Confidence Intervals for $\mu_X - \mu_Y$:

Goals of the lab.
To give students an understanding of making a confidence interval for the mean difference between two paired populations.

Background:
This lab discusses the difference of the means when those two populations are independent. That is, the first population and second population are unrelated. There are two cases to consider: The small sample case and the large sample case.

I) Small samples (n<20 and m<20)
Relevant Formula:

$$(X - Y) \pm t_{n,m} \frac{s_p}{\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

where

$$s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}}$$

Example:
Suppose we want to create a 95% CI for the difference of independent population means. Summary info for the two samples follows:

$\bar{X} = 18.3$, $s_X = 3.4$, $n = 15$.

$\bar{Y} = 12.6$, $s_Y = 2.9$, $m = 16$.

Since $n < 20$ and $m < 20$, and $s_X$ is approximately $s_Y$ then we can use the following procedure to make our CI.

$$s_p = \sqrt{\frac{(15-1)3.4^2 + (16-1)2.9^2}{15+16-2}} = 3.15$$

Then,

$$(X - Y) \pm t_{15+16\,2.045\,0.025} \frac{s_p}{\sqrt{\frac{1}{n} + \frac{1}{m}}} = 5.7 \pm 2.31 = (3.39, 8.01)$$

So we are 95% confident that the difference $\mu_X - \mu_Y$ is between 3.39 and 8.01. Since 0 (zero) is not inside the confidence interval then we can conclude with 95% confidence that there is a difference between the means of these two populations.
II) Large Samples (n≥20 and m≥20)

Relevant Formula:

\[
(X - Y) \pm z^* \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}
\]

Example:

Suppose we want to create a 90% CI for the difference of independent population means. Summary information from the two samples follows:

\( \bar{X} = 16.37, s_X = 7.89, n = 34. \)
\( \bar{Y} = 17.12, s_Y = 6.33, m = 32. \)

Since n ≥ 20 and m ≥ 20, and \( s_X \) is approximately \( s_Y \), then we can use the following formula to make our CI.

\[
(\bar{X} - \bar{Y}) \pm z^* \sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}
\]

\[
= (16.37 - 17.12) \pm 1.645 \sqrt{\frac{7.89^2}{34} + \frac{6.33^2}{32}}
\]

\[
= -0.75 \pm 2.89
\]

\[
= (-3.64, 2.19)
\]

So we are 90% confident that the difference \( \mu_X - \mu_Y \) is between –3.64 and 2.19. Since 0 (zero) is inside the confidence interval, then we can conclude with 90% confidence that there is no difference between the means of these two populations.