Seminar 9
How to conduct and interpret the paired-sample t-test?

Quantitative Research Methods
Dr. Naser Alziyadat

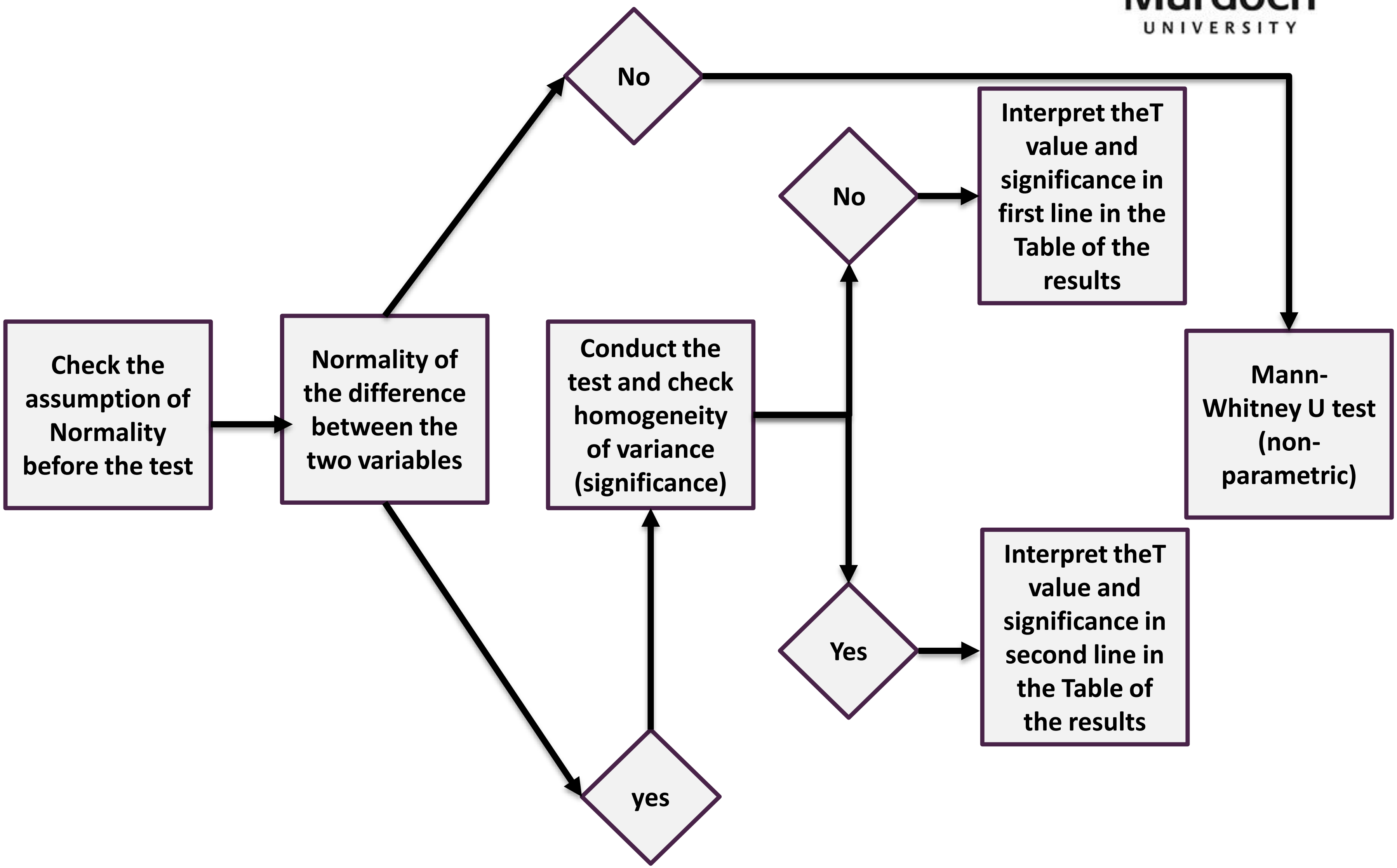
Murdoch
UNIVERSITY

Review

T-tests

- There are different types available in SPSS. We are going to cover two of them:
  - independent-samples t-test, used when you want to compare the mean scores of two different groups of people or conditions; and
  - paired-samples t-test, used when you want to compare the mean scores for the same group of people on two different occasions, or when you have matched pairs.
  - If you have more than two groups, or conditions, you will need to use analysis of variance instead (Seminar 10)

How to decide on independent sample T test



Review

Independent-samples t-test

- Independent-samples t-test, used when you want to compare the mean scores of two different groups of people or conditions.
- Example
  - survey.sav data file.
  - Explores sex differences in self-esteem scores.
  - The two variables used are
    - SEX (with males coded as 1, and females coded as 2) and
    - TSLFEST, which is the total score that participants recorded on a ten-item self-esteem scale.
- Research Question
  - Is there a significant difference in the mean self-esteem scores for males and females?
- Non-parametric alternative: Mann-Whitney Test

Normality assessment

- Analyse-Descriptive Statistics-Explore.
- Move the variable “Sex” to the Factor list.
- Move the variable “TSLFEST” to the dependent list.
- Click on plots- de-select Stem-and-Leaf>> Select Histogram>>Select Normality plots with tests.
- Are both total self esteem for males and total self esteem for females normally distributed?

Normality results

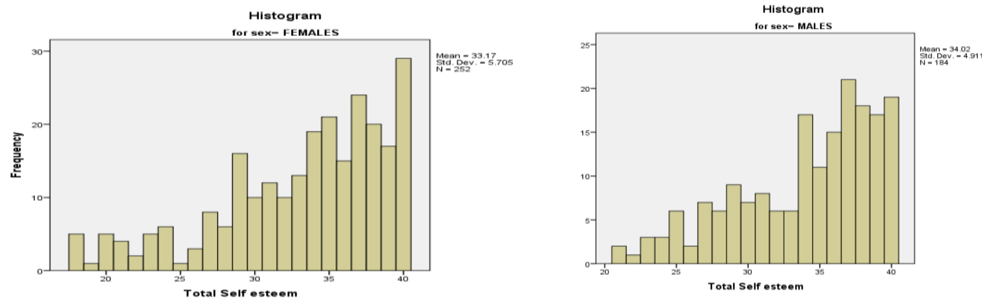
Descriptives

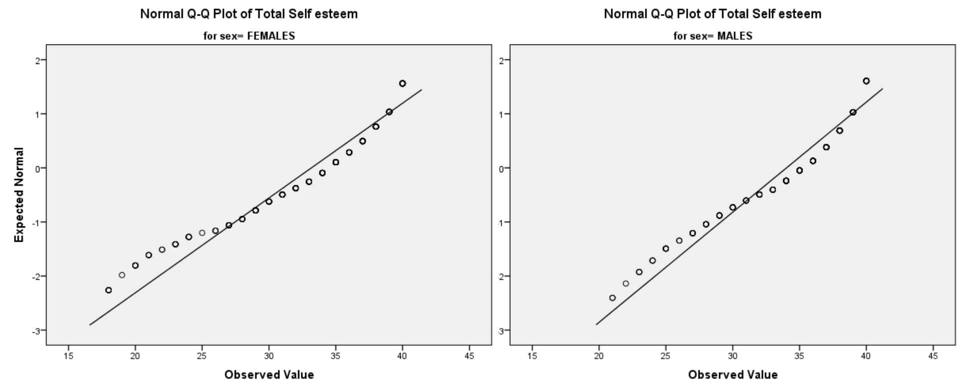
	sex			Statistic	Std. Error
Total Self esteem	MALES	Mean		34.02	.362
		95% Confidence Interval for Mean	Lower Bound	33.31	
			Upper Bound	34.74	
		5% Trimmed Mean		34.31	
		Median		35.00	
		Variance		24.120	
		Std. Deviation		4.911	
		Minimum		21	
		Maximum		40	
		Range		19	
		Interquartile Range		8	
		Skewness		-.758	.179
		Kurtosis		-.378	.356
	FEMALES	Mean		33.17	.359
		95% Confidence Interval for Mean	Lower Bound	32.47	
			Upper Bound	33.88	
		5% Trimmed Mean		33.57	
		Median		34.50	
		Variance		32.551	
		Std. Deviation		5.705	
		Minimum		18	
		Maximum		40	
		Range		22	
		Interquartile Range		8	
		Skewness		-.887	.153
		Kurtosis		.066	.306

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	sex	Statistic	df	Sig.	Statistic	df	Sig.
Total Self esteem	MALES	.146	184	.000	.916	184	.000
	FEMALES	.133	252	.000	.912	252	.000

a. Lilliefors Significance Correction

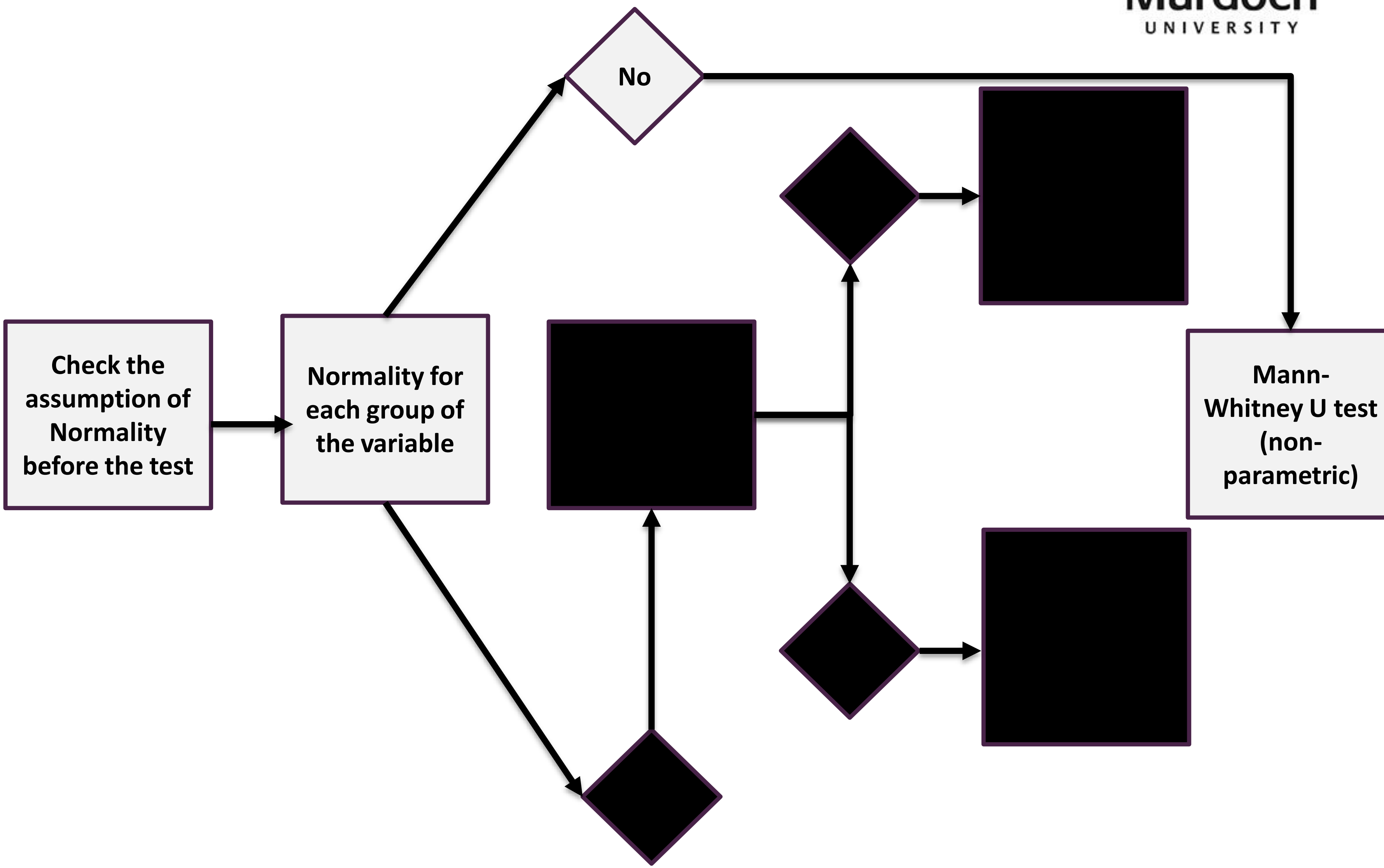




Criterion	Male	female
Skewness	-.758	-.887
KS, SW	< 0.05	< 0.05
Mean-median	Not large	A bit large
Histogram	Skewed	Skewed
Q-Q	Skewed	Skewed

Normality has been violated for both of them. Wee need to go for Mann-Whitney U test (Session 11).

How to decide on independent sample T test



Review

Procedure

Murdoch UNIVERSITY

Procedure for independent-samples t-test

1. From the menu at the top of the screen click on: **Analyze**, then click on **Compare means**, then on **Independent Samples T-test**.
2. Move the dependent (continuous) variable (e.g. total self-esteem) into the area labelled **Test variable**.
3. Move the independent variable (categorical) variable (e.g. sex) into the section labelled **Grouping variable**.
4. Click on **Define groups** and type in the numbers used in the data set to code each group. In the current data file 1=males, 2=females; therefore, in the **Group 1** box, type 1; and in the **Group 2** box, type 2.
5. Click on **Continue** and then **OK**.

Review

Interpretation of the results

Have a look at the mean and standard deviation for each of your groups.
Check the number of people in each group (N). Always check these values first. Do they seem right? Are the N values for males and females correct? Or are there a lot of missing data? If so, find out why. Perhaps you have entered the wrong code for males and females (0 and 1, rather than 1 and 2). Check with your codebook.

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
Total self-esteem	MALES	184	34.02	4.91	.36
	FEMALES	252	33.17	5.71	.36

Review

Interpretation of the results

Check Levene's test for equality of variances. This tests whether the variance (variation) of scores for the two groups (males and females) is the same.

* If your Sig. value is larger than .05 (e.g. .07, .10), you should use the *first line* in the table, which refers to **Equal variances assumed**.

* If the significance level of Levene's test is p=.05 or less (e.g. .01, .001), this means that the variances for the two groups (males/females) are not the same. Therefore your data violates the assumption of equal variance. In this case, you should use the information in the *second line* of the t-test table, which refers to **Equal variances not 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
Total self-esteem	Equal variances assumed	3.506	.062	1.622	434	.105	.85	.52	-.18	1.87
	Equal variances not assumed			1.661	422.349	.098	.85	.51	-.16	1.85

Review

Interpretation of the results

Murdoch UNIVERSITY

Assess significance of the difference

If the value in the **Sig. (2-tailed)** column is *equal or less* than .05 (e.g. .03,.01, .001), then there is a significant difference in the mean scores on your dependent variable for each of the two groups. If the value is *above* .05 (e.g. .06, .10), there is no significant difference between the two groups.

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
Total self-esteem	Equal variances assumed	3.506	.062	1.622	434	.105	.85	.52	-.18	1.87
	Equal variances not assumed			1.661	422.349	.098	.85	.51	-.16	1.85

Review

Interpretation of the results

Murdoch UNIVERSITY

Calculating the effect size for independent-samples t-test

Effect size statistics provide an indication of the **magnitude of the differences between your groups** (not just whether the difference could have occurred by chance).

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
Total self-esteem	MALES	184	34.02	4.91	.36
	FEMALES	252	33.17	5.71	.36

$$\text{Eta squared} = \frac{t^2}{t^2 + (N1 + N2 - 2)}$$

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
Total self-esteem	Equal variances assumed	3.506	.062	1.622	434	.105	.85	.52	-.18	1.87
	Equal variances not assumed			1.661	422.349	.098	.85	.51	-.16	1.85

Presenting the results

An independent-samples t-test was conducted to compare the self-esteem scores for males and females. There was no significant difference in scores for males (M=34.02, SD=4.91) and females [M=33.17, SD=5.71); t(434)=1.62, p=.11]. The magnitude of the differences in the means was very small (eta squared=.006).

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
Total self-esteem	MALES	184	34.02	4.91	.36
	FEMALES	252	33.17	5.71	.36

$$\text{Eta squared} = \frac{1.62^2}{1.62^2 + (184 + 252 - 2)}$$

$$\text{Eta squared} = .006$$

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
Total self-esteem	Equal variances assumed	3.506	.062	1.622	434	.105	.85	.52	-.18	1.87
	Equal variances not assumed			1.661	422.349	.098	.85	.51	-.16	1.85

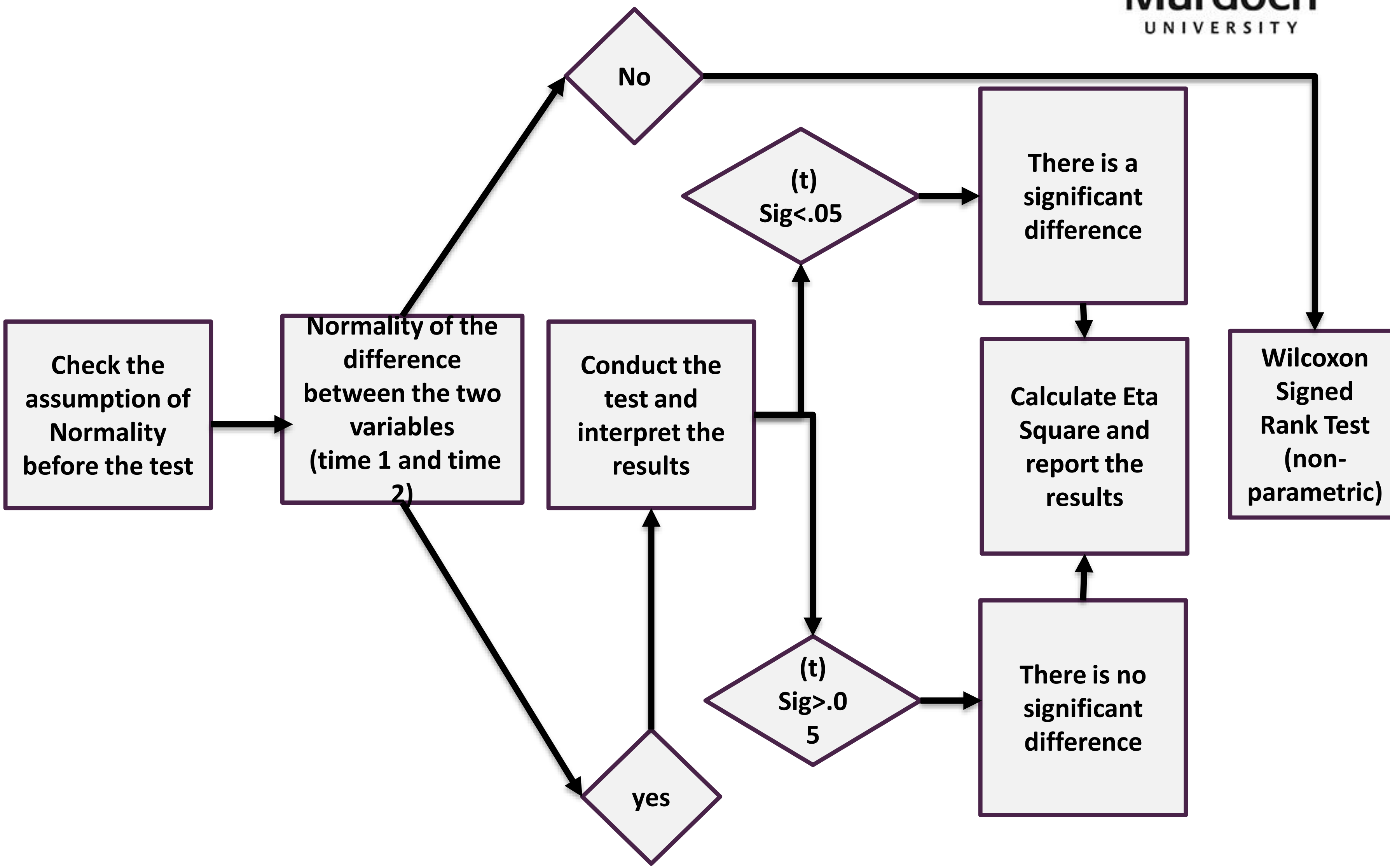
Paired-samples t-test

- Also referred to as repeated measures.
- Used when you have only one group of people (or companies, or machines etc.) and you collect data from them on two different occasions, or under two different conditions.
- Pre-test/post-test experimental designs are an example of the type of situation where this technique is appropriate.
- Example
  - experim.sav data file.
  - We will be exploring the impact of an intervention designed to increase students' confidence in their ability to survive a compulsory statistics course. Students were asked to complete a Fear of Statistics Test (FOST) both before (Time 1) and after the intervention (Time 2). The two variables from the data file that I will be using are: FOST1 (scores on the Fear of Statistics Test at Time 1) and FOST2 (scores on the Fear of Statistics Test at Time 2).

Paired-samples t-test

- Research question:
  - Is there a significant change in participants' fear of statistics scores following participation in an intervention designed to increase students' confidence in their ability to successfully complete a statistics course?
  - Does the intervention have an impact on participants' fear of statistics scores?
- A paired-samples t-test will tell you whether there is a statistically significant difference in the mean scores for Time 1 and Time 2.
- Additional assumption (to the general ones covered previously):
  - The difference between the two scores obtained for each subject should be normally distributed.
  - Non-parametric alternative
  - Wilcoxon Signed Rank Test

How to decide on paired-sample T test



Assessing the normality

- We need to create a new variable that represents the difference between the two variables.
- Procedure: Transform menu>>>Compute variables>>>Move the first variable to the Numeric expression>> then add minus>>> move the second variable to the numeric expression>>> Give a name to your new variable in the “Target Variable”>>> klick OK>
- Now follow the procedure of normality assessment to assess your new variable.

Normality

Descriptives[a]

			Statistic	Std. Error
Diffeence	Mean		-2.6667	.49441
	95% Confidence Interval for Mean	Lower Bound	-3.6779	
		Upper Bound	-1.6555	
	5% Trimmed Mean		-2.5185	
	Median		-2.0000	
	Variance		7.333	
	Std. Deviation		2.70801	
	Minimum		-10.00	
	Maximum		1.00	
	Range		11.00	
	Interquartile Range		3.25	
	Skewness		-.725	.427
	Kurtosis		.340	.833

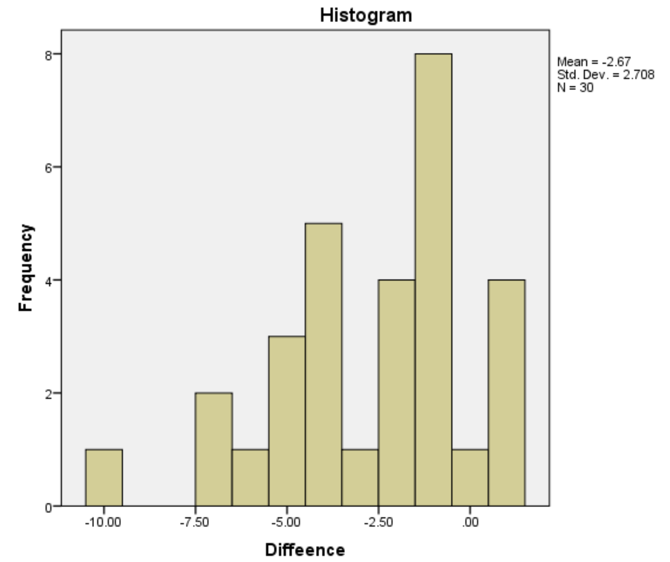
a. Diffeence2 is constant. It has been omitted.

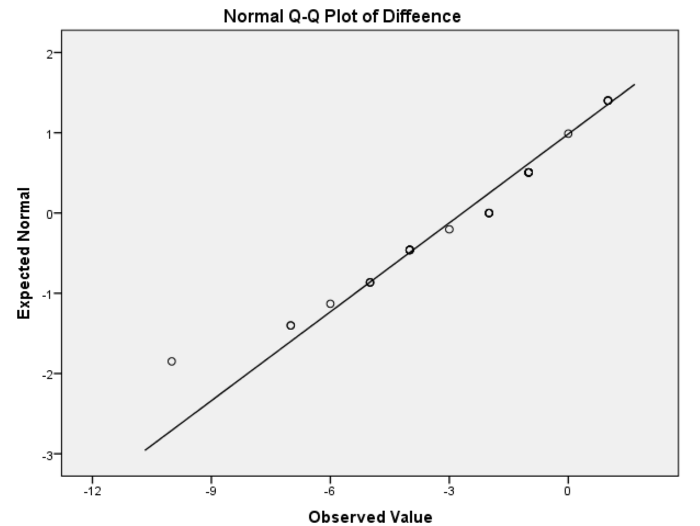
Tests of Normality[b]

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Diffeence	.164	30	.038	.931	30	.052

a. Lilliefors Significance Correction

b. Diffeence2 is constant. It has been omitted.





What is your decision?

Paired-samples t-test: Procedure

Let's assume the difference was normally distributed.

1. From the menu at the top of the screen click on: **Analyze**, then click on **Compare Means**, then on **Paired Samples T-test**.
2. Click on the two variables that you are interested in comparing for each subject (e.g. fost1: fear of stats time1, fost2: fear of stats time2).
3. With both of the variables highlighted, move them into the box labelled **Paired Variables** by clicking on the arrow button. Click on **OK**.

Interpreting the results

Have a look at the mean and standard deviation for each of your groups.
Check the number of people in each group (N). Always check these values first. Do they seem right? Are the N values for males and females correct? Or are there a lot of missing data? If so, find out why. Perhaps you have entered the wrong code for males and females (0 and 1, rather than 1 and 2). Check with your codebook.

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	fear of stats time1	40.17	30	5.16	.94
	fear of stats time2	37.50	30	5.15	.94

Interpreting the results

Check for the significance: if this value is less than .05 (e.g. .04, .01, .001), then you can conclude that there is a significant difference between your two scores. Therefore, we can conclude that there is a significant difference in the Fear of Statistics Test scores at Time 1 and at Time 2.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	fear of stats time1 - fear of stats time2	2.67	2.71	.49	1.66	3.68	5.394	29	.000

Interpreting the results

Compare the means: the mean Fear of Stats score at Time 1 was 40.17 and the mean score at Time 2 was 37.50. Therefore, we can conclude that there was a significant decrease in Fear of Statistics Test scores from Time 1 (prior to the intervention) to Time 2 (after the intervention).

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	fear of stats time1	40.17	30	5.16	.94
	fear of stats time2	37.50	30	5.15	.94

Calculate the size of effect

$$\text{Eta squared} = \frac{(5.39)^2}{(5.39)^2 + 30 - 1}$$

$$= \frac{29.05}{29.05 + 30 - 1}$$

$$= .50$$

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	fear of stats time1 - fear of stats time2	2.67	2.71	.49	1.66	3.68	5.394	29	.000

Reporting the results

A paired-samples t-test was conducted to evaluate the impact of the intervention on students' scores on the Fear of Statistics Test (FOST). There was a statistically significant decrease in FOST scores from Time 1 (*M*=40.17, *SD*=5.16) to Time 2 [*M*=37.5, *SD*=5.15, *t*(29)=5.39, *p*<.0005]. The eta squared statistic (.50) indicated a large effect size.

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	fear of stats time1	40.17	30	5.16	.94
	fear of stats time2	37.50	30	5.15	.94

$$\text{Eta squared} = \frac{(5.39)^2}{(5.39)^2 + 30 - 1} = \frac{29.05}{29.05 + 30 - 1} = .50$$

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	fear of stats time1 - fear of stats time2	2.67	2.71	.49	1.66	3.68	5.394	29	.000