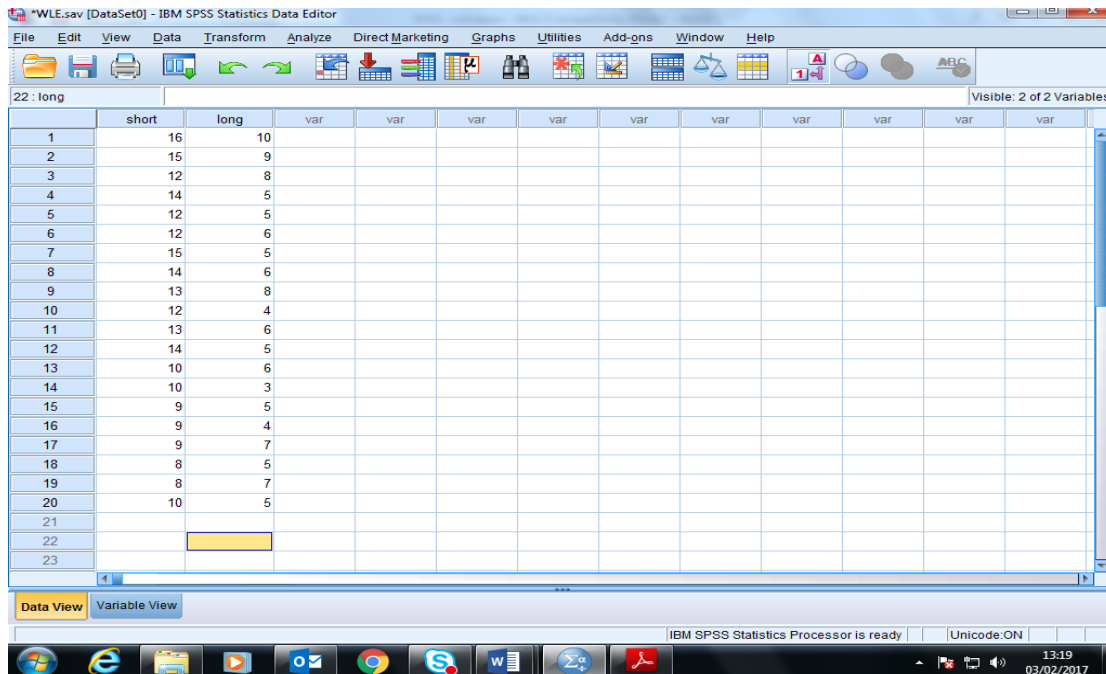


Introductory Research Methods and Statistics

The WLE task

Wilcoxon test

We will analyse the results of the WLE experiment using the Wilcoxon test. We have two scores from each participant, the long words score and the short words score, all scores recorded in number of words recalled. Hence we will have two columns of data (I have called them *long* and *short*) so your data file should look something like this (substituting your own class data).

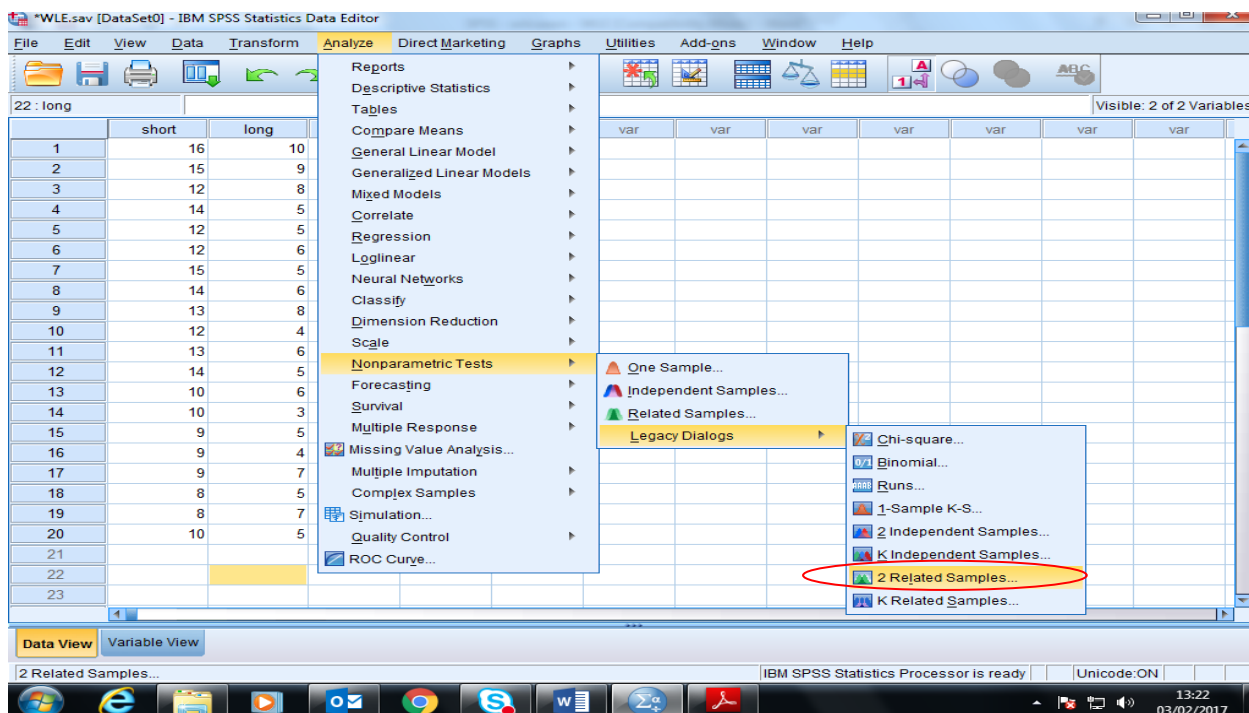


The screenshot shows the IBM SPSS Statistics Data Editor window. The dataset is named 'WLE.sav [DataSet0]'. The data is displayed in 'Data View' with two columns: 'short' and 'long'. The 'short' column contains values ranging from 8 to 16, and the 'long' column contains values ranging from 4 to 10. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready' and 'Unicode: ON'.

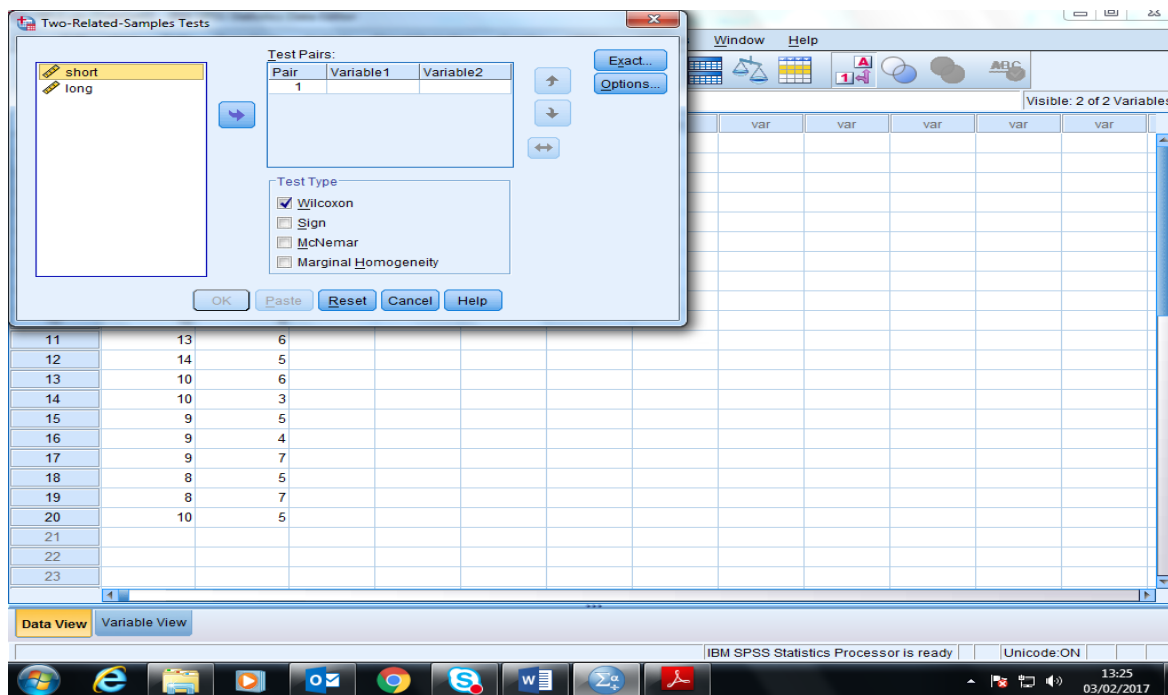
	short	long
1	16	10
2	15	9
3	12	8
4	14	5
5	12	5
6	12	6
7	15	5
8	14	6
9	13	8
10	12	4
11	13	6
12	14	5
13	10	6
14	10	3
15	9	5
16	9	4
17	9	7
18	8	5
19	8	7
20	10	5
21		
22		
23		

Next to conduct the analysis:

1. Go to the **Analyze** menu, and from this select **Nonparametric tests, Legacy Dialogs** then **2 Related samples**.

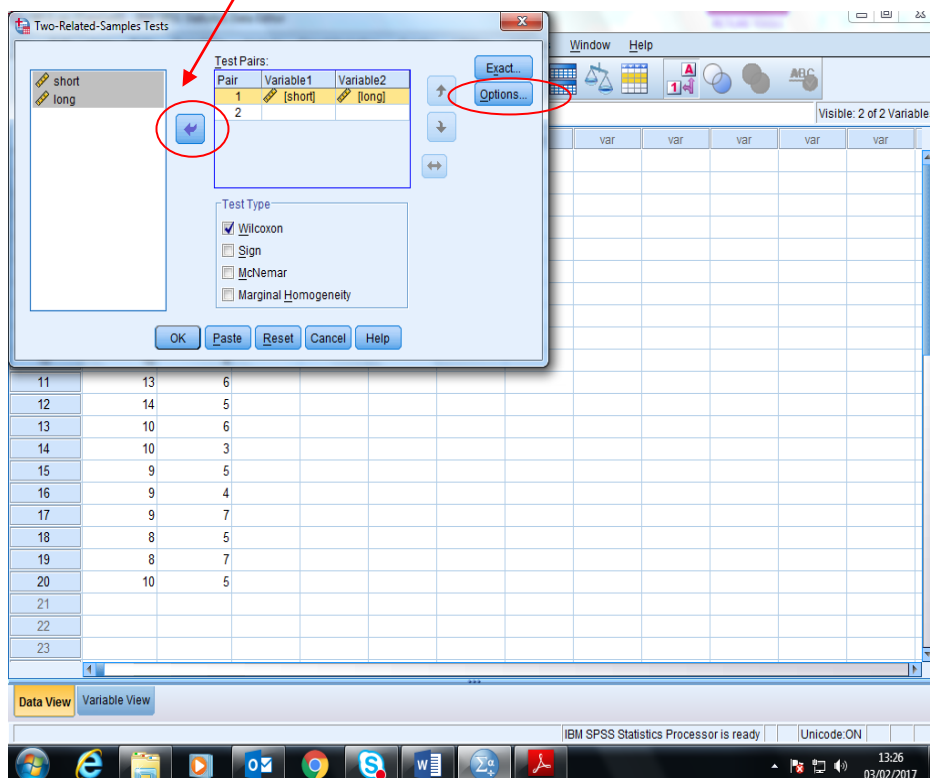


2. This produces the following dialogue box:



3. Click the mouse over the two variable names. Then click the arrow button to move them into the **Paired Variables** box:

Make sure that the Wilcoxon button is selected.



If the Exact test is needed, click on **Exact** button in the Two-Related Samples Tests dialog box, and the following Exact Tests dialog box will appear. Check the **Exact** bullet and enter the amount of time allowed for computing the exact sampling distribution of the test statistic (leave on 5mins).

4. Click on **Options** and then select the **Descriptive box and Quartiles**
5. **Click continue** and then **OK** button and SPSS will perform the analysis for you. The first table in the output will look as follows:

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	25th	Percentiles 50th (Median)	75th
short	20	11.75	2.489	8	16	9.25	12.00	14.00
long	20	5.95	1.761	3	10	5.00	5.50	7.00

This table allows you to inspect the mean scores for each condition. The prediction is that recall will be better for short words compared to long words. This is the case in the above example. As you have used a nonparametric test it is most likely that you should use the quartiles information to describe both your groups. What do your results indicate about the outcome of the experiment?

Ranks

		N	Mean Rank	Sum of Ranks
long - short	Negative Ranks	20 ^a	10.50	210.00
	Positive Ranks	0 ^b	.00	.00
	Ties	0 ^c		
	Total	20		

a. long < short

b. long > short

c. long = short

The second table above shows that the negative mean rank is more than the positive mean rank. This suggests that the recall in the long words condition was poorer than in the short words condition.

The final table shows the results of the Wilcoxon test to tell us whether or not there is a significant difference between conditions.

Test Statistics^a

	long - short
Z	-3.926 ^b
Asymp. Sig. (2-tailed)	.000
Exact Sig. (2-tailed)	.000
Exact Sig. (1-tailed)	.000
Point Probability	.000

P values

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

For your result to be significant, the **significance value should be less than or equal to .05**. Also note that by default, SPSS supplies us with a 2-tailed value, but the experimental hypothesis of our experiment was 1-tailed. In order to make your probability 1-tailed, divide the sig. value provided by SPSS by 2. In any case, the value in the above example is very low and much less than .05. It does not mean a zero percent probability of obtaining the result by chance, just that the probability is so low, it is off the scale! Hence in the above example we can conclude that the difference between the conditions is statistically significant.

Reporting the t-test results

Indicate the Z value degrees of freedom and whether or not your sig. Value was less than the significance level employed (in this case .05). **When you report p values – report $p = .007$, if the p value is .000 report $p < .001$ (so anything other than .000 use =).**

To report the above example:

A Wilcoxon signed-rank test showed a significant difference between the long and short words conditions ($Z = -3.926, p < .001$). Median recall scores for short words was significantly better than for long words (12 and 5.5, respectively).