Example 1:

A comparison of men and women on job satisfaction.

Men	Women
7	1
8	10
6	3
5	4
6	1
5	1
6	2
9	3
8	5
1	4
7	3
2	5
4	6
6	4
7	2
8	5
9	1
7	4

t-Test: Two-Sample Assuming Equal Variances

	Variable 1	*Variable 2*
Mean	6.166666667	3.555555556
Variance	4.735294118	5.08496732
Observations	18	18
Pooled Variance	4.910130719	
Hypothesized Mean Difference	0	
df	34	
t Stat	**3.53508679**	
P(T<=t) one-tail	0.000599612	
t Critical one-tail	1.690923455	
P(T<=t) two-tail	0.001199224	
t Critical two-tail	2.032243174	

A company has been accused by some employees of not treating women well. The company makes the claim that job satisfaction is the same for men and women at the firm. A researcher decides to test this. She randomly samples 36 employees (18 men and 18 women) and asks them to complete a job satisfaction attitude scale. Scores on this attitude scale range from a low of 0 ("not at all satisfied") to a high of 10 ("extremely satisfied"). The average job satisfaction score for men is 6.17 (rounded) and 3.56 for women. This is a difference of about 2.61 in satisfaction. It appears that men have more job satisfaction than women at this firm. The company claims that a sample of 36 is quite small given the fact that 5,000 people work at the company and they are asserting that the difference is entirely due to sampling error. Given the Excel printout above, what do you think?

Answer: If a two-tail test was done, the probability of getting the sample evidence given that there is no difference in the population means of job satisfaction scores for men and women is .0012. In another words, if men and women feel the same about working in this firm, there is only 12 chances out of 10,000 of getting the sample evidence we obtained (or one showing an even larger difference). Statisticians usually test at an alpha of .05 so we are going to reject the null hypothesis.

The difference between the two sample means was 2.61. If men and women in the firm actually have the same job satisfaction (i.e., the difference between the population means is actually 0), the likelihood of getting a difference between two sample means of men and women of 2.61 or greater is .0012. This is why we reject H_O that the two population means are the same.

Conclusion: There is a statistically significant difference between the average job satisfaction scores of men and women at this firm.

Example 2:

A comparison of men and women and how much they spend on wine per year.

Women	Men
$100.00	$107.00
$250.00	$240.00
$890.00	$880.00
$765.00	$770.00
$456.00	$409.00
$356.00	$500.00
$876.00	$800.00
$740.00	$900.00
$231.00	$1,000.00
$222.00	$489.00
$555.00	$800.00
$666.00	$890.00
$876.00	$770.00
$10.00	$509.00
$290.00	$100.00
$98.00	$102.00
$56.00	$134.00

t-Test: Two-Sample Assuming Equal Variances

	Variable 1	*Variable 2*
Mean	437.4705882	552.9412
Variance	97784.13971	104221.6
Observations	17	17
Pooled Variance	101002.8493	
Hypothesized Mean Difference	0	
df	32	
t Stat	**-1.05928794**	
P(T<=t) one-tail	0.148699644	
t Critical one-tail	1.693888407	
P(T<=t) two-tail	0.297399288	
t Critical two-tail	2.036931619	

A marketer wants to determine whether men and women spend different amounts on wine. (It is well known that men spend considerably more on beer.) A researcher decides to test this. She randomly samples 34 people (17 women and 17 men) and finds that the average amount spent on wine (in a year) by women is $437.47. The average amount spent by men is $552.94**. Given the Excel printout above, is the difference statistically significant?**

Answer: If a two-tail test was done, the probability of getting the sample evidence given that there is no difference in the population means of job satisfaction scores for men and women is .30 (rounded from .297399288). In another words, if men and women spend the same on wine, there is a 30% chance of getting the sample evidence we obtained (or even a larger difference). Statisticians usually test at an alpha of .05 so we do not have evidence to reject the null hypothesis. Conclusion: There is no statistically significant difference between men and women on how much they spend on wine consumption.

The calculated t-statistic is -1.059287941. Why is it negative?
Answer: The amount spent on wine by women is less than that spent by men (although the difference is not statistically significant). If you make men the first variable the t-value will be positive but the results will be exactly the same (the t-distribution is symmetric).

What would the calculated t-value have to be for us to reject it?
Answer: If a two-tail test is being done, the critical value of t is 2.036931619. To reject the null hypothesis, we would need a calculated t-value of more than 2.036931619 or less than -2.036931619.