Using SPSS to perform *t*-tests:

This handout explains how to perform the two types of *t*-test in SPSS. It's very easy, as long as you keep in mind that the two types of *t*-test require the data to be entered in quite different ways. This makes sense if you remember that SPSS generally treats each *column* in the spreadsheet as being a separate variable, and each *row* in the spreadsheet as being the data for a single participant.

(a) The repeated-measures *t*-test:

We use this when we have two conditions, and we want to know if the mean performance on one condition is significantly different from the mean performance in the other condition. It's used with repeated-measures designs: i.e., each participant does both conditions in the experiment.

Suppose we wanted to know if Prozac affects driving ability. Ten participants have their driving performance tested twice on a sheep farm. We count how many sheep they collide with, during each of two 30-minute sessions.

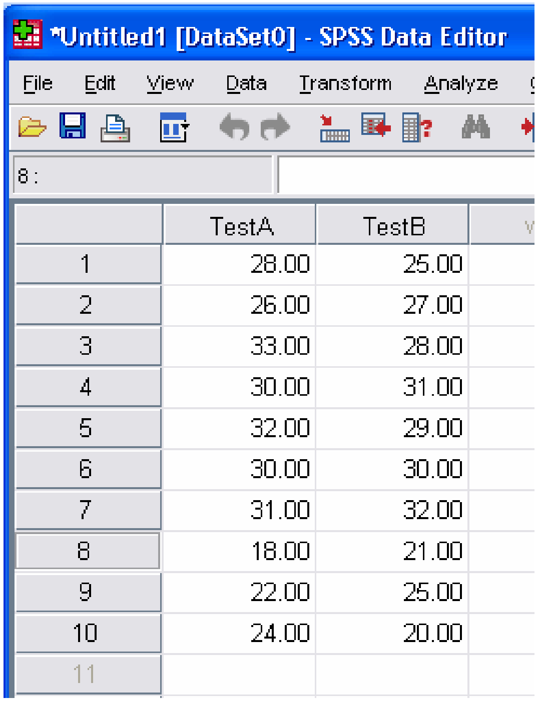
Test A occurs after they have taken Prozac. (This is our experimental condition).

Test B occurs while they are drug-free. (This is our control condition).

Each participant thus provides two scores (one for each session). In practice, to avoid order effects, five people would do condition A then B, and five would do condition B then A.

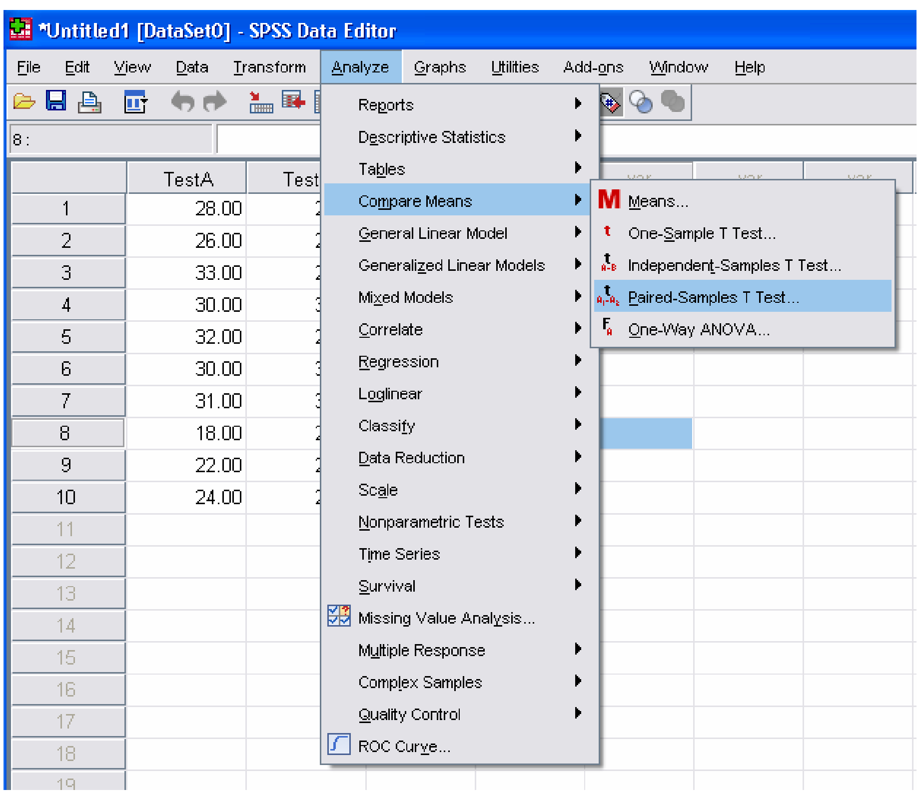
Step 1:

Input the data like this. Each row is one participant's data. We have two columns. The first column shows a participant's score for condition A, and the other column shows their score for condition B. Go to "variable view" and label the columns, to make life easier when it comes to interpreting the output. So here, participant number one hit 28 sheep during test A (when he had taken Prozac) and 25 sheep during test B (when he was drug-free). Participant two hit 26 sheep in test A, and 27 sheep in test B; and so on.

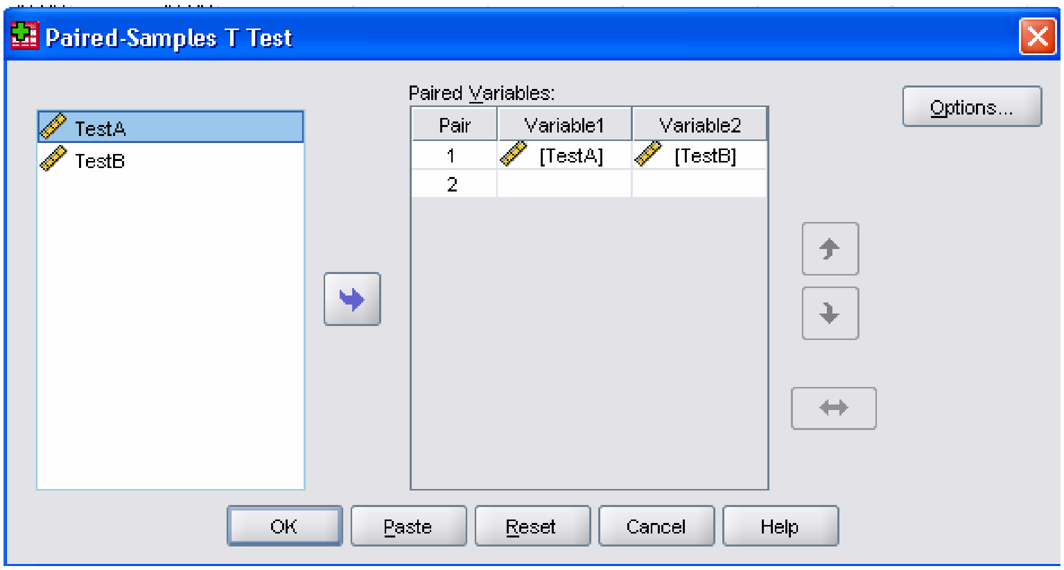


Step 2:

On the top menu of SPSS, go to "Analyze"; click on "Compare Means"; and then click on "Paired Samples T test".



The following dialogue box comes up:



Click on each of the two variables that you want to compare (Test A and Test B in this case) and then click on the arrow, to move the variables into the "Paired Variable(s)" box. It doesn't matter which way round they are entered. Then click on "OK" to run the test.

Step 3:
Here's the output you should get.

(a) The following box gives you the descriptive statistics for each condition - the mean, the number of scores in each condition (worth glancing at in order to check that SPSS has used the right data), the standard deviation, and the standard error.

Thus the drivers hit an average of 27.4 sheep when they were under the influence of Prozac, and 26.8 sheep when they were drug-free. The standard deviations tell us how much the scores were spread out around these means, and the standard errors give us an estimate of how much variation we are likely to get in these means if we repeated the same experiment many times. With experience, you can look at these data and get an idea of whether the t-test is likely to show a significant difference between the conditions. Here, the difference between the conditions is quite small in relation to the size of the s.d. and the s.e., suggesting that there is probably no significant difference between the two conditions - Prozac hasn't had much of an effect on driving ability.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	TestA	27.4000	10	4.83506	1.52898
	TestB	26.8000	10	4.04969	1.28062

(b) SPSS also performs a Pearson's correlation between performance in the two conditions. You can ignore this part of the output. In this case, it's a highly significant correlation, suggesting that people who hit lots of sheep in one condition also did so in the other condition, and vice versa.

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	TestA & TestB	10	.799	.006

(c) The next table gives us the results of the t-test. There's lots of information here, so let's take it a bit at a time, starting from the left side of the table.

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	TestA - TestB	.60000	2.91357	.92135	-1.48424	2.68424	.651	9	.531

First, the table shows the mean difference between the two conditions, and the standard deviation and standard error associated with that difference. The mean difference in this case is 0.6: the drivers hit just over half a sheep more when they were under the influence of Prozac than when they were drug-free. (i.e. their driving was very slightly worse when they were drugged).

Next, the table shows a "95% confidence interval" for the difference between the means. If we did the same experiment many times over, what sort of difference might we expect to get between the two conditions? In this case, the confidence interval tells us that on 95% of occasions, the difference between the two conditions would be somewhere between -1.4842 and 2.6842. In other words, sometimes performance on test A would be a little *better* than performance on test B, and sometimes performance on test A would be a little *worse* than performance on test B. The fact that the confidence interval straddles zero (i.e., ranges from a negative value to a positive value for the difference) tells us that our particular, observed difference could easily be in entirely the opposite direction if we did the experiment again. The advantage of confidence intervals is that they stop you getting fixated on thinking about "*the* difference" between the conditions: if you did the experiment again, you would almost certainly get a different value for the difference, and in this case the effects of the drug are so weak that you could easily get a difference next time that was the opposite way round, so that it looked as if Prozac *improved* driving ability!

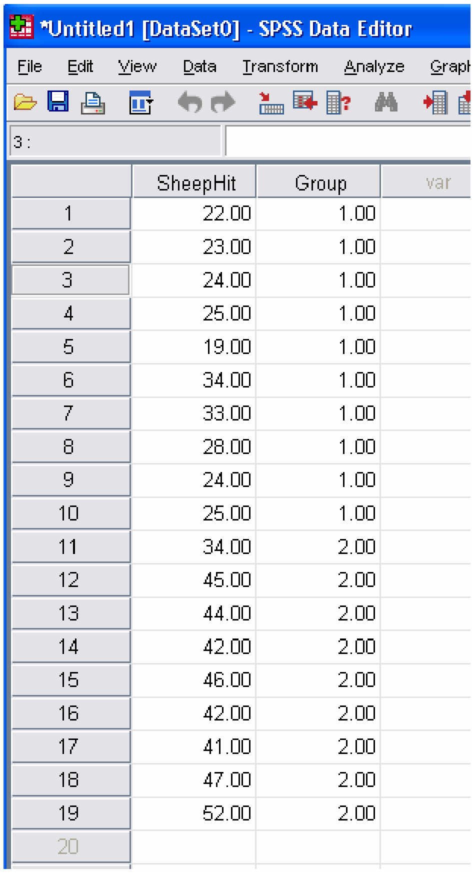
The final part of the table gives you the value of *t*, its degrees of freedom, and its exact significance level. Our obtained *t* of 0.65 is very small, and in fact is highly likely to have occurred by chance. (With 9 df, values of *t* this small will occur by chance .53 of the time, i.e. very often). Our conclusion would therefore be that Prozac had no detectable effect on driver's ability to avoid hitting sheep.

(b) The independent-measures *t*-test:

We use this when we have two conditions, and we want to know if the mean performance on one condition is significantly different from the mean performance in the other condition. It's used with independent-measures designs: i.e., there are two different groups of participants, one for each condition in the experiment.
Suppose we did the "Prozac" experiment again, but this time we used two different groups of participants. One group (group 1) drive without having taken any drug, and the other group (group 2) drive while under the influence of Prozac,

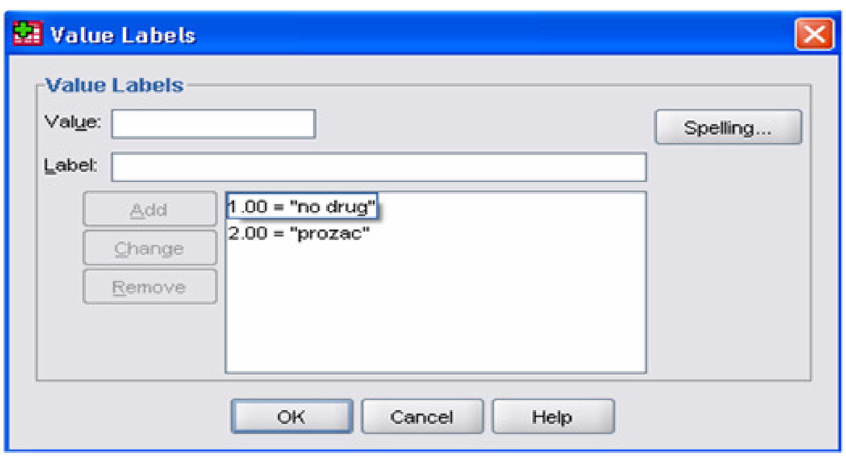
Step 1:

Input the data.
You need two columns. One column contains the data (i.e. each participant's score is on a separate row). The other one contains code numbers that tell SPSS which condition the adjacent data belong to (i.e. whether the participant was in condition 1 or 2). It doesn't actually matter what code numbers you use: I generally use 0 and 1, or 1 and 2, but you could use any numbers you like in theory.
Thus, in the example below, you can see that the first participant hit 22 sheep, and she was in the "drug free" condition, as shown by the "1" in the column to the right of her score. the next participant hit 25 sheep, and was also in the "drug-free" condition, as shown by the "1" next to his score; and so on. We have 10 people in the "drug-free" condition and 9 in the "Prozac" condition. (For an independent-measures t-test, it is desirable, but by no means essential, to have the same number of participants in each group).

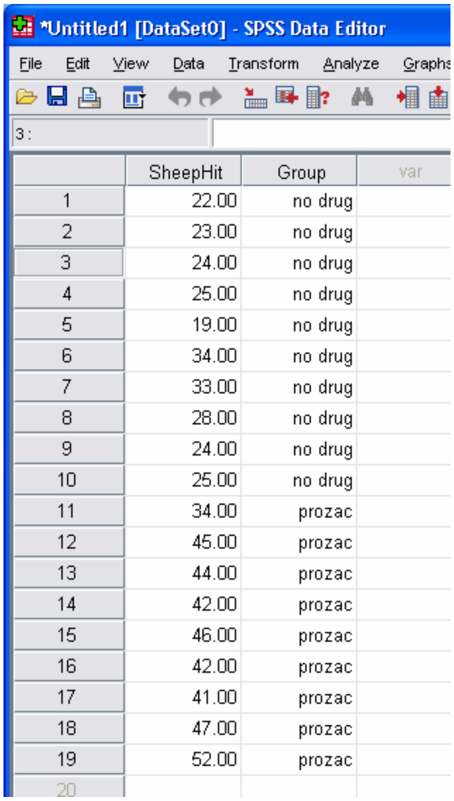


	SheepHit	Group	var
1	22.00	1.00	
2	23.00	1.00	
3	24.00	1.00	
4	25.00	1.00	
5	19.00	1.00	
6	34.00	1.00	
7	33.00	1.00	
8	28.00	1.00	
9	24.00	1.00	
10	25.00	1.00	
11	34.00	2.00	
12	45.00	2.00	
13	44.00	2.00	
14	42.00	2.00	
15	46.00	2.00	
16	42.00	2.00	
17	41.00	2.00	
18	47.00	2.00	
19	52.00	2.00	
20			

To make life easier for yourself, go to "variable view" and for the "group" variable, click on "values". Label conditions 1 and 2 with more meaningful names, like so:

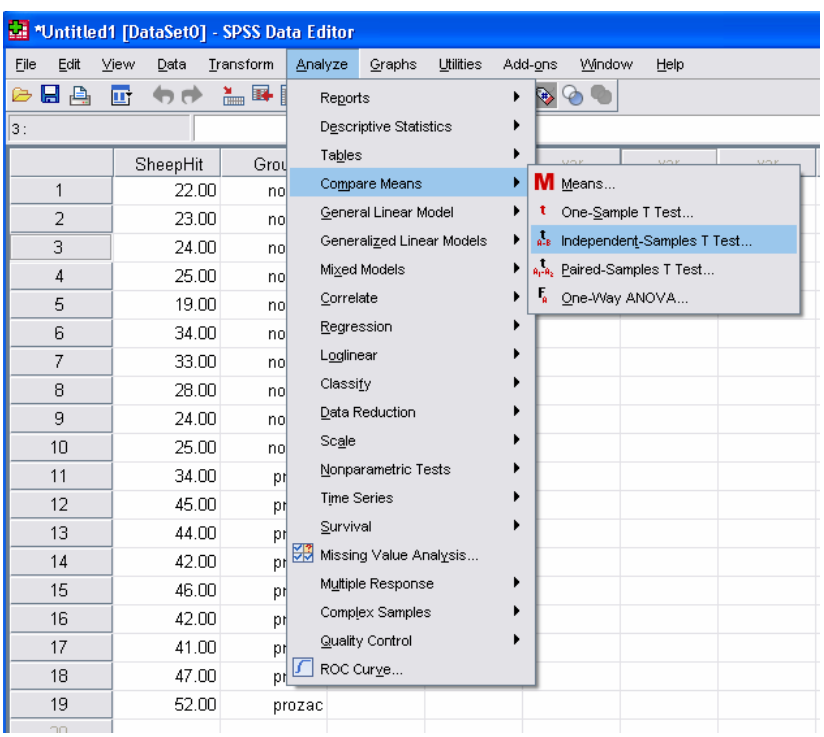


When you return to "data view", the spreadsheet should now look like this:

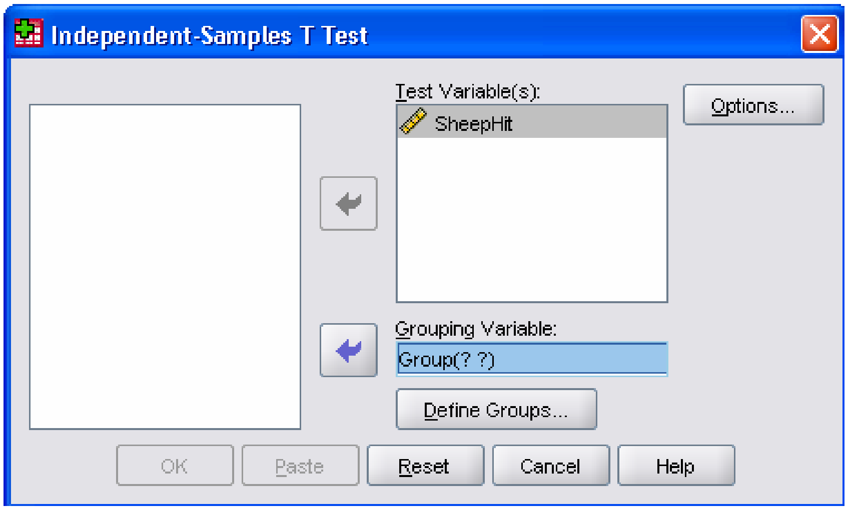


Step 2:

Go to the top menu; click on "Analyze", then "Compare Means"; and finally "Independent Samples T test".

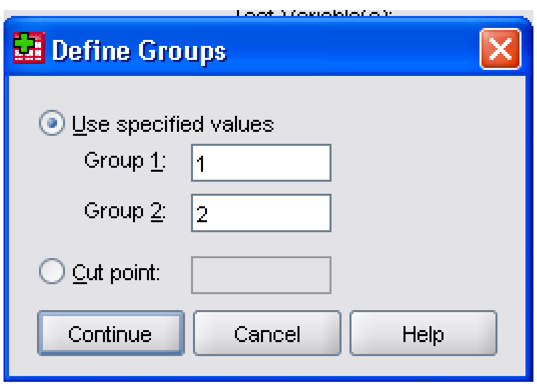


The following dialogue box pops up.



Click on your test variable (the one with the scores in it - "SheepHit" in this case), and then click on the upper of the two buttons with arrows on, to transfer your test variable into the box headed "Test Variable(s)".
Then click on your grouping variable (the one with the code values in it - "Group" in this case) and then click on the lower button.

SPSS will show the variable name, with "(??)" after it. Click on the "Define Groups..." button, and tell SPSS that your code numbers are 1 and 2 (or whatever other numbers you perversely chose to use).



Then click on "OK" to run the test.

Step 3:

Here's the output.

First, you get the descriptive statistics for each group - the number of participants in each group; the mean score for each group; and its standard deviation and standard error.

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
SheepHit	no drug	10	25.7000	4.71522	1.49108
	prozac	9	43.6667	4.92443	1.64148

The box containing the results of the t-test contains a lot of information, so let's take it a bit at a time. There are actually three separate statistical tests here.

(a) Levene's test:

Both kinds of t-test are *parametric* tests - they make certain assumptions about the nature of the data on which they are being used. For an independent-measures t-test, one of these assumptions is that the two groups show *homogeneity of variance*. In other words, it's assumed that the spread of scores is roughly similar within each group (so that both means are equally representative of their respective groups). SPSS checks whether or not this assumption is actually true by performing the Levene's test. This test enables us to decide whether or not the variances in the two conditions are significantly different from each other. It is one statistical test that you usually do *not* want to be significant, as if it is, you have a problem!

The lefthand side of the table shows the results of the Levene's test. The test statistic is called "F", and the bigger it is, the less likely it is to have occurred by chance. To the right of it, in the column headed "Sig.", is the probability of getting an F-value this large by chance. If this *p* is less than .05, it means that the difference between the two variances is so large that it is unlikely to have occurred by chance; your data show inhomogeneity of variance.

(b) Two flavours of t-test:

What do you do with this information?

(1) If the Levene's test is significant (i.e. the significance level is less than 0.05), strictly speaking, you should not perform a t-test on these data. You should use a non-parametric test instead, in this case the Mann-Whitney test.

In practice, SPSS goes ahead and does a t-test anyway, using a different version of the t-test formula. This is in the row of the table that's labelled "equal variances not assumed". (You'll notice that the t-value is slightly smaller, and the d.f. look a bit funny).

(2) If the Levene's test is *not* significant, it's OK to use the t-test (assuming of course that the data satisfy the other requirements for a t-test, i.e. that you've got interval or ratio data, and that the data in each group are roughly normally distributed). In our case, F = 0.005, a very small value, and the associated *p*-value is .945, which is well above 0.05. We have no reason to believe that our data suffer from inhomogeneity of variance, and so we can use the version of the t-test that's shown in the row labelled "Equal variances assumed".

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SheepHit	Equal variances assumed	.005	.945	-8.121	17	.000	-17.96667	2.21225	-22.63411	-13.29923
	Equal variances not assumed			-8.102	16.602	.000	-17.96667	2.21761	-22.65397	-13.27936

So, our obtained value of *t* is -8.12, with 17 d.f., and this is so unlikely to have occurred by chance that the significance level is shown as ".000" - it's gone off the scale because SPSS can only show numbers to 3 significant digits. *t* has a negative value because of the way in which I coded the two groups: had I coded "no drug" with "2" and "Prozac" with "1", then I would have obtained the same value of *t*, but as a positive value. We could write this result as follows: "Prozac had a highly significant effect on driving ability, as reflected in the number of sheep hit within a 30-minute driving session ($t(17) = -8.10$, $p <.0005$)".

The rest of the table gives the mean difference between the two conditions, which in this case was quite large - on average 17.97 more sheep were hit by the Prozac group than by the undrugged drivers), and the 95% confidence interval associated with this difference. Here, the CI(95%) is from -22.63 to -13.30. In other words, if we did the same experiment many times over, the size of the difference between the two groups would vary somewhat, but on 95% of tests, the Prozac drivers would hit many more sheep than the undrugged drivers.