Dr Silia Vitoratou

Department: Biostatistics and Health Informatics

Topic materials:
Silia Vitoratou

Contributions:
Zahra Abdula

Improvements:
Nick Beckley-Hoelscher
Kim Goldsmith
Sabine Landau

Institute of Psychiatry, Psychology and Neuroscience

Module Title: Introduction to Statistics

Session Title: Equality of medians (non-parametric tests)

Topic title: Comparing groups II (non-parametric methods)

Learning Outcomes

- Learn when and how to use the **non-parametric** tests for equality of medians.
- Understand the assumptions of the various test of equality of medians.
- Be able to conduct these test in a statistical software.

Previously on ‘Introduction to Statistics’…..

Based on the **type** of data and on the hypotheses, we use different statistical tests.

Hypotheses testing	Means	Proportions
one group *versus* a pre-defined value	one sample t-test	one sample χ^2-test
one group *versus* another group	two independent samples t-test	(two independent samples) Pearson’s χ^2-test
one group (twice or) *versus* another matched group	two paired samples t-test	(two paired samples) McNemar test

Previously on 'Introduction to Statistics'.....

Based on the **type** of data and on the hypotheses, we use different statistical tests.

Hypotheses testing	Means	Proportions
one group *versus* a pre-defined value	is the mean weight equal to μ_0=66kg?	is the proportion of women in the sample π_0=50%?
one group *versus* another group	is the mean 'weight before' equal across genders?	proportion of women who exercised before =proportion of men who exercised before ?
one group (twice or) *versus* another matched group	is the mean 'weight before' equal to the mean 'weight after'?	proportion of those who exercised before =proportion of those who exercised after ?

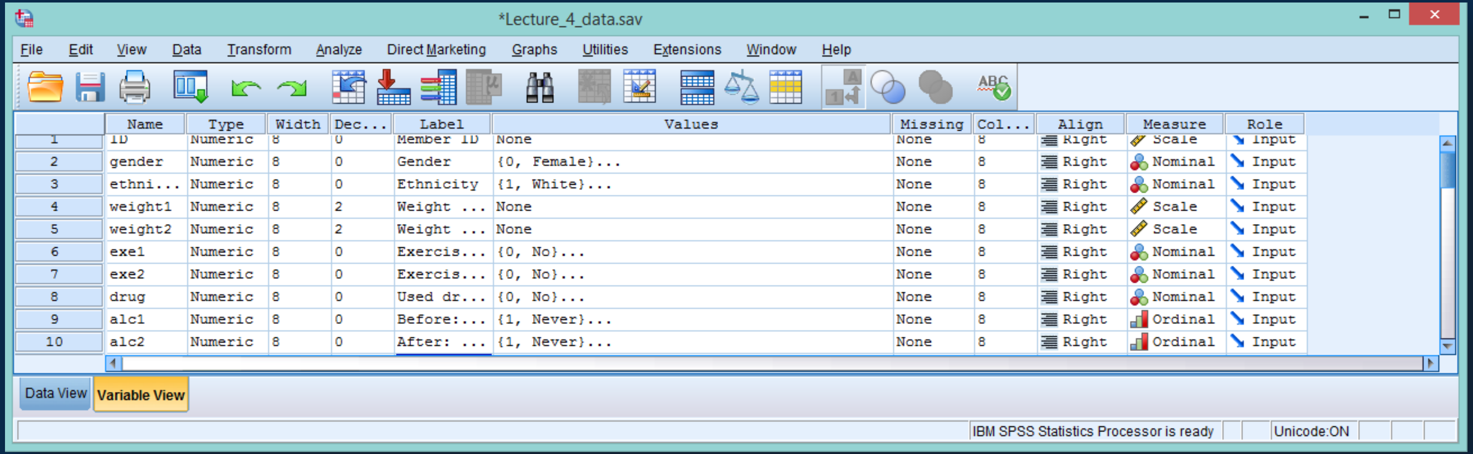
What if the assumptions do not hold?

Parametric and Non-Parametric Tests

Numerical data	Normality assumed	Normality not assumed
Hypotheses testing	parametric	Non-parametric
one group *versus* a pre-defined value	one sample t-test	Wilcoxon signed rank
one group *versus* another group	two independent samples t-test	Mann-Whitney (Wilcoxon sum rank)
one group (twice or) *versus* another matched group	two paired samples t-test	Wilcoxon signed rank for paired samples

SPSS Slide

Download the data that we are going to use during the lecture. The dataset is the **lecture_5_data.sav**. We have data for 300 individuals.



	Name	Type	Width	Dec...	Label	Values	Missing	Col...	Align	Measure	Role
1	ID	Numeric	8	0	Member ID	None	None	8	Right	Scale	Input
2	gender	Numeric	8	0	Gender	{0, Female}...	None	8	Right	Nominal	Input
3	ethni...	Numeric	8	0	Ethnicity	{1, White}...	None	8	Right	Nominal	Input
4	weight1	Numeric	8	2	Weight ...	None	None	8	Right	Scale	Input
5	weight2	Numeric	8	2	Weight ...	None	None	8	Right	Scale	Input
6	exe1	Numeric	8	0	Exercis...	{0, No}...	None	8	Right	Nominal	Input
7	exe2	Numeric	8	0	Exercis...	{0, No}...	None	8	Right	Nominal	Input
8	drug	Numeric	8	0	Used dr...	{0, No}...	None	8	Right	Nominal	Input
9	alc1	Numeric	8	0	Before:...	{1, Never}...	None	8	Right	Ordinal	Input
10	alc2	Numeric	8	0	After: ...	{1, Never}...	None	8	Right	Ordinal	Input

Same scenarios as in lecture 4, but this time the assumptions are violated)

- **gender**: 1-male, 0-female and **ethnicity** : 1-white, 2-black, 3-Asian, 4-other
- **weight1**: their weight when they entered the programme (in kg)
- **weight2**: their weight by the end of the programme (in kg)
- **exe1**: info if they regularly exercised (1-yes, 0-no) when they entered the programme
- **exe2**: info if they regularly exercised (1-yes, 0-no) by the end of the programme
- **drug**: if they have ever used drugs to lose weight (1-yes, 0-no)
- **alc1**: more than 2 units of alcohol, before (1:never, 2: sometimes, 3:always)
- **alc2**: more than 2 units of alcohol, after (1:never, 2: sometimes, 3:always)

Wilcoxon Signed Rank Test

The Wilcoxon signed rank test is the non-parametric analogue of the one sample t-test.

When to use

- Skewed continuous data
- Ordinal (interval) or discrete data

Null hypothesis

- H_0: Median equals a certain pre-defined value
- H_a: Median is different than a certain pre-defined value

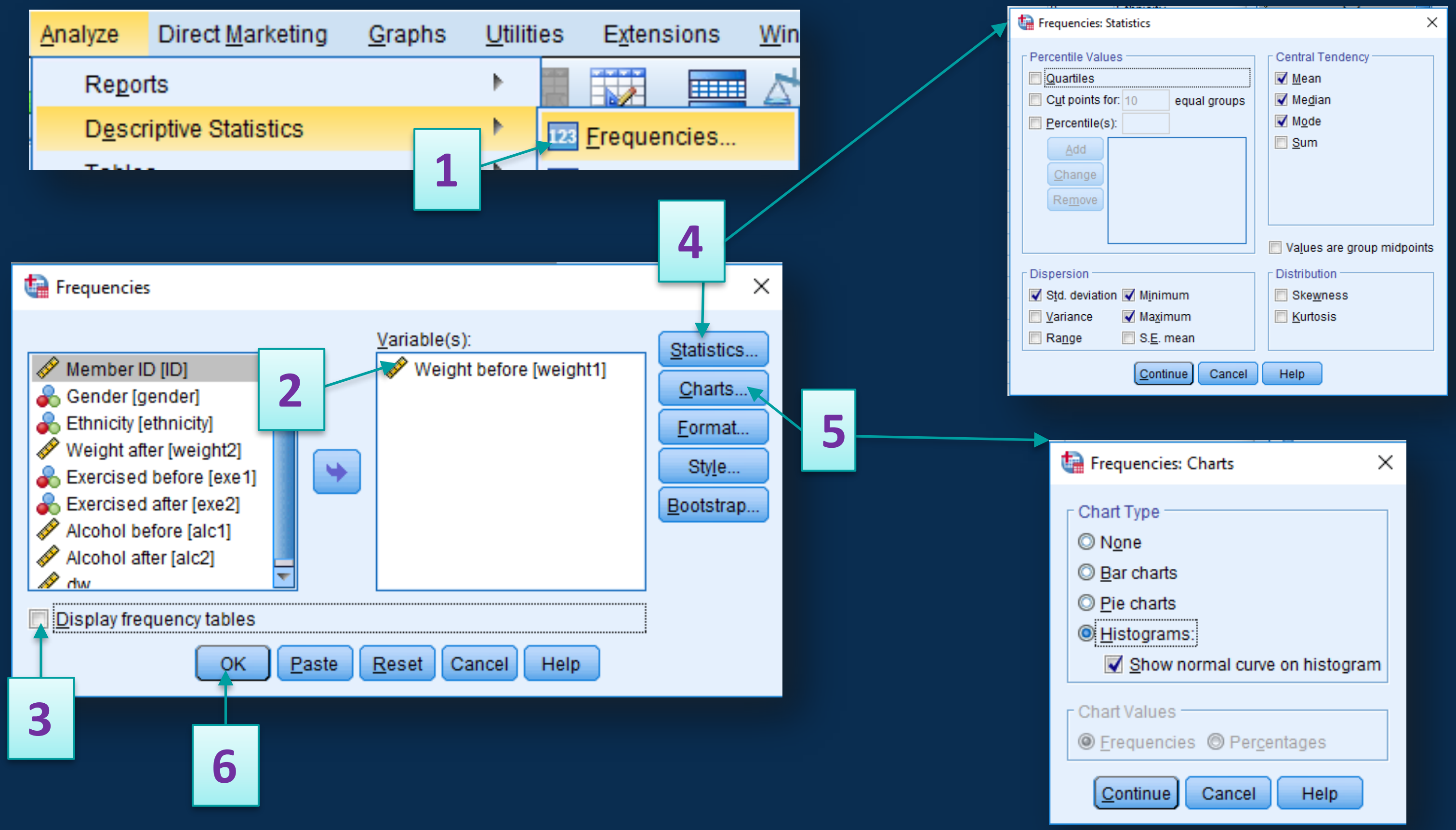
Assumptions

- The observations are randomly and independently drawn
- At least interval data

SPSS Slide: 'how to'

According to the researchers, in the population from which our data were sampled from the median weight of the people is **66kg**. Do our data support this?

<u>Step 1</u>: Check the suitability of the data, here what type of distribution has 'weight1' ?



In 'Statistics' ask for descriptive statistics

In 'Charts' ask for a Histogram

Output & Interpretation Slide

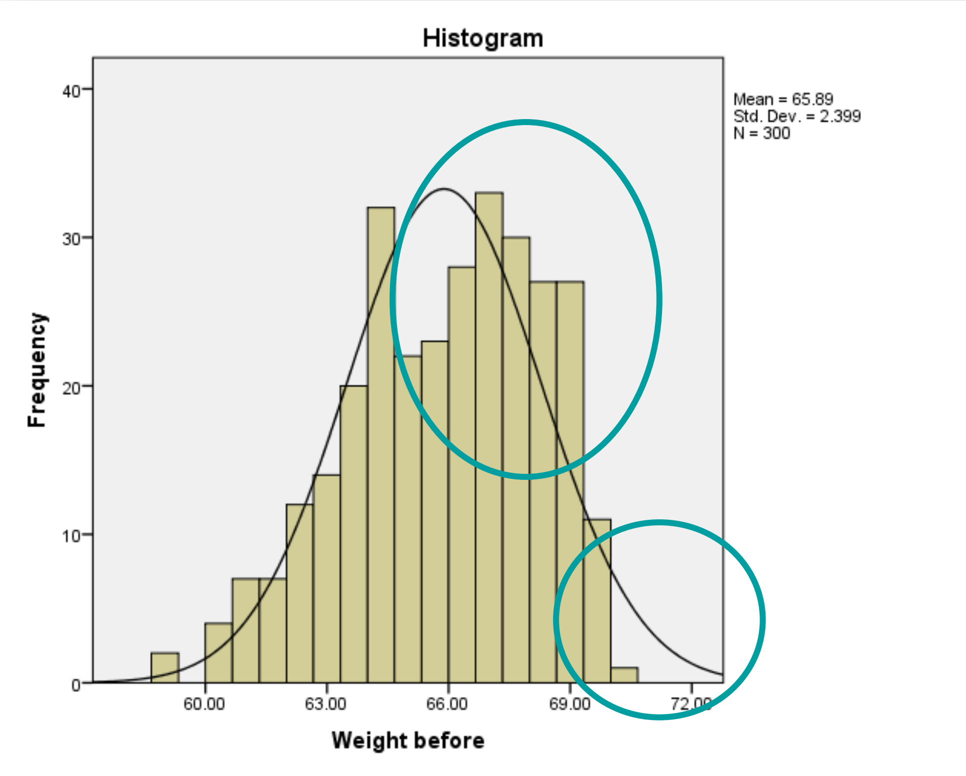
Step 1: Check the suitability of the data, here: what type of distribution has 'weight1' ?

Statistics

Weight before

N	Valid	300
	Missing	0
Mean		65.8856
Median		66.2450
Mode		62.55[a]
Std. Deviation		2.39880
Minimum		58.70
Maximum		70.04

a. Multiple modes exist. The smallest value is shown



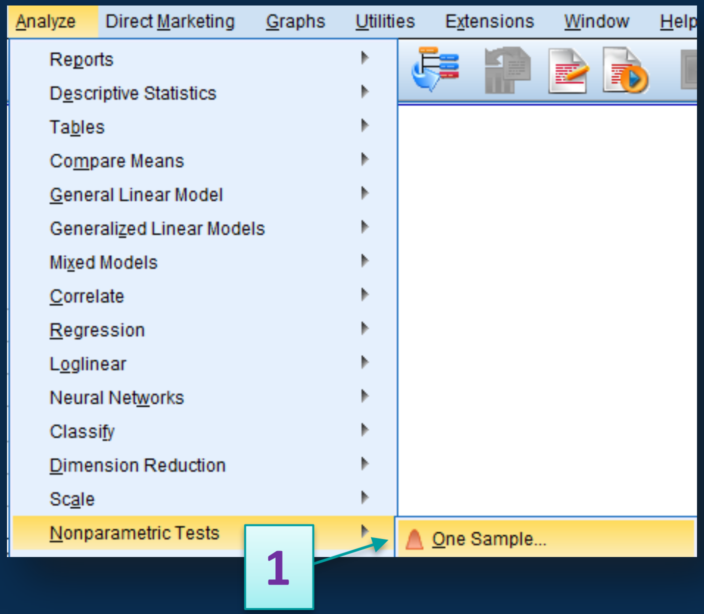
'weight1' is a negatively skewed variable, we can conclude it is not normally distributed. Therefore it is best not to rely on its mean and standard deviation, and use a non-parametric test.

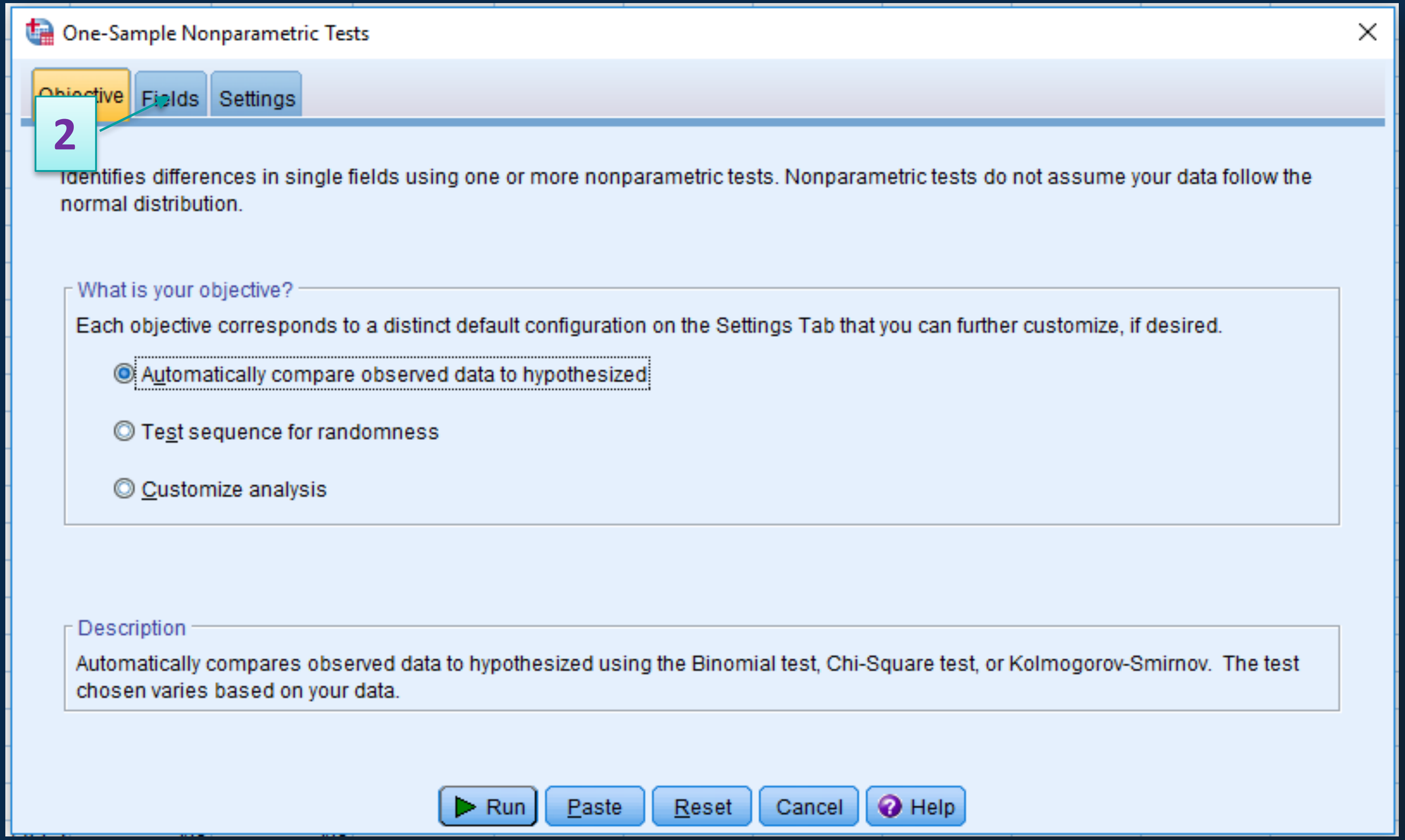
H_0: Median = 66kg H_a: Median ≠ 66kg

SPSS Slide: 'how to'

Step 2: Use the appropriate test, here: 'Wilcoxon Signed Rank test'.

Analyse -> Nonparametric tests-> 'One sample'



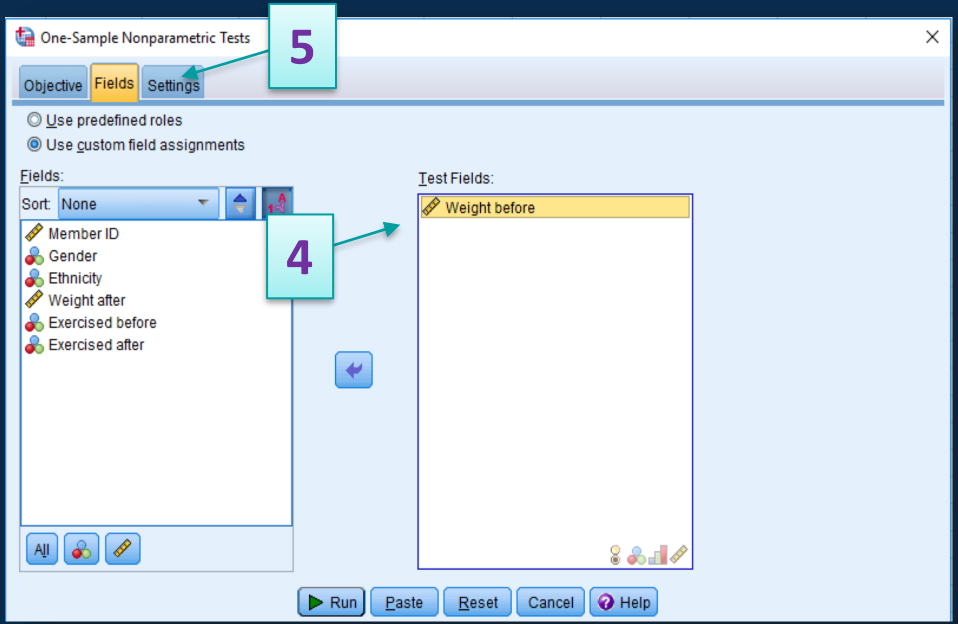


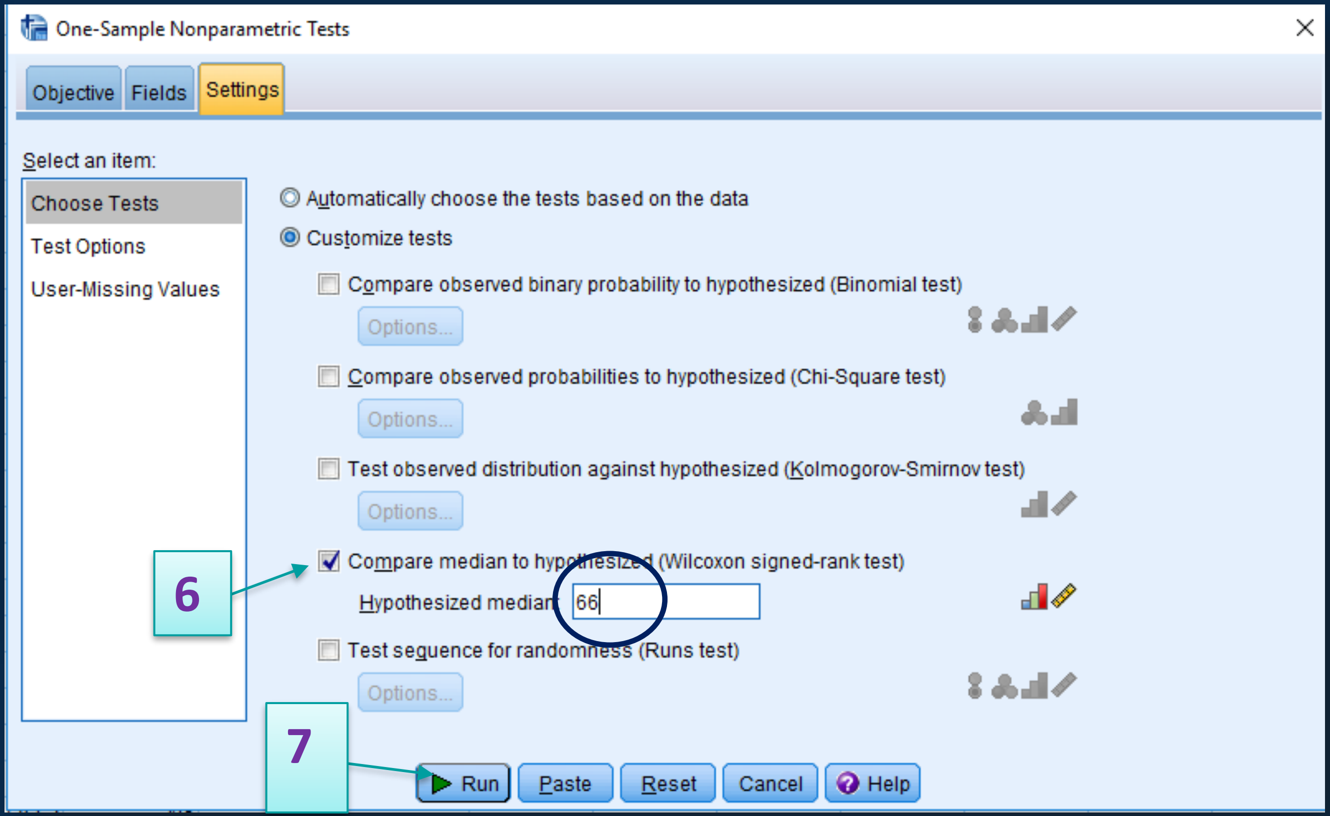
Click on the 'Fields' tab '

SPSS Slide: ‘how to’

Step 2: Use the appropriate test, here: ‘Wilcoxon Signed Rank test’.

Analyse -> Nonparametric tests-> ‘One sample’





Add the variable of interest (weight before) in the ‘Test Fields’ box

Add in the hypothesized median: 66kg

Output and Interpretation Slide

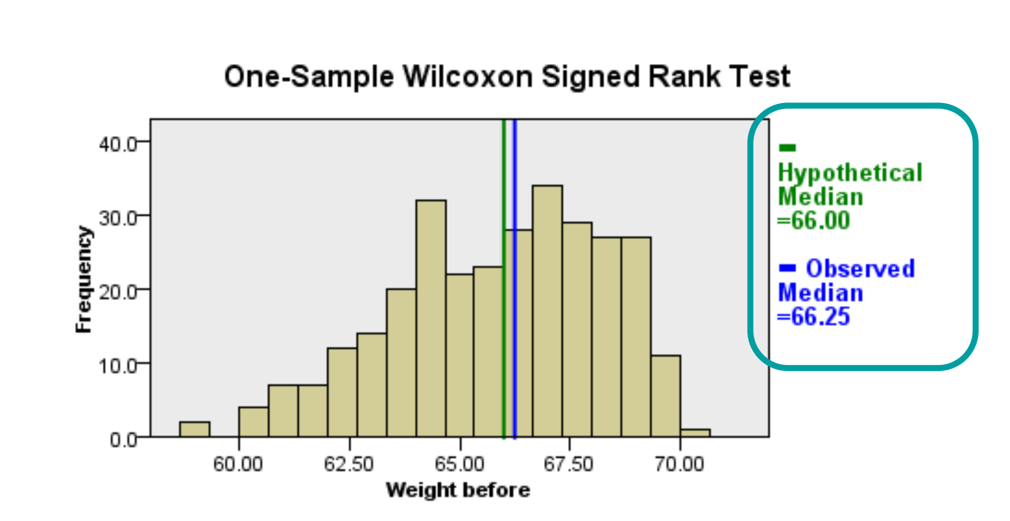
SPSS prints a table with all the information we need

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of Weight before equals 66.00.	One-Sample Wilcoxon Signed Rank Test	.874	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

But if you double click this table, you are able to see more useful details:



Total N	300
Test Statistic	22,187.000
Standard Error	1,496.243
Standardized Test Statistic	-.159
Asymptotic Sig. (2-sided test)	.874

The one-sample Wilcoxon signed-rank test indicated that the median was not significantly different than 66kg ($Z = -0.159$, $p = 0.874$).

Mann – Whitney U Test

The Mann-Whitney U test is the non-parametric analogue of the two independent samples t-test.

When to use

Skewed continuous data

Ordinal (interval) or discrete data

Null hypothesis

H_0: the two distributions are equal

H_a: the two distributions are not equal

Assumptions

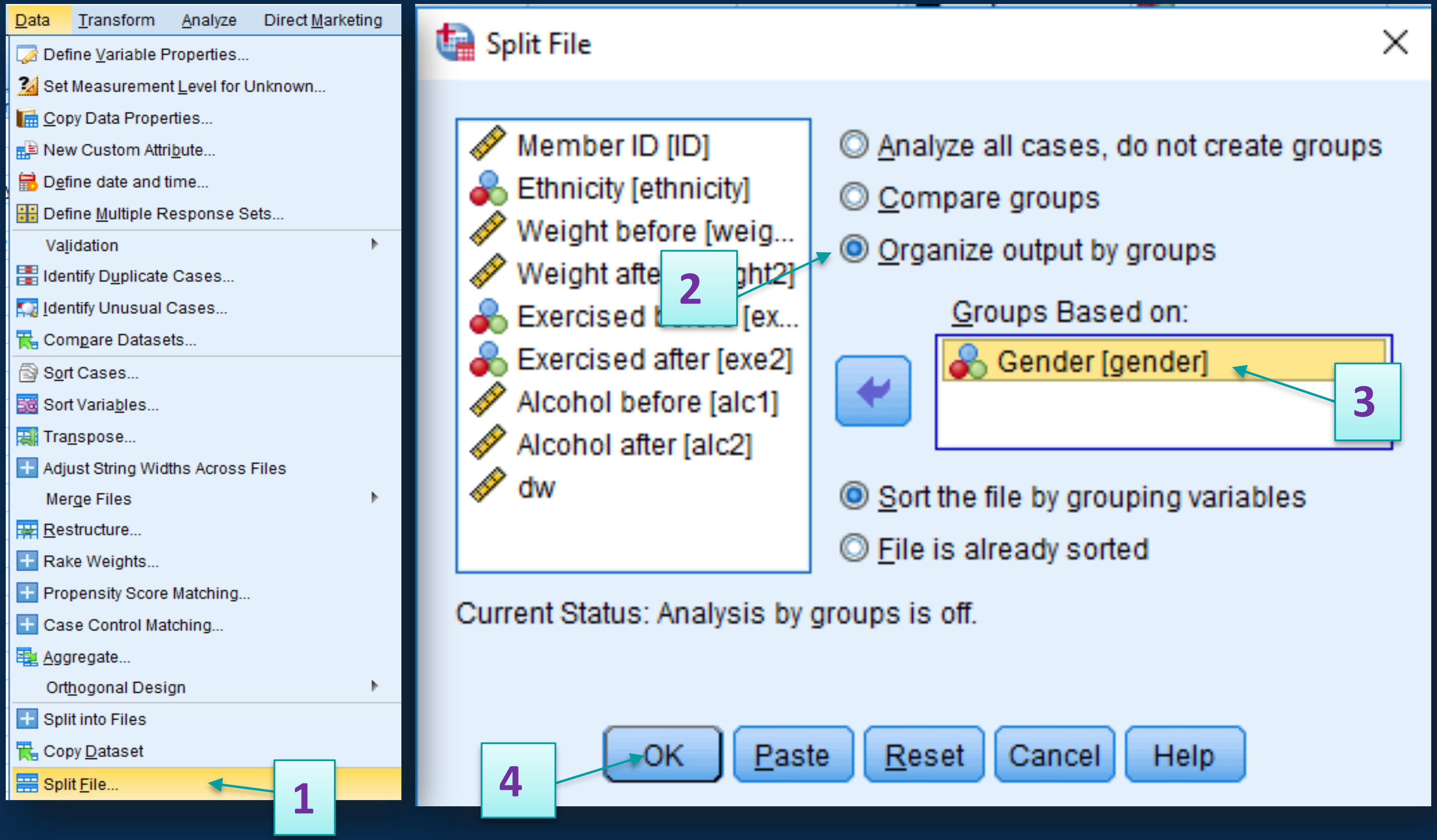
The observations are randomly and independently drawn

At least interval data

SPSS Slide: 'how to'

The next question is whether the 'weight before' was different across genders.

Step 1: Check the suitability of the data, here: what type of distribution has 'weight1', for each gender ?



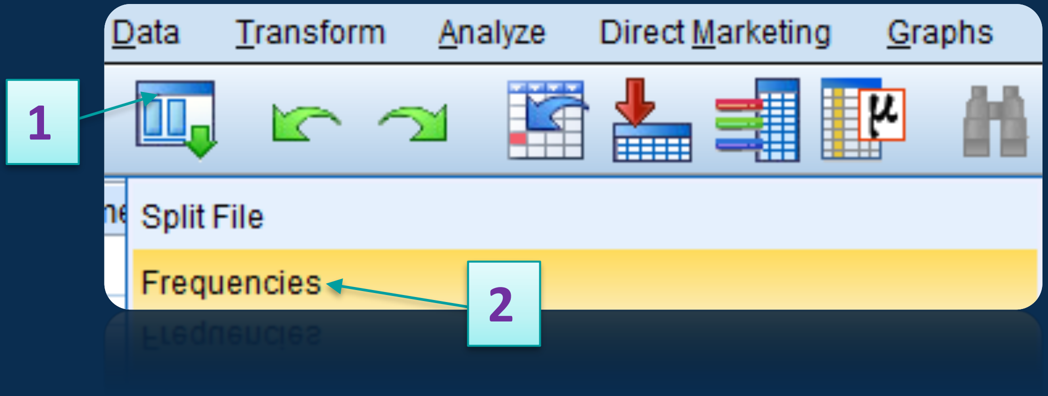
Go to 'Data' to use the 'Split File' function

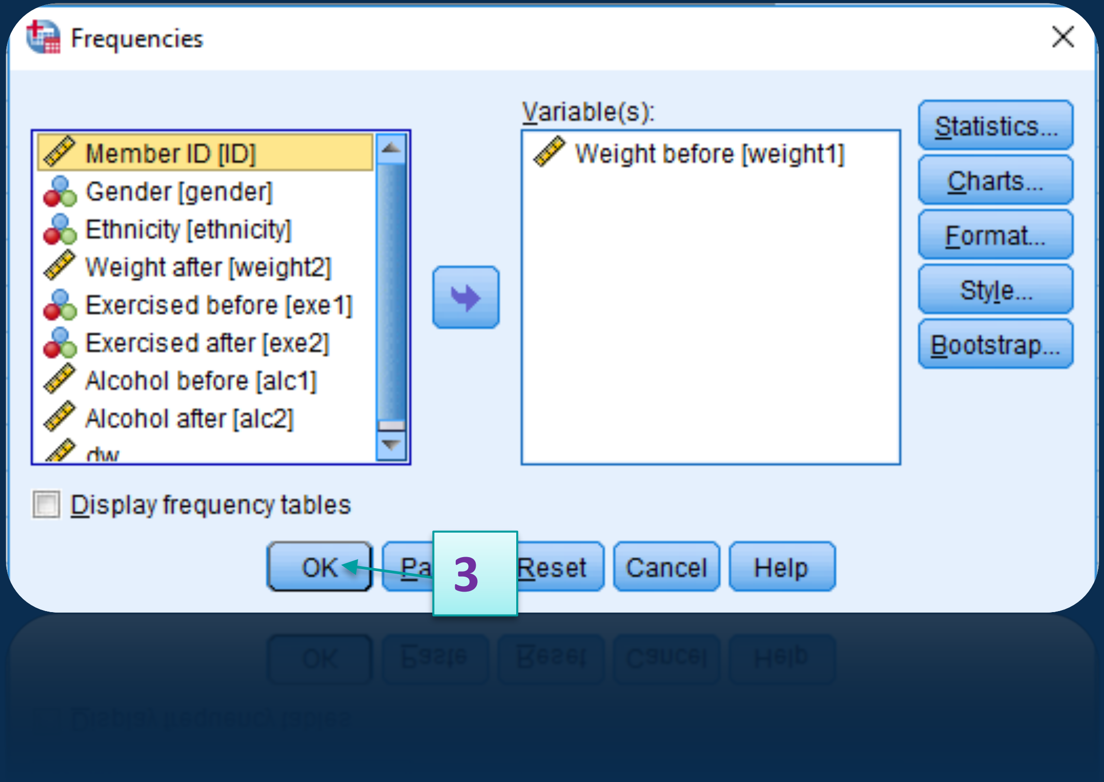
Split the file by groups (gender)

Click on 'OK'

SPSS Slide: ‘how to’

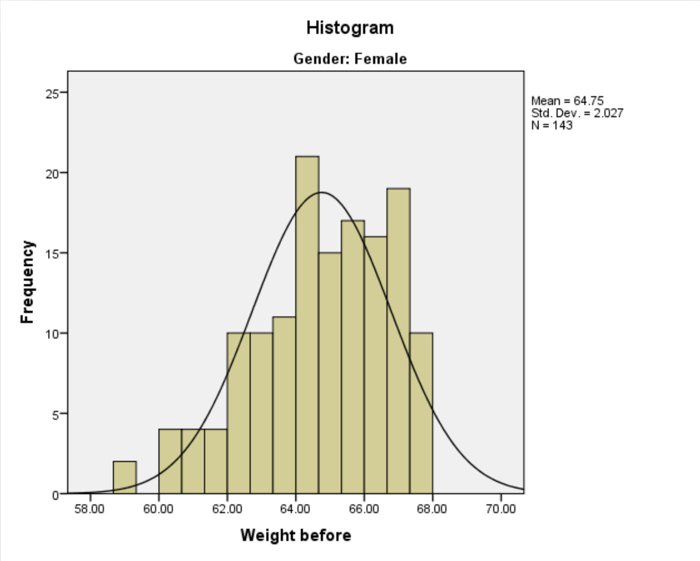
Step 1: Check the suitability of the data, here what type of distribution has ‘weight1’, for each gender ?
SPSS is now ready to show us the frequencies for each gender separately. Use the recall button.

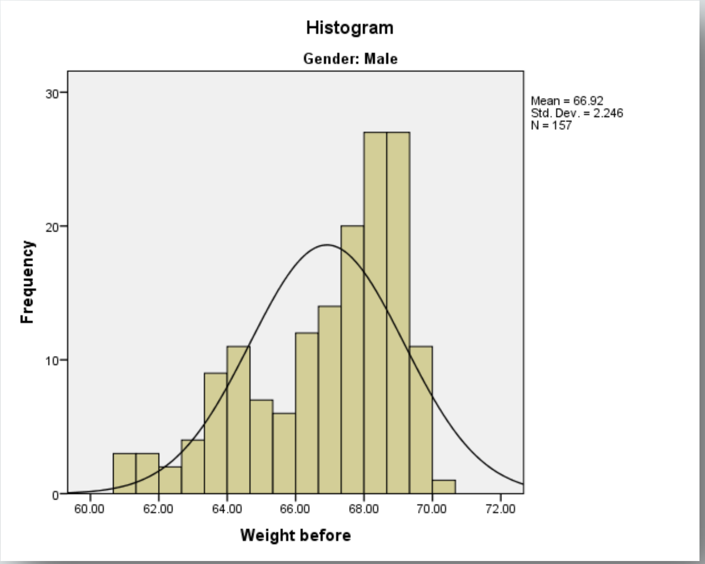




Or click on the ‘Analyse Tab’→ ‘Descriptive Statistics’ → ‘Frequencies’
Add the variable of interest (weight1) into the ‘Variable(s)’ box
In ‘Charts’ choose to display histograms
Click on ‘OK.

Output & Interpretation Slide





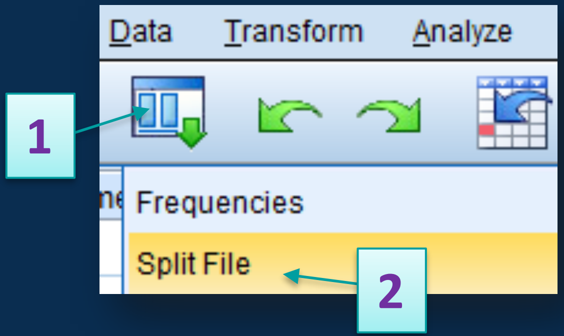
Both distributions are severely skewed (left, negative). Therefore we should use the 'Mann – Whitney test' for the hypotheses:

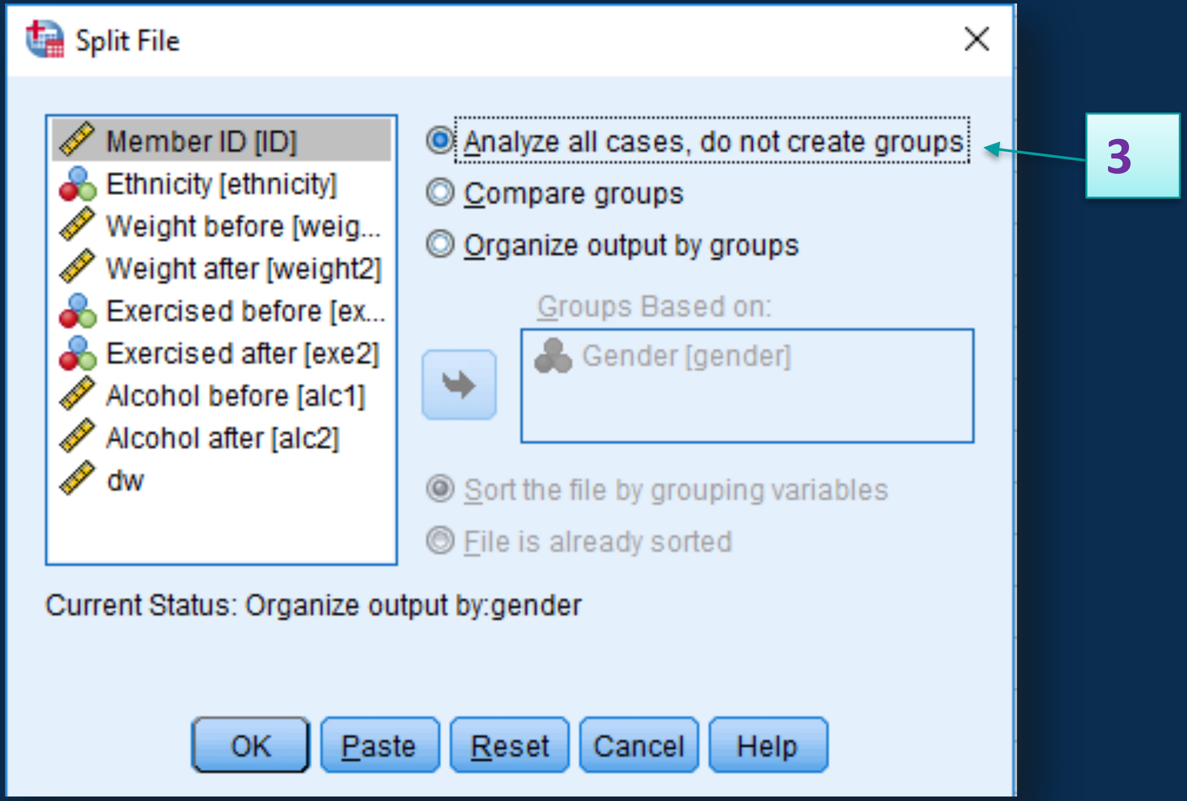
H_0: the two distributions are equal

H_a: the two distributions are different

SPSS Slide: ‘how to’

Before proceeding with the test, use the ‘recall button’ to go back to the ‘split file’ and re-unite the data.

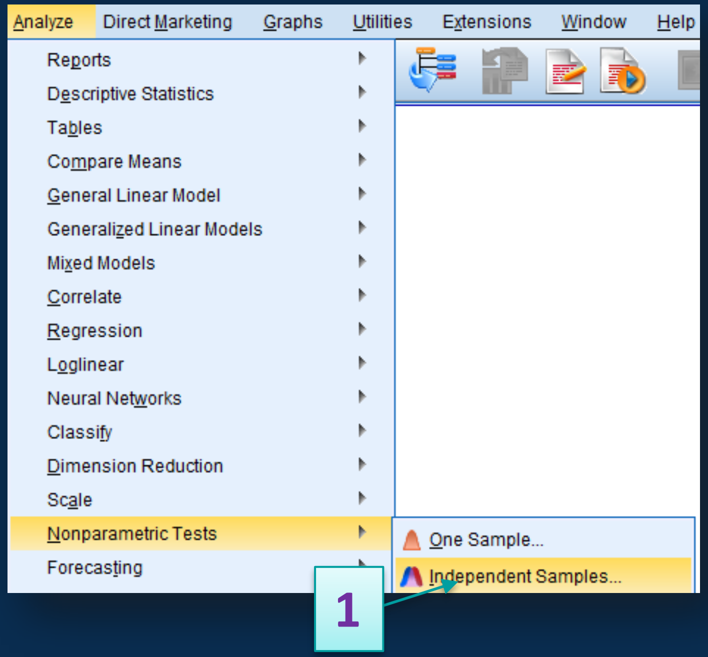


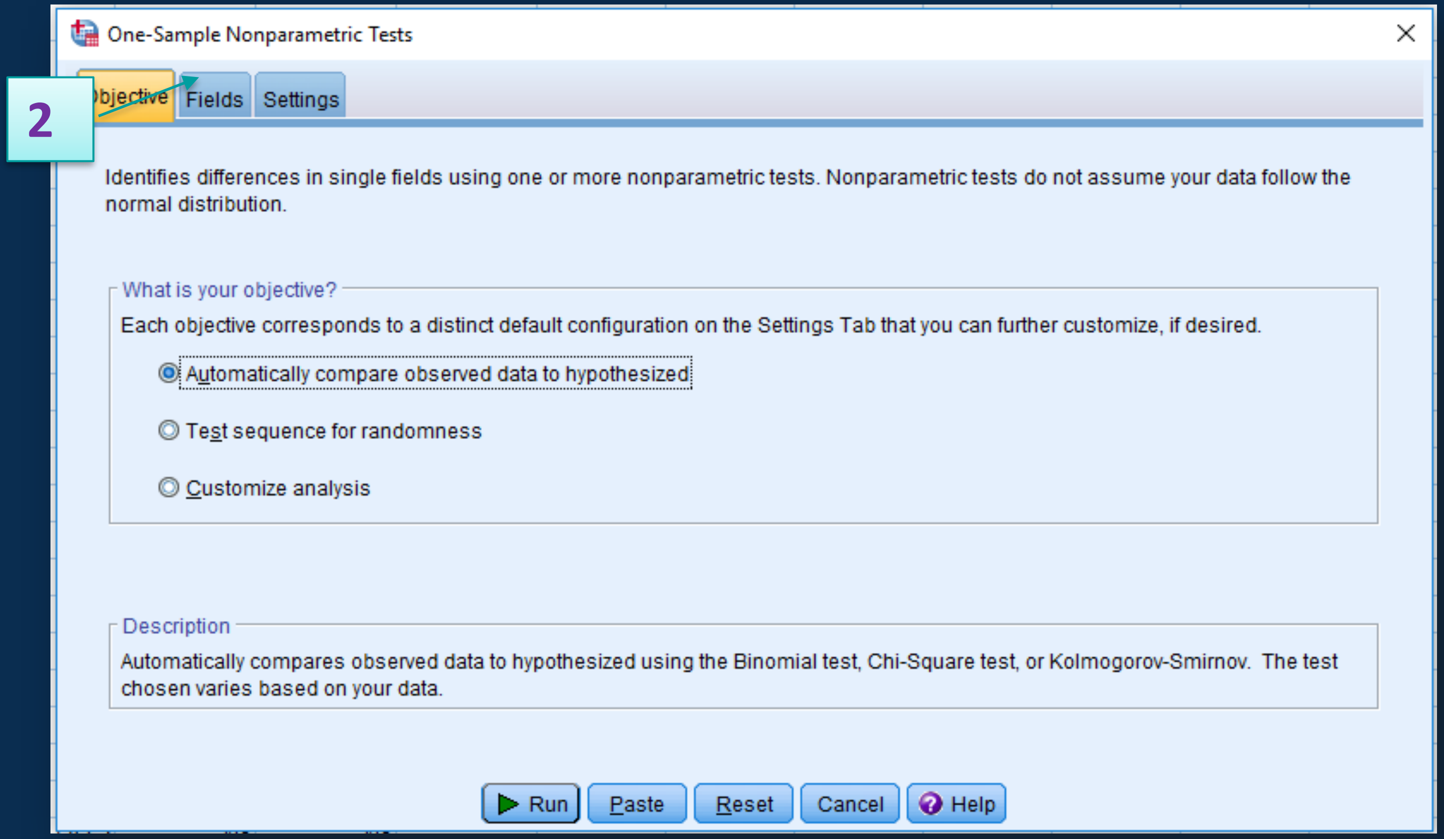


Go to ‘Data’ to use the ‘Split File’ function
‘Click on Analyse all cases’
Click on ‘OK’

SPSS Slide: ‘how to’

Step 2: Use the appropriate test, here: ‘Mann – Whitney test’.
Analyse -> Nonparametric tests-> ‘independent samples’

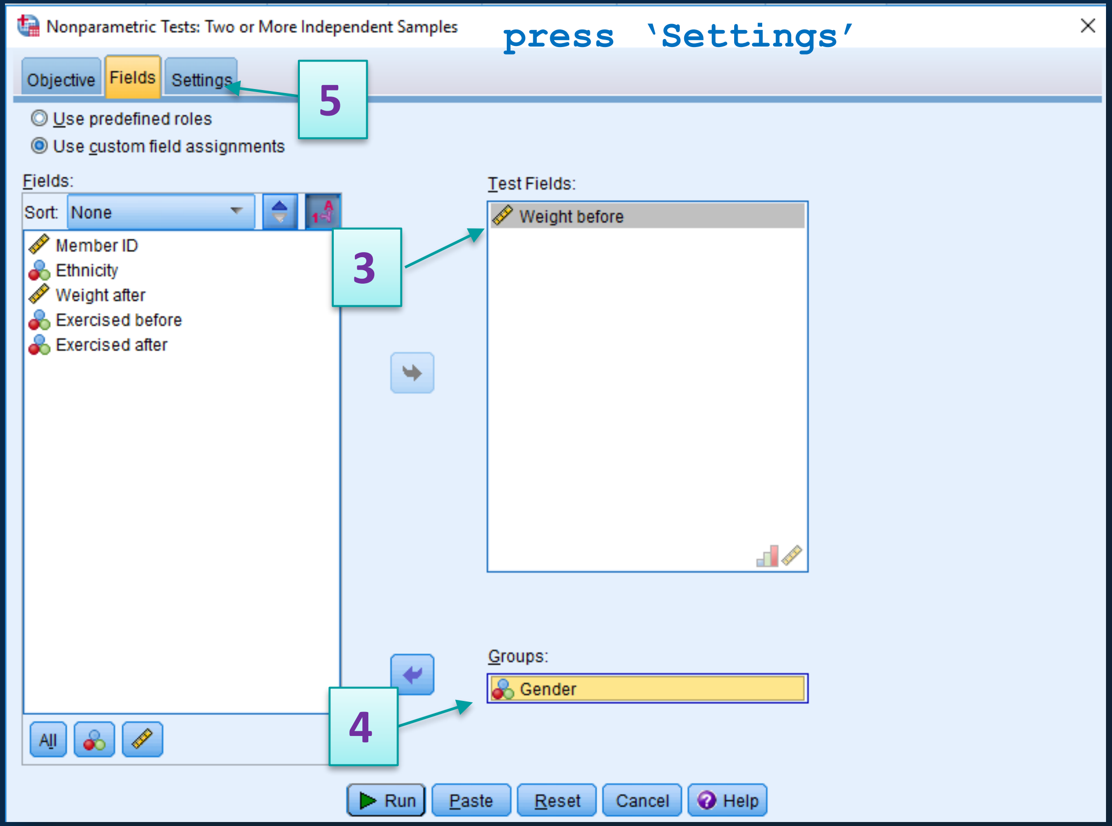




Click on the ‘Fields’ tab ’

SPSS Slide: ‘how to’

Analyse -> Nonparametric tests-> ‘independent samples’



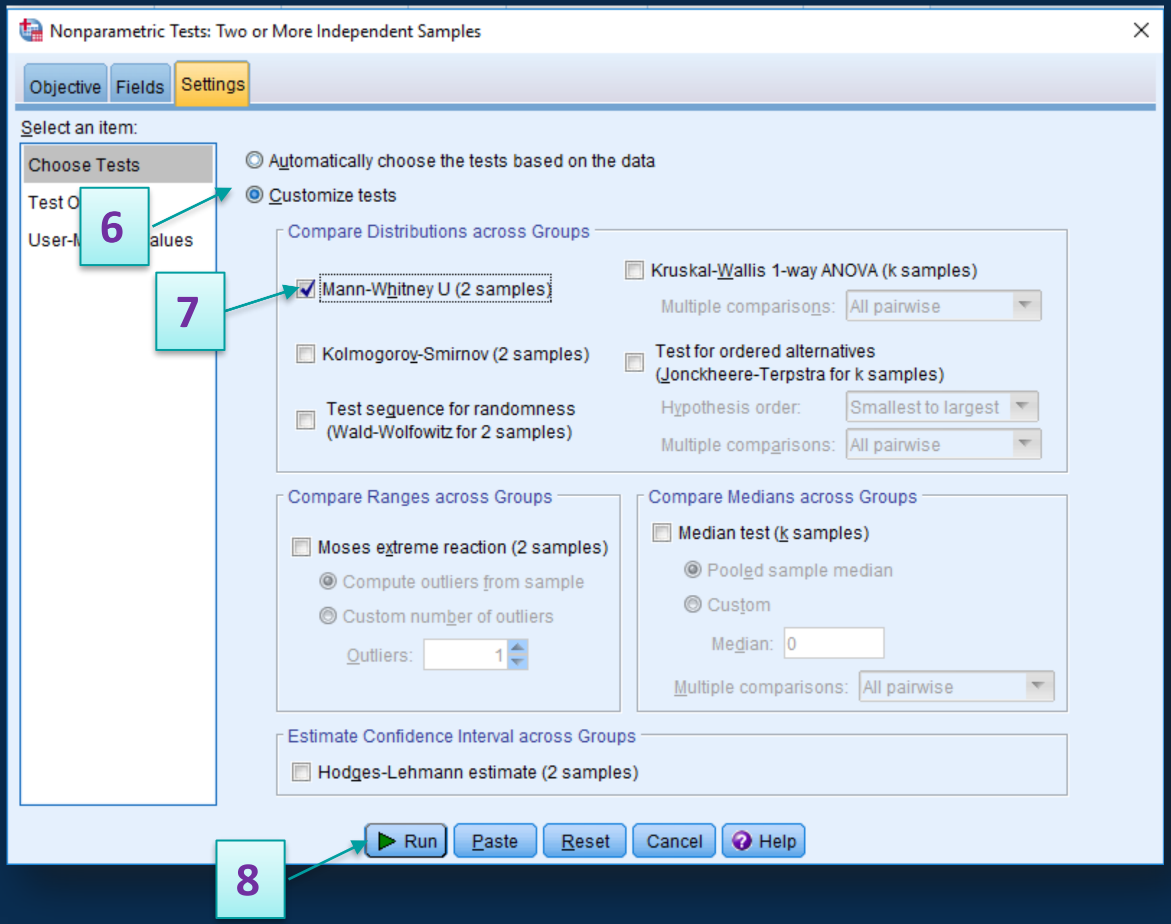
Add the variable of interest (weight before) in the ‘Test Fields’ box

Add the grouping variable (gender) in the ‘Groups’ box

Click on the ‘Settings’ tab

SPSS Slide: ‘how to’

Analyse -> Nonparametric tests-> ‘independent samples’



Choose the ‘Customised tests’ option
Choose the ‘Mann Whitney U’ test
Click on ‘Run’

Output & Interpretation Slide

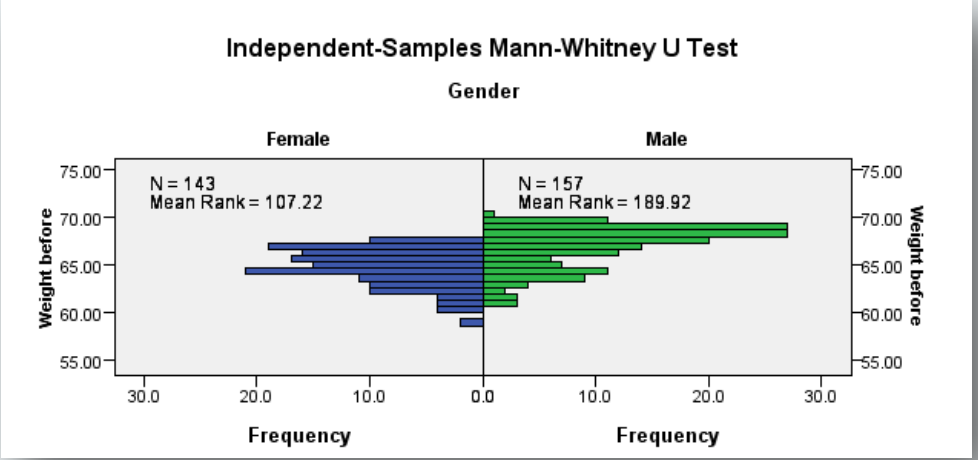
SPSS prints a table with all the information we need

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Weight before is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.

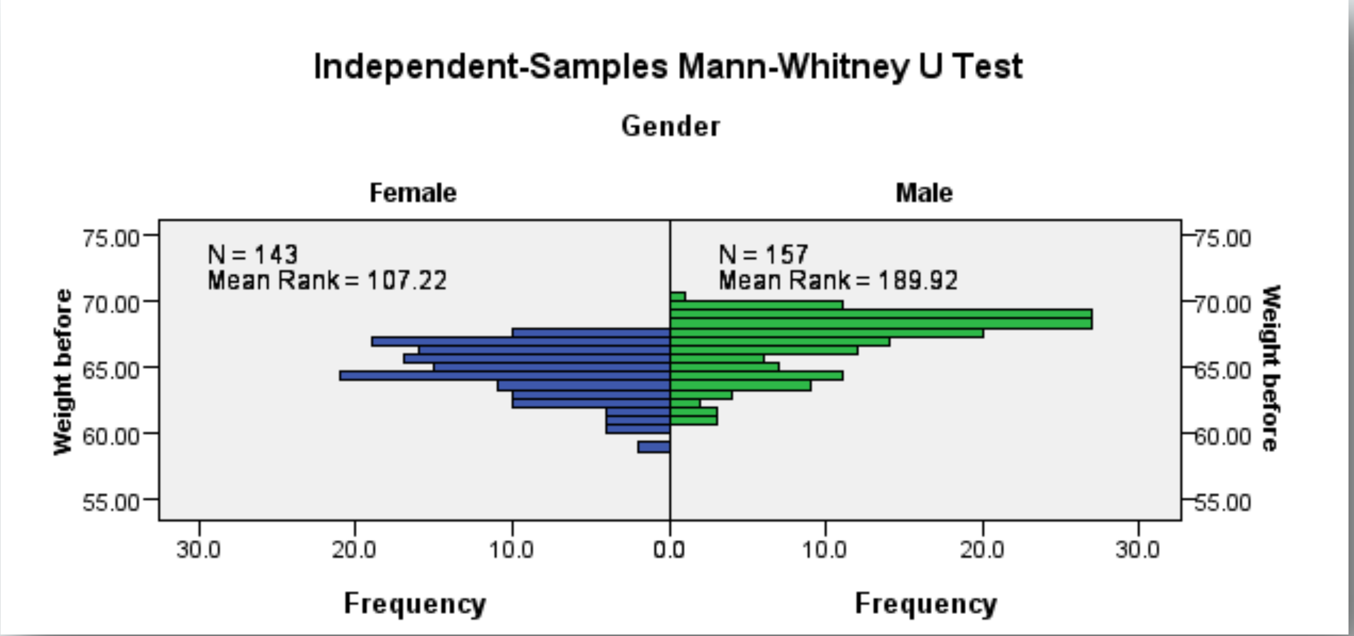
Asymptotic significances are displayed. The significance level is .05.

But if you double click this table, you are able to see more useful details:



Total N	300
Mann-Whitney U	17,414.500
Wilcoxon W	29,817.500
Test Statistic	17,414.500
Standard Error	750.425
Standardized Test Statistic	8.247
Asymptotic Sig. (2-sided test)	.000

Output & Interpretation Slide



Total N	300
Mann-Whitney U	17,414.500
Wilcoxon W	29,817.500
Test Statistic	17,414.500
Standard Error	750.425
Standardized Test Statistic	8.247
Asymptotic Sig. (2-sided test)	.000

The distribution of 'weight before' was statistically different across genders (Mann-Whitney U= 17,414.5, $p < 0.001$), with men's weight tending to be higher than women's, before the program.

To see which gender had higher values, the best strategy is to check the descriptive indices (slide 16).

Wilcoxon Matched-Pair Signed Rank Test

The Wilcoxon Matched – Pair Signed Rank test is the non-parametric analogue of the paired sample t-test

When to use

Skewed continuous data

Ordinal (interval) or discrete data

Null hypothesis

H_0: Median of the paired differences equals zero

H_a: Median of the paired differences is different than zero

Assumptions

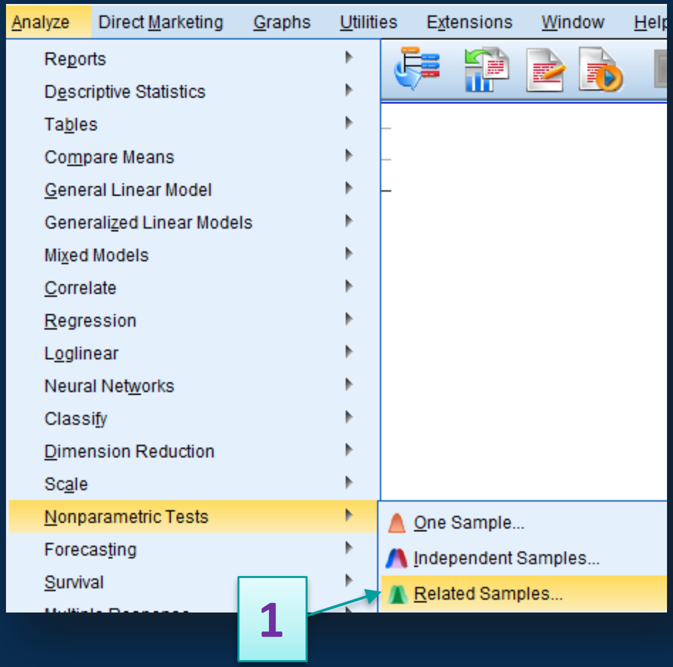
The pairs of observations are randomly and independently drawn

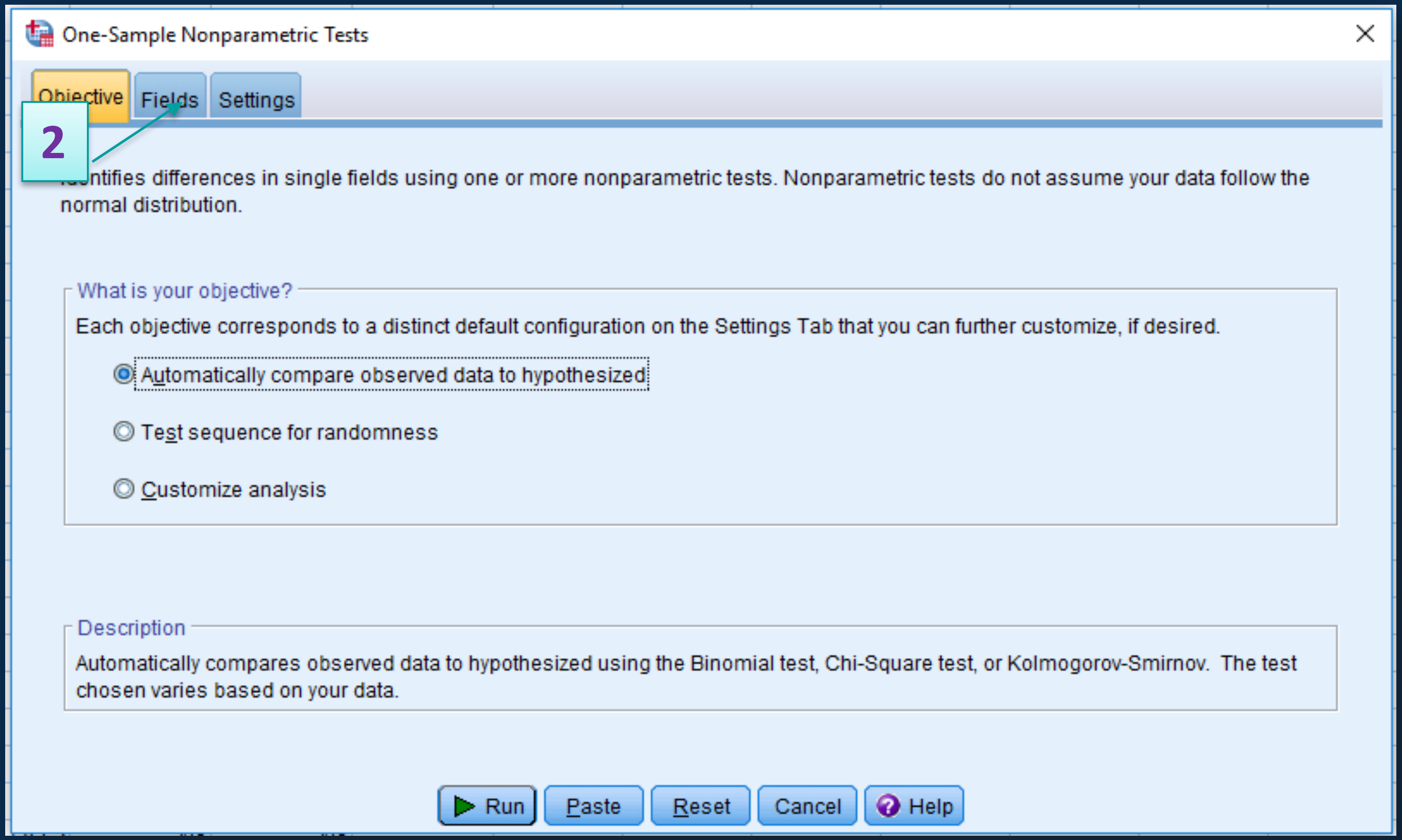
At least interval data

The two samples need to be dependent observations of the cases, i.e. they are paired or matched

SPSS Slide: ‘how to’

Step 1: Use the appropriate test, here ‘related samples Wilcoxon signed rank test’.
Analyse -> nonparametric tests -> ‘related samples’

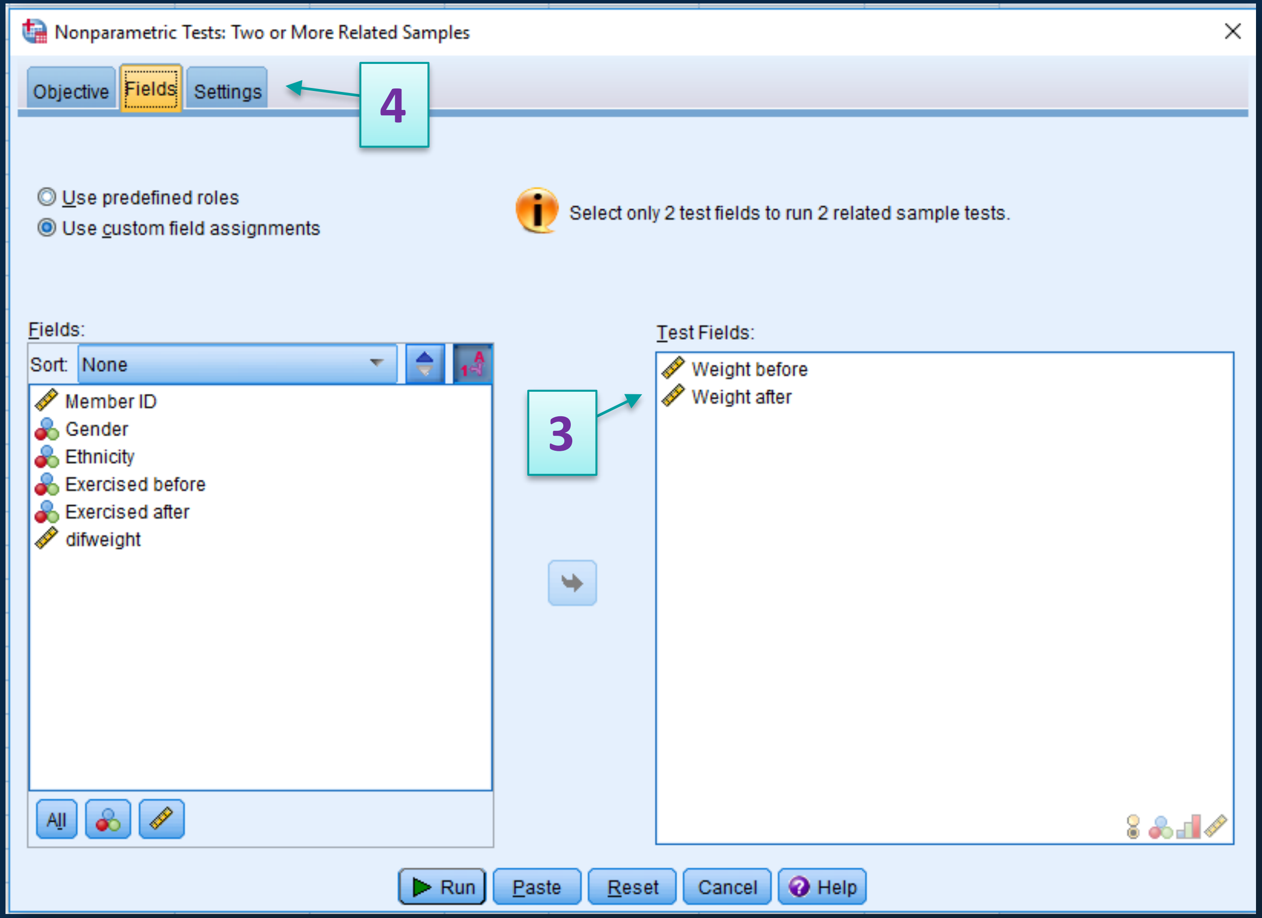




Click on the ‘Fields’ tab ’

SPSS Slide: ‘how to’

Analyse -> nonparametric tests -> ‘related samples’

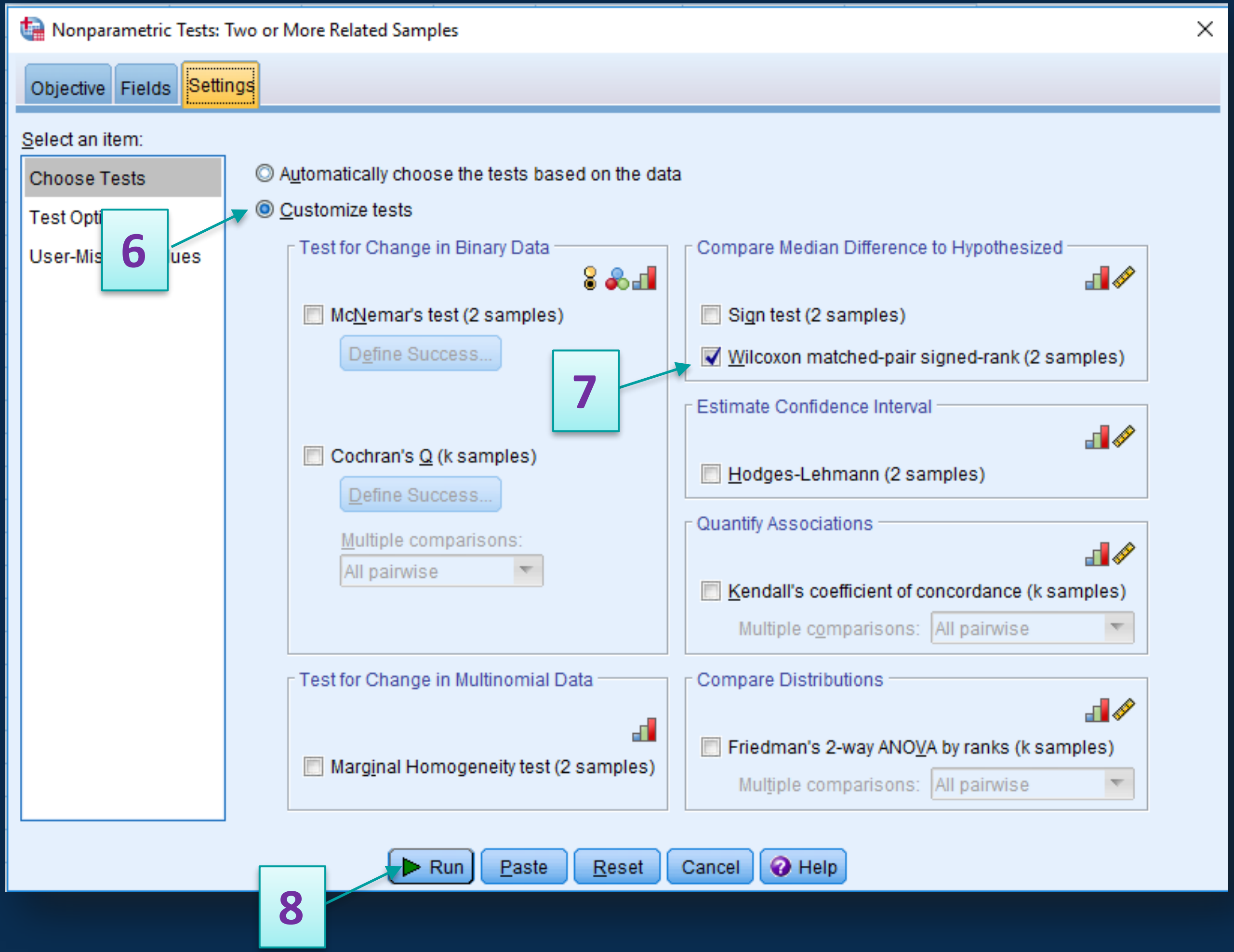


Add the variables of interest (weight before and weight after) into the ‘Test fields’ box

Click on ‘Settings’

SPSS Slide: ‘how to’

Analyse -> nonparametric tests -> ‘related samples’



Click on ‘Customise tests’
Choose ‘Wilcoxon matched-pair signed-rank’
Click on ‘Run’

Output & Interpretation Slide

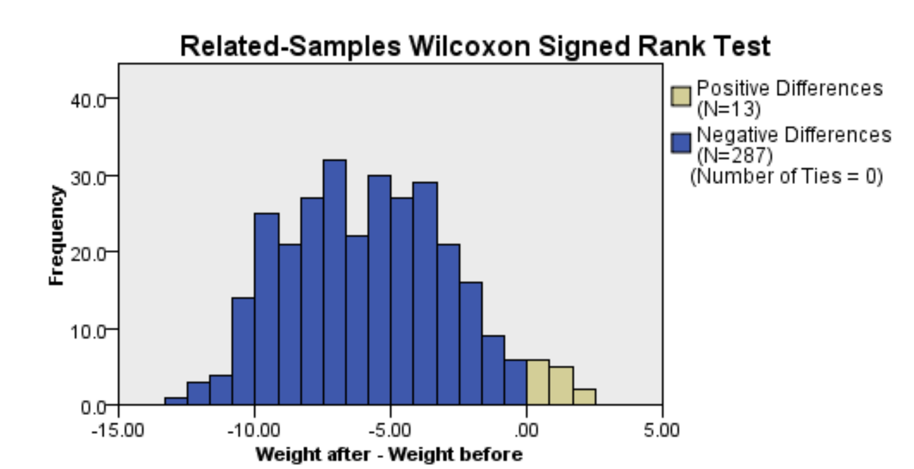
SPSS prints a table with all the information we need

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Weight before and Weight after equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.

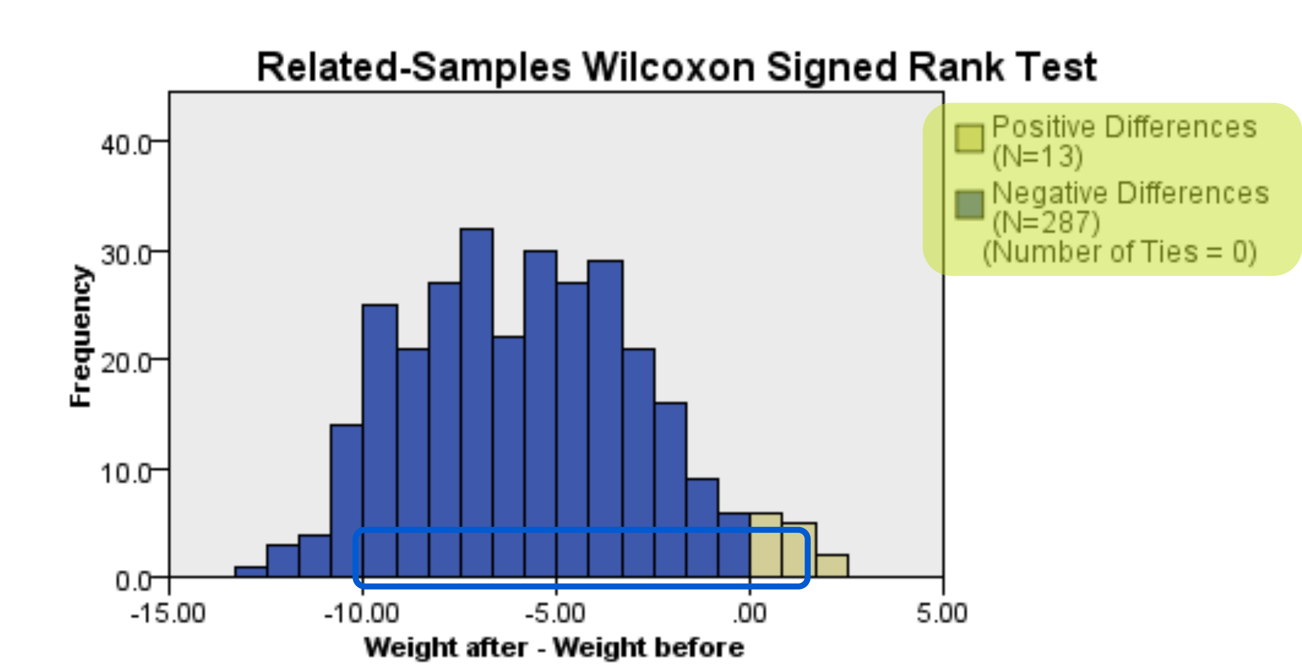
Asymptotic significances are displayed. The significance level is .05.

But if you double click this table, you are able to see more useful details:



Total N	300
Test Statistic	197.000
Standard Error	1,503.748
Standardized Test Statistic	-14.881
Asymptotic Sig. (2-sided test)	.000

Output & Interpretation Slide



Total N	300
Test Statistic	197.000
Standard Error	1,503.748
Standardized Test Statistic	-14.881
Asymptotic Sig. (2-sided test)	.000

The median difference between the ‘weight after’ and the ‘weight before’ was significantly different than zero (Wilcoxon rank sum Z = -14.88, p <0.001). The weight decreases significantly after the programme.

Parametric and Non-Parametric Tests

Numerical data	Normality assumed	Normality not assumed
Hypotheses testing	Means	Medians
one group *versus* a pre-defined value	one sample t-test	Wilcoxon signed rank
one group *versus* another group	two independent samples t-test	Mann-Whitney (Wilcoxon sum rank)
one group (twice or) *versus* another matched group	two paired samples t-test	Wilcoxon signed rank for paired samples

Parametric and Non-Parametric Tests for Numerical Data: Comparison

Parametric tests

They assume approximately normally distributed data

Not suitable for small sample sizes (less than N=30)

Powerful

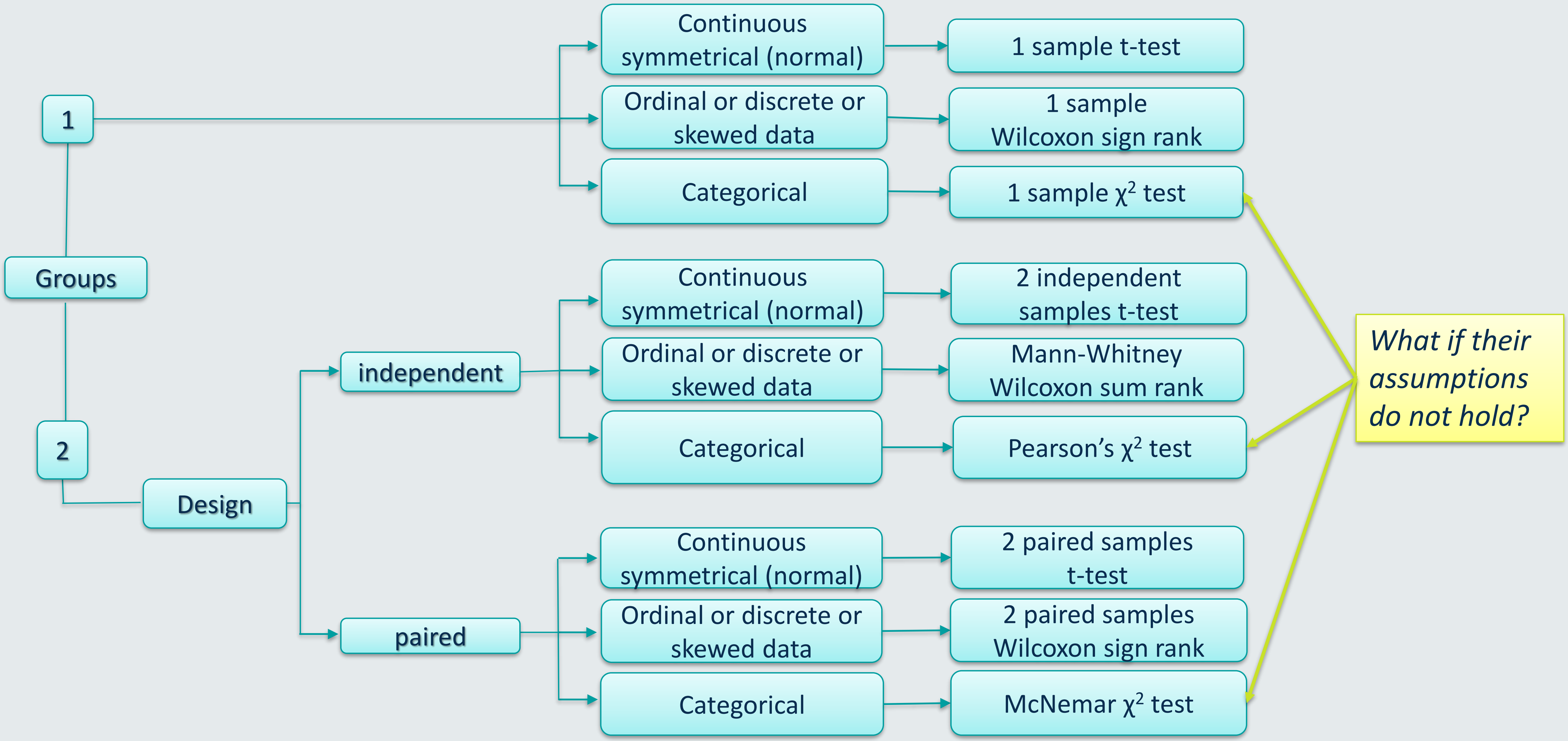
Non-parametric tests

They do not assume approximately normally distributed data

Suitable for small sample sizes (less than N=30)

Less powerful

The Tests in a Flow Chart…



Knowledge Test

Match the scenario with the correct test.

Tom wants to test if mothers' reported ADHD scores for children are higher than those reported by fathers.

Tom wants to test if boys' ADHD scores are higher than those of girls.

Tom wants to test if children's ADHD scores are higher than 30.

1 sample
Wilcoxon signed rank

Mann-Whitney
(Wilcoxon sum rank)

2 paired samples
Wilcoxon signed rank

Reflection

Reflecting on your field of study

Write down three examples from your research that would require the use of each of the three non-parametric tests.

Reference List

- **Agresti and Finlay (2009) Statistical Methods for the Social Sciences, 4th Edn, Pearson Hall, Upper Saddle River, NJ.**
  - Comparison of Two Groups, Ch 7, pages 183-209
  - Analyzing Association between Categorical Variables, Ch 8, pages 221-239

- **Field (2005) Discovering Statistics using SPSS, 2nd Edn, Sage, London.**
  - Comparing Two Means, Ch 7
  - Categorical Data, Ch 16

Thank you

Please contact <u>your module leader</u> or <u>the course lecturer of your programme</u>, or visit the module's <u>forum</u> for any questions you may have.

If you have comments on the materials (spotted typos or missing points) please contact Dr Vitoratou:

Silia Vitoratou, PhD
Psychometrics & Measurement Lab,
Department of Biostatistics and Health Informatics
IoPPN, King's College London, SE5 8AF, London, UK
silia.vitoratou@kcl.ac.uk

For any other comments or remarks on the module structure, please contact one of the three module leaders of the Biostatistics and Health Informatics department:

Zahra Abdula: **zahra.abdulla@kcl.ac.uk**

Raquel Iniesta: **raquel.iniesta@kcl.ac.uk**

Silia Vitoratou: **silia.vitoratou@kcl.ac.uk**

Institute of Psychiatry, Psychology and Neuroscience

Dr Silia Vitoratou

Department: Biostatistics and Health Informatics

Topic materials:
Silia Vitoratou

Contributions:
Zahra Abdulla

Improvements:
Nick Beckley-Hoelscher
Kim Goldsmith
Sabine Landau

Module Title: Introduction to Statistics

Session Title: Equality of proportions (exact tests)

Topic title: Comparing groups II (non-parametric methods)

Learning Outcomes

- Learn when and how to use the **exact** tests for equality of proportions.
- Understand when assumptions have been violated to warrant the use of exact tests.
- Be able to conduct these tests in a statistical software.

Parametric and Non-Parametric Tests

Numerical data | **Normality assumed** | **Normality not assumed**

Hypotheses testing	parametric	Non-parametric
one group *versus* a pre-defined value	one sample t-test	Wilcoxon sign rank
one group *versus* another group	two independent samples t-test	Mann-Whitney (Wilcoxon sum rank)
one group (twice or) *versus* another matched group	two paired samples t-test	Wilcoxon sign rank

Previously on ‘Introduction to Statistics’.....

Categorical data	Assumptions hold	Assumptions do not hold
Hypotheses testing	Chi-squared tests	Exact tests
one group *versus* a pre-defined value	one sample χ^2-test	Binomial exact test
one group *versus* another group	Pearson’s χ^2-test	Fisher’s Exact test
one group (twice or) *versus* another matched group	McNemar test	Binomial exact test

Equality of Proportions: Exact Tests

For associations between categorical data, so far we have been considering chi-square (χ^2) tests. However, we can trust the results of the test only if the assumptions hold

Up to 20% cells can have expected count less than 5.
The minimum expected count is larger than 1.

If these assumptions are not satisfied, then test statistic of the chi-square is not reliable

$$\sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

and we use instead the '**exact**' tests.

One Sample Chi-Square Test - Reminder

When to use

To test if according to the current data, the proportion in the population equals a certain, pre-specified, value.

Hypotheses:

H_0: the proportion in the population equals a certain pre-specified value
H_a: the proportion in the population is different than a certain pre-specified value

Assumptions:

- The observations are randomly and independently drawn
- The number of cells with expected frequencies less than 5, are less than 20%
- The minimum expected frequency is at the very least 1.

SPSS Slide: ‘how to’

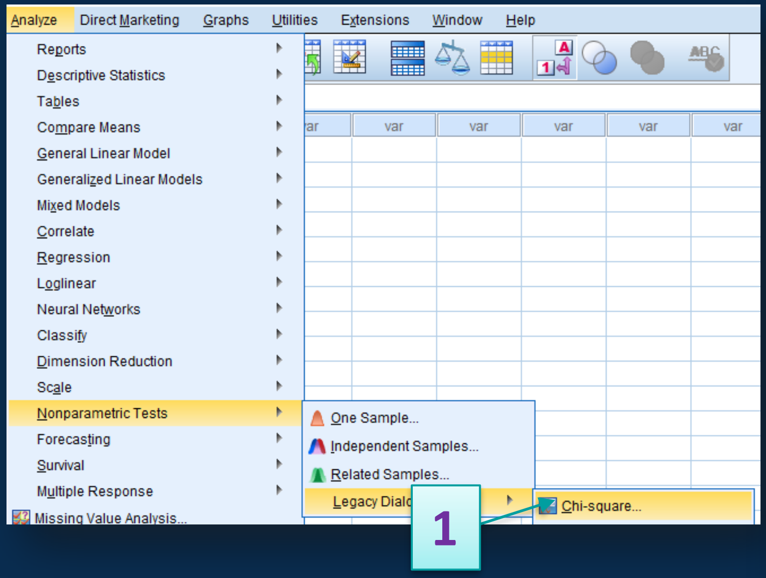
The programme developers had hoped that no more than 1% of the population has used drugs. Is that what our data tell us?

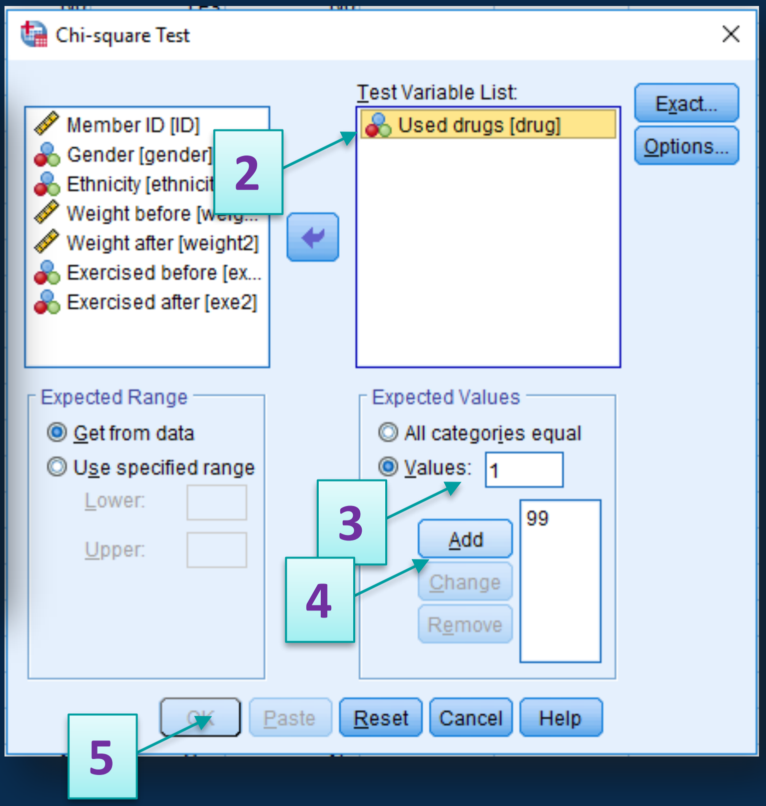
H_0: π=1%

H_a: π≠1%

<u>Step 1</u>: Use the appropriate test, here ‘one-sample chi-square t-test’.

Analyse -> nonparametric tests -> legacy dialogs -> ‘Chi-square’





Add the variable of interest (used drugs) in to the ‘Variables box’. Choose to Test ‘values’

**Check the coding in the dataset as ‘0’ coding is ‘not used drugs’ ‘add’ in 99 (100%- 1%) first.
Then as ‘1’ coding is ‘used drugs’ ‘add’ in 1 (to represent the 1% in the question).
Click ‘OK’**

Output and Interpretation Slide

Step 2: Check the suitability of the data, do the assumptions of the chi-square test hold?

Used drugs

	Observed N	Expected N	Residual
No	290	297.0	-7.0
Yes	10	3.0	7.0
Total	300		

Test Statistics

	Used drugs
Chi-Square	16.498[a]
df	1
Asymp. Sig.	.000

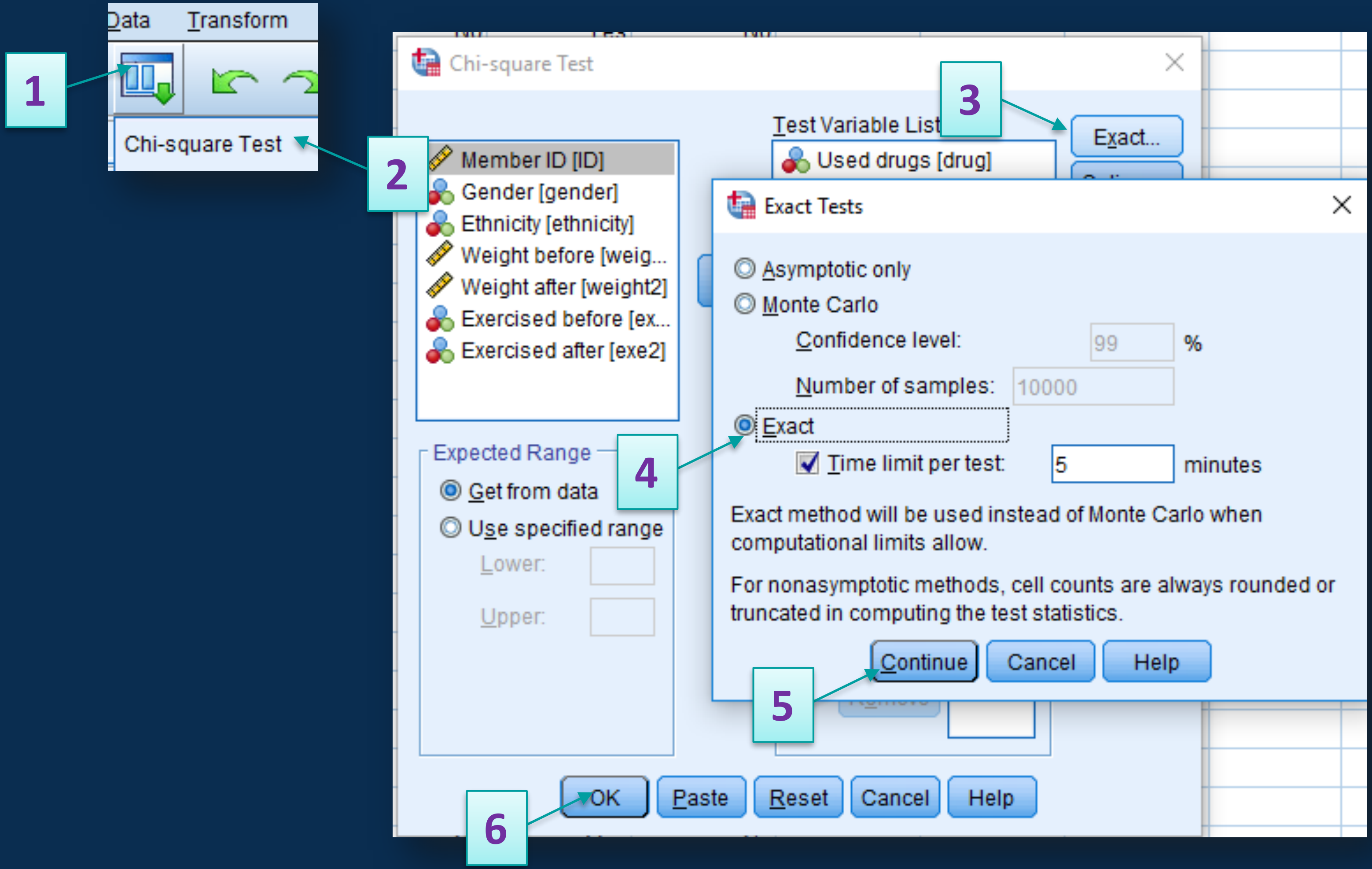
a. 1 cells (50.0%) have expected frequencies less than 5. The minimum expected cell frequency is 3.0.

Only up to 20% of the cells are allowed to have expected frequencies less than 5.

Thus, one assumption is violated!

SPSS Slide: 'how to'

Use the recall button to go back to the chi-square, and this time click the '**exact**' tab.



Select the 'Exact' option
Click 'Continue'
Click 'OK'

Output & Interpretation Slide

Used drugs

	Observed N	Expected N	Residual
No	290	297.0	-7.0
Yes	10	3.0	7.0
Total	300		

Test Statistics

	Used drugs
Chi-Square	16.498[a]
df	1
Asymp. Sig.	.000
Exact Sig.	.001
Point Probability	.001

a. 1 cells (50.0%) have expected frequencies less than 5. The minimum expected cell frequency is 3.0.

Based on our sample, the expected proportion of people who have used drugs to lose weight is different (larger) than 1% (exact p-value=0.001).

Pearson's Chi-Square Test

When to use

To test if, according to the current data, the proportions in the population of one variable change based on another variable.

Hypotheses:

H_0: there is no association between the two variables
H_a: there is an association between the two variables

Assumptions:

- The observations are randomly and independently drawn
- The number of cells with expected frequencies less than 5, are less than 20%
- The minimum expected frequency is at the very least 1.
- The observations are not paired

SPSS Slide: 'how to'

The next question is: do men exercise more than women prior to entering the programme?

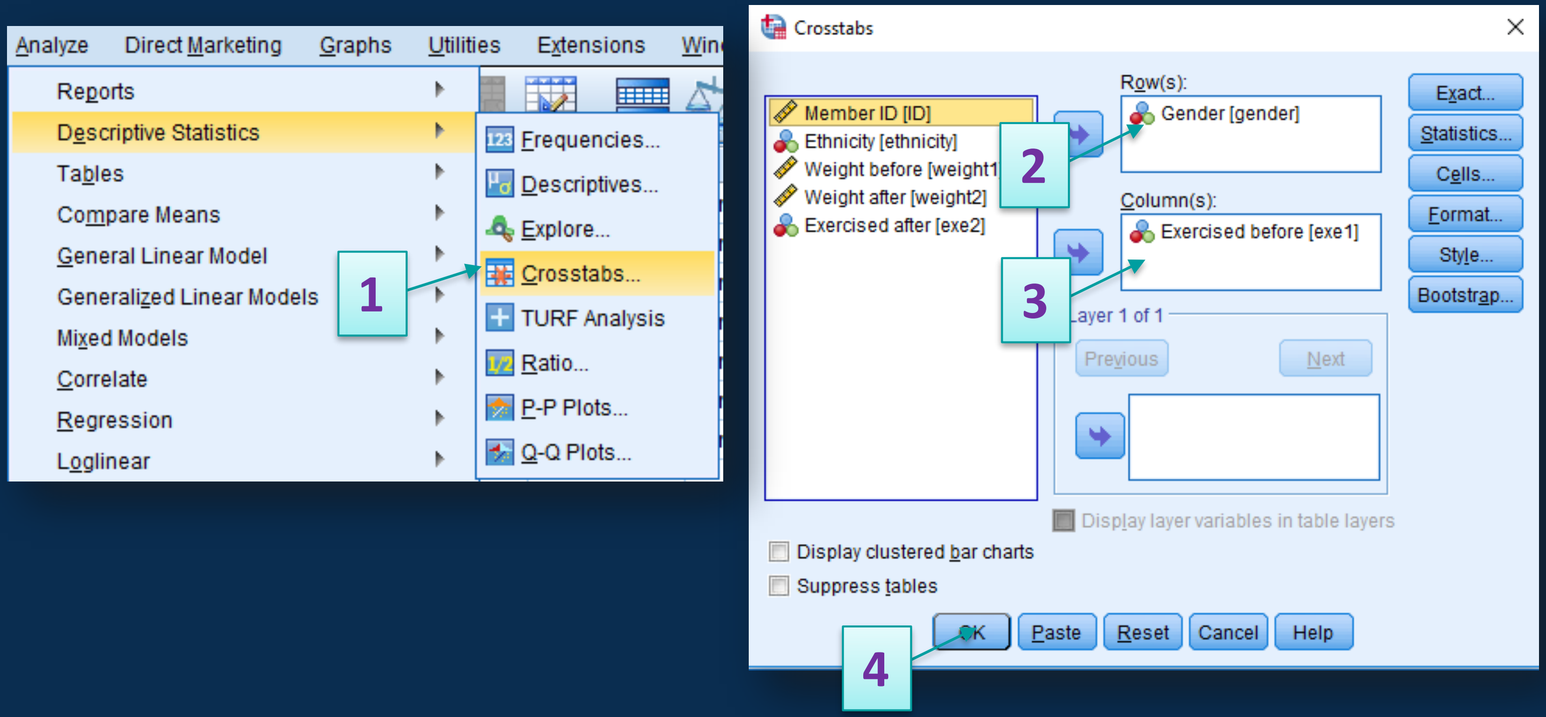
Are the proportions of those exercised before the programme, different for men and women?

<u>Step 1</u>: Use the appropriate test, here: 'Pearson's chi-square test'.

Analyse -> Descriptive Statistics-> Crosstabs

H_0: $\pi_m = \pi_f$

H_a: $\pi_m \neq \pi_f$

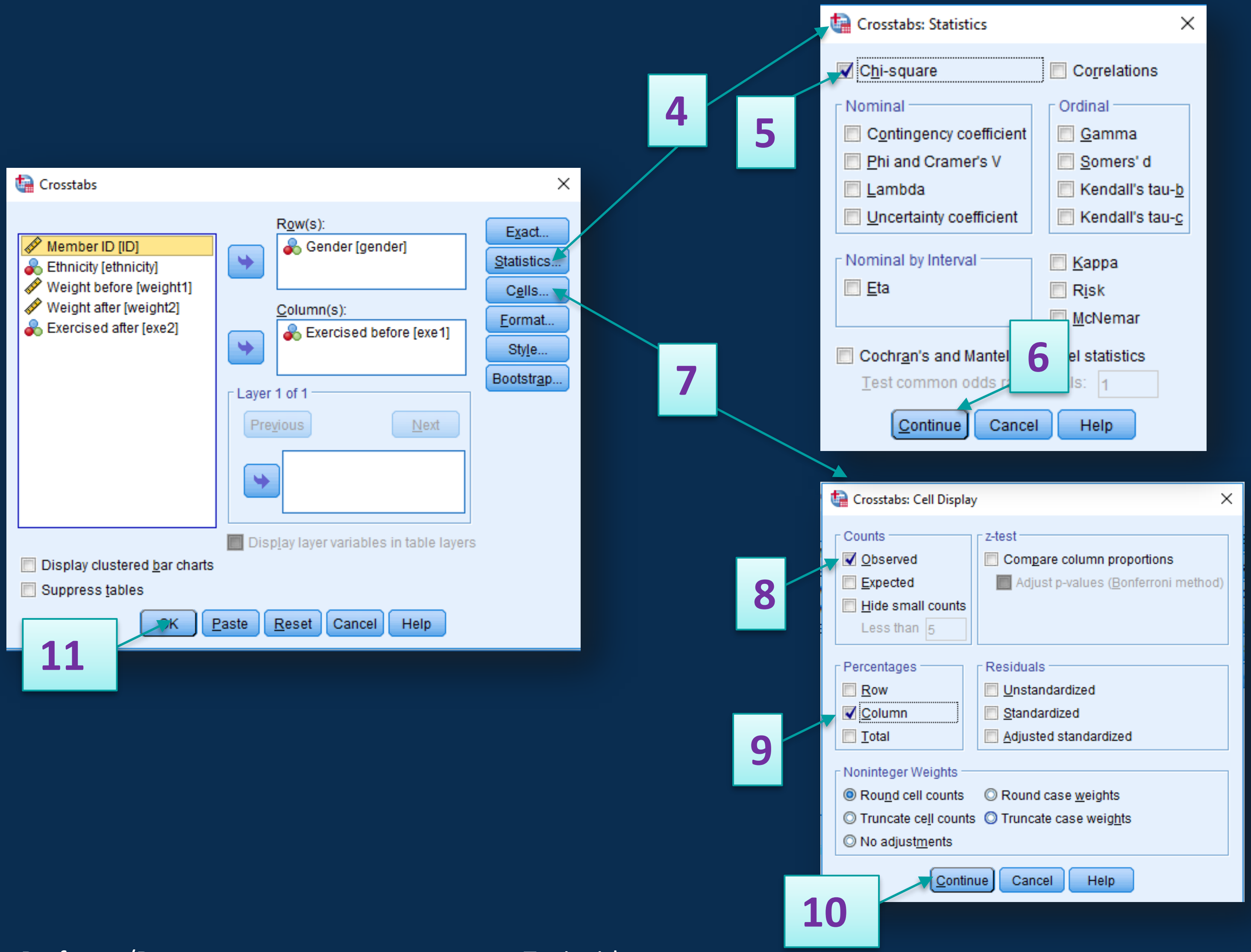


Add the variable of interest (gender) in to the 'Rows box'.

Add the second variable interest (Exe before) in the 'columns box'

SPSS Slide: ‘how to’

Step 1: Use the appropriate test, here: ‘Pearson’s chi-square test’.



Output and Interpretation Slide

Unfortunately, due to a technical problem, most of the data we had in 'exercise before' were accidentally deleted. We only have info for 29 people.

Gender * Exercised before Crosstabulation

			Exercised before		
			No	Yes	Total
Gender	Female	Count	17	3	20
		% within Exercised before	100.0%	25.0%	69.0%
	Male	Count	0	9	9
		% within Exercised before	0.0%	75.0%	31.0%
Total		Count	17	12	29
		% within Exercised before	100.0%	100.0%	100.0%

<u>Step 2</u>: Check the suitability of the data, do the assumptions of the chi-square test hold?

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	18.488[a]	1	.000		
Continuity Correction[b]	15.149	1	.000		
Likelihood Ratio	22.428	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	17.850	1	.000		
N of Valid Cases	29				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 3.72.

b. Computed only for a 2x2 table

- Only up to 20% of the cells are allowed to have expected frequencies less than 5.
- **Thus, one assumption is violated!**

Output and Interpretation Slide

Gender * Exercised before Crosstabulation

			Exercised before		
			No	Yes	Total
Gender	Female	Count	17	3	20
		% within Exercised before	100.0%	25.0%	69.0%
	Male	Count	0	9	9
		% within Exercised before	0.0%	75.0%	31.0%
Total		Count	17	12	29
		% within Exercised before	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	18.488[a]	1	.000		
Continuity Correction[b]	15.149	1	.000		
Likelihood Ratio	22.428	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	17.850	1	.000		
N of Valid Cases	29				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 3.72.

b. Computed only for a 2x2 table

Among females, the proportion of those who exercised before the programme was lower than those who did not exercise before the programme (25% versus 100%, respectively). This difference was statistically significant according to Fisher's exact test (exact p<0.001).

Therefore, we conclude that men tend to exercise (before the programme) more often than women, in the population. The variables 'gender' and 'exe1' are related.

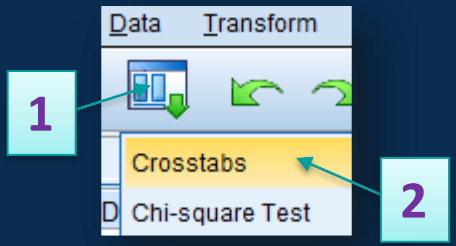
SPSS Slide: ‘how to’

Neither Pearson’s chi-square, nor Fisher’s exact test are restricted to 2x2 tables. We may have 2 groups but more than two categories in the categorical variable.

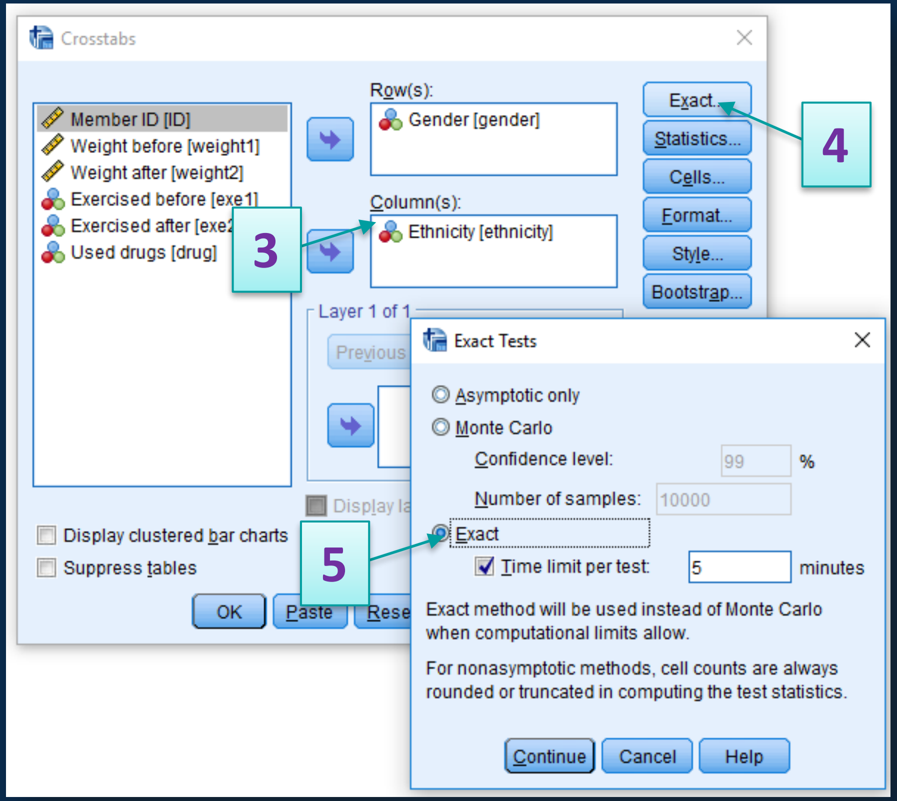
The next question is: is ethnicity (more than 2 categories) associated with gender?

<u>Step 1</u>: Use the appropriate test, here: ‘Pearson’s chi-square test’.

Use the recall button and simply replace ‘exe1’ with ‘ethnicity’



While Fisher’s exact test appears automatically in 2x2 tables, but that is NOT the case in 2xP tables. We need to go the ‘**<u>exact</u>**’ tab (as we did for the one sample previously).



Interpretation Slide

Step 2: Check the suitability of the data, do the assumptions of the chi-square test hold?

Gender * Ethnicity Crosstabulation

			Ethnicity				
			White	Black	Asian	Other	Total
Gender	Female	Count	53	44	46	0	143
		% within Ethnicity	42.7%	49.4%	58.2%	0.0%	47.7%
	Male	Count	71	45	33	8	157
		% within Ethnicity	57.3%	50.6%	41.8%	100.0%	52.3%
Total		Count	124	89	79	8	300
		% within Ethnicity	100.0%	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	12.136[a]	3	.007	.006		
Likelihood Ratio	15.219	3	.002	.002		
Fisher's Exact Test	12.663			.004		
Linear-by-Linear Association	.801[b]	1	.371	.394	.204	.035
N of Valid Cases	300					

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 3.81.

b. The standardized statistic is -.895.

According to Fisher's exact test (exact p=0.004), there was an association between gender and ethnicity in our sample. For example, for women the highest percentage was Asian ethnicity (58%).

McNemar Test

When to use

To test if, according to the current data, the proportions in the population of a variable change based on another matched variable.

Hypotheses:

H_0: there is no association between the two (paired) variables
H_a: there is an association between the two (paired) variables

Assumptions:

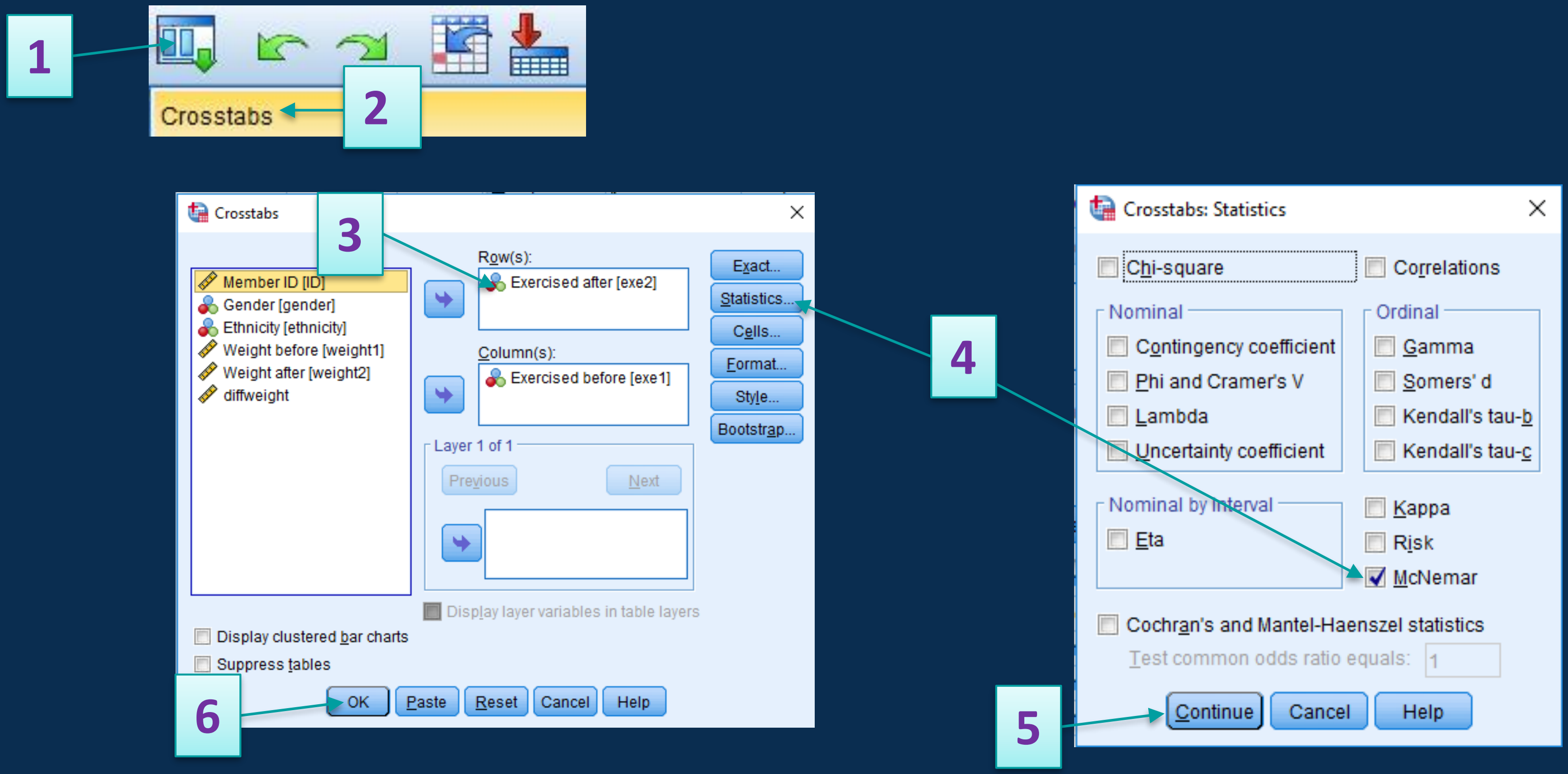
- The observations are randomly and independently drawn
- There are at least 25 observations in the discordant cells
- The data are paired

SPSS Slide: 'how to'

The next question is: do people exercise more after the programme than before?

Are the proportions of those exercised before the programme, different of those exercised after the programme?

<u>Step 1</u>: Use the appropriate test, here: 'McNemar chi-square test'.



Output and Interpretation Slide

Step 1: Use the appropriate test, here: 'McNemar test'.

Exercised after * Exercised before Crosstabulation

Count

		Exercised before No	Exercised before Yes	Total
Exercised after	No	9	7	16
	Yes	8	5	13
Total		17	12	29

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		1.000[a]
N of Valid Cases	29	

a. Binomial distribution used.

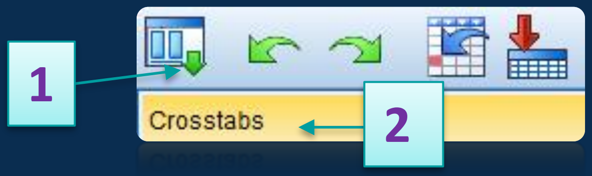
Step 2: Check the suitability of the data:

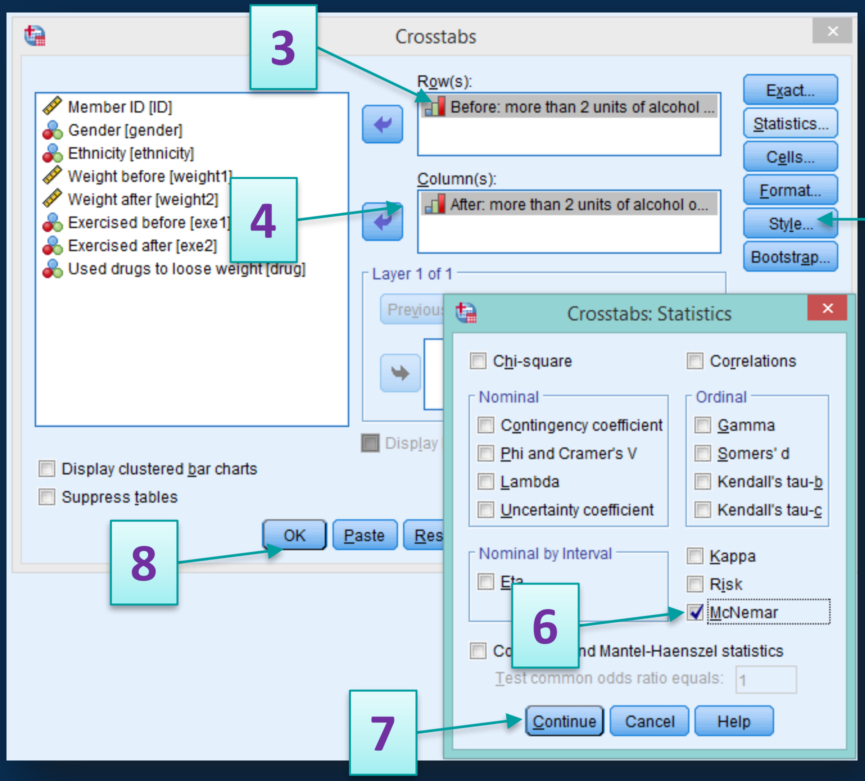
For McNemar's test, we needed at least 25 discordant observations, and paired data. We only have 15 observations. SPSS instead of printing the asymptotic p-value prints the exact test. However, SPSS always prints the exact p-value (binomial), even if the assumptions hold! (SPSS only does this).

Therefore, in SPSS only, the McNemar p-value is always valid, as long as we have paired categorical data.

SPSS Slide: 'how to'

We can compute the McNemar test for PxP tables as well.





Analyse-> Descriptive Statistics->Crosstabs

Place before variable in 'row' box and after variable in 'column' box.

In 'Statistics' choose McNemar

Click 'continue'

Click 'ok'

Output and Interpretation Slide

Step 1: Use the appropriate test, here: 'McNemar test'.

Before: more than 2 units of alcohol on a weekend * After: more than 2 units of alcohol on a weekend Crosstabulation

			After: more than 2 units of alcohol on a weekend			
			Never	Sometimes	Always	Total
Before: more than 2 units of alcohol on a weekend	Never	Count	65	3	0	68
		% within Before: more than 2 units of alcohol on a weekend	95.6%	4.4%	0.0%	100.0%
	Sometimes	Count	9	148	0	157
		% within Before: more than 2 units of alcohol on a weekend	5.7%	94.3%	0.0%	100.0%
	Always	Count	0	13	62	75
		% within Before: more than 2 units of alcohol on a weekend	0.0%	17.3%	82.7%	100.0%
Total		Count	74	164	62	300
		% within Before: more than 2 units of alcohol on a weekend	24.7%	54.7%	20.7%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
McNemar-Bowker Test	16.000	2	.000
N of Valid Cases	300		

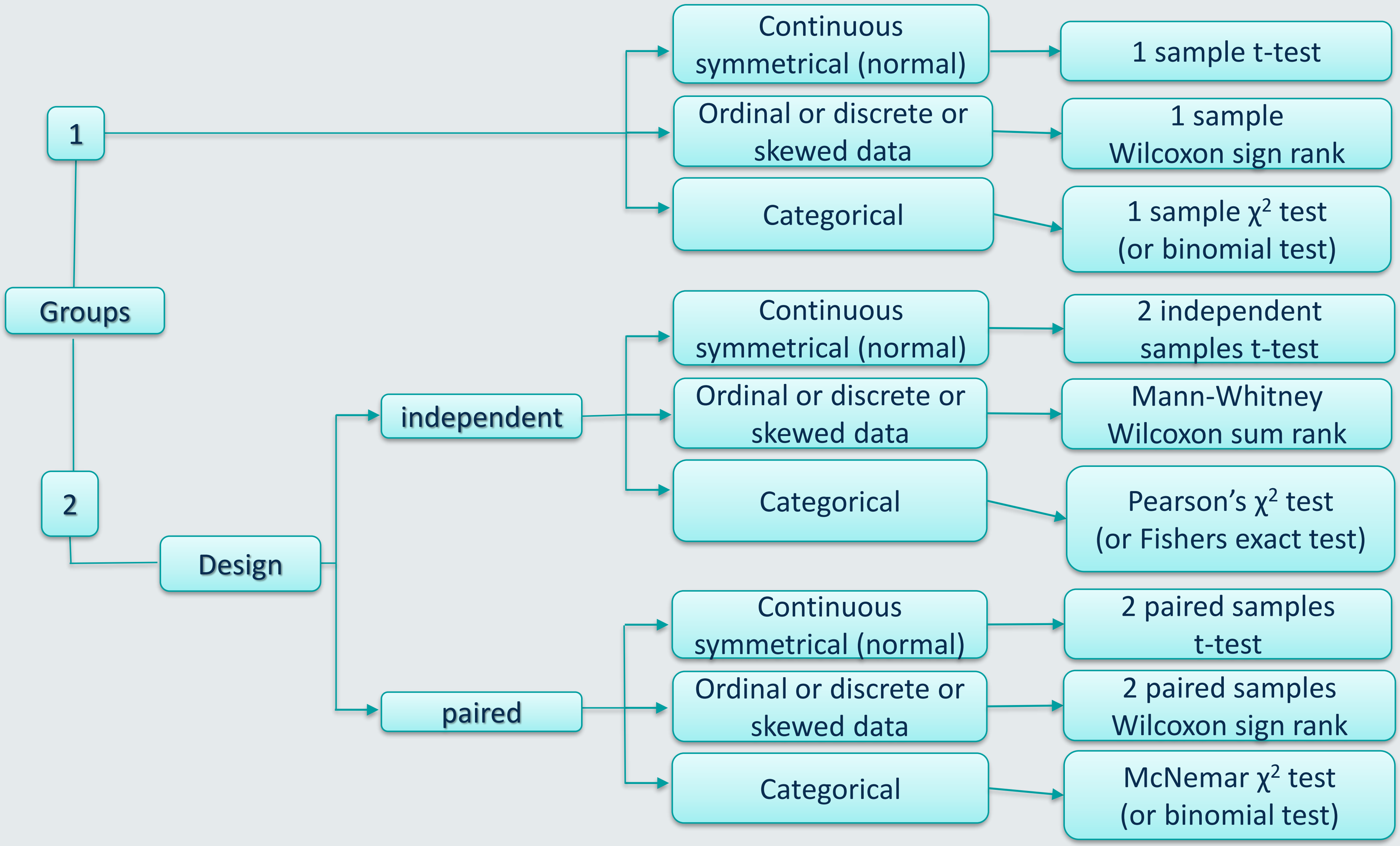
SPSS automatically prints the McNemar – Bowker test, which is the extension of the original test for pxp tables.

There is an association between the amount of alcohol consumed before and the amount of alcohol consumed after the programme (McNemar Browker χ^2=16.0, df=2, p<0.001).

Cheat Sheet

	Numerical data		Categorical data	
Hypotheses testing	parametric	non-parametric	chi-squared tests	exact tests
one group *versus* a pre-defined value	one sample t-test	Wilcoxon sign rank	one sample χ^2-test	One sample Binomial exact test
one group *versus* another group	two independent samples t-test	Mann-Whitney (Wilcoxon sum rank)	Pearson's χ^2-test	Fisher's Exact test
one group (twice or) *versus* another matched group	two paired samples t-test	Wilcoxon sign rank	McNemar test	Binomial exact test

Cheat Sheet in a Flow Chart…



Knowledge Test

The assumptions have been violated. Match the scenario with the correct test.

Tom wants to test if boys' proportions in ADHD high/low classification groups are different than those of girls.

Tom wants to test if mothers' reported ADHD high/low classification for children are different than those reported by fathers.

Tom wants to test if children's high classification ADHD proportion is higher than 50%.

One sample Binomial exact test

Fisher's exact test

McNemar χ^2 test

Reflection

Reflecting on your own field of study.

Write down three examples from your research that would require the use of each of the three exact tests.

Reference List

Agresti and Finlay (2009) Statistical Methods for the Social Sciences, 4th Edn, Pearson Hall, Upper Saddle River, NJ.

- Comparison of Two Groups, Ch 7, pages 183-209
- Analyzing Association between Categorical Variables, Ch 8, pages 221-239

Field (2005) Discovering Statistics using SPSS, 2nd Edn, Sage, London.

- Comparing Two Means, Ch 7
- Categorical Data, Ch 16

Thank you

Please contact your module leader or the course lecturer of your programme, or visit the module's forum for any questions you may have.

If you have comments on the materials (spotted typos or missing points) please contact Dr Vitoratou:

Silia Vitoratou, PhD
Psychometrics & Measurement Lab,
Department of Biostatistics and Health Informatics
IoPPN, King's College London, SE5 8AF, London, UK
silia.vitoratou@kcl.ac.uk

For any other comments or remarks on the module structure, please contact one of the three module leaders of the Biostatistics and Health Informatics department:

Zahra Abdula: **zahra.abdulla@kcl.ac.uk**

Raquel Iniesta: **raquel.iniesta@kcl.ac.uk**

Silia Vitoratou: **silia.vitoratou@kcl.ac.uk**