In this lab we will explore the difference between paired and two independent samples t-procedures.

Note: this lab is long. To compensate, Lab 20 will be shorter and both labs will be due at the same time.

Researchers were interested in "proving" statistically that the pockets in women's jeans are a sham. For each brand, researchers recorded the measurements of the front and back pocket in 20 pairs of straight or boot-cut women's jeans and the same measurements for an equivalent-style pair of men's jeans from the same brand at the same store.

Assume that these 20 brands constitute a random sample of all brands of jeans that offer both men's and women's jeans (and that there are way more than 20 brands, making this a sample from a very large population).

Question #1 Given the above description, which type of t-test (one sample, matched pairs or two independent samples) is most appropriate? Why?

It would be most appropriate to use a matched pair design because you have two objects coming from a given source. In this case it is the men and women's jeans from each retailer.

This lab requires two datasets found on Titanium, straight_jeans and straight_jeans2. These two datasets contain the same information but recorded in two different ways. Import both files to R Studio.

Question #2 What is an observational unit in the straight_jeans dataset (that is, what does one row in the dataset represent)? What is an observational unit in the straight_jeans2 dataset?

For the dataset for <u>straight jeans</u>, each observational unit was either a numerical or categorical variable. This meant what kind of jeans we were taking notes on, and if they were for men or women.

For the dataset for <u>straight jeans2</u>, we see numerical variables. These are descriptive qualities of the jeans being observed and are purely measurements.

For this lab, we will focus on the width of the back pocket. The variable maxWidthBack represents the width (in cm) of a back pocket at the opening and the variable minWidthBack represents the width (in cm) of the same pocket at the point before it starts to taper.

In this part of the lab, we are going to learn code to "facet" histograms. What this means is to produce a single figure with multiple histograms, where each histogram represents the distribution of the same variable at a different level of a factor variable.

```
> library(ggplot2)
> ggplot(straight_jeans, aes(x = maxWidthBack)) +
+   geom_histogram(binwidth = 0.5, center = 12.25) +
+     facet_grid(menWomen ~ .) + labs(#finish this here)
```

Question #3 Paste below your figure (two total histograms). What differences do you notice between men's and women's jeans based on the variable *maxWidthBack*?

Men seem to be more normally distributed and women seem to be more skewed left.
Please see first attached graph after the lab questions.

Now, copy and edit the code to produce the same set of histograms for the variable *minWidthBack*.

Question #4 Paste below your figure (two total histograms). What differences do you notice between men's and women's jeans based on the variable *minWidthBack*?

There does not seem to be much variability when it comes how wide women's jeans are. The data is clustered in the middle whereas for men it looks more flush and uniform. Please see second graph after lab questions.

Next, we are going to perform a two-sample *t* hypothesis test (with significance level 0.05). Although the researchers want to prove that the pockets in women's jeans are a sham, we'll pretend like we don't know beforehand which gender should have narrower pockets and use a two-sided test instead (H_a: μ_{women} – μ_{men} ≠ 0 or H_a: μ_{men} – μ_{women} ≠ 0).

The code for this test uses an argument known as a <u>formula</u>, which we saw earlier in the lab when faceting. When used as the argument to a model or hypothesis test function, the *left hand side* of the formula represents the <u>response</u> variable and the *right hand side* represents the explanatory variable(s). Thus we perform the t-test:

```
> t.test(maxWidthBack ~ menWomen, data = straight_jeans)
```

Question #5 Paste below the output for the test for *maxWidthBack*. Report and interpret the t-statistic and p-value for this test.

Welch Two Sample t-test

data: maxWidthBack by menWomen
t = 1.3522, df = 37.272, p-value = 0.1845
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1793123 0.8993123
sample estimates:
  mean in group men mean in group women
        13.845   13.485

Given our p-value is 0.1845 and is greater than the significance level of 0.05, we fail to reject the null hypothesis. In this case it is that women's jeans are a sham. Going forward, we can still make that assumption. For our t-statistic we had a value of 1.3522.

Copy and edit the code to perform a two-sample t-test for *minWidthBack*.

Question #6 Paste below the output for the test for *minWidthBack*. Report and interpret the t-statistic p-value for this test.

Welch Two Sample t-test

data: minWidthBack by menWomen
t = 1.9244, df = 37.999, p-value = 0.06181
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.02857202 1.12857202
sample estimates:
  mean in group men mean in group women
        12.345   11.795

Given that our p-value is 0.06181, we can again fail to reject the null hypothesis. For our t-statistic it is a value of 1.9244.

Now let's look at what happens if we use matched pairs t-procedures instead. The straight_jeans2 dataset provides data in a format that will be easier to work with in the matched pairs framework.

First, we are going to create two new variables representing the paired differences.

```
> library(dplyr)
> jeans_diff <- straight_jeans2 %>% mutate(maxWidthBackDiff =
maxWidthBackMens – maxWidthBackWomens, minWidthBackDiff = minWidthBackMens
– minWidthBackWomens)# note: difference is mens - womens
```

Now create a histogram of each new variable.

```
> library(ggplot2)
> ggplot(jeans_diff, aes(x = maxWidthBackDiff)) +
+   geom_histogram(binwidth = 0.5, center = 0.25) +
+ labs(#finish this here)
```

Question #7 Paste below your histogram for *maxWidthBackDiff*. After matching for brand, does one gender appear to have wider pockets based on the maximum width?

**Please see the third page after the lab questions and responses.
It appears that men have wider pockets than women do.**

Question #8 Paste below your histogram for *minWidthBackDiff*. After matching for brand, does one gender appear to have wider pockets based on the minimum width?

**Please see the fourth page after the lab questions and responses.
Men will have wider pockets because there are more positive values.**

Now let's do the actual hypothesis test in the matched pairs framework.

```
> t.test(jeans_diff$maxWidthBackDiff)
```

Question #9 Paste below the output of this test. Report and interpret the t-statistic and p-value for this test.

        One Sample t-test

data: jeans_diff$maxWidthBackDiff
t = 1.2936, df = 19, p-value = 0.2113
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.2224556 0.9424556
sample estimates:
mean of x
   0.36

Given our p-value is 0.2113 and is greater than the significance level of 0.05, we fail to reject the null hypothesis. In this case, this means that men's pocket sizes are larger. For our t-statistic we had a value of 1.3522.

Copy and edit the code to do a paired t-test involving the minimum back pocket width ($minWidthBackDiff$).

Question #10 Paste below the output of this test. Report and interpret the t-statistic and p-value for this test.
        One Sample t-test

data: jeans_diff$minWidthBackDiff

t = 2.415, df = 19, p-value = 0.02598

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

 0.07332131 1.02667869

sample estimates:

mean of x

   0.55

Given our p-value is 0.02598 and is less than the significance level of 0.05, we reject the null hypothesis. In this case, this means the women's pocket sizes are inherently smaller than men's pocket sizes. For our t-statistic, we had a value of 2.415.

Question #11 Compare your answer to **Question #9** with your answer to **Question #5** and your answer to **Question #10** with your answer to **Question #6**. Did the values of the t-statistics change when we switched frameworks? What about the p-values?

The values did slightly change but not much.