# **Introduction:**

The single sample t-test is used to determine whether the mean of a single sample is significantly different from some predetermined population mean. The participants are one group of people who have been tested on a single dependent variable. For example, you might examine if the incoming freshman class (the sample) has a different mean than the overall university average on GPA (the population mean).

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

In order to run a single sample t-test, there are several assumptions that need to be considered. You will see these assumptions again throughout the semester!

* Assumption #1: You have **one dependent variable** that is measured at the **continuous** (i.e., **ratio** or **interval**) level. Examples of **continuous variables** include revision time (measured in hours), intelligence (measured using IQ score), exam performance (measured from 0 to 100), weight (measured in kg), and so forth.
* Assumption #2: The distribution of the dependent variable should be approximately normally distributed. The assumption of normality is necessary for statistical significance testing using any type of t-test. However, t-tests are often considered "robust" to violations of normality. This means that some violation of this assumption can be tolerated and the test will still provide valid results. Therefore, you will often hear of this test only requiring *approximately* normal data. Furthermore, as sample size increases, the distribution can be very non-normal and, thanks to the Central Limit Theorem, t-tests can still provide valid results.
* Assumption #3: There should be no significant outliers in the sample. If there are any scores that are unusual for your sample, in that their value is extremely small or large compared to the other scores, these scores are called outliers (e.g., 8 participants in a group scored between 60-75 out of 100 in a difficult math test, but one participant scored 98 out of 100). Outliers can have a large negative effect on your results because they can exert a large influence (i.e., change) on the mean and standard deviation for that group, which can affect the statistical test results. Outliers are more important to consider when you have smaller sample sizes, as the effect of the outlier will be greater.

## **Null and Alternative Hypotheses:**

Sometimes, you will be required in your work to explicitly state the null and alternative hypotheses, and to then state which was accepted or rejected at the end of the experiment. A common null hypothesis for a single sample t-test is:

H0: The sample mean will the same as the population (*M* = µ).

And the research hypothesis is:

H1: The sample mean will be different from the population mean (*M* ≠ µ).

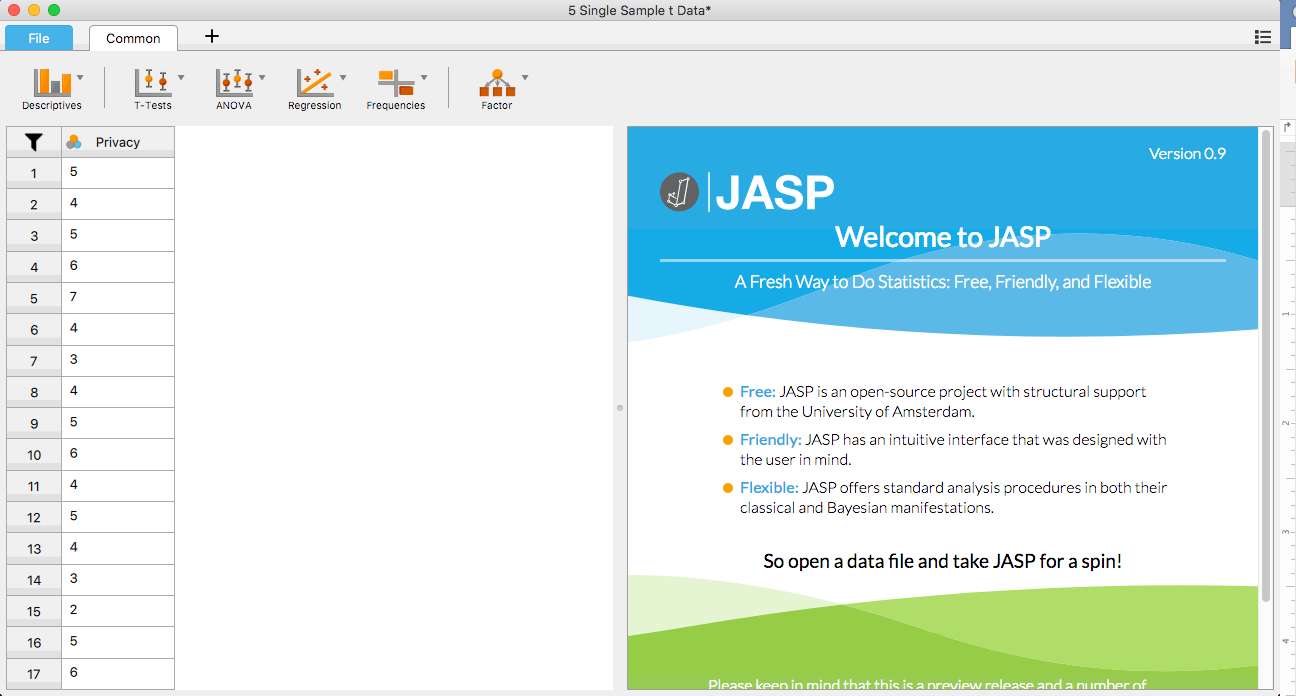
These guides will mostly cover two-tailed hypothesis testing, where you do not pick a direction that you expect the test to be found (i.e., the scores will be higher than the population mean). Please be aware that your instructor may choose to have you examine a one-tailed or directional test. Depending on your results, you can then state whether to reject, or fail to reject, the null hypothesis, and whether to reject or accept the research hypothesis. You will be shown how to do this later in this guide.

## **Example:**

A researcher has gathered data on the recent issues with privacy and the use of social media. 40 participants were asked to rate their concern with online privacy on a 1 (*not at all concerned*) to 7 (*very much concerned*) scale. This data was then compared to a national average concern rating of around a neutral score where µ = 4.12.

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 Single Sample t Data.



## **Check your assumptions:**

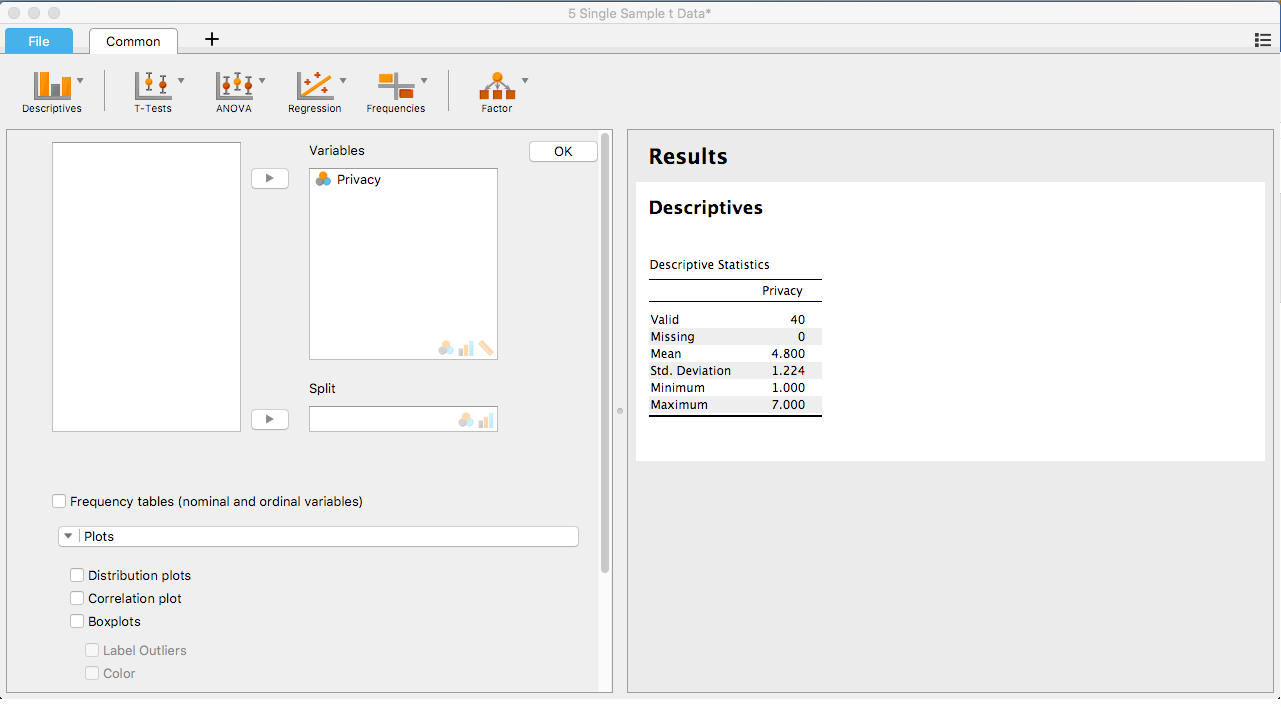
Is the dependent variable at least scale (ratio or interval)? Yes, we are using interval style data.

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

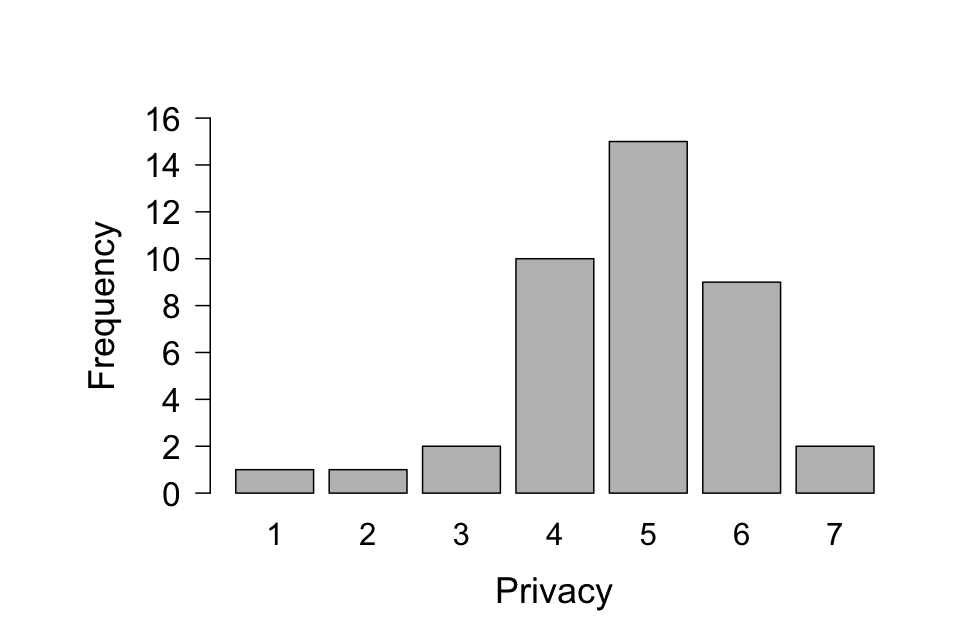
To examine if any data might be considered an outlier, we can use the Descriptives  options you learned about previously. Click Descriptives 🡪 Descriptive Statistics.



In this window, we want to click on Privacy and click the arrow  to move it over to the right hand side under Variables. Click on the plots options:  to see more available options.



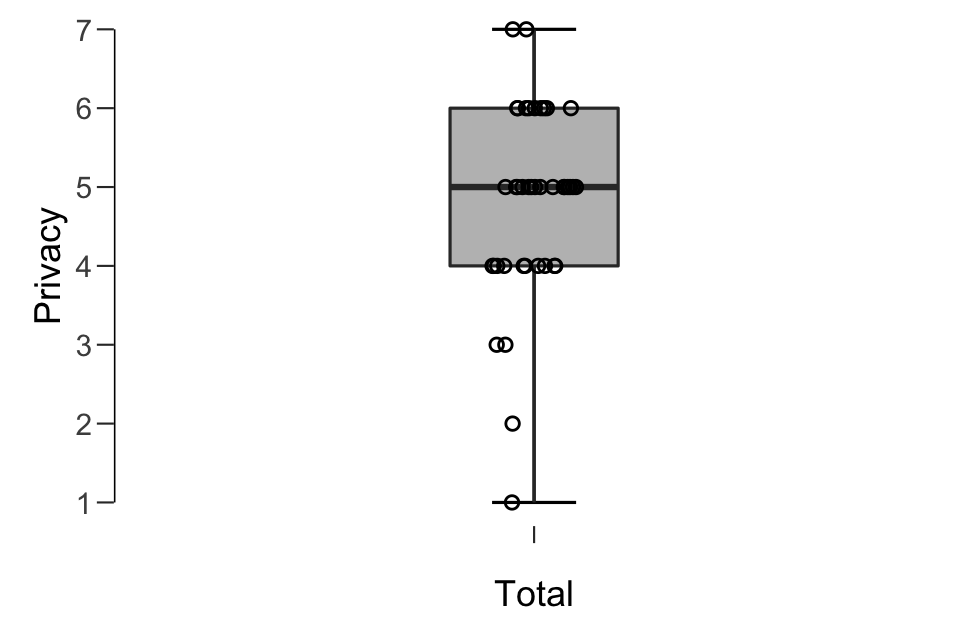
Here we can look at two different options to see if any participants scores are very different from other participants scores. First, click on Distribution plots. 



A histogram will appear. We can tell that most participants scores are between 4 and 6, with a few in the very low range of 1 and 2. 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.

#### Boxplots

##### Privacy



Quartile - 1.5 \* IQR

Quartile + 1.5 \* IQR

Third Quartile

First Quartile

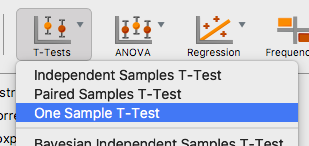
Median

A boxplot indicates the following parts (labeled on top of the diagram above, which you will not see in JASP). Remember that IQR is the interquartile range or the third quartile minus the first quartile. 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.

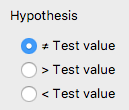
Click on t-tests  🡪 One sample t-test



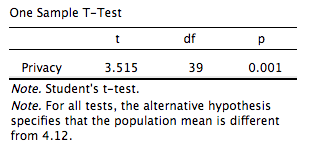
In this window, we want to click on Privacy and click the arrow  to move it over to the right hand side under Variables. The default population mean is 0 in JASP, which is listed under Test value: .

You will want to change the zero to the population mean from the example, which is 4.12 . **This step is very important, so do it or your *t*-test values will be wrong!**

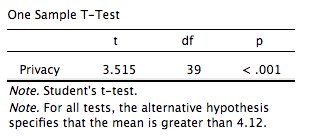
Under Hypothesis, we can pick the type of test we are interested in (two-tailed/non-directional or one-tailed/directional).



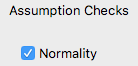
When you pick a value here 🡪 notice that it changes the test options on the right underneath the *t*-test box:



Versus:



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



### Assumption Checks

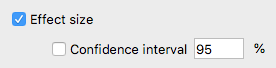
| **Test of Normality (Shapiro-Wilk)** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **W** | | **p** | |
| Privacy |  | 0.905 |  | 0.003 |  |
|  | | | | | |
| Note.  Significant results suggest a deviation from normality. | | | | | |

If your data is normally distributed (i.e., the assumption of normality is met), the significance level (value in the *p* column) should be more than .05 (i.e., *p* > .05). If your data is not normally distributed (i.e., the assumption of normality is violated), the significance level will be less than .05 (i.e., *p* < .05). The null hypothesis of the Shapiro-Wilk test is that your data's distribution is equal to a normal distribution and the alternative hypothesis is that your data's distribution is not equal to a normal distribution. Thus, if you reject the null hypothesis (*p* < .05), this means that your data's distribution is not equal to a normal distribution and if you fail to reject the null hypothesis, your data is normally distributed.

Here we see that our data is NOT normally distributed because *p* < .05. However, as mentioned earlier, *t*-tests are robust to violations of normality, especially with larger sample sizes.

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

Now, we can finish out running the t-test by clicking on a few more options. On the right hand side, you will want to add effect size (*d*):



You can also add the descriptive statistics by clicking on descriptives: .

## **Interpretation and Reporting:**

**One Sample T-Test**

| **One Sample T-Test** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **t** | | **df** | | **p** | | **Cohen's d** | |
| Privacy |  | 3.515 |  | 39 |  | 0.001 |  | 0.556 |  |
|  | | | | | | | | | |
| *Note.*  Student's t-test. | | | | | | | | | |
| *Note.*  For the Student t-test, effect size is given by Cohen's *d* . | | | | | | | | | |
| *Note.*  For all tests, the alternative hypothesis specifies that the population mean is different from 4.12. | | | | | | | | | |

In the first box, you will find your *t* value that tells you the test statistic for the difference between the sample and population mean (remember that t-values are the difference between mean and population mean divided by the standard error 🡪 it’s the difference on the top and the error on the bottom).

Second are the degrees of freedom (df, here sample N-1), followed by the *p* value. Let’s say you set your alpha value (or Type I error rate) at *p* < .05. You can also use other values, such as *p* < .10 or *p* < .01, but *p* < .05 is a very popular value. If *p* < .05 in our results, this means that the sample and the population are different and this test is statistically significant. Alternatively, if p > .05, you do not have a statistically significant mean difference between the sample and population mean. In this example, the statistical significance level is stated as .001, which means p = .001 (i.e. p < .05).

It is important to remember that the level of significance (*p*-value) does not indicate the strength or importance of the mean difference between the related groups, only the likelihood of a mean difference as large or larger as the one you observed, given that the null hypothesis is true. For example, if this example had produced a p-value of .012 (*p* = .012), this does not mean it is twice as 'strong' or 'important' as *p* = .024. In layman's terms, the *p*-value is simply trying to inform you whether the mean difference in the sample and population you studied is not a 'fluke' and it really is likely that you would expect to see a mean difference like the one in your study in the population (not just in your sample).

The last value reported here is the *d* statistic, which gives you an idea of the effect size or magnitude of the difference in means. An effect size is an attempt to provide a measure of the practical significance of the result. The importance of the value of Cohen's *d* (as reported by Cohen (1998)) is as follows:

|  | **Effect Size** | **Strength** |
| --- | --- | --- |
|  | .2 | small |
|  | .5 | medium |
|  | .8 | large |
| Table: Interpretation of values of Cohen's *d*. | | |

As the effect size, *d*, is 0.56 you can conclude that there is a medium effect. However, one of the major weaknesses of using effect sizes is that guidelines indicating the importance of the effect size are subject-specific and there are not, as yet, comprehensive guidelines for interpretation of the strength of an effect size. Moreover, there is usually more than one measure of effect size (i.e., different types and calculations of effect size). You should report the value of the effect size with the result of the hypothesis test.

To report this in APA style, you might use:

The participants rated their privacy concerns as significantly different than the population, *t*(39) = 3.52, *p* = .001, *d* = 0.56.

Remember the following rules:

* Format: *t*(df) = t value, *p* = p value, *d* = effect size value.
* If p is listed as .000, then use *p* < .001.
* Two decimals for all statistics except p values, which use three decimals.

**Assumption Checks**

| **Test of Normality (Shapiro-Wilk)** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **W** | | **p** | |
| Privacy |  | 0.905 |  | 0.003 |  |
|  | | | | | |
| *Note.*  Significant results suggest a deviation from normality. | | | | | |

As described earlier, we would note here that we did not meet the assumption of normality because *p* < .05. Remember that assumptions tests are backwards – you do not want them to be significant.

To report this in APA style, you might use:

The assumption of normality was not met, as assessed by Shapiro-Wilk's test (*p* = .003).

**Descriptives**

| **Descriptives** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **N** | | **Mean** | | **SD** | | **SE** | |
| Privacy |  | 40.00 |  | 4.800 |  | 1.224 |  | 0.193 |  |
|  | | | | | | | | | |

Last, we always want to include our descriptive statistics for the study. We might report these like this:

The participants in this study rated their concerns about privacy as *M* = 4.80 (*SD* = 1.22).

## **Reporting All Together:**

A single sample t-test was used to determine if participants (*N* = 40) privacy concerns were different than previous studies on this topic (µ = 4.12). No outliers were found when examining a boxplot. The participants in this study rated their concerns about privacy as *M* = 4.80 (*SD* = 1.22). The assumption of normality was not met, as assessed by Shapiro-Wilk's test (*p* = .003). The participants rated their privacy concerns as significantly different than the population, *t*(39) = 3.52, *p* = .001, *d* = 0.56.