brown bats can vary in weight by 10–15g. However, considering that pups are born weighing only a few grams and don't even reach 10g until they are a couple weeks old, it wouldn't make sense to make the cut off for weight this high.

Then, I would perform two 1-tailed t-tests with 95% CI and compare to my margin of non-inferiority. Importantly, I would need to look at the 95% CI to base my conclusions. For example, the 95% CI crossing 0 would suggest that there is no difference between males and females. If the CI falls above 0, it would suggest a positive effect, and below 0 would suggest a negative effect. What I would consider a positive vs. negative effect would be based on which group I am comparing to which (e.g. do females weigh more than males? If the 95% CI falls above 0, it would suggest that females do weigh more than males).

As we've discussed, I'm not just interested in whether the CI is positive or negative. Like I mentioned, I want to compare to my margin. If the 95% CI falls outside of my margin, I would consider it to be a big effect (i.e. a big enough difference between the two groups that, based on my cut off decision, it is significant and should "mean something").