# **Introduction:**

The independent-samples t-test is used to determine if a difference exists between the means of two independent groups on a continuous dependent variable. More specifically, it will let you determine whether the difference between these two groups is statistically significant. This test is also known by a number of different names, including the independent t-test, independent-measures t-test, between-subjects t-test and unpaired t-test.

For example, you could use the independent-samples t-test to determine whether (mean) salaries, measured in US dollars, differed between males and females (i.e., your dependent variable would be "salary" and your independent variable would be "gender", which has two groups: "males" and "females"). You could also use an independent-samples t-test to determine whether (mean) reaction time, measured in milliseconds, differed in under 21 year olds versus those 21 years old and over (i.e., your dependent variable would be "reaction time" and your independent variable would be "age group", split into two groups: "under 21 year olds" and "21 years old and over").

# **Assumptions of the Independent Sample t-test:**

When you choose to analyze your data using an independent-samples t-test, a critical part of the process involves checking to make sure that the data you want to analyze can actually be analyzed using this test. In fact, the independent-samples t-test has six assumptions that you have to consider.

* Assumption #1: You have **one dependent variable that is measured at the continuous (i.e., ratio or interval) level.**
* **Assumption #2:** Your independent variable is categorical with two separate groups of participants.
* Assumption #3: You have independence of observations, which means that each person’s score is not related to (or independent) from every other person’s score.
* Assumption #4: There should be no significant outliers in the two groups of your independent variable in terms of the dependent variable. Therefore, in this example, you need to investigate whether engagement has no outliers for each group of gender (i.e., you are testing whether the engagement score is outlier free for both the "Male" and "Female" groups). Due to the effect that outliers can have on your results, you have to choose whether to include these in your data.
* Assumption #5: Your dependent variable should be approximately normally distributed for each group of the independent variable. Also, it should be noted that if the distributions are all skewed in a similar manner (e.g., all moderately negatively skewed), this is not as troublesome as when compared to the situation where you have groups that have differently-shaped distributions (e.g., the distribution of Group A is moderately 'positively' skewed, whilst the distribution of Group B is moderately 'negatively' skewed). Therefore, in this example, you need to investigate whether engagement is normally distributed. There are many different methods available to test this assumption. Other methods include skewness and kurtosis values and histograms.
* Assumption #6: You have homogeneity of variances (i.e., the variance is equal in each group of your independent variable). The assumption of homogeneity of variances states that the population variance for each group of your independent variable is the same. If the sample size in each group is similar, violation of this assumption is not often too serious. However, if sample sizes are quite different, the independent-samples t-test is sensitive to the violation of this assumption. Either way, JASP uses Levene's test of equality of variances, and you can ask for two differently-calculated independent-samples t-tests, which will give you a valid result irrespective of whether you met or violated this assumption (i.e., JASP provides an independent-samples t-test that is calculated normally (with pooled variances) and another for when the assumption is violated that uses separate variances (i.e., non-pooled variances) and the Welch-Satterthwaite correction to the degrees of freedom).

Since it is not uncommon for the data you have collected to violate (i.e., fail) one or more of these assumptions, we show you different ways to proceed. This could include (a) making corrections to your data so that it no longer violates the assumptions, (b) using an alternative statistical test, or (c) proceeding with your analysis even when your data violates certain assumptions.

## **Null and Alternative Hypotheses:**

The null hypothesis is:

H0: µ = µ Scores are equal in the population. There is no difference between Groups/Conditions.

The alternative hypothesis is:

HA: µ ≠ µ Scores are not equal in the population. There is a difference between Groups/Conditions.

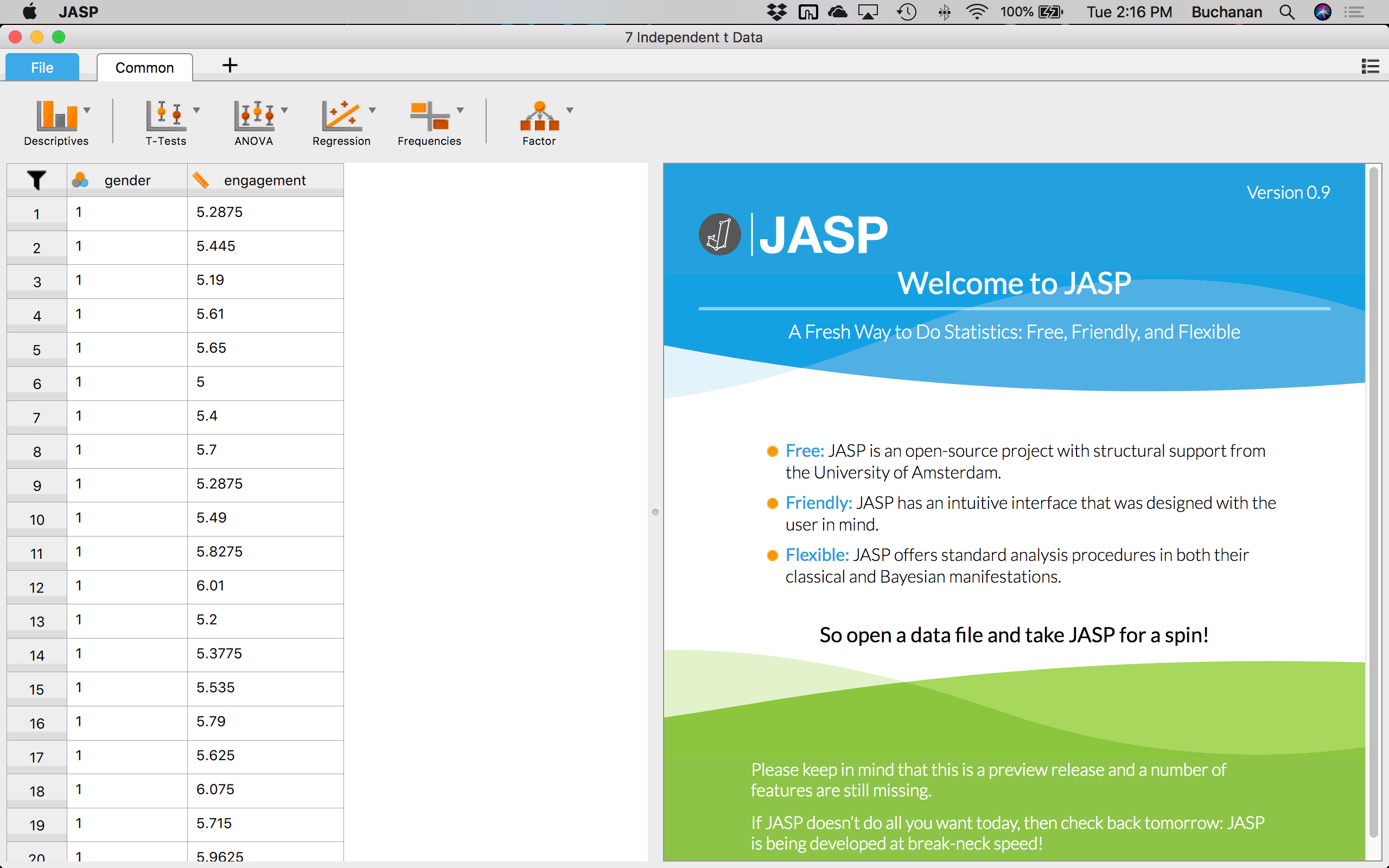
## **Example:**

An Advertising Agency is commissioned to create a TV advert to promote a new product. Since the product is designed for men and women, the TV advert has to appeal to men and women equally. Before the company that commissioned the Advertising Agency spends $250,000 across a number of TV networks, it wants to make sure that the TV advert created by the Advertising Agency appeals equally to men and women. More specifically, the company wants to know whether the way that men and women engage with the TV advert is the same. To achieve this, the TV advert is shown to 20 men and 20 women, who are then asked to fill in a questionnaire that measures their engagement with the advertisement. The questionnaire provides an overall engagement score.

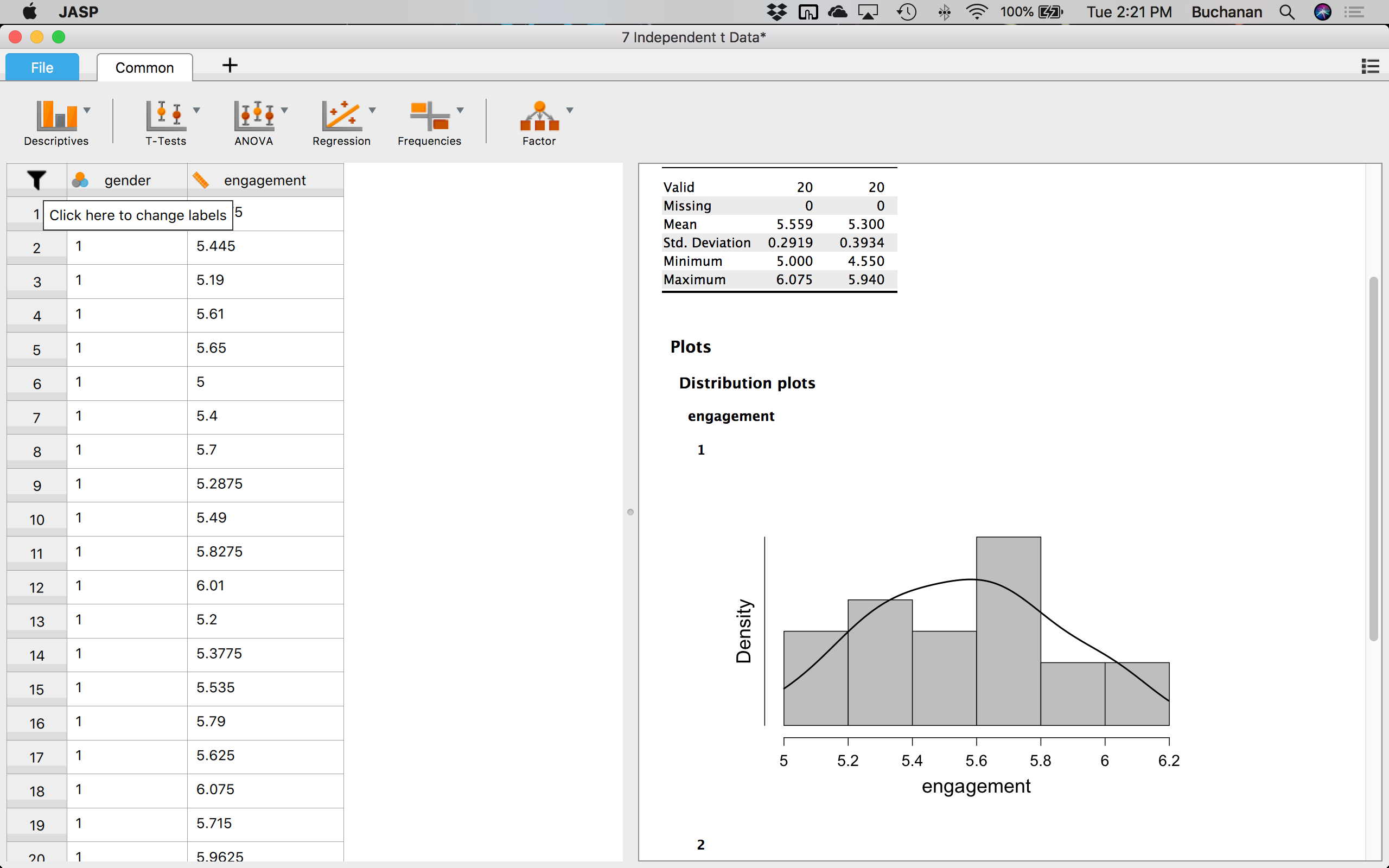
This overall engagement score is the dependent variable, which we have labelled engagement in SPSS Statistics. Our independent variable, which we have labelled gender in SPSS Statistics, contains two groups: "Male" and "Female". In variable terms, the Advertising Agency would like to know if the independent variable, gender, has an effect on the dependent variable, engagement (or expressed another way, if there are differences in engagement between levels of gender). In other words, is the mean engagement score different for males and females? Since the Advertising Agency needs the advertisement to be similarly engaging, they hope there is no difference!

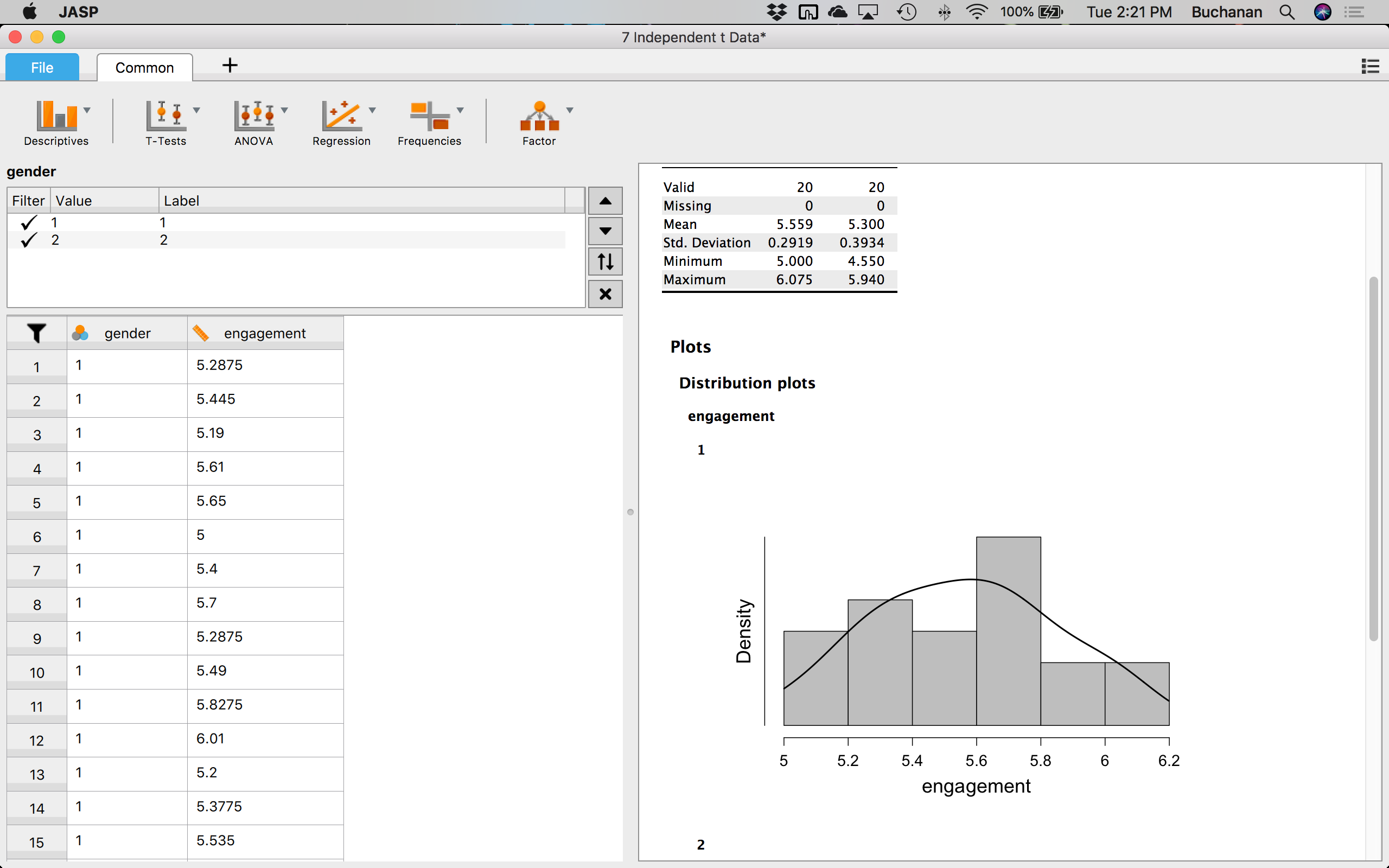
To get started, open the dataset for this example in JASP. Remember, you can always use the previous help guides for greater detail in case you do not remember how to do something.

File 🡪 Open 🡪 Computer 🡪 Browse 🡪 Pick the Independent t Data.

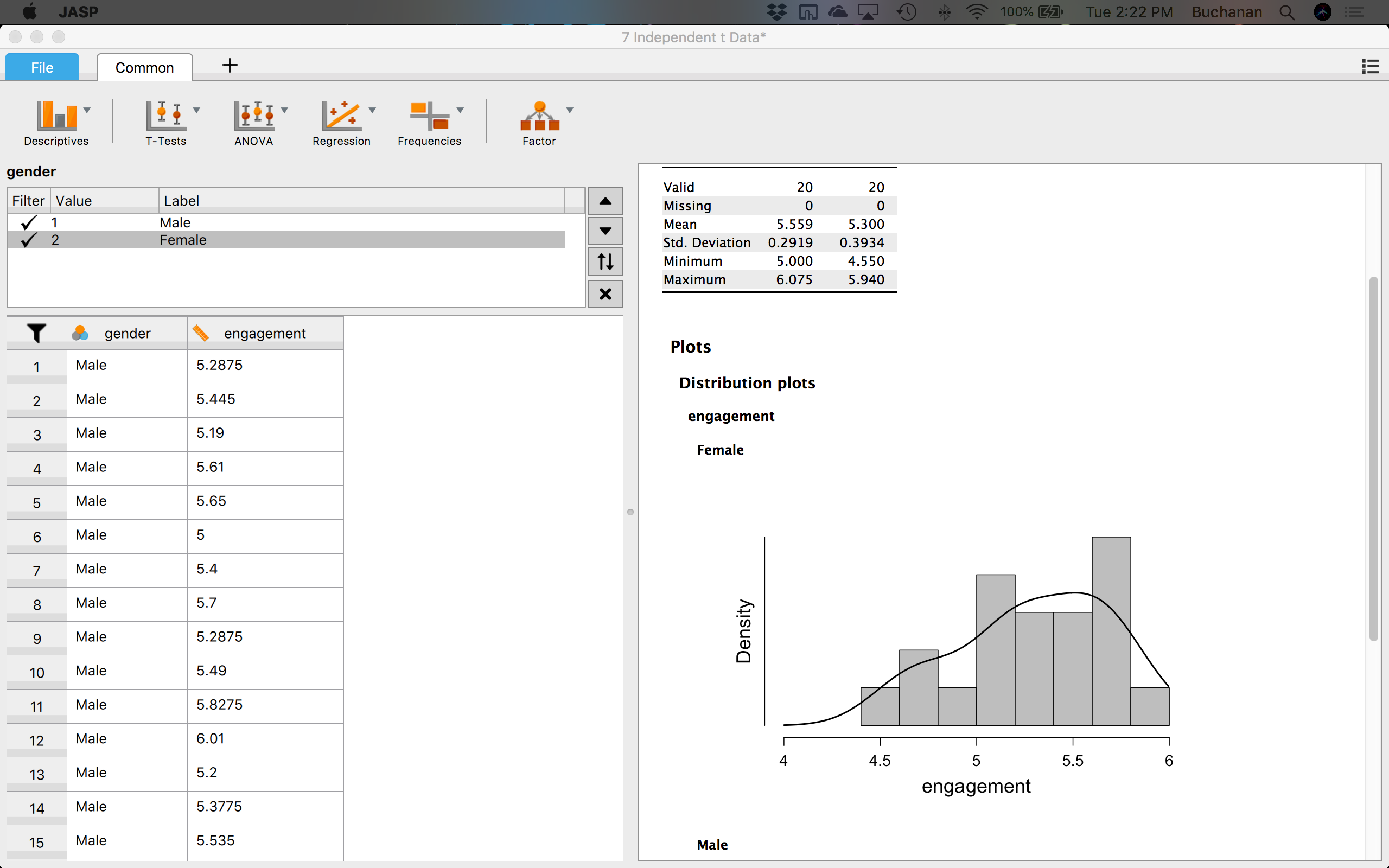


## Before we get started, we can fix our gender column. At the moment, it says 1 and 2. We could go into Excel and rename all of those to Male and Female. Or we can use the JASP option to give them labels. If you hover over gender, you will see a note saying “click here to change labels”. Click on it.





We want to change group 1 to be Male. Click on the 1 under the Label column to start typing. Then do the same thing and edit the 2 under the Label column to be Female.



As you edit these labels, the data will update to say Male and Female under gender. To close the window click the on the X button .

## **Check your assumptions:**

**Is the dependent variable at least scale (ratio or interval)?**

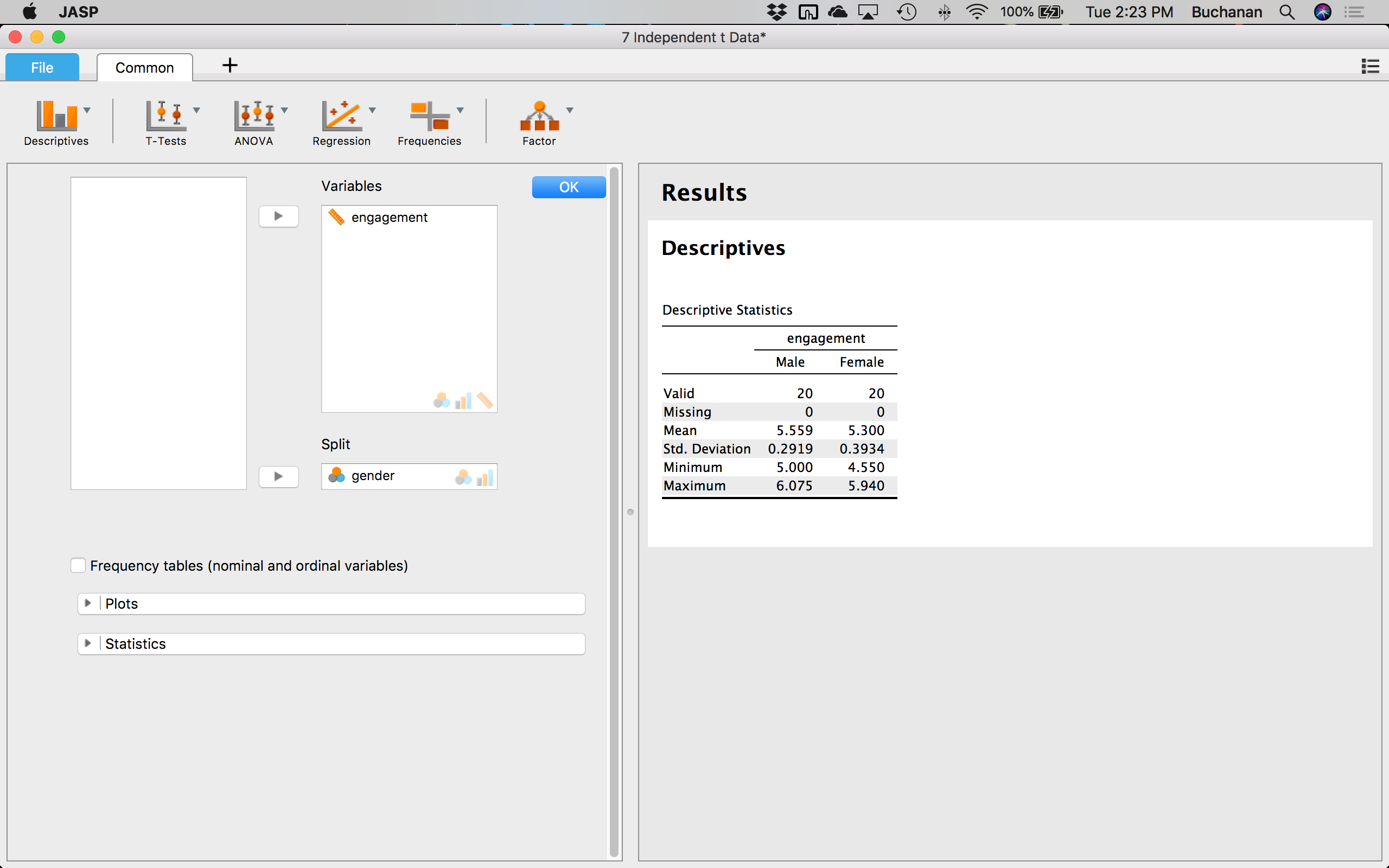
Yes, we are using ratio style data.

**Are there any outliers in the sample?**

To examine if any data might be considered an outlier, we can use the Descriptives  options you learned about previously. Click Descriptives 🡪 Descriptive Statistics.



In this window, we want to click on engagement and click the arrow  to move it over to the right hand side under Variables. In this window, you will see a new  Split option. This option will allow you to separate out descriptives and outlier analysis by group! Move the gender variable into the Split box.



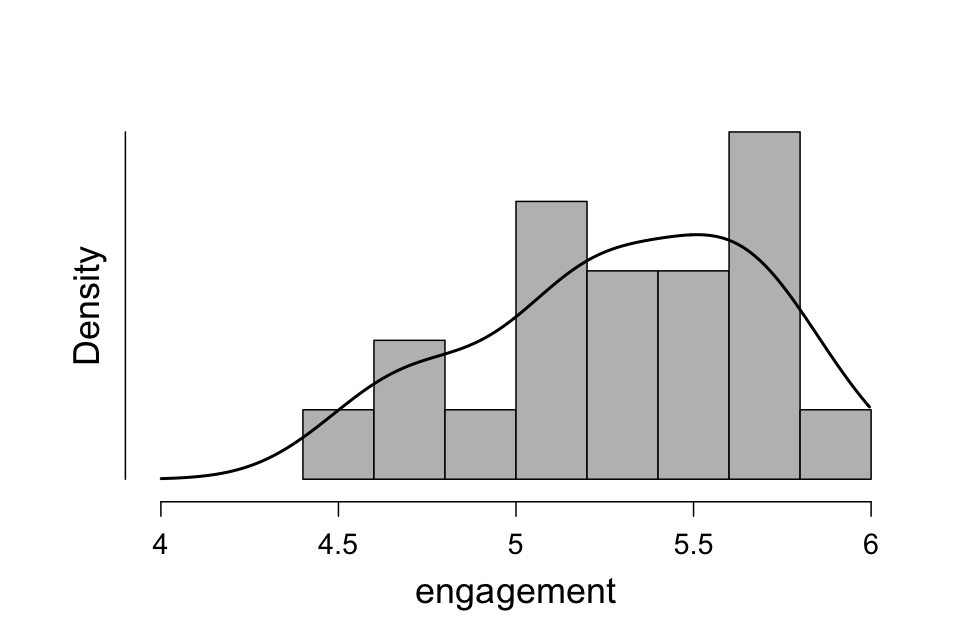
Click on the plots options:  to see more available options.

Here we can look at two different options to see if any participants scores are very different from other participants scores. First, click on Distribution plots. 

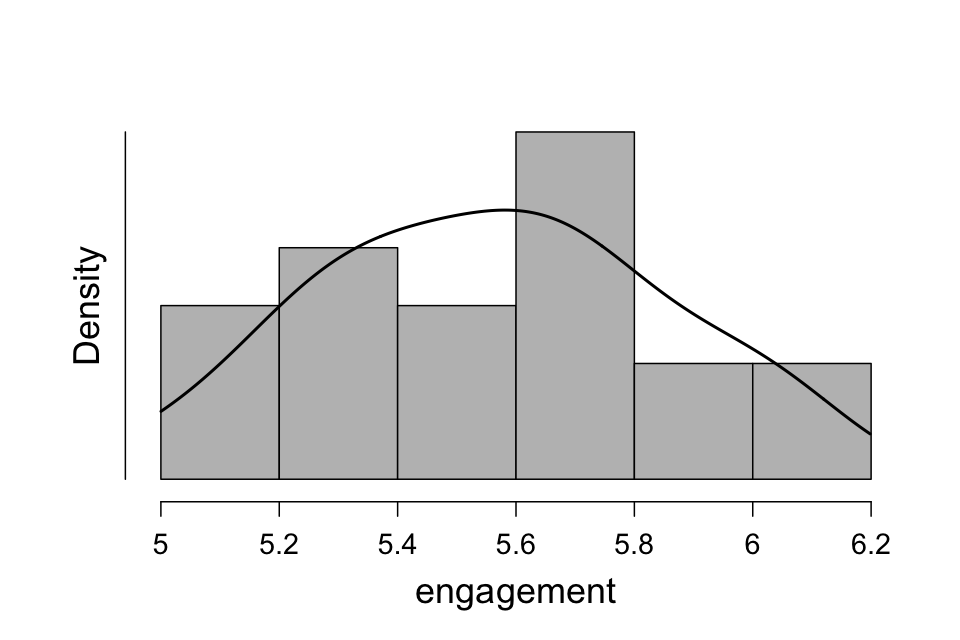
#### Distribution plots

##### engagement

###### Female



###### Male

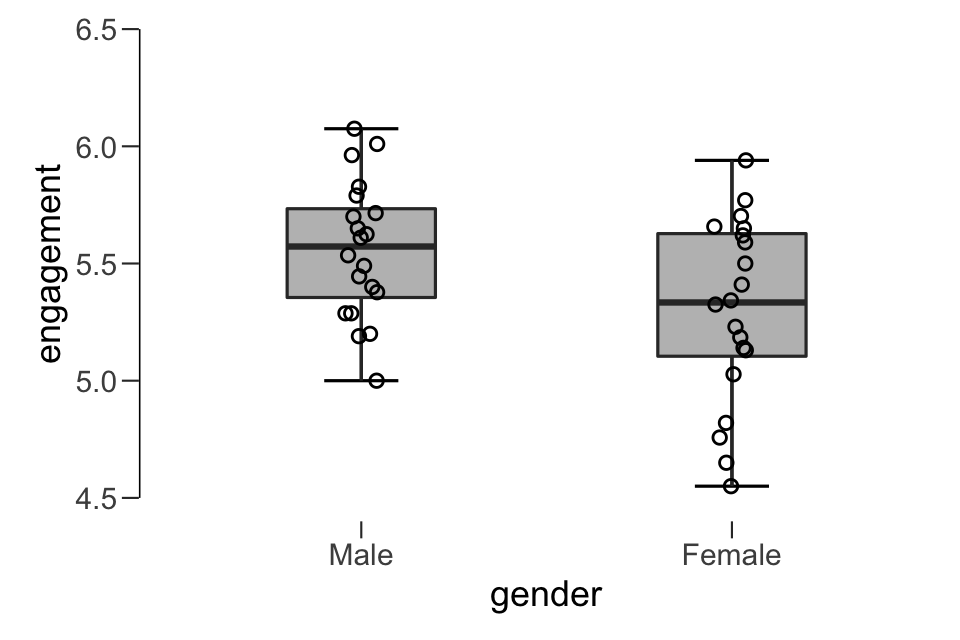


Two histograms will appear – one for each group with the labels we added earlier to help us distinguish which graph is which (notice that it ran them in alphabetical order!). We can tell that most people are fairly close together, so probably not any outliers.

Another option would be to select Box Plots , Label Outliers , and Jitter Element . You will see outliers labeled with a special symbol, and Jitter Element allows you to see all the participants scores as dots on the plot.

##### 

##### engagement

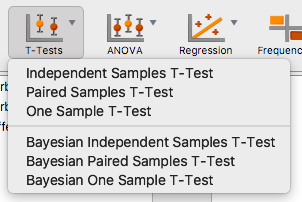


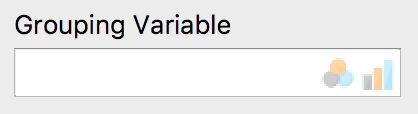
##### If we had outliers, they would be outside the top and bottom lines and would be marked with a star.

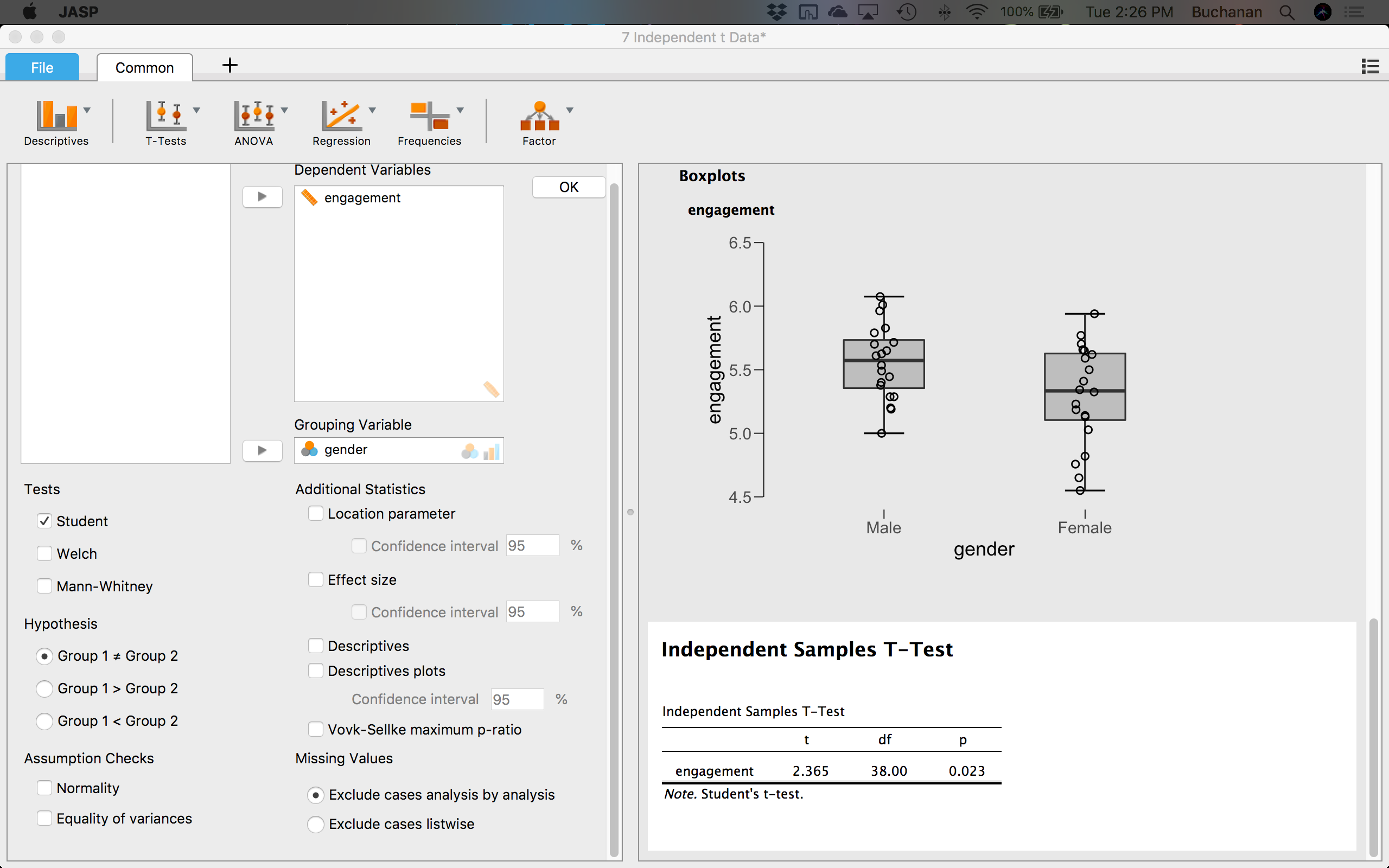
**Is the dependent variable normally distributed?**

We can view the histogram created earlier to look at if the data appears normal, but we might also consider using the Shapiro-Wilk test to determine if the data is normal. To get this test, we need to run the actual t-test to see that output.

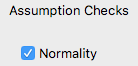
Click on t-tests  🡪 Independent Samples t Test



In this window, we want to click on engagement and click the arrow  to move it over to the right hand side under Variables. Then click on gender and move it into the Grouping Variable box. 



To get the normality assumption test, click on Normality, under Assumptions:



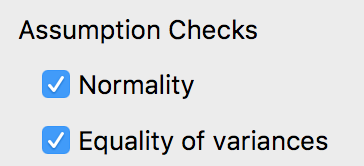
### Assumption Checks

| **Test of Normality (Shapiro-Wilk)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | **W** | | **p** | |
| engagement |  | Male |  | 0.983 |  | 0.970 |  |
|  |  | Female |  | 0.961 |  | 0.560 |  |
|  | | | | | | | |
| Note.  Significant results suggest a deviation from normality. | | | | | | | |

The Shapiro-Wilk test ran for each group separately, so we can tell if the assumptions were met for each sample. We see that our data is normally distributed because *p* > .05 for both groups. Remember, t-tests are robust to violations of normality, especially with larger sample sizes.

**Do you have homogeneity of variances?**

Underneath normality, you will now see a new option for Equality of Variances, and click that option to view Levene’s test.

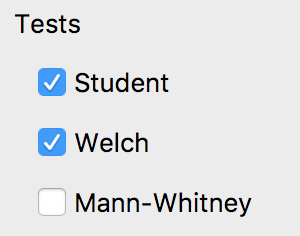


| **Test of Equality of Variances (Levene's)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **F** | | **df** | | **p** | |
| engagement |  | 1.922 |  | 1 |  | 0.174 |  |
|  | | | | | | | |

An important assumption of the independent-samples t-test is that the two group's variances are equal in the population. Failure to adhere to this assumption (i.e., if variances are unequal), will generally increase the chance of making a Type I error. Equality of variances is often referred to as homogeneity of variances. The same rules apply for Levene’s as the Shapiro-Wilk test, we do not want to find a significant result (*p* < .05). In this example, the significance value is ".174" (i.e., p = .174). If the population variance of both groups is equal, this test will return a p-value greater than 0.05 (i.e., p > .05), indicating that you have met the assumption of homogeneity of variances. However, if the test returns a p-value less than 0.05 (p < .05), the population variances are unequal and you have violated the assumption of homogeneity of variances. In this example, the population variances of the engagement scores for both groups are equal because p = .174 (i.e., p > .05). Therefore, the assumption of homogeneity of variances is met.

You interpret your results after running an independent-samples t-test depends on whether your data met or violated the assumption of homogeneity of variances.

* Homogeneity of variances was met: If your data has met the assumption of homogeneity of variances, you simply need to interpret the 'standard' independent-samples t-test output.
* Homogeneity of variances was violated: If your data has violated the assumption of homogeneity of variances, you can still continue with your analysis uisng a modified t-test. This modified t-test, often referred to as the **unequal variance t-test**, **separate variances t-test** or **Welch t-test** after its creator (Welch, 1947), can accommodate unequal variances and still deliver a valid test result. To get the Welch test, simply click on Welch under tests:



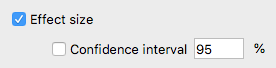
Here’s what that will look like (not shown below because we met the assumption):

| **Independent Samples T-Test** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Test** | | **Statistic** | | **df** | | **p** | | **Cohen's d** | |
| engagement |  | Student |  | 2.365 |  | 38.00 |  | 0.023 |  | 0.748 |  |
|  |  | Welch |  | 2.365 |  | 35.05 |  | 0.024 |  | 0.748 |  |
|  | | | | | | | | | | | |

## **The t-test and effect size:**

Now, we can finish out running the t-test by clicking on a few more options.

You will want to add effect size (*d*):



You can also add the descriptive statistics of each measurement individually by clicking on descriptives: .

## **Interpretation and Reporting:**

**Independent Samples T-Test**

| **Independent Samples T-Test** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **t** | | **df** | | **p** | | **Cohen's d** | |
| engagement |  | 2.365 |  | 38.00 |  | 0.023 |  | 0.748 |  |
|  | | | | | | | | | |
| *Note.*  Student's t-test. | | | | | | | | | |

When the assumption of homogeneity of variances is met, you can interpret the results from the standard t-test that uses pooled variances in its calculations and requires no modification to the degrees of freedom. This is, for the most part, advantageous when we are trying to detect the mean difference between groups.

You are presented with the observed t-value (the "**t**" column), the degrees of freedom (the "**df**" column), and the statistical significance (p-value). If p < .05, this means that the mean difference between the two groups is statistically significant. Alternatively, if p > .05, you do not have a statistically significant mean difference between the two groups. In this example, p = .023 (i.e. p < .05). Therefore, it can be concluded that males and females have statistically significantly different mean engagement scores. Or, phrased another way, the mean difference in engagement score between males and females is statistically significant. What this result means is that there is a 23 in 1,000 chance (2.3%) of getting a mean difference at least as large as the one obtained if the null hypothesis was true (the null hypothesis stating that there is no difference between the group means). Remember, the independent-samples t-test is testing whether the means are equal in the population.

Last, we want to add and interpret our effect size, since our *p* value does not tell us the importance of the results. The *d* value here is 0.75, which is close to large effect size given traditional interpretations. You might report all this information in APA style like this:

There was a statistically significant difference in engagement scores between males and females, with males scoring higher than females, *t*(38) = 2.37, *p* = .023, *d* = 0.75.

**Assumption Checks**

| **Test of Normality (Shapiro-Wilk)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | **W** | | **p** | |
| engagement |  | Male |  | 0.983 |  | 0.970 |  |
|  |  | Female |  | 0.961 |  | 0.560 |  |
|  | | | | | | | |
| *Note.*  Significant results suggest a deviation from normality. | | | | | | | |

Here’s how you might report Shapiro-Wilk in APA style:

Engagement scores for each level of gender were normally distributed, as assessed by Shapiro-Wilk's test (Male p = .970, Female *p* = .560).

| **Test of Equality of Variances (Levene's)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **F** | | **df** | | **p** | |
| engagement |  | 1.922 |  | 1 |  | 0.174 |  |
|  | | | | | | | |

Here’s how you might report Levene’s test in APA style:

There was homogeneity of variances for engagement scores for males and females, as assessed by Levene's test for equality of variances (*p* = .174).

**Descriptives**

| **Group Descriptives** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Group** | | **N** | | **Mean** | | **SD** | | **SE** | |
| engagement |  | Male |  | 20 |  | 5.559 |  | 0.292 |  | 0.065 |  |
|  |  | Female |  | 20 |  | 5.300 |  | 0.393 |  | 0.088 |  |
|  | | | | | | | | | | | |

This table includes the sample size (N), mean, standard deviation, and standard error for your two groups – "Male" and "Female" – which will help you get a "feel" for your data and will be used when you report your results. You can make an initial interpretation of your data, and you might report that in APA style:

There were 20 male and 20 female participants. The advertisement was more engaging to male viewers (*M* = 5.56, *SD* = 0.29) than female viewers (*M* = 5.30, *SD* = 0.39).

## **Reporting All Together:**

There were 20 male and 20 female participants. An independent-samples t-test was analyzed to determine if there were differences in engagement to an advertisement between males and females. There were no outliers in the data, as assessed by inspection of a boxplot. Engagement scores for each level of gender were normally distributed, as assessed by Shapiro-Wilk's test (Male p = .970, Female *p* = .560), and there was homogeneity of variances, as assessed by Levene's test for equality of variances (*p* = .174). The advertisement was more engaging to male viewers (M = 5.56, SD = 0.35) than female viewers (M = 5.30, SD = 0.35). This was indicated by a large and significant difference between groups, *t*(38) = 2.37, *p* = .023, *d* = 0.75.