Workshop 11
How to conduct and interpret Mann–Whitney U and Kruskal–Wallis tests?

Based on Pallant 2016

Dr. Naser Alziyadat

Murdoch UNIVERSITY

When to use Mann-Whitney U Test

- This test is the non-parametric alternative to the t-test for independent samples.
- The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either

  Ordinal; or

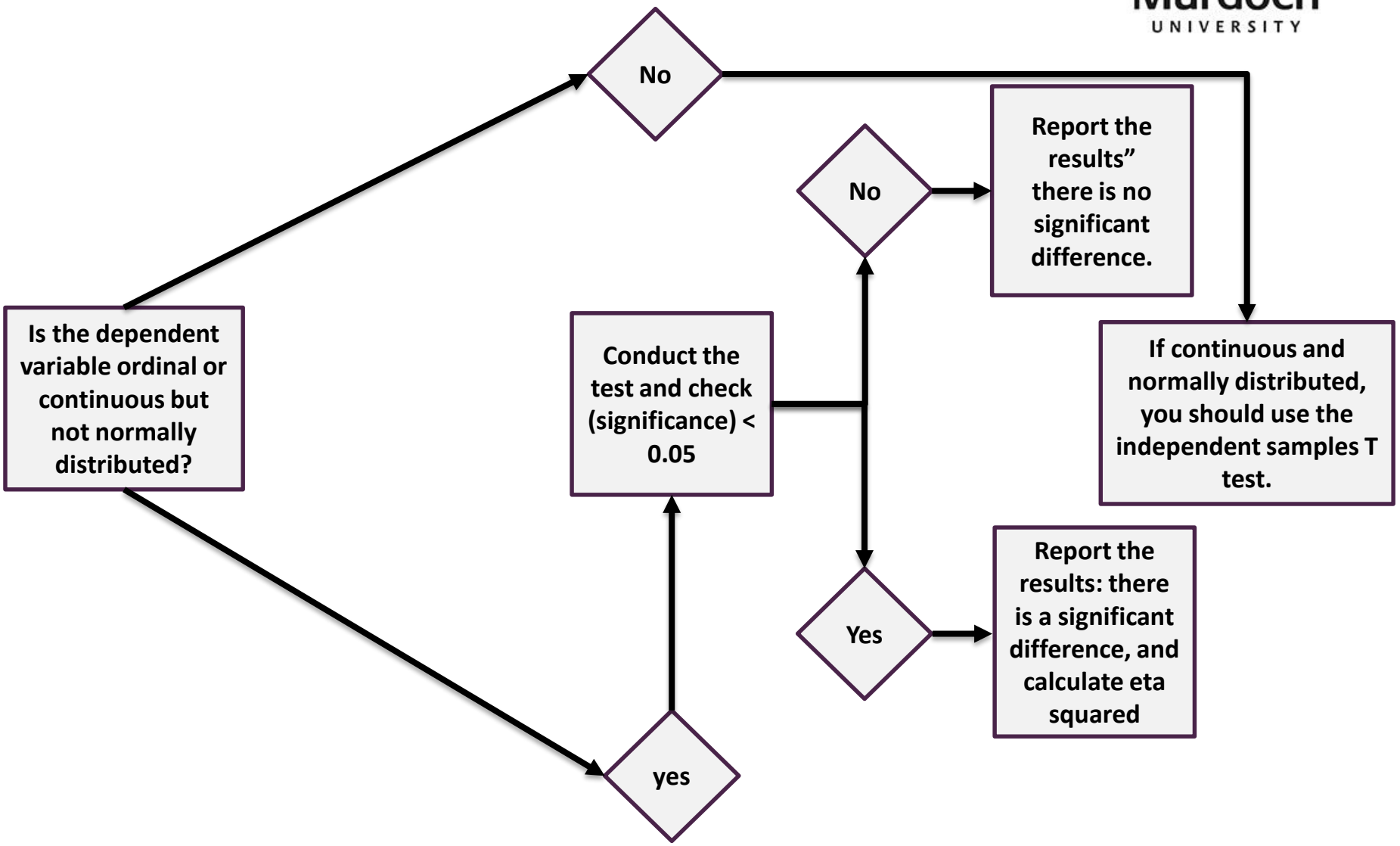
  continuous, but not normally distributed.

Mann-Whitney U Test

Murdoch UNIVERSITY

- This technique is used to test for differences between two independent groups on a continuous measure.
- Instead of comparing means of the two groups, as in the case of the t-test, the Mann-Whitney U Test actually compares medians.
- It converts the scores on the continuous variable to ranks, across the two groups.
- It then evaluates whether the ranks for the two groups differ significantly.
- As the scores are converted to ranks, the actual distribution of the scores does not matter.

How to decide on Mann-Whitney U test



Example

- Example
  - survey.sav data file.
  - Explore 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?

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

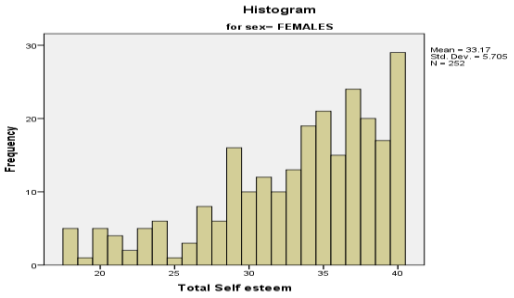
Descriptives

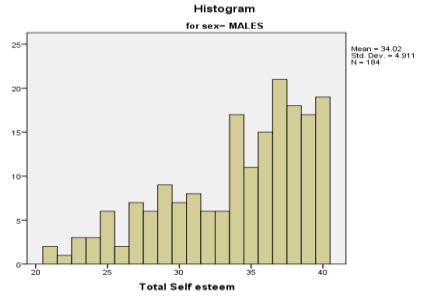
	sex			Statistic	Std. Error
Total Self esteem	MALES	Mean		34.02	.362
		95% Confidence Interval for Mean	Lower Bound	33.31	
			Upper Bound	34.74	
		5% Trimmed Mean		34.31	
		Median		35.00	
		Variance		24.120	
		Std. Deviation		4.911	
		Minimum		21	
		Maximum		40	
		Range		19	
		Interquartile Range		8	
		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

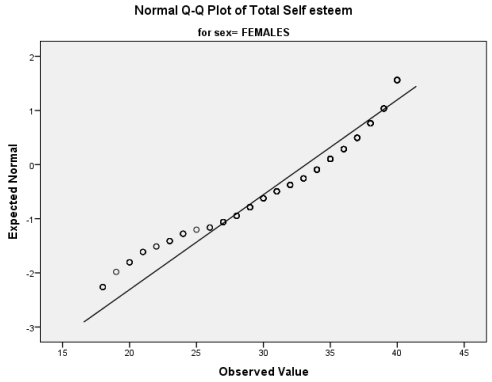
Tests of Normality

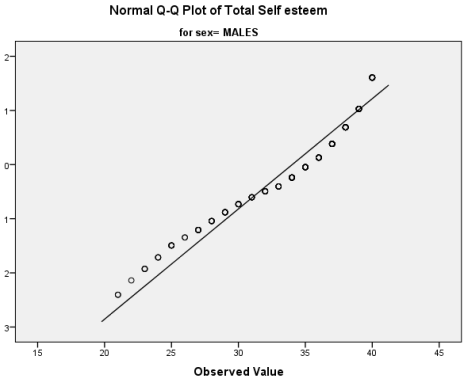
		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	sex	Statistic	df	Sig.	Statistic	df	Sig.
Total Self esteem	MALES	.146	184	.000	.916	184	.000
	FEMALES	.133	252	.000	.912	252	.000

a. Lilliefors Significance Correction





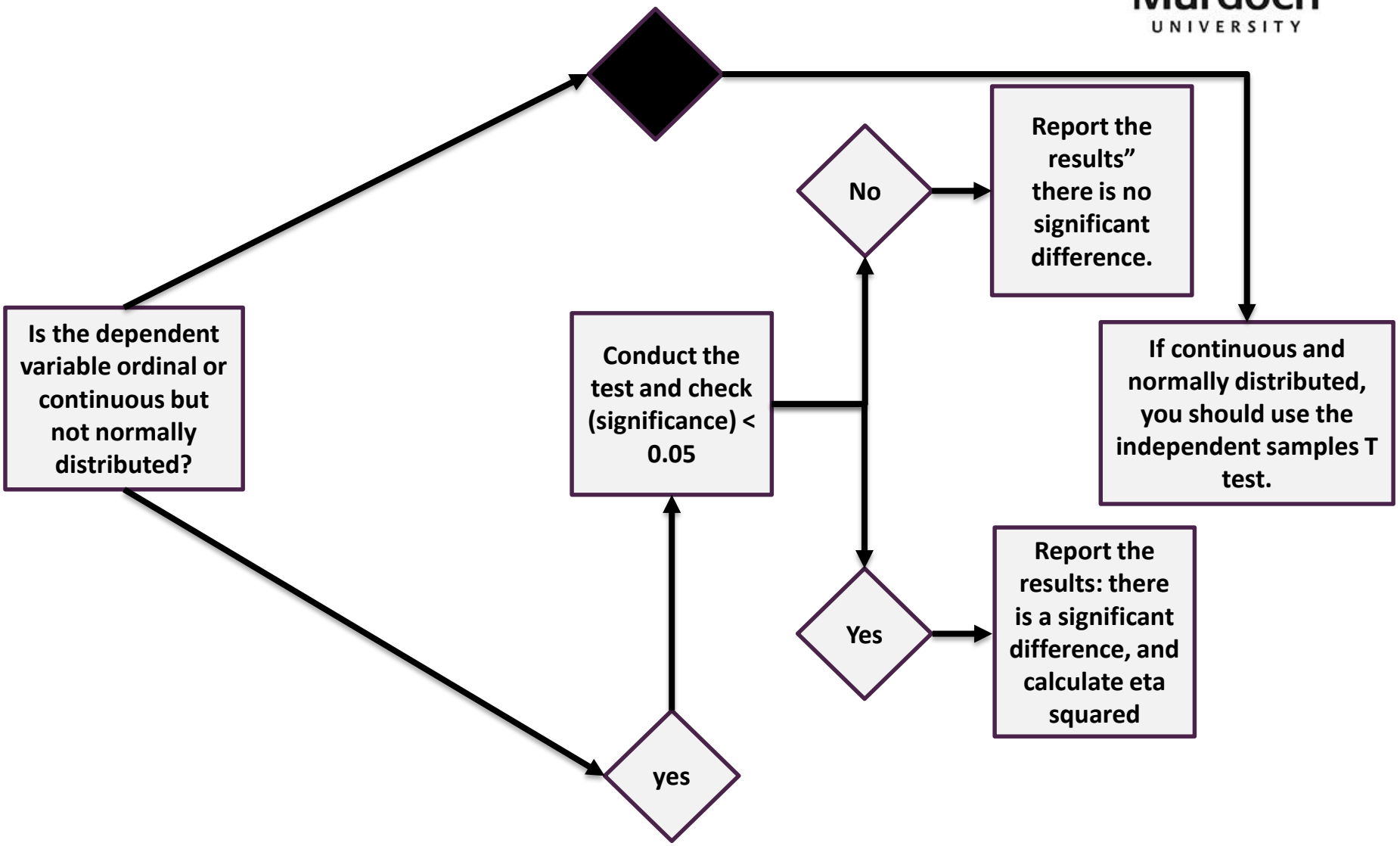




Criterion	Male	female
Skewness	-.758	-.887
KS, SW	< 0.05	< 0.05
Mean-median	Not large	A bit large
Histogram	Skewed	Skewed
Q-Q	Skewed	Skewed

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

How to decide on Mann-Whitney U test



Mann-Whitney U Test

Murdoch UNIVERSITY

Data file: Survey, on LMS

Procedure for Mann-Whitney U Test

1. From the menu at the top of the screen click on: **Analyze**, then click on **Non-parametric Tests**, then on **2 Independent Samples**.
2. Click on your continuous (dependent) variable (e.g. total self-esteem) and move it into the **Test Variable List** box.
3. Click on your categorical (independent) variable (e.g. sex) and move into **Grouping Variable** box.
4. Click on **Define Groups** button. Type in the value for Group 1 (e.g. 1) and for Group 2 (e.g. 2). These are the values that were used to code your values for this variable (see your codebook). Click on **Continue**.
5. Make sure that the **Mann-Whitney U** box is ticked under the section labelled **Test Type**. Click on **OK**.

Interpreting the Results

Murdoch UNIVERSITY

Ranks

	sex	N	Mean Rank	Sum of Ranks
Total Self esteem	MALES	184	227.14	41794.00
	FEMALES	252	212.19	53472.00
	Total	436		

Test Statistics[a]

	Total Self esteem
Mann-Whitney U	21594.000
Wilcoxon W	53472.000
Z	-1.227
Asymp. Sig. (2-tailed)	.220

a. Grouping Variable: sex

Interpretation of output from Mann-Whitney U Test

The two values that you need to look at in your output are the Z value and the significance level, which is given as Asymp. Sig (2-tailed). If your sample size is larger than 30, SPSS will give you the value for a Z-approximation test which includes a correction for ties in the data. In the example given above, the Z value is –1.23 (rounded) with a significance level of p=.22. The probability value (p) is not less than or equal to .05, so the result is not significant. There is no statistically significant difference in the self-esteem scores of males and females.

The results of the Mann-Whitney test

Ranks

	sex	N	Mean Rank	Sum of Ranks
Total Self esteem	MALES	184	227.14	41794.00
	FEMALES	252	212.19	53472.00
	Total	436		

Test Statistics[a]

	Total Self esteem
Mann-Whitney U	21594.000
Wilcoxon W	53472.000
Z	-1.227
Asymp. Sig. (2-tailed)	.220

a. Grouping Variable: sex

Report the results according to an academic style.
This link helps you
https://acquia-prod.oswego.edu/psychology/sites/acquiaprod.oswego.edu.psychology/files/mwu.pdf

Another example

- Explore the difference of the ordinal variable ‘level of education’ across ‘sex.
- If the variable is ordinal, no need to assess the normality.

Kruskal-Wallis Test

- This technique is the non-parametric alternative to a one-way between-groups analysis of variance.
- Scores are converted to ranks and the mean rank for each group is compared. This is a ‘between groups’ analysis, so different people must be in each of the different groups.

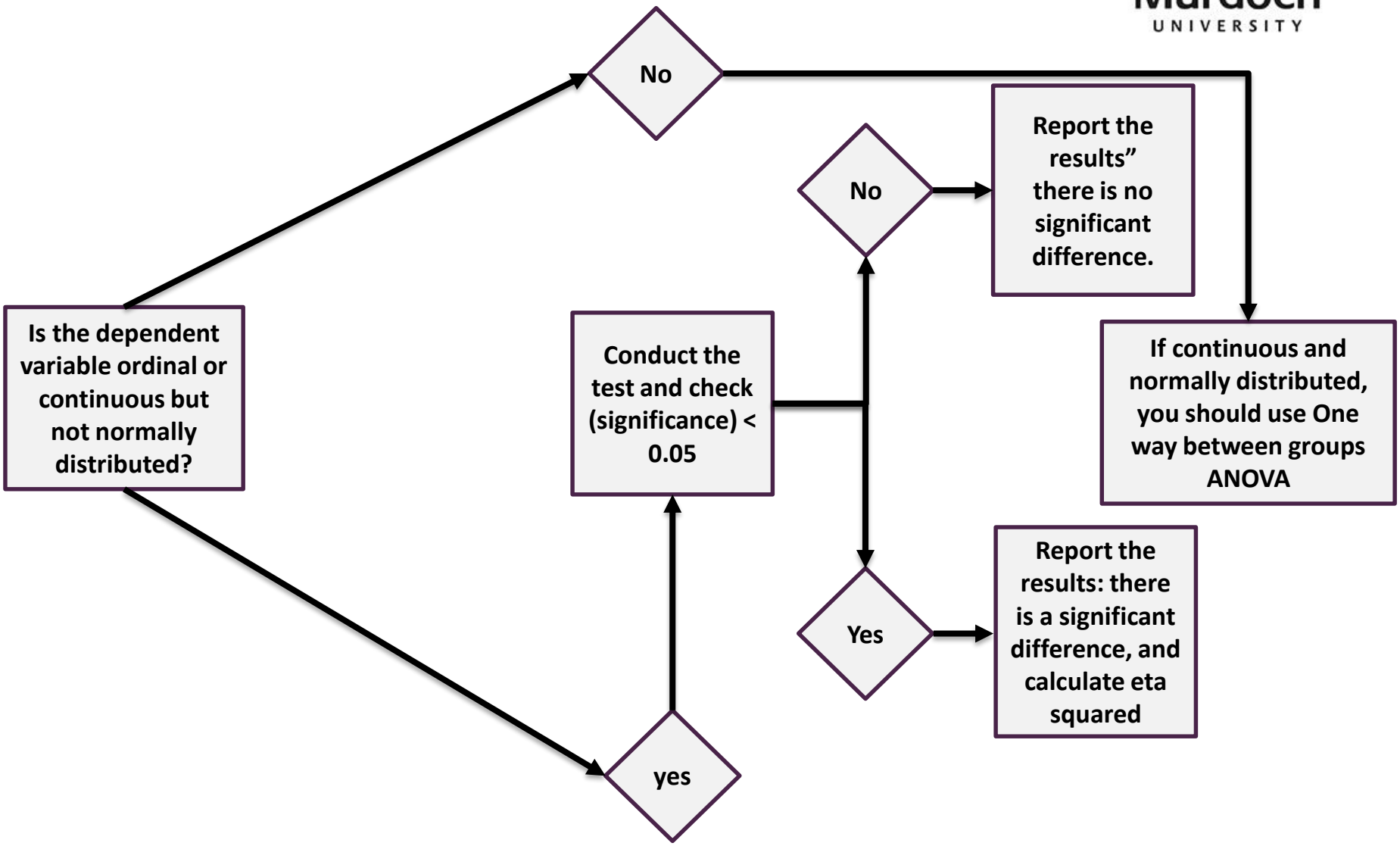
When to use Kruskal-Wallis test

- This test is the non-parametric alternative to the t-test for One Way between Groups ANOVA.
- The Kurskal Wallis test is used to compare differences between three or more independent groups when the dependent variable is either

  Ordinal; or

  continuous, but not normally distributed.

How to decide on Kruskal-Wallis test



Example

- Example from survey.sav data file.
- Research Question
  - Is there a difference in optimism scores for young, middle-aged and old subjects?
  - Check the dependent variable: is it continuous or ordinal?
  - Continuous then it should be heavily skewed.
  - Conduct normality assessment.

File name	Variable name	Variable label	Coding instructions
survey.sav	Toptim	Total optimism	Total score on the Optimism scale. Scores can range from 6 to 30 with high scores indicating higher levels of optimism.
	Agegp3	Agegp3	This variable is a recoded variable, dividing age into three equal groups (see instructions for how to do this in Chapter 8): Group 1: 18–29 = 1 Group 2: 30–44 = 2

Normality assessment

- Analyse-Descriptive Statistics-Explore.
- Move the variable “Agep3” to the Factor list.
- Move the variable “Toptim” to the dependent list.
- Click on plots- de-select Stem-and-Leaf>> Select Histogram>>Select Normality plots with tests.
- Are the data for each of the age groups normally distributed?

Let’s assume (consider) the variable Toptim is heavily skewed.

Kruskal-Wallis Test

Murdoch UNIVERSITY

Procedure for Kruskal-Wallis Test

1. From the menu at the top of the screen click on: **Analyze**, then click on **Non-parametric Tests**, then on **K Independent Samples**.
2. Click on your continuous (dependent variable) (e.g. total optimism) and move it into the **Test Variable List** box.
3. Click on your categorical (independent variable) (e.g. agegp3) and move it into the **Grouping Variable** box.
4. Click on the **Define Range** button. Type in the first value of your categorical variable (e.g., 1) in the **Minimum** box. Type the largest value for your categorical variable (e.g. 3) in the **Maximum** box. Click on **Continue**.
5. In the **Test Type** section make sure that the **Kruskal-Wallis H** box is ticked. Click on **OK**.

Interpreting the Results

Ranks

	age 3 groups	N	Mean Rank
Total Optimism	18 - 29	147	198.18
	30 - 44	153	216.05
	45+	135	241.80
	Total	435	

Test Statistics[a,b]

	Total Optimism
Chi-Square	8.573
df	2
Asymp. Sig.	.014

a. Kruskal Wallis Test

b. Grouping Variable: age 3 groups

The main pieces of information you need from this output are: Chi-Square value, the degrees of freedom (df) and the significance level (presented as Asymp. Sig.). If this significance level is a value less than .05 (e.g. .04, .01, .001), then you can conclude that there is a statistically significant difference in your continuous variable across the three groups. You can then inspect the Mean Rank for the three groups presented in your first output table. This will tell you which of the groups had the highest overall ranking that corresponds to the highest score on your continuous variable.

In the output presented above the significance level was .01 (rounded). This is less than the alpha level of .05, so these results suggest that there is a difference in optimism levels across the different age groups. An inspection of the mean ranks for the groups suggest that the older group (45+) had the highest optimism scores, with the younger group reporting the lowest.

The results of the Kruskal-Wallis test

Report the results according to an academic style.
This link helps you
https://www.slideshare.net/plummer48/reporting-a-kruskal-wallis-test

After reporting you must know between which groups was the difference, thus, you have to conduct Mann-Whitney U tests between each pairs to explore the differences between groups. Do it now, and compare it with your ANOVA results.

Another example

- The explore the differences of the level of education across age groups ‘Agep3’.
- If the variable is ordinal, no need to test the normality.