Learning Goals:

- Under appropriate conditions, conduct a hypothesis test about a difference between two population means. State a conclusion in context.
- Construct a confidence interval to estimate a difference in two population means (when conditions are met). Interpret the confidence interval in context.

Introduction:

In this activity we will use StatCrunch to conduct inference procedures (hypothesis test and confidence interval) for a difference in population means.

To use T-procedures, samples must be randomly selected and independent, or, in the case of an experiment, randomization must be used to assign treatments.

Conditions for use of the t-test that we learned earlier in Unit 10 must be met for both samples.

The hypotheses for two population means are similar to those for two population proportions. The null hypothesis, H_0, is a statement of "no effect" or "no difference."

H_0: $\mu_1 - \mu_2 = 0$, which is the same as H_0: $\mu_1 = \mu_2$

The alternative hypothesis, H_a, takes one of the following three forms:

- H_a: $\mu_1 - \mu_2 < 0$, which is the same as H_a: $\mu_1 < \mu_2$
- H_a: $\mu_1 - \mu_2 > 0$, which is the same as H_a: $\mu_1 > \mu_2$
- H_a: $\mu_1 - \mu_2 \neq 0$, which is the same as H_a: $\mu_1 \neq \mu_2$

If conditions are met, we will use StatCrunch to find a P-value based on a T-model determined by degrees of freedom $(n_1 - 1) + (n_2 - 1)$.

The confidence interval estimates the difference between two population means or the difference between means coming from two treatments. The confidence interval has the usual form: *sample statistic ± margin of error*, where the margin of error is based on the standard error. For those who like formulas, here is the formula for the confidence interval:

$$(\bar{x}_1 - \bar{x}_2) \pm T_c \cdot \sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}$$

If conditions are met, we will use StatCrunch, instead of this formula, to find confidence intervals in this module based on a T-model determined by degrees of freedom $(n_1 - 1) + (n_2 - 1)$.

Swimming With Dolphins Lifts Depression: Study

Taking a dip with dolphins can be a tremendous therapy for people with depression, according to a study published in the weekly British Medical Journal (BMJ).

Nature lovers - biophiles, to give them their scientific name - have long argued that interaction with animals can soothe a troubled mind but this claim has always been anecdotal, lacking the scientific data to back it up.

Seeking to find out more, psychiatrists Christian Antonioli and Michael Reveley at Britain's University of Leicester, recruited 30 people in the United States and Honduras who had been diagnosed with mild or moderate depression.

The severity of their symptoms was calculated according to established yardsticks for mental health, the Hamilton and Beck scales, which are based on interviews and questionnaires with the patient.

The volunteers were required to stop taking any anti-depressant drugs and psychotherapy for four weeks.

Half of the group was then randomly selected to play, snorkel and take care of dolphins each day at an institute for marine sciences in Honduras.

The other half was assigned to a programme of outdoor activities, also at the institute, that included swimming and snorkelling at a coral reef, but without the dolphins.

Two weeks later, both groups had improved, but especially so among patients who had been swimming with the dolphins.

Measurable symptoms of depression in the dolphin group had fallen by half and by two-thirds according to the two scales - twice as much as in the non-dolphin group.

In addition, a self-rating measurement of anxiety symptoms, the Zung scale, found a fall of more than 20 percent among the dolphin group, compared with a decline of 11 percent among the non-dolphin groups.

"To the best of our knowledge, this is the first randomized, single blind, controlled trial of animal-facilitated therapy with dolphins," say Antonioli and Reveley.

Source: http://www.banderasnews.com/0511/hb-swimwithdolphins.htm

The news article summarizes the actual results from the 30 subjects and the 15 in the treatment group that swam with dolphins had greater improvements than the 15 in the control group that swam and snorkeled. But, are the differences statistically significant? Or would we expect to see the differences described when comparing two identical treatments? In other words, could the differences be due to expected variability arising naturally when comparing two groups? We will investigate these questions in groups.

Each group will investigate **one** option and present a short report to the class using the document camera. Details are on the next page.

AFTER the class presentations, in the box below you will write an overall conclusion summarizing the findings of this study for a general audience … something that could be added to newspaper article about this study.

The researchers used three different tools for assessing depression: Hamilton scale, Beck inventory, and Zung self-rating. Higher scores on these scales indicate more severe depression or anxiety.

We do not have access to the raw data, so we will assume that the conditions for use of the T-model are met despite the small sample sizes.

Options:

Option 1: Based on the Hamilton depression scale, is the difference in improvement between the treatment group and control group statistically significant? Conduct a hypothesis test to answer this question.

Option 2: How big is the treatment effect based on the Hamilton depression scale? Estimate the size of the treatment effect with a confidence interval.

Option 3: Based on the Beck depression inventory, is the difference in improvement between the treatment group and control group statistically significant? Conduct a hypothesis test to answer this question.

Option 4: How big is the treatment effect based on the Beck depression inventory? Estimate the size of the treatment effect with a confidence interval.

Option 5: Based on the Zung self-rating anxiety scale, is the difference in improvement between the treatment group and control group statistically significant? Conduct a hypothesis test to answer this question.

Option 6: How big is the treatment effect based on the Zung self-rating anxiety scale? Estimate the size of the treatment effect with a confidence interval.

Group work:

1) Psychologists have developed a variety of tools to measure depression or anxiety levels. As you can imagine, measuring a mental state is harder and less precise than measuring physical conditions, like height or weight.

   For this reason, in this study, the researchers used three different surveys that are commonly used by the mental health professionals.

   Review the survey associated with the scale or inventory that is relevant to your research question. Your instructor will provide an easy way to access these links.

   DO NOT complete the survey because this will take too much time. Just review the survey and discuss in your group how to best summarize the survey during your presentation to the class. You might also want to pull a few survey questions as examples.

   Hamilton depression scale (0-38):
   http://healthnet.umassmed.edu/mhealth/HAMD.pdf

   Beck depression inventory (0-60):
   http://www.hr.ucdavis.edu/asap/pdf_files/Beck_Depression_Inventory.pdf

   Zung self-rating anxiety scale (20-80):
   https://www.mnsu.edu/comdis/isad16/papers/therapy16/sugarmanzunganxiety.pdf

2) Here are the results of the study excerpted from the British Medical Journal. Source : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1289317/

	Number of participants	Change = mean of differences (before – after)	SD of differences (before – after)
Hamilton rating scale for depression			
Treatment group	15	7.27	3.47
Control group	15	3.60	3.36
Beck depression inventory			
Treatment group	15	13.40	7.58
Control group	15	6.07	7.28
Zung self-rating anxiety scale			
Treatment group	15	9.80	7.32
Control group	15	5.73	5.76

The researchers used both matched-pairs and two-sample T-procedures.

- Using each of the 3 scales, they rated each subject both before and after the experiment. They then calculated the difference (before minus after) to represent the effect of the treatment on the subject according to each scale; this is the matched-pairs part.
- The researchers then calculated the mean and standard deviation of these differences for the control group and for the treatment group; this is the two-sample part. These are the means and standard deviations in the table.

For example, suppose subject #1 is randomly assigned to the treatment group. His Hamilton rating was 14 before the experiment and 8 after the experiment. His change (an improvement) is 6 (14-8=6). His change score is one of the 15 numbers used to calculate the mean of 7.27 and SD of 3.47 shown in the table.

Subject #1 will also have a Beck "before" and "after" score. The difference between these two scores will be one of the 15 numbers used to calculate the mean of 13.40 and SD of 7.58 shown in the table, etc.

a) StatCrunch requires you to enter a mean, standard deviation and sample size for each of the two groups. Circle the information in the table you will use.

b) Write a sentence to explain what each of these numbers tells us.

3) Conduct a two-sample T-procedure in StatCrunch that addresses your research question. (Open a StatCrunch spreadsheet. Choose Stat, T Stats, Two sample, With summary)

Group presentation:

- State your research question and briefly explain the scale used in your investigation (e.g. Hamilton).
- Identify the appropriate mean improvement for the treatment group and for the control group from the table of results. Explain what these numbers tell us. Do the same for the standard deviations.
- Complete the StatCrunch template below that matches your inference procedure (hypothesis test or confidence interval).
- State a conclusion that references the context of the study.
- Be prepared to explain the meaning of each of the numbers in the StatCrunch print-out.

Research question:

Sample results and their meaning:

Two sample T-procedure:

μ_1 :

μ_2 :

$\mu_1 - \mu_2$: Difference between two means

Inference Results:

For a hypothesis test:

$H_0 : \mu_1 - \mu_2$

$H_A : \mu_1 - \mu_2$

(with pooled variances)

Difference	Sample Diff.	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$					

For a 95% confidence interval:

Difference	Sample Diff.	Std. Err.	DF	L. Limit	U. Limit
$\mu_1 - \mu_2$					

Conclusion: