Instructions for doing two-sample t-test in Excel

(1) If you do not see *Data Analysis* in the menu, this means you need to use *Add-ins* and make sure that the box in front of Analysis ToolPak is checked. Instructions on how to use *Add-ins* are in *Assignments*.

(2) Go to *Data Analysis—t-test: Two Sample Assuming Equal Variances* in menu. [If you know that the variances are not equal, then you may use the t-test that assumes unequal variances. The problem of heteroscedasticity (unequal variances) is beyond the scope of this course so let's assume equal variances.]

(3) You will have to indicate where the two variables are located, e.g., b5:b17 and d5:d18.
Variable 1 Range: b5:b17
Variable 2 Range: d5:d18

The default is Alpha = .05 so leave this alone.

You have to indicate where you want the output to appear. You will probably want the output to appear either on the same page or on another worksheet. If you want the output to appear on the same page, then check the circle in front of *Output Range* and indicate where the output should go. If the data appear in rows, say, b5: b17 and d5:d18, then your output should not appear in the first 18 rows. I would indicate **a19** next to *Output Range*. Of course, you can check the circle in front of *New Worksheet Ply* and your output will appear on another worksheet. This may be a good idea if you are afraid that the output is too large to appear on the same page as the input.

Example: A researcher wants to determine which drug is more effective against lung cancer. A total of 27 patients have been randomly assigned to two groups. The data represents how many years each patient lived after taking the drug.

Drug X	Drug Y
3	**6**
4	**5**
6	**7**
5	**9**
7	**3**
2	**5**
1	**5**
4	**8**
8	**6**
2	**4**
6	**5**
4	**7**
3	**8**
	12

t-Test: Two-Sample Assuming Equal Variances

	Variable 1	*Variable 2*
Mean	**4.230769231**	**6.428571429**
Variance	**4.358974359**	**5.340659341**
Observations	**13**	**14**
Pooled Variance	**4.869450549**	
Hypothesized Mean Difference	**0**	
df	**25**	
t Stat	**-2.58584468**	
P(T<=t) one-tail	**0.007965466**	
t Critical one-tail	**1.708140189**	
P(T<=t) two-tail	**0.015930932**	
t Critical two-tail	**2.05953711**	

Those taking Drug X live, on average, another 4.23 years and those taking Drug Y live another 6.43 years longer. If Drug Y costs considerably more, it makes sense to test in order to make sure that we are not looking at sampling error.

The t-test is significant; the t-value is 2.5858. The critical value (two-tail test) is 2.0595. The probability of getting the sample evidence (the input data) is .015953711. In other words, there is less than a 2% chance of getting the sample evidence if the null hypothesis (that the two groups have equal population means) is true. Therefore, we reject H_0 if we are testing at the .05 level. Individuals taking Drug Y will live longer than those taking Drug X.

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. 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.

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. 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. Statisticians usually test at an alpha of .05 so we are going to reject the null hypothesis.

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