

What is the Independent Samples t Test and how is it used to compare two sets of data?

Authored by
stats writer

June 24, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the Independent Samples t Test and how is it used to compare two sets of data?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=149913>

The Independent Samples t Test is a statistical method used to compare the means of two independent groups of data. It is commonly used to determine if there is a significant difference between the means of two groups. This test is appropriate when the two groups being compared are unrelated and have different participants. It is used to analyze numerical data and determine if the difference between the two groups is due to chance or if there is a significant difference between the two populations. This test is often used in research studies to determine the effectiveness of a treatment or intervention by comparing the outcomes of a treatment group to a control group. It provides a reliable and objective way to measure the significance of differences between two groups and is an important tool in data analysis.

Independent Samples t Test

The Independent Samples t Test compares the means of two independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. The Independent Samples t Test is a parametric test.

This test is also known as:

Independent t Test
Independent Measures t Test
Independent Two-sample t Test
Student t Test
Two-Sample t Test
Uncorrelated Scores t Test
Unpaired t Test
Unrelated t Test

The variables used in this test are known as:

Dependent variable, or test variable
Independent variable, or grouping variable

Common Uses

The Independent Samples t Test is commonly used to test the following:

Statistical differences between the means of two groups
Statistical differences between the means of two interventions
Statistical differences between the means of two change scores

Note: The Independent Samples t Test can only compare the means for two (and only two) groups. It cannot make comparisons among more than two groups. If you wish to compare the means across more than two groups, you will likely want to run an ANOVA.

Data Requirements

Your data must meet the following requirements:

Dependent variable that is continuous (i.e., interval or ratio level)
Independent variable that is categorical *and* has exactly two categories
Cases that have values on both the dependent and

independent variables Independent samples/groups (i.e., independence of observations)

There is no relationship between the subjects in each sample. This means that:

Subjects in the first group cannot also be in the second group No subject in either group can influence subjects in the other group No group can influence the other group Violation of this assumption will yield an inaccurate p value Random sample of data from the population Normal distribution (approximately) of the dependent variable for each group

Non-normal population distributions, especially those that are thick-tailed or heavily skewed, considerably reduce the power of the test Among moderate or large samples, a violation of normality may still yield accurate p values Homogeneity of variances (i.e., variances approximately equal across groups)

When this assumption is violated and the sample sizes for each group differ, the p value is not trustworthy. However, the Independent Samples t Test output also includes an approximate t statistic that is not based on assuming equal population variances. This alternative statistic, called the Welch t Test statistic¹, may be used when equal variances among populations cannot be assumed. The Welch t Test is also known as Unequal Variance t Test or Separate Variances t Test. No outliers

Note: When one or more of the assumptions for the Independent Samples t Test are not met, you may want to run the nonparametric Mann-Whitney U Test instead.

Researchers often follow several rules of thumb:

Each group should have at least 6 subjects, ideally more. Inferences for the population will be more tenuous with too few subjects. A *balanced* design (i.e., same number of subjects in each group) is ideal. Extremely unbalanced designs increase the possibility that violating any of the requirements/assumptions will threaten the validity of the Independent Samples t Test.

¹ Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. *Biometrika*, 34(1-2), 28-35.

Hypotheses

The null hypothesis (H₀) and alternative hypothesis (H₁) of the Independent Samples t Test can be expressed in two different but equivalent ways:

H₀: $\mu_1 = \mu_2$ ("the two population means are equal")

H₁: $\mu_1 \neq \mu_2$ ("the two population means are not equal")

OR

H₀: $\mu_1 - \mu_2 = 0$ ("the difference between the two population means is equal to 0")

H1: $\mu_1 - \mu_2 \neq 0$ ("the difference between the two population means is not 0")

where μ_1 and μ_2 are the population means for group 1 and group 2, respectively. Notice that the second set of hypotheses can be derived from the first set by simply subtracting μ_2 from both sides of the equation.

Levene's Test for Equality of Variances

Recall that the Independent Samples *t* Test requires the assumption of *homogeneity of variance* -- i.e., both groups have the same variance. SPSS conveniently includes a test for the homogeneity of variance, called **Levene's Test**, whenever you run an independent samples *t* test.

The hypotheses for Levene's test are:

H0: $\sigma_1^2 - \sigma_2^2 = 0$ ("the population variances of group 1 and 2 are equal")

H1: $\sigma_1^2 - \sigma_2^2 \neq 0$ ("the population variances of group 1 and 2 are not equal")

This implies that if we reject the null hypothesis of Levene's Test, it suggests that the variances of the two groups are not equal; i.e., that the homogeneity of variances assumption is violated.

The output in the Independent Samples Test table includes two rows: **Equal variances assumed** and **Equal variances not assumed**. If Levene's test indicates that the variances are equal across the two groups (i.e., p-value large), you will rely on the first row of output, **Equal variances assumed**, when you look at the results for the actual Independent Samples *t* Test (under the heading *t*-test for Equality of Means). If Levene's test indicates that the variances are not equal across the two groups (i.e., p-value small), you will need to rely on the second row of output, **Equal variances not assumed**, when you look at the results of the Independent Samples *t* Test (under the heading *t*-test for Equality of Means).

The difference between these two rows of output lies in the way the independent samples *t* test statistic is calculated. When equal variances are assumed, the calculation uses pooled variances; when equal variances cannot be assumed, the calculation utilizes un-pooled variances and a correction to the degrees of freedom.

Test Statistic

The test statistic for an Independent Samples *t* Test is denoted *t*. There are actually two forms of the test statistic for this test, depending on whether or not equal variances are assumed. SPSS produces both forms of the test, so both forms of the test are described here. **Note that the null and alternative hypotheses are identical for both forms of the test statistic.**

Equal variances assumed

When the two independent samples are assumed to be drawn from populations with identical population variances (i.e., $\sigma_1^2 = \sigma_2^2$), the test statistic t is computed as:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

with

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Where

\overline{x}_1 = Mean of first sample

\overline{x}_2 = Mean of second sample

n_1 = Sample size (i.e., number of observations) of first sample

n_2 = Sample size (i.e., number of observations) of second sample

s_1 = Standard deviation of first sample

s_2 = Standard deviation of second sample

s_p = Pooled standard deviation

The calculated t value is then compared to the critical t value from the t distribution table with degrees of freedom $df = n_1 + n_2 - 2$ and chosen confidence level. If the calculated t value is greater than the critical t value, then we reject the null hypothesis.

Note that this form of the independent samples t test statistic assumes equal variances.

Because we assume equal population variances, it is OK to "pool" the sample variances (s_p). However, if this assumption is violated, the pooled variance estimate may not be accurate, which would affect the accuracy of our test statistic (and hence, the p -value).

Equal variances not assumed

When the two independent samples are assumed to be drawn from populations with unequal variances (i.e., $\sigma_1^2 \neq \sigma_2^2$), the test statistic t is computed as:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where

\overline{x}_1 = Mean of first sample

\overline{x}_2 = Mean of second sample

n_1 = Sample size (i.e., number of observations) of first sample

(n_{2}) = Sample size (i.e., number of observations) of second sample

(s_{1}) = Standard deviation of first sample

(s_{2}) = Standard deviation of second sample

The calculated t value is then compared to the critical t value from the t distribution table with degrees of freedom

$$df = \frac{\left(\frac{s_{1}^2}{n_{1}} + \frac{s_{2}^2}{n_{2}} \right)^2}{\frac{1}{n_{1}-1} \left(\frac{s_{1}^2}{n_{1}} \right) + \frac{1}{n_{2}-1} \left(\frac{s_{2}^2}{n_{2}} \right)}$$

and chosen confidence level. If the calculated t value > critical t value, then we reject the null hypothesis.

Note that this form of the independent samples t test statistic does not assume equal variances. This is why both the denominator of the test statistic and the degrees of freedom of the critical value of t are different than the equal variances form of the test statistic.

Data Set-Up

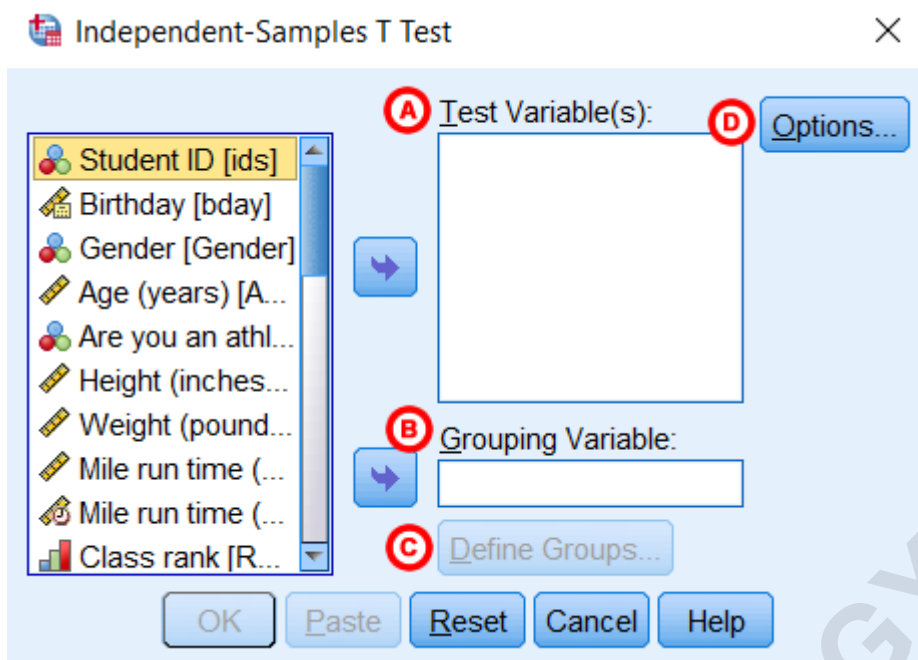
Your data should include two variables (represented in columns) that will be used in the analysis. The independent variable should be categorical and include exactly two groups. (Note that SPSS restricts categorical indicators to numeric or short string values only.) The dependent variable should be continuous (i.e., interval or ratio). SPSS can only make use of cases that have nonmissing values for the independent and the dependent variables, so if a case has a missing value for either variable, it cannot be included in the test.

The number of rows in the dataset should correspond to the number of subjects in the study. Each row of the dataset should represent a unique subject, person, or unit, and all of the measurements taken on that person or unit should appear in that row.

Run an Independent Samples t Test

To run an Independent Samples t Test in SPSS, click **Analyze > Compare Means > Independent-Samples T Test**.

The Independent-Samples T Test window opens where you will specify the variables to be used in the analysis. All of the variables in your dataset appear in the list on the left side. Move variables to the right by selecting them in the list and clicking the blue arrow buttons. You can move a variable(s) to either of two areas: **Grouping Variable** or **Test Variable(s)**.



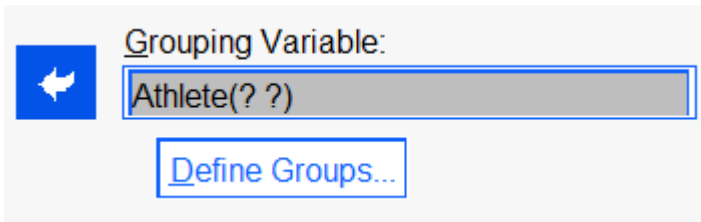
A Test Variable(s): The dependent variable(s). This is the continuous variable whose means will be compared between the two groups. You may run multiple t tests simultaneously by selecting more than one test variable.

B Grouping Variable: The independent variable. The categories (or groups) of the independent variable will define which samples will be compared in the t test. The grouping variable must have at least two categories (groups); it may have more than two categories but a t test can only compare two groups, so you will need to specify which two groups to compare. You can also use a continuous variable by specifying a cut point to create two groups (i.e., values at or above the cut point and values below the cut point).

C Define Groups: Click **Define Groups** to define the category indicators (groups) to use in the t test. If the button is not active, make sure that you have already moved your independent variable to the right in the **Grouping Variable** field. You must define the categories of your grouping variable before you can run the Independent Samples *t* Test procedure.

You will not be able to run the Independent Samples *t* Test until the levels (or cut points) of the grouping variable have been defined. The OK and Paste buttons will be unclickable until the levels have been defined.

You can tell if the levels of the grouping variable have not been defined by looking at the Grouping Variable box: if a variable appears in the box but has two question marks next to it, then the levels are not defined:

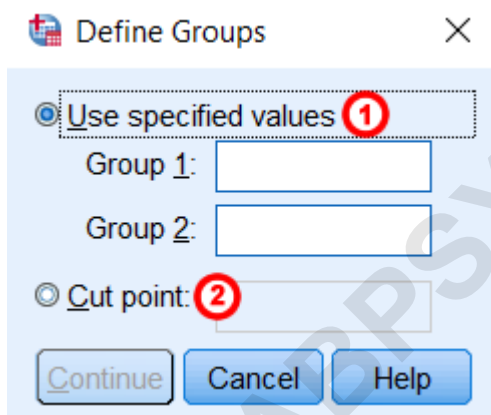


Options: The Options section is where you can set your desired confidence level for the confidence interval for the mean difference, and specify how SPSS should handle missing values.

When finished, click **OK** to run the Independent Samples t Test, or click **Paste** to have the syntax corresponding to your specified settings written to an open syntax window. (If you do not have a syntax window open, a new window will open for you.)

Define Groups

Clicking the Define Groups button (C) opens the Define Groups window:



1 Use specified values: If your grouping variable is categorical, select **Use specified values**. Enter the values for the categories you wish to compare in the **Group 1** and **Group 2** fields. If your categories are numerically coded, you will enter the numeric codes. If your group variable is string, you will enter the exact text strings representing the two categories. If your grouping variable has more than two categories (e.g., takes on values of 1, 2, 3, 4), you can specify two of the categories to be compared (SPSS will disregard the other categories in this case).

Note that when computing the test statistic, SPSS will subtract the mean of the Group 2 from the mean of Group 1. Changing the order of the subtraction affects the sign of the results, but does not affect the magnitude of the results.

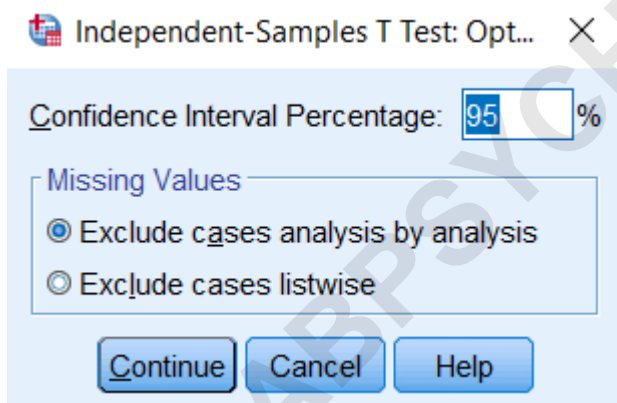
2Cut point: If your grouping variable is numeric and continuous, you can designate a *cut point* for dichotomizing the variable. This will separate the cases into two categories based on the cut point. Specifically, for a given cut point x , the new categories will be:

Group 1: All cases where grouping variable $>x$ Group 2: All cases where grouping variable $< x$

Note that this implies that cases where the grouping variable is equal to the cut point itself will be included in the "greater than or equal to" category. (If you want your cut point to be included in a "less than or equal to" group, then you will need to use Recode into Different Variables or use DO IF syntax to create this grouping variable yourself.) Also note that while you can use cut points on any variable that has a numeric type, it may not make practical sense depending on the actual measurement level of the variable (e.g., nominal categorical variables coded numerically). Additionally, using a dichotomized variable created via a cut point generally reduces the power of the test compared to using a non-dichotomized variable.

Options

Clicking the Options button (D) opens the Options window:



The Confidence Interval Percentage box allows you to specify the confidence level for a confidence interval. Note that this setting does NOT affect the test statistic or p-value or standard error; it only affects the computed upper and lower bounds of the confidence interval. You can enter any value between 1 and 99 in this box (although in practice, it only makes sense to enter numbers between 90 and 99).

The Missing Values section allows you to choose if cases should be excluded "analysis by analysis" (i.e. pairwise deletion) or excluded listwise. This setting is not relevant if you have only specified one dependent variable; it only matters if you are entering more than one dependent (continuous numeric) variable. In that case, excluding "analysis by analysis" will use all nonmissing values for a given variable. If you exclude "listwise", it will only use the cases with nonmissing

values for all of the variables entered. Depending on the amount of missing data you have, listwise deletion could greatly reduce your sample size.

Example: Independent samples T test when variances are not equal

Problem Statement

In our sample dataset, students reported their typical time to run a mile, and whether or not they were an athlete. Suppose we want to know if the average time to run a mile is different for athletes versus non-athletes. This involves testing whether the sample means for mile time among athletes and non-athletes in your sample are statistically different (and by extension, inferring whether the means for mile times in the population are significantly different between these two groups). You can use an Independent Samples t Test to compare the mean mile time for athletes and non-athletes.

The hypotheses for this example can be expressed as:

H0: $\mu_{\text{non-athlete}} - \mu_{\text{athlete}} = 0$ ("the difference of the means is equal to zero")

H1: $\mu_{\text{non-athlete}} - \mu_{\text{athlete}} \neq 0$ ("the difference of the means is not equal to zero")

where μ_{athlete} and $\mu_{\text{non-athlete}}$ are the population means for athletes and non-athletes, respectively.

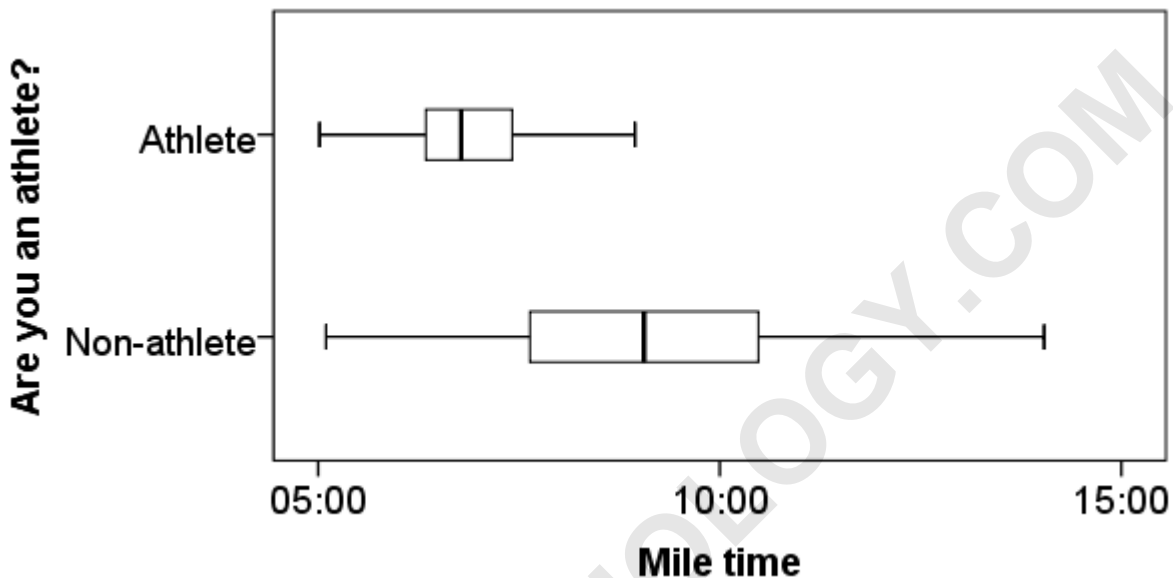
In the sample data, we will use two variables: Athlete and MileMinDur. The variable Athlete has values of either "0" (non-athlete) or "1" (athlete). It will function as the independent variable in this T test. The variable MileMinDur is a numeric duration variable (h:mm:ss), and it will function as the dependent variable. In SPSS, the first few rows of data look like this:

Athlete	MileMinDur
0	0:06:39
1	0:08:16
0	0:10:26
0	0:07:22
0	0:06:25

Before the Test

Before running the Independent Samples t Test, it is a good idea to look at descriptive statistics and graphs to get an idea of what to expect. Running Compare Means (**Analyze > Compare**

Means > Means) to get descriptive statistics by group tells us that the standard deviation in mile time for non-athletes is about 2 minutes; for athletes, it is about 49 seconds. This corresponds to a variance of 14803 seconds for non-athletes, and a variance of 2447 seconds for athletes¹. Running the Explore procedure (**Analyze > Descriptives > Explore**) to obtain a comparative boxplot yields the following graph:



If the variances were indeed equal, we would expect the total length of the boxplots to be about the same for both groups. However, from this boxplot, it is clear that the spread of observations for non-athletes is much greater than the spread of observations for athletes. Already, we can estimate that the variances for these two groups are quite different. It should not come as a surprise if we run the Independent Samples t Test and see that Levene's Test is significant.

Additionally, we should also decide on a *significance level* (typically denoted using the Greek letter alpha, α) before we perform our hypothesis tests. The significance level is the threshold we use to decide whether a test result is significant. For this example, let's use $\alpha = 0.05$.

¹When computing the variance of a duration variable (formatted as hh:mm:ss or mm:ss or mm:ss.s), SPSS converts the standard deviation value to seconds before squaring.

Running the Test

To run the Independent Samples t Test:

Click **Analyze > Compare Means > Independent-Samples T Test**. Move the variable Athlete to the **Grouping Variable** field, and move the variable MileMinDur to the **Test Variable(s)** area. Now Athlete is defined as the independent variable and MileMinDur is defined as the dependent

variable. Click **Define Groups**, which opens a new window. **Use specified values** is selected by default. Since our grouping variable is numerically coded (0 = "Non-athlete", 1 = "Athlete"), type "0" in the first text box, and "1" in the second text box. This indicates that we will compare groups 0 and 1, which correspond to non-athletes and athletes, respectively. Click **Continue** when finished. Click **OK** to run the Independent Samples t Test. Output for the analysis will display in the Output Viewer window.

Syntax

```
T-TEST GROUPS=Athlete(0 1)
/MISSING=ANALYSIS
/VARIABLES=MileMinDur
/CRITERIA=CI(.95).
```

Output

Tables

Two sections (boxes) appear in the output: **Group Statistics** and **Independent Samples Test**. The first section, **Group Statistics**, provides basic information about the group comparisons, including the sample size (n), mean, standard deviation, and standard error for mile times by group. In this example, there are 166 athletes and 226 non-athletes. The mean mile time for athletes is 6 minutes 51 seconds, and the mean mile time for non-athletes is 9 minutes 6 seconds.

Group Statistics

Are you an athlete?		N	Mean	Std. Deviation	Std. Error Mean
Mile time	Non-athlete	226	0:09:06	0:02:01.668	0:00:08.093
	Athlete	166	0:06:51	0:00:49.464	0:00:03.839

The second section, **Independent Samples Test**, displays the results most relevant to the Independent Samples t Test. There are two parts that provide different pieces of information: (A) Levene's Test for Equality of Variances and (B) t-test for Equality of Means.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means C						
		A		B		Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
		F	Sig.	t	df				Lower	Upper
Mile time	Equal variances assumed	102.98	.000	13.475	390	.000	0:02:14	0:00:10	0:01:55	0:02:34
	Equal variances not assumed			15.047	315.846	.000	0:02:14	0:00:08	0:01:57	0:02:32

A Levene's Test for Equality of of Variances: This section has the test results for Levene's Test. From left to right:

F is the test statistic of Levene's test *Sig.* is the p-value corresponding to this test statistic.

The *p*-value of Levene's test is printed as ".000" (but should be read as $p < 0.001$ -- i.e., *p* very small), so we we reject the null of Levene's test and conclude that the variance in mile time of athletes is significantly different than that of non-athletes. **This tells us that we should look at the "Equal variances not assumed" row for the *t* test (and corresponding confidence interval) results.** (If this test result had not been significant -- that is, if we had observed $p > \alpha$ -- then we would have used the "Equal variances assumed" output.)

B t-test for Equality of Means provides the results for the actual Independent Samples t Test. From left to right:

t is the computed test statistic, using the formula for the equal-variances-assumed test statistic (first row of table) or the formula for the equal-variances-not-assumed test statistic (second row of table) *df* is the degrees of freedom, using the equal-variances-assumed degrees of freedom formula (first row of table) or the equal-variances-not-assumed degrees of freedom formula (second row of table) *Sig (2-tailed)* is the p-value corresponding to the given test statistic and degrees of freedom *Mean Difference* is the difference between the sample means, i.e. $x_1 - x_2$; it also corresponds to the numerator of the test statistic for that test *Std. Error Difference* is the standard error of the mean difference estimate; it also corresponds to the denominator of the test statistic for that test

Note that the mean difference is calculated by subtracting the mean of the second group from the mean of the first group. In this example, the mean mile time for athletes was subtracted from the mean mile time for non-athletes (9:06 minus 6:51 = 02:14). The sign of the mean difference corresponds to the sign of the *t* value. The positive *t* value in this example indicates that the mean mile time for the first group, non-athletes, is significantly greater than the mean for the second

group, athletes.

The associated p value is printed as ".000"; double-clicking on the p-value will reveal the unrounded number. SPSS rounds p-values to three decimal places, so any p-value too small to round up to .001 will print as .000. (In this particular example, the p-values are on the order of 10-40.)

Confidence Interval of the Difference: This part of the *t*-test output complements the significance test results. Typically, if the CI for the mean difference contains 0 within the interval -- i.e., if the lower boundary of the CI is a negative number and the upper boundary of the CI is a positive number -- the results are not significant at the chosen significance level. In this example, the 95% CI is , which does not contain zero; this agrees with the small *p*-value of the significance test.

Decision and Conclusions

Since $p < .001$ is less than our chosen significance level $\alpha = 0.05$, we can reject the null hypothesis, and conclude that the that the mean mile time for athletes and non-athletes is significantly different.

Based on the results, we can state the following:

There was a significant difference in mean mile time between non-athletes and athletes ($t_{315.846} = 15.047, p < .001$). The average mile time for athletes was 2 minutes and 14 seconds lower than the average mile time for non-athletes.