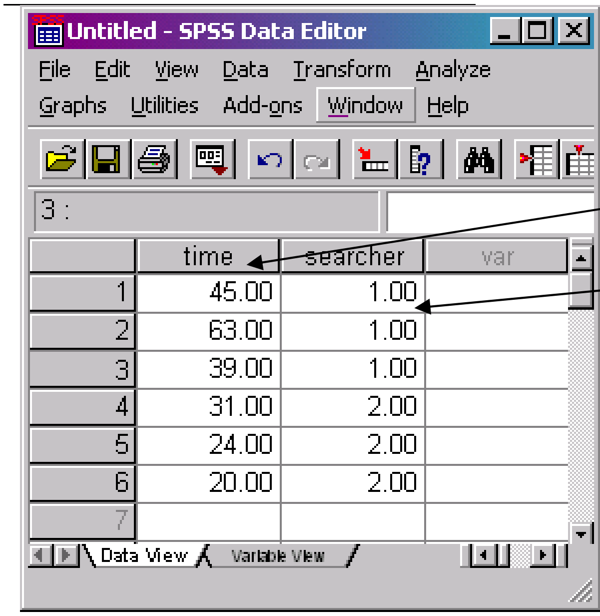
t-test and Oneway ANOVA set up and interpretation in SPSS

For your final exam there are basically three different tests you will need to be able to run independently in SPSS, factorial ANOVA is just an extension of the ideas of Oneway ANOVA.

1. Independent samples t-test (problems similar to #8, pg. 205 in your text, see text for details)

Fellow Humans	Pigeons
45	31
63	24
39	20

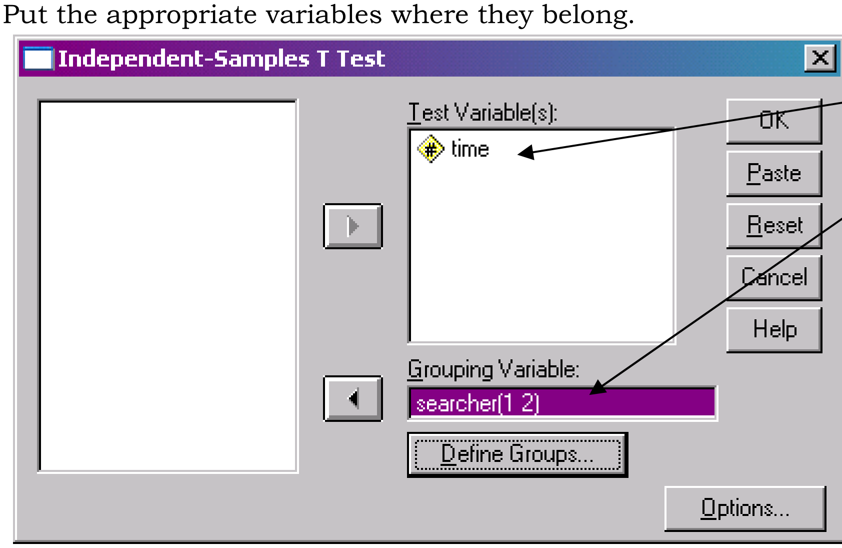


Dependent variable generally in column 1. For the independent samples design all groups data will be entered in one continuous column in the data editor

Independent variable (grouping variable) entered as simply 1 & 2 to identify each group's members. In this example, 1 identifies fellow humans, 2 identifies pigeons. Setting up value labels for the grouping variable will reduce the possibility of confusion.

The ANOVA designs we will be using will be set up exactly the same way except that there will be more than two groups. Column 1 will be your dependent variable and column 2 your independent variable (or factor). After your data is entered, On the SPSS menu click on ANALYZE-> Compare Means -> Independent samples t test to get to the correct dialog box to set up your analysis.

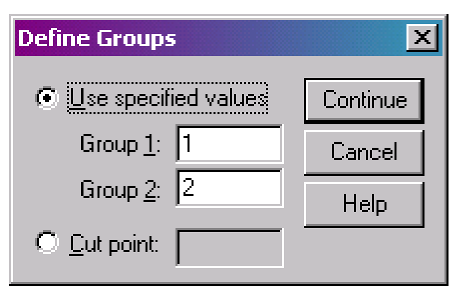
Put the appropriate variables where they belong.



Dependent variable

Independent variable

Click the "Define groups" button to tell SPSS the values assigned to your groups, in this case 1 & 2 for humans & pigeons, respectively.



SPSS Output from Independent samples t-test
First, group statistics give us means and standard deviations for both groups, as well as standard error of the mean for each group separately.

Group Statistics

	searcher	N	Mean	Std. Deviation	Std. Error Mean
time	1.00 human	3	49.0000	12.49000	7.21110
	2.00 pigeon	3	25.0000	5.56776	3.21455

The actual t-test results: will be able to safely ignore most of this table. Levene's test for equality of variances is testing the assumption that variances are equal across groups (ignore this). Of the two rows of statistics below, we will only interested in the top row. The values we will need to pull out of this table will be the values in the top row for t, df, and Sig. (2-tailed).

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
time	Equal variances assumed	2.639	.180	3.040	4	.038	24.00000	7.89515	2.07956	45.92044
	Equal variances not assumed			3.040	2.765	.062	24.00000	7.89515	-2.38019	50.38019

Thus, our hypotheses are:
H_0: $\mu_{humans} = \mu_{pigeons}$
H_1: $\mu_{humans} \neq \mu_{pigeons}$

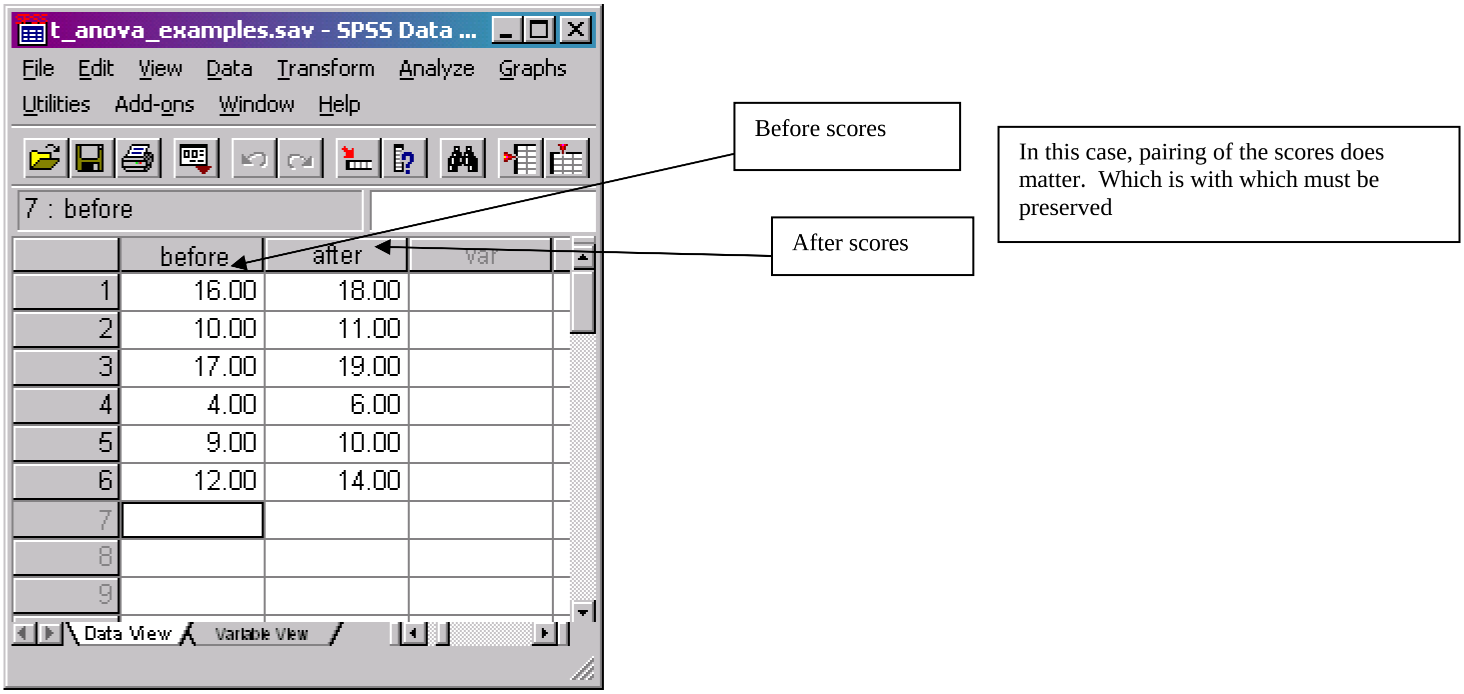
If the null hypothesis was true, we would have a .038 probability of observing a difference between human and pigeon times as large as the difference that we observed.

In this case, our null hypothesis is that there is no difference in how quickly humans and pigeons can spot someone lost in the ocean. The data have produced a t_{obs} of 3.040, with df=4, p = .038. Thus, we reject the null hypothesis and conclude that pigeons can locate a lost person in the ocean significantly faster than a person can.

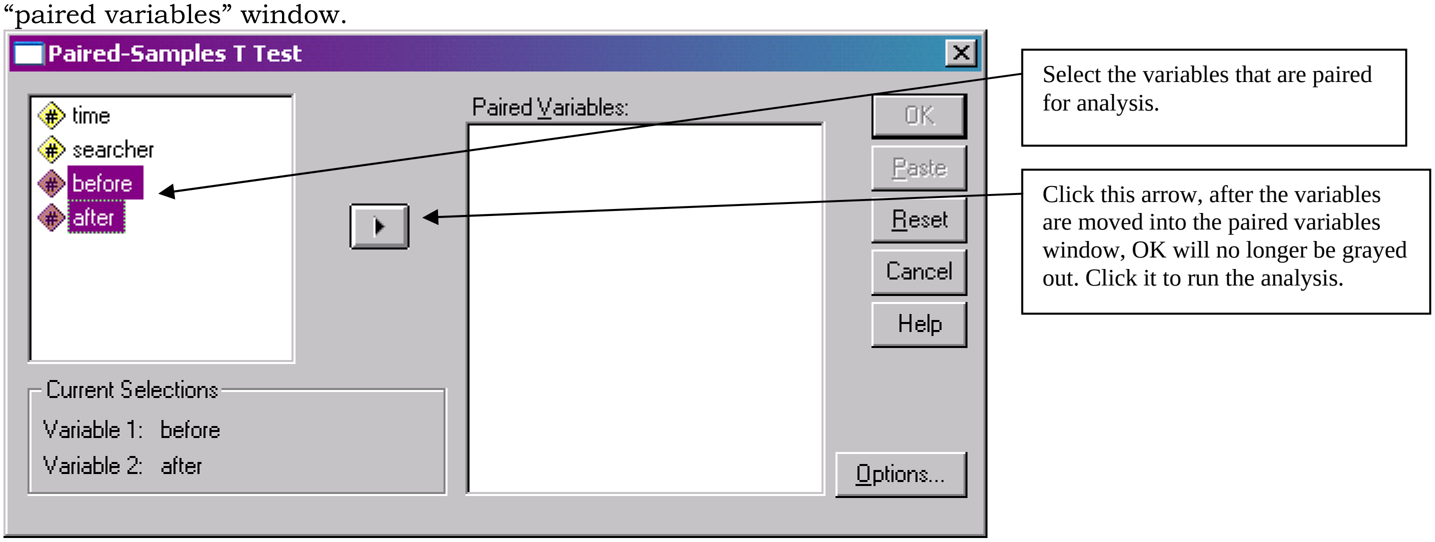
2. Correlated samples t-test (problems similar to #13, pg. 210 in your text, see text for details of problem and data)

Before	After
16	18
10	11
17	19
4	6
9	10
12	14

For the correlated samples design, each variable must be in it's own column (or data from each point in time for repeated measures). Each row must remain exactly intact. That is, in this example, each before-after pair of scores.



To assign variables for a paired (correlated) t test click ANALYZE-> Compare Means -> Paired samples t test and enter variables, you have to select two variables, either using the mouse, or hold down the ctrl key while clicking the two variables, once the paired variables are selected, click the arrow to move them into the "paired variables" window.



Output from paired samples t test

First, SPSS provides the descriptive statistics for our variables:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	11.3333	6	4.80278	1.96073
	after	13.0000	6	4.97996	2.03306

Next, the correlation between the variables is presented along with the probability of observing the correlation in the data if the population correlation was actually zero. In this case, the $p < .001$ (Sig. = .000)

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	before & after	6	.995	.000

Finally, the actual t-test and associated information. In the area of the table labeled "paired differences" all of the information presented pertains to the difference between groups. The last three columns (to the right of the "paired differences" section present the t statistic observed in the data (t), the associated degrees of freedom (df), and the probability of observing a t as large or larger if the null hypothesis was true [Sig. (2-tailed)].

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	before - after	-1.66667	.51640	.21082	-2.20859	-1.12474	-7.906	5	.001

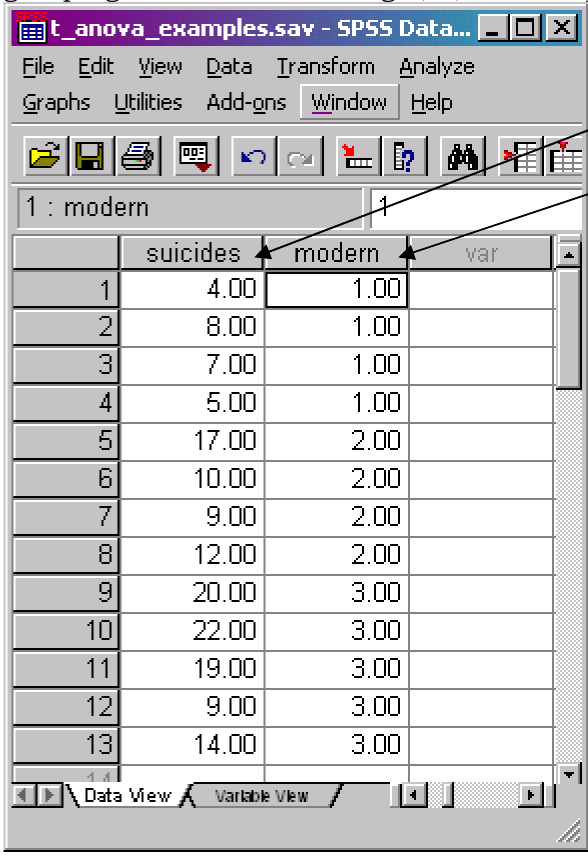
In the case of this problem, the null hypothesis is that there are no differences in aggressive acts while playing before and after seeing other children playing with newer, nicer toys than the ones the subjects were given to play with.
Thus, our hypotheses are:
H_0: $\mu_{before} = \mu_{after}$
H_1: $\mu_{before} \neq \mu_{after}$
Our research hypothesis is that observing the children playing with nicer newer toys will make our subjects more likely to behave aggressively.

If the null hypothesis was true, we would have a 0.1% probability of observing a difference between before and after scores as large as the difference that we observed, thus the null is not tenable and we conclude that observing the children with the nicer toys lead to the increase in aggressive behavior.

3. Oneway ANOVA (Analysis of variance) (Problems such as #8 on pg. 237-238 of the text – see for details of data and question). Degree of Modernization is the IV and suicide rates per 100,000 people are the DV.

Low	Medium	High
4	17	20
8	10	22
7	9	19
5	12	9
		14

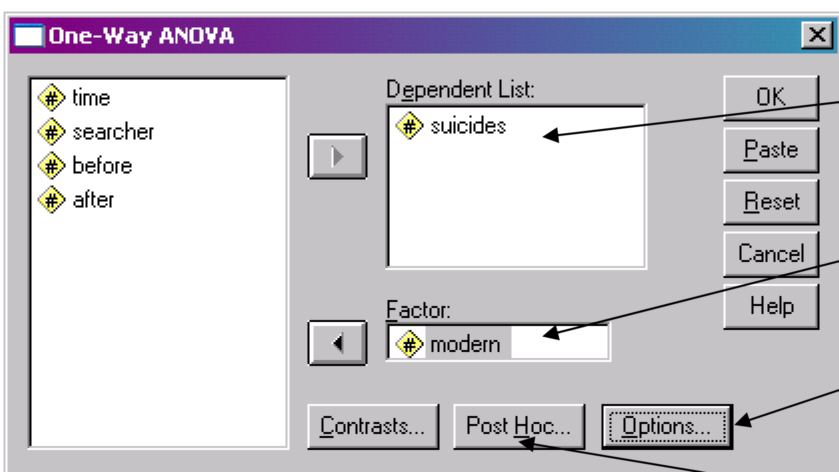
Just like the independent samples t-test set up in SPSS, put all of the three group's data in one column and then differentiate between groups via a grouping variable. I am using 1, 2, & 3 to represent low, medium, and high "modernized" countries.



Put your dependent variable in column 1. This will be the variable that was measured in the study

Independent variable in column two. SPSS refers to the independent variable as the "factor" This will be the variable that simply identifies which cases belong to which groups.

After your data is entered, On the SPSS main menu click on ANALYZE-> Compare Means -> Oneway ANOVA to get to the correct dialog box to set up your analysis.

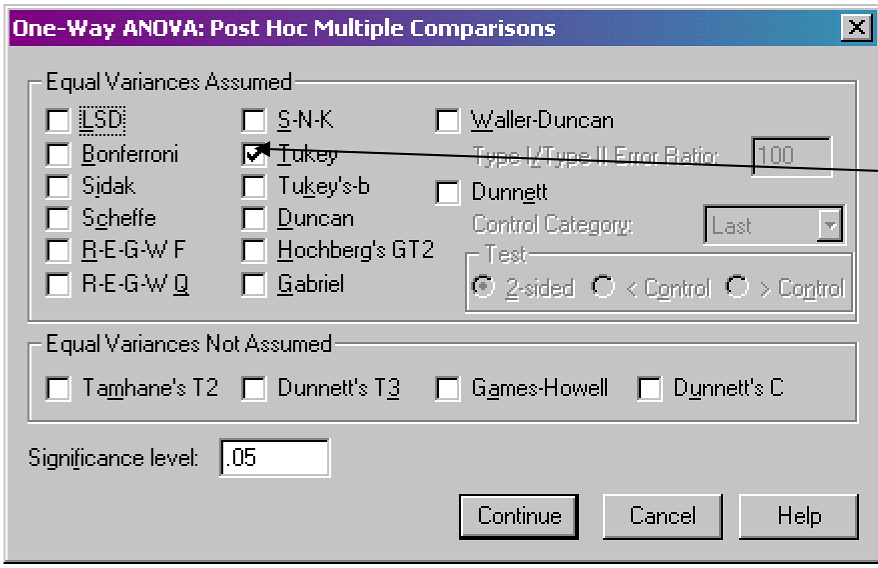


Dependent variable is whatever was measured in the study. In this case, suicide rates

Independent variable (SPSS refers to as "factor(s)". In this case, degree of modernization of the country in question.

Click Options button and check the box to obtain Descriptives (to get means and standard deviations for all groups)

Click Post Hoc to request multiple comparison tests (see next page)



Check here to request the Tukey HSD test, as our author mentions, there are a lot of different tests available, each of these little check boxes lets us request a different one.

The following output is obtained from SPSS Oneway ANOVA with Descriptives and Post-Hocs (Tukey HSD) requested
First we get descriptive statistics for each group separately and the overall sample.

Descriptives

suicides

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00 low	4	6.0000	1.82574	.91287	3.0948	8.9052	4.00	8.00
2.00 medium	4	12.0000	3.55903	1.77951	6.3368	17.6632	9.00	17.00
3.00 high	5	16.8000	5.26308	2.35372	10.2650	23.3350	9.00	22.00
Total	13	12.0000	5.90198	1.63691	8.4335	15.5665	4.00	22.00

Next is the ANOVA table. Be sure you are able to match the information in this table with the various parts of the calculations by hand that are outlined in the text.
In this case, the written summary would be $F(2,10)=8.161$, $p < .01$.

ANOVA

suicides

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	259.200	2	129.600	8.161	.008
Within Groups	158.800	10	15.880		
Total	418.000	12			

Our hypotheses are:
H_0: $\mu_{low} = \mu_{medium} = \mu_{high}$
H_1: not($\mu_{low} = \mu_{medium} = \mu_{high}$)
In this case, since our F statistic calculated from the data would arrive by chance if the null hypothesis was true with a probability of .008, we will reject the null hypothesis and conclude that there are differences of some sort in countries suicide rates based on the degree of modernization in the country, but we don't know which of the three means might be different from which. By simply inspecting the means in the descriptives table we got with the ANOVA output, we should suspect that the difference between the low and high modernization groups might be significant, since that is the biggest difference, but how about the difference between the low and the medium groups? This is the information we get from our multiple comparison test.

The table of results for the Tukey HSD tests – the table shows all of the possible pairs of comparisons that can be made, and as the legend shows, the mean differences that are statistically significant at the $p < .05$ level are marked with an asterisk in that column. Thus, the only difference that is statistically significant is between low and high modernization groups.

Multiple Comparisons

Dependent Variable: suicides

Tukey HSD

(I) modern	(J) modern	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00 low	2.00 medium	-6.00000	2.81780	.133	-13.7244	1.7244
	3.00 high	-10.80000*	2.67320	.006	-18.1280	-3.4720
2.00 medium	1.00 low	6.00000	2.81780	.133	-1.7244	13.7244
	3.00 high	-4.80000	2.67320	.220	-12.1280	2.5280
3.00 high	1.00 low	10.80000*	2.67320	.006	3.4720	18.1280
	2.00 medium	4.80000	2.67320	.220	-2.5280	12.1280

*. The mean difference is significant at the .05 level.