

# **PUBH620 Biostatistics**

## **Life in Statistics**

### **Week 3**

#### **T-tests and Chi-square tests**

Lecture notes by Dr Brandon Cheong, Dr Peter Mahoney  
and Dr Michael Taylor

# ELECTRONIC WARNING NOTICE

Commonwealth of Australia

*Copyright Act 1968*

**Form of notice for paragraph 135KA (a) of the *Copyright Act*  
1968**

## **Warning**

This material has been copied and communicated to you by or on behalf of *Australian Catholic University under Part VA of the Copyright Act 1968 (the **Act**)*.

The material in this communication may be subject to copyright under the Act. Any further copying or communication of this material by you may be the subject of copyright or performers' protection under the Act.

Do not remove this notice.

# In this lecture...

(clicking the links below will direct you to the topic page)

## [3.1 T-tests](#)

[3.1.1 One sample t-test](#)

[3.1.2 Independent samples t-test](#)

[3.1.3 Paired sample t-test](#)

## [3.2 Chi-square tests](#)

[3.2.1 Chi-square test for goodness of fit](#)

[3.2.2 Chi-square test of association](#)

[3.2.3 Chi-square test of association with Yate's correction method](#)

## [3.3 APA statistical reporting](#)

# Topic Learning Objectives (TLOs)

- Comparison of means: establishing difference between mean values for groups: independent, paired or for a sample against a standard/test value.
- Understand t-tests and be able to apply them in a public health context.
- Comparison of proportions: observing difference in proportions (frequencies) between categories.
- Understand chi-square tests and their application to public health.

## 3.1 T-tests

- A t-test is a commonly used test to compare the means of two groups.
- We will learn three types of t-tests in this unit: (1) one-sample, (2) independent and (3) paired t-tests.
- Because we are comparing means, data must be continuous to use a t-test. In other words, data is measured on a ratio or interval scale.

While we will learn how to calculate some test statistics by hand, this is purely for your knowledge and theory. It is highly advisable that you learn to use SPSS to do your analysis. SPSS software is used in the “real-world” globally and developing skills in it will be of great benefit to you in any sort of health industry (not just public health).

Every hypothesis test we learn in this unit will have a set of assumptions. These assumptions need to be satisfied before proceeding with the test. **There are some exceptions however, some assumptions become less relevant with larger sample sizes ( $n > 40$ ).** ie. It is important to check the assumptions of the statistical test before conducting it however, violation of the normality assumption will not cause major issues for larger sample sizes.

Please read the article below:

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10(2), 486.

## 3.1.1 One sample t-test

The one sample t-test is used to compare a sample mean with a pre-determined population mean or constant value (test value in SPSS).

The comparison that is done is used to see if there is statistical significance between the means of the two groups.

Eg. The recommended LDL cholesterol level should be under 100 mg/dL (5.5 mmol/L) according to the Journal of the American College of Cardiology (JACC). You have a sample of 55 patients who have experienced 1 or 2 strokes in their life and you measure their cholesterol levels in mg/dL.

129	115	124	105	123	103	122	96	116	127	129
92	85	100	120	107	123	105	113	114	95	112
122	129	106	105	88	93	123	112	91	90	112
108	103	122	90	108	94	115	89	110	130	116
101	123	106	119	112	121	98	126	126	97	85

## 3.1.1 One sample t-test

Immediately, you might be able to see that majority of the LDL cholesterol numbers are greater than 100 and you could say already that there was a statistical difference however, we need to test this properly. In practice, you might be working with much larger sample sizes! Also note that the sample size is greater than 40 but we will still check the normality assumption to ensure the data are not too deviated away from a normal distribution (and to teach you about some of the tests used for checking normality).

The first thing we do is check the assumptions of the one sample t-test:

Assumptions of the one sample t-test:

1. Data is measured on a ratio or interval scale (continuous data).
2. The sample data should follow a normal distribution (or as close to).

As data is continuous, we have satisfied the first assumption but we still need to check the second assumption of normality. There are two main ways we can do this in SPSS. We already know one – the Kolmogorov-Smirnov test, the other one is the Shapiro-Wilk test.

## 3.1.1 One sample t-test

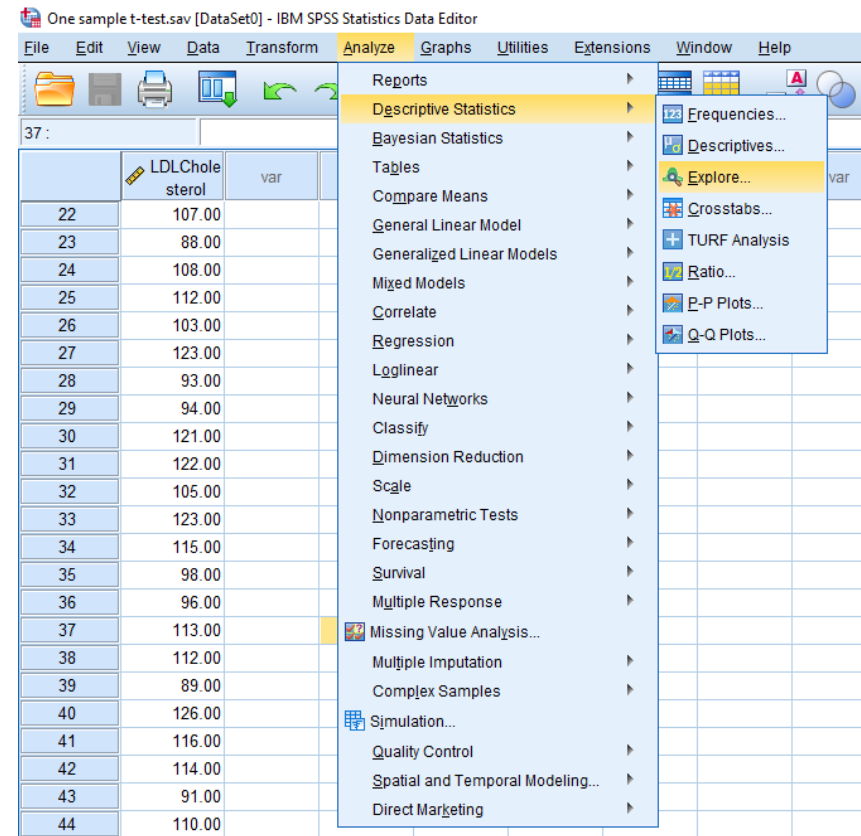
The easiest way to get both tests in the same table is to use the **<Explore>** option in SPSS under descriptive statistics.

You can access this by going:

**<Analyze>**

**<Descriptive Statistics>**

**<Explore>**

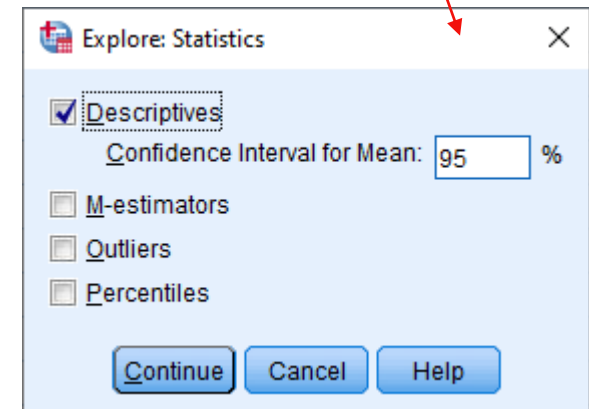
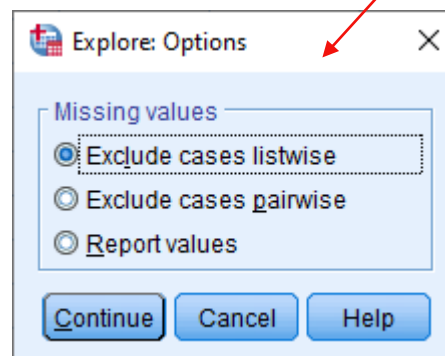
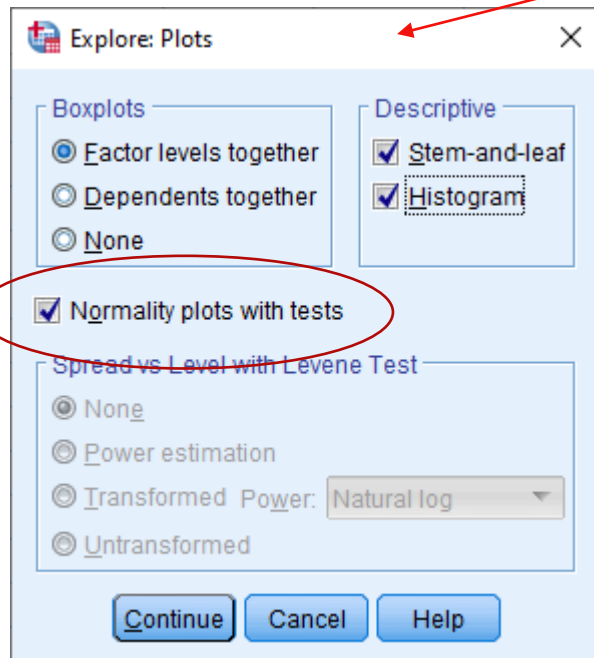
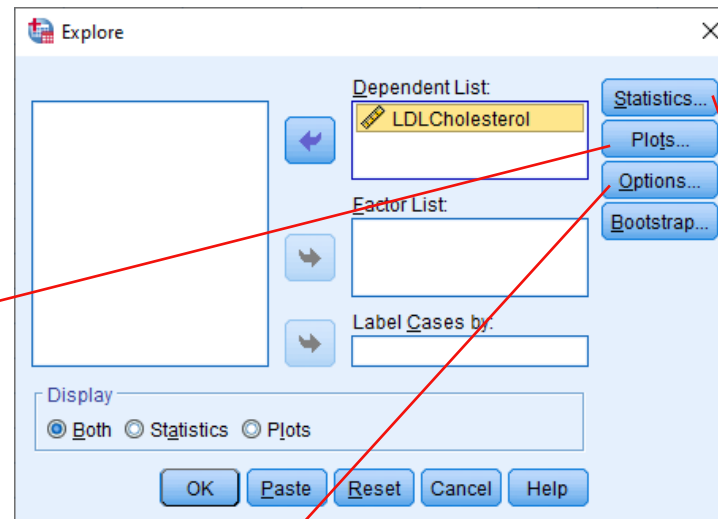




## 3.1.1 One sample t-test

Explore: Normality test

You will see this dialogue box →  
 Have a look at the extra options  
 and select boxes as shown here.

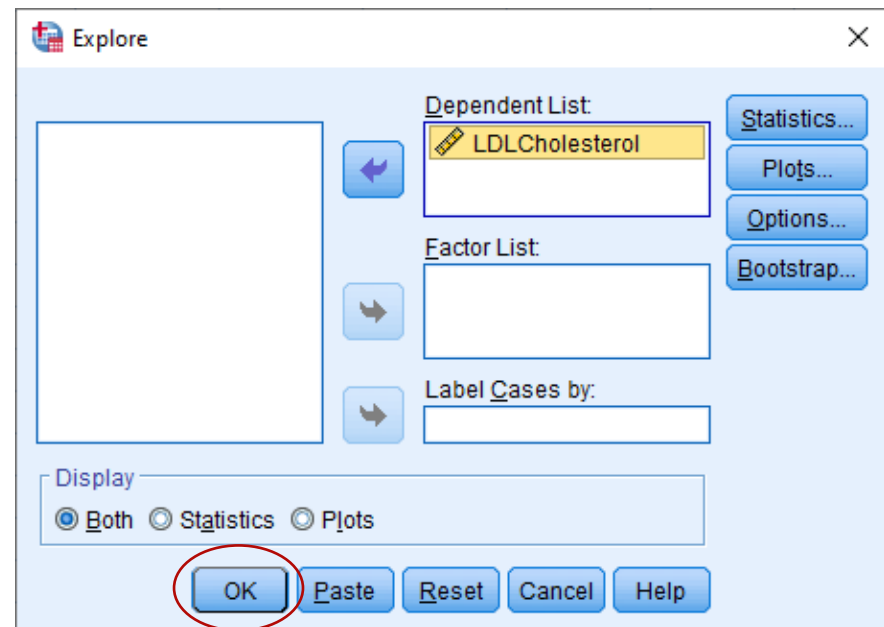


## 3.1.1 One sample t-test

Explore: Normality test

Click OK and you will get:

- Case processing summary
- Descriptives
- Tests of Normality
- Histogram
- Stem-and-Leaf plot
- Normal Q-Q Plot
- Detrended Normal Q-Q Plot



## 3.1.1 One sample t-test

### Case processing summary

Shows how many cases were analysed by SPSS. It also shows how many cases were not included in the analysis due to having missing data. If there is no missing data, number of cases analysed should equal the sample size.

Case Processing Summary						
	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
<u>LDLCholesterol</u>	55	100.0%	0	0.0%	55	100.0%

Here we see that there are 55 cases ( $N = 55$ ) and all cases are accounted for (Total Percent = 100.0% or Cases Missing Percent = 0.0%).

## 3.1.1 One sample t-test

### Descriptives

Contains the summary information such as mean LDL cholesterol level and standard deviation etc. 5% Trimmed Mean represents the mean that has been modified to remove outliers in the top and bottom 5% of the data set.

The mean LDL cholesterol level is 109.55 and standard deviation is 13.18.

Skewness and Kurtosis gives us an indication of the shape of the distribution and can also be used to help determine normality. The Kolmogorov-Smirnov test and Shapiro-Wilk test are more commonly used.

### Descriptives

		Statistic	Std. Error
<u>LDL Cholesterol</u>	Mean	109.5455	1.77749
	95% Confidence Interval for Mean		
	Lower Bound	105.9818	
	Upper Bound	113.1091	
	5% Trimmed Mean	109.7626	
	Median	112.0000	
	Variance	173.771	
	Std. Deviation	13.18222	
	Minimum	85.00	
	Maximum	130.00	
	Range	45.00	
	Interquartile Range	24.00	
	Skewness	-.225	.322
	Kurtosis	-1.096	.634

## 3.1.1 One sample t-test

### Tests of Normality

Remember the null hypothesis for the test of normality?

$H_0$ : The sample data are normal

$H_1$ : The sample data are not normal

Therefore if  $p > 0.05$ , we accept the  $H_0$  and have a normal distribution.

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
<u>LDL Cholesterol</u>	.100	55	.200*	.951	55	.025

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

We can clearly see that the KS test has a p-value (Sig.) of 0.200, which is  $> 0.05$  however, the SW test has a p-value of  $0.025 < 0.05$ . We need to confirm if this data set is normal and there are other ways that we can do this.

## 3.1.1 One sample t-test

### Other ways to confirm normality

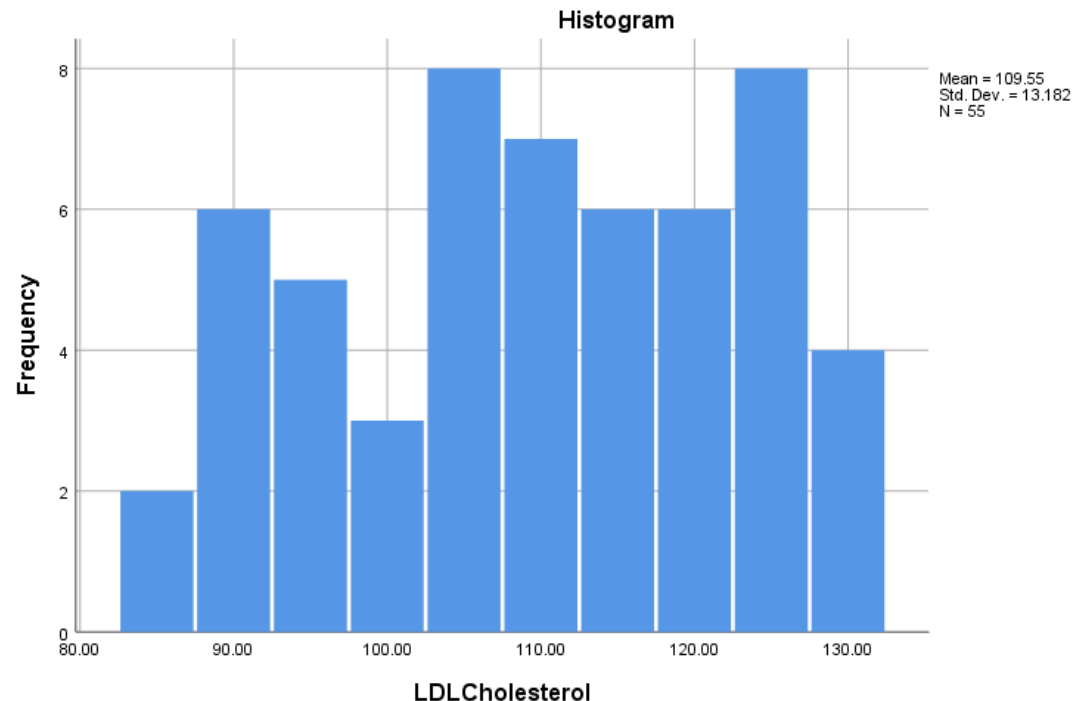
To confirm that our dataset is normal, we can use a histogram, Normal Q-Q plot and Detrended Normal Q-Q plot.

#### Histogram

The vertical axis represents Frequency and the horizontal axis represents the LDL cholesterol levels.

Here we can see that there are 6 patients that have 90 mg/dL as their LDL cholesterol level.

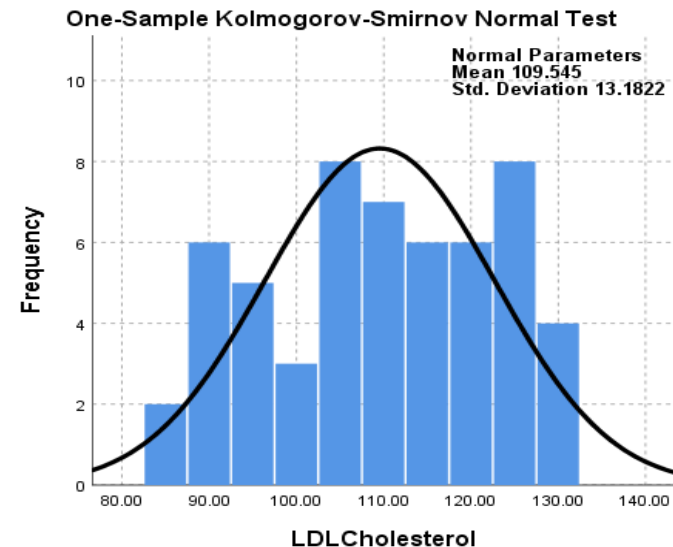
The histogram is used to view the overall distribution of the dataset. On first glance it may not look like this data is normal but let's overlay a normal curve over it using the KS test in SPSS.



## 3.1.1 One sample t-test

### Histogram

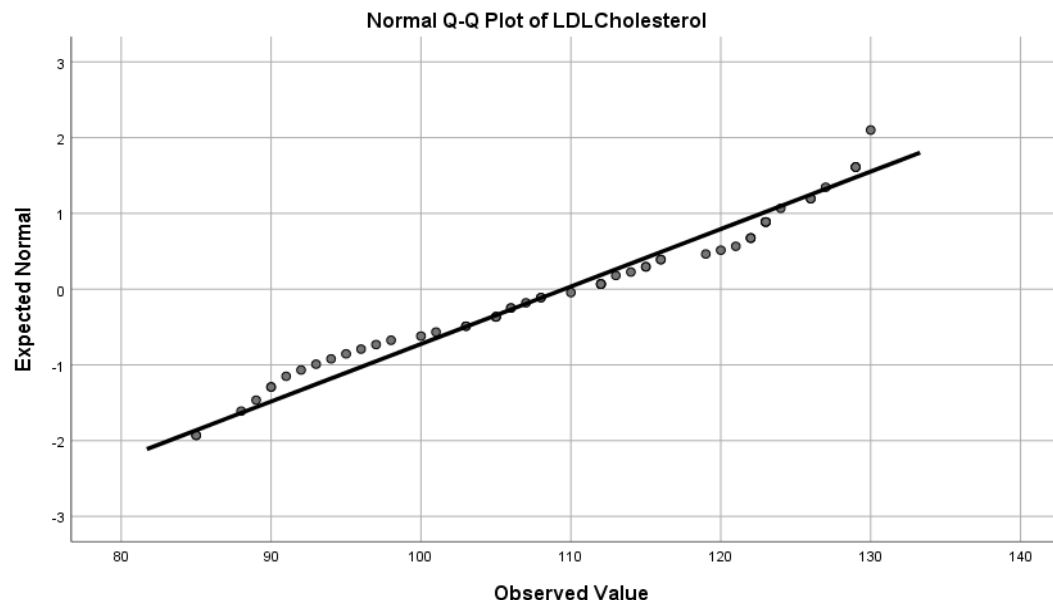
Here we can see that most of the histogram bars fit inside the normal curve and that the highest frequencies hover around the mean so this gives us a good indication that our dataset is in fact normal. Sometimes being “approximately normal” is considered having a normal dataset. That’s why it is so important to test your data properly for normality!



## 3.1.1 One sample t-test

### Normal Q-Q (Quantile-Quantile) Plot

This graph plots the observed data against the expected values if the dataset was normal. If the dataset is normal, the points on the graph will be as close as possible to the line of best fit.



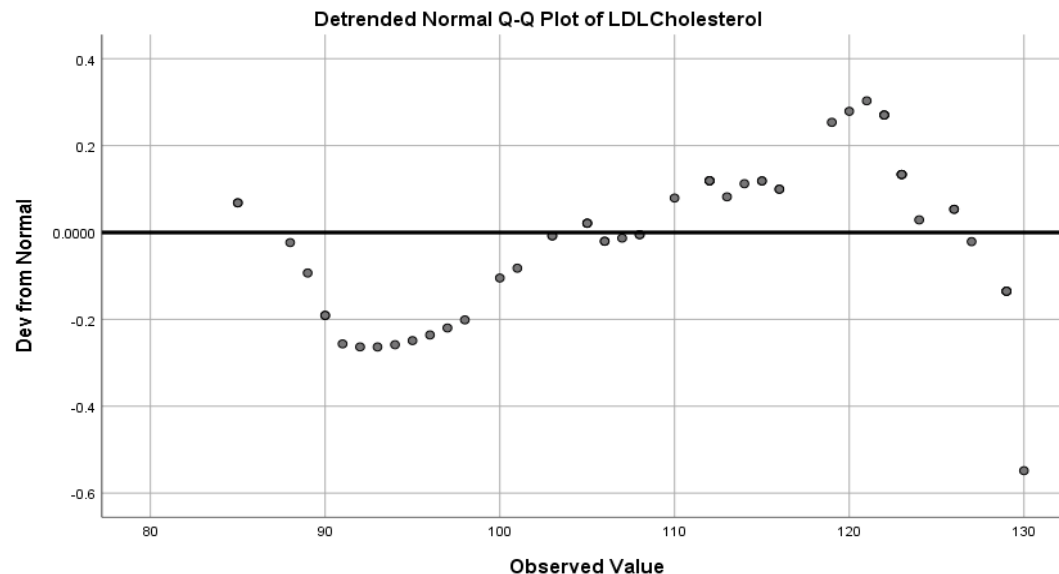
While some points diverge a little from the line of best fit, most of the points are clustered around it hence also showing that we have a normal dataset.



## 3.1.1 One sample t-test

### Detrended Normal Q-Q Plot

This graph plots the deviations from the line of best fit in the normal Q-Q plot. If the dataset is normal, we should see an even amount of points above and below the horizontal line.



We can see that there is roughly an even amount of points above and below the horizontal line showing that we have a normal distribution.

## 3.1.1 One sample t-test

Establish the null hypothesis,  $H_0$ .

After checking the assumptions, the next step is to establish the null hypothesis  $H_0$ .

Remember  $H_0$  is a testable statement that can be falsified, hence our  $H_0$  will be:

$H_0$ : there is no difference in the means of LDL cholesterol levels of the 55 patients and the recommended level of 100 mg/dL.

The alternate hypothesis would therefore be the opposite...

$H_1$ : there is a difference in the means of the LDL cholesterol levels of the 55 patients and the recommended level of 100 mg/dL.

## 3.1.1 One sample t-test

Calculating the test statistic for a one sample t-test

Once both assumptions have been satisfied we can proceed to calculate the test statistic.

Formula for one sample t-test:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

First, let's list out our parameters:

$$\bar{x} = 109.55$$

$$\mu = 100$$

$$s = 13.18$$

$$n = 55$$

Therefore, substituting into the equation we get:

$$t = \frac{109.55 - 100}{\frac{13.18222}{\sqrt{55}}} = 5.37$$

So what do we do now?

## 3.1.1 One sample t-test

Calculating the test statistic for a one sample t-test

Remember the t distribution table? By convention we choose a two-tailed test where  $\alpha = 0.05$  (more common) or  $0.01$ .

We see that highlighted values in the table are 1.96 and 2.58 respectively, these are known as the critical t values for a degrees of freedom of infinity (remembering our sample size was 55 and greater than 30).

On the previous slide,  $t = 5.37$  is known as the calculated t and if this value is greater than the critical t then we reject the null hypothesis and say we have statistical significance. So for  $\alpha = 0.05$ ,  $t > \text{critical } t$  therefore  $p < 0.05$ .

This is the method you would use if you were calculating test statistics by hand. More often than not you are encouraged to just use SPSS.

Degrees of Freedom	Alpha Level for one-tailed tests			
	0.05	0.025	0.01	0.005
	Alpha Level for two-tailed tests			
	0.1	0.05	0.02	0.01
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.269
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1.701	2.048	2.467	2.763
29	1.699	2.043	2.462	2.756
30	1.697	2.042	2.457	2.750
$\infty$	1.645	1.960	2.326	2.576

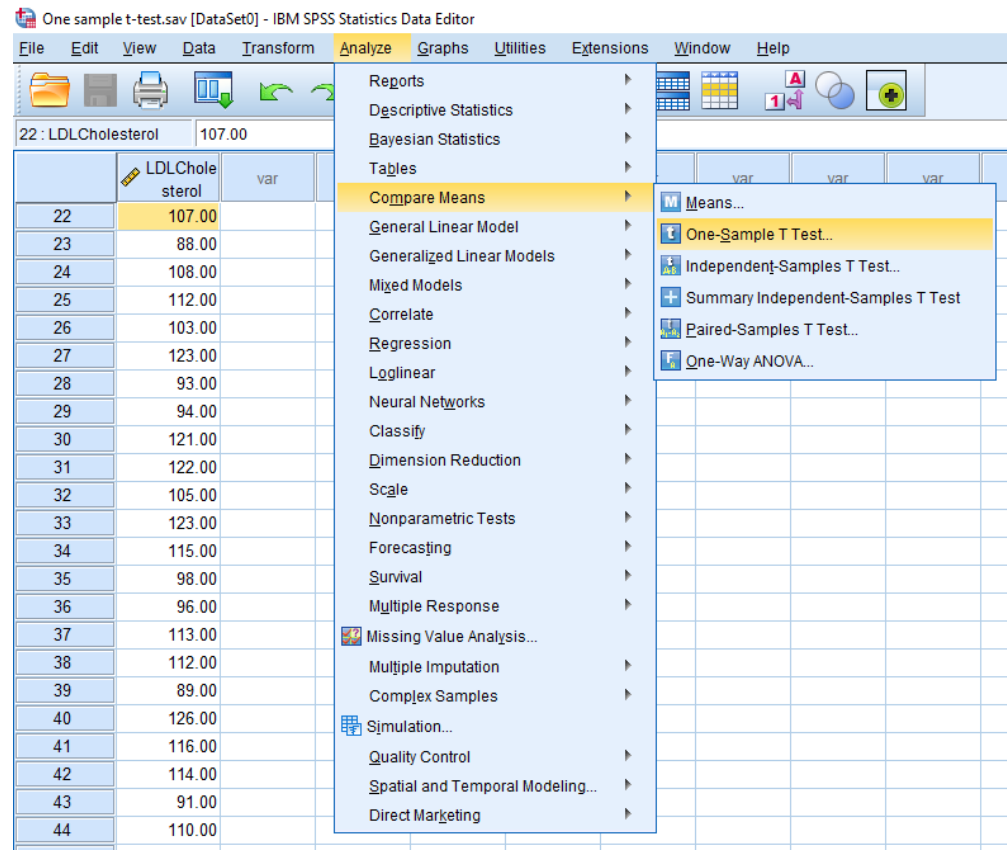
# 3.1.1 One sample t-test

One sample t-test in SPSS

<Analyze>

<Compare Means>

<One-Sample T Test>



## 3.1.1 One sample t-test

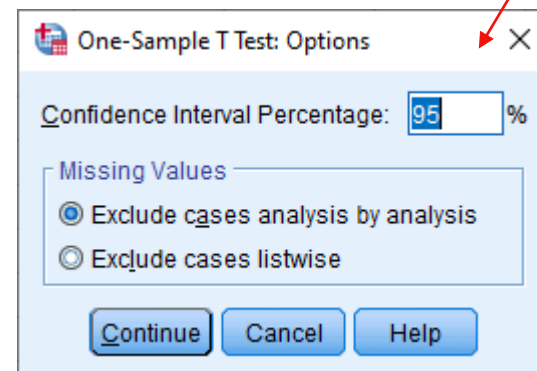
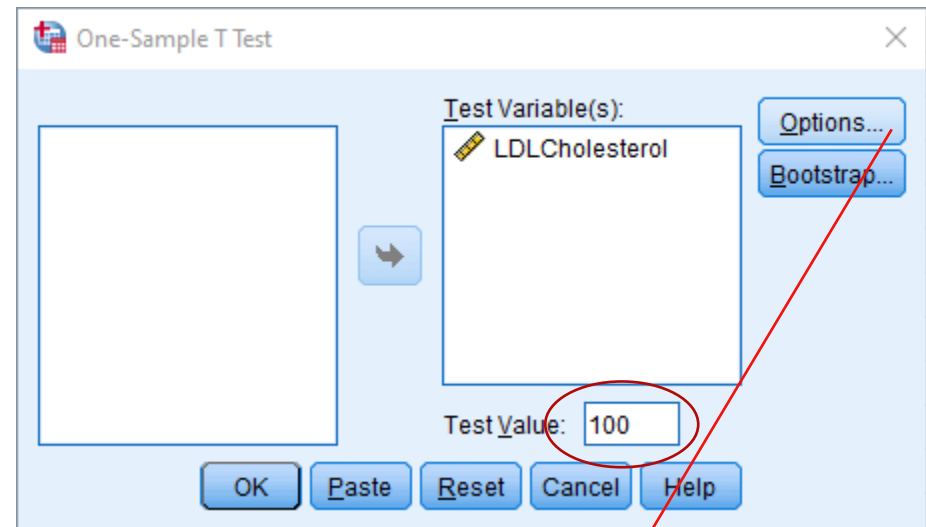
### One sample t-test in SPSS

You will see this dialogue box →

As our population mean is 100 mg/dL, input this as the Test Value.

SPSS will compare our sample's mean with this test value and determine whether there is a statistical difference or not.

If there is statistical significance, we reject the  $H_0$  that there is a difference in the means of the sample and recommended level and  $p < 0.05$ .



## 3.1.1 One sample t-test

One sample t-test in SPSS

Click OK and SPSS will generate the output of the one sample t-test.

The one-sample statistics table shows your descriptives such as sample size, mean, standard deviation and standard error of mean.

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
LDLCholesterol	55	109.5455	13.18222	1.77749

The One-Sample Test table shows the test statistic, degrees of freedom, p-value (Sig. (2-tailed)), mean difference and 95% confidence interval.

One-Sample Test					
Test Value = 100					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference
					Lower Upper
LDLCholesterol	5.370	54	.000	9.54545	5.9818 13.1091

The mean difference tells us how many points the 55 patients were above the recommended level, which was 9.55. Dividing this value by the sample standard deviation gives us what is called the effect size or Cohen's d. This is an indication of the magnitude difference between the sample mean and test value. As a guide, d = 0.20 is small, d = 0.50 is medium and d = 0.80 is large. d = 0.72 therefore indicating a medium magnitude of difference (the sample mean is around 0.72 standard deviations higher than the recommended level).

As we had calculated,  $t = 5.37$  and  $p < .001$ , therefore telling us that there is a statistical difference between our sample mean and the test value. Note that SPSS is limited when it comes to expressing more than three decimal places and so p-value = 0.000 should be interpreted as p-value  $< 0.001$ .

In reporting these results, it is very important that the proper format is used.

## 3.1.1 One sample t-test

Reporting results of t-tests (APA format)

### General approach:

1. What test was used to do the analysis?
2. What were the means and SDs (descriptives)?
3. What tests were used to check assumptions (KS test, SW test, QQ plot or detrended QQ plot) *as required*?
4. What were the statistical parameters (test statistic, degrees of freedom, p-value, Cohen's d and 95% CI)?

**Note that p-values should always be expressed as the exact value unless  $p < .001$ . ie. write  $p = .03$  instead of  $p < .05$ .**

Reporting the results of a statistical test is the most important step in hypothesis testing. Please make sure you know how to do it well!



## 3.1.1 One sample t-test

Complete reporting of the one sample t-test (APA format)

### Results

A one sample *t* test was used to compare the LDL Cholesterol levels of 55 patients ( $M = 109.55$ ,  $SD = 13.18$ ) against the recommended level of 100 mg/dL.

Visualisation of normal Q-Q and detrended normal Q-Q plots were used to confirm that the patients' LDL cholesterol levels were normally distributed.

The patients scored 9.55 points, 95% CI [105.98, 113.11], above the recommended level. This difference was found to be highly statistically significant,  $t(54) = 5.37$ ,  $p < .001$ , and medium  $d = 0.72$

## 3.1.2 Independent samples t-test

An independent samples t-test is used to compare the means between two independent groups.

The comparison that is done is used to see if there is statistical significance between the means of the two independent groups. The independent samples t-test is also known as the student's t-test.

Eg. Are the contestants on the 2014 version of the Biggest Loser as overweight/obese as those from the 2013 version?

Bodyweight is measured in kg – continuous data

Two independent groups – 2013 and 2014 cohorts, changing the order of one weight value in 2014 won't affect any of the values in 2013 hence they are independent.

Establish the null hypothesis,  $H_0$ .

$H_0$ : there is no difference in the means of bodyweights between The Biggest Loser 2013 and 2014 cohorts.

$H_1$ : there is a difference in the means of the bodyweights between The Biggest Loser 2013 and 2014 cohorts.

## 3.1.2 Independent samples t-test

### Assumptions of the independent samples t-test

1. Data are measured on interval or ratio scale.
2. Each participant (contestant) can only participate once in the research and cannot influence the participation of others.
3. Each group of values (weights) need to be normally distributed.
4. Homogeneity of variance – each group should have an equal amount of variability among each set of values (weights).

Assumptions 1 and 2 have already been satisfied, we have data on a ratio scale (weight measured in kg) and the two groups are independent of each other.

Use SPSS to check the remaining two assumptions however, note that assumption 4 is embedded in the statistical independent t-test and is called the Levene's test.

2013	Starting	2014	Starting
Contestant	Weight (kg)	Contestant	Weight (kg)
Katie	136.8	Kerry	141.6
Robyn	112.1	Craig	183.4
Todd	112.2	Jane	121.7
Gerald	159.1	Caitlin	134.4
Amber	111.6	Toni	124.2
Richard	139.7	Sharon	109.8
Janet	106.6	Cal	131.6
Kirsten	98.1	Natalie	117.6
Brett	187.9	Shannon	97.2
Mandy	113.3	Katrina	131.1
Sam	129.4	Mary	181.5
Jess	134.1	Matt	122.8
Chris	152.3	Cameron	146.4
Mark	162.9	Rodger	125.6
Cher	172.2	<b>Mean</b>	<b>133.5</b>
Anita	117.6	<b>St.dev</b>	<b>24.1</b>
Rosemary	100.1	<b>n</b>	<b>14</b>
<b>Mean</b>	<b>132.1</b>		
<b>St.dev</b>	<b>26.8</b>		
<b>n</b>	<b>17</b>		

## 3.1.2 Independent samples t-test

Assumptions of the independent samples t-test

Setting your data up in SPSS

Data needs to be set up like so for an independent samples t-test.

Group 1 = 2013 cohort

Group 2 = 2014 cohort

Contestant and Group names can be labelled in the “Variable View” tab.

Independent samples t-test.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Contestant	Bodyweights	Groups
1	1	136.8	1
2	2	112.1	1
3	3	112.2	1
4	4	159.1	1
5	5	111.6	1
6	6	139.7	1
7	7	106.6	1
8	8	98.1	1
9	9	187.9	1
10	10	113.3	1
11	11	129.4	1
12	12	134.1	1
13	13	152.3	1
14	14	162.9	1
15	15	172.2	1
16	16	117.6	1
17	17	100.1	1
18	18	141.6	2
19	19	183.4	2
20	20	121.7	2
21	21	134.4	2
22	22	124.2	2
23	23	109.8	2
24	24	131.6	2
25	25	117.6	2
26	26	97.2	2
27	27	131.1	2
28	28	181.5	2
29	29	122.8	2
30	30	146.4	2
31	31	125.6	2

## 3.1.2 Independent samples t-test

Assumptions of the independent samples t-test

Use Values column to label your groups and contestants.

Independent samples t-test.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Contestant	Numeric	8	0		None	None	20	Right	Nominal	Input
2	Bodyweights	Numeric	8	1		None	None	16	Right	Scale	Input
3	Groups	Numeric	8	0		None	None	16	Right	Nominal	Input

Value Labels

Value: 7

Label: Janet

Add  
 Change  
 Remove

- 1 = "Katie"
- 2 = "Robyn"
- 3 = "Todd"
- 4 = "Gerald"
- 5 = "Amber"
- 6 = "Richard"

OK Cancel Help

Value Labels

Value:

Label:

Add  
 Change  
 Remove

- 26 = "Shannon"
- 27 = "Katrina"
- 28 = "Mary"
- 29 = "Matt"
- 30 = "Cameron"
- 31 = "Rodger"

OK Cancel Help

Value Labels

Value:

Label:

Add  
 Change  
 Remove

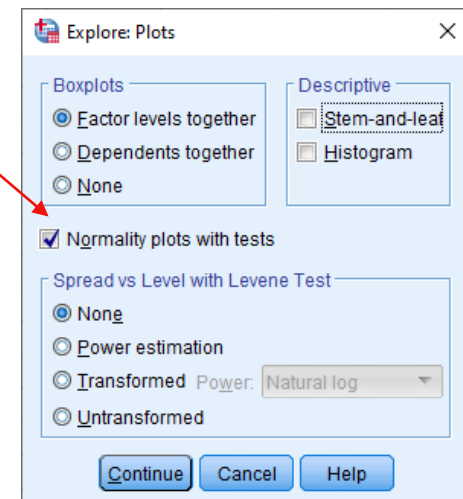
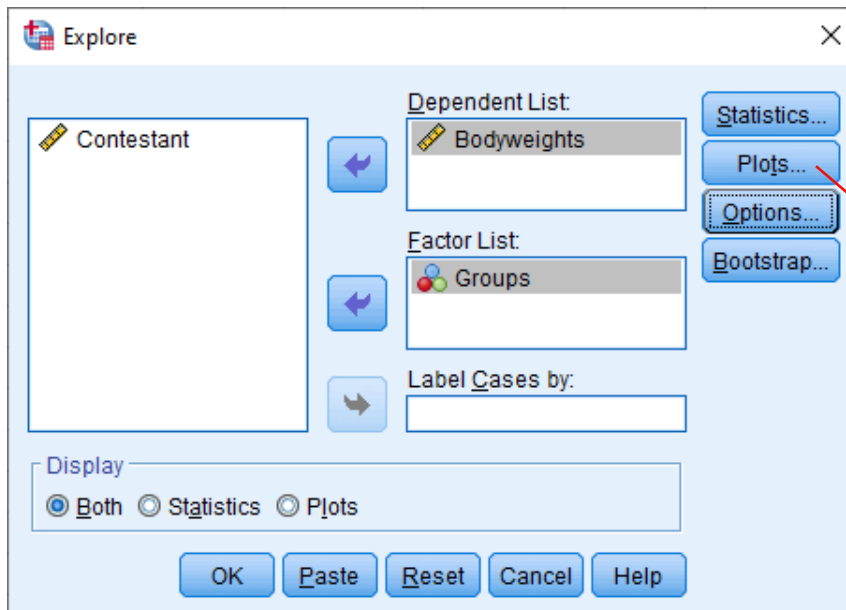
- 1 = "2013 Cohort"
- 2 = "2014 Cohort"

OK Cancel Help

## 3.1.2 Independent samples t-test

Assumptions of the independent samples t-test

Using <**Explore**> descriptive statistics and check “Normality plots with tests” to check assumption 3 that both groups are normally distributed.



## 3.1.2 Independent samples t-test

### Assumptions of the independent samples t-test

Tests of Normality							
		Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Groups	Statistic	df	Sig.	Statistic	df	Sig.
Bodyweights	2013 Cohort	.176	17	.166	.929	17	.213
	2014 Cohort	.199	14	.137	.885	14	.068

a. Lilliefors Significance Correction

Both normality tests show p-values (Sig.) > 0.05 for each group, therefore we can conclude that both groups are normally distributed. Assumption 3 satisfied.

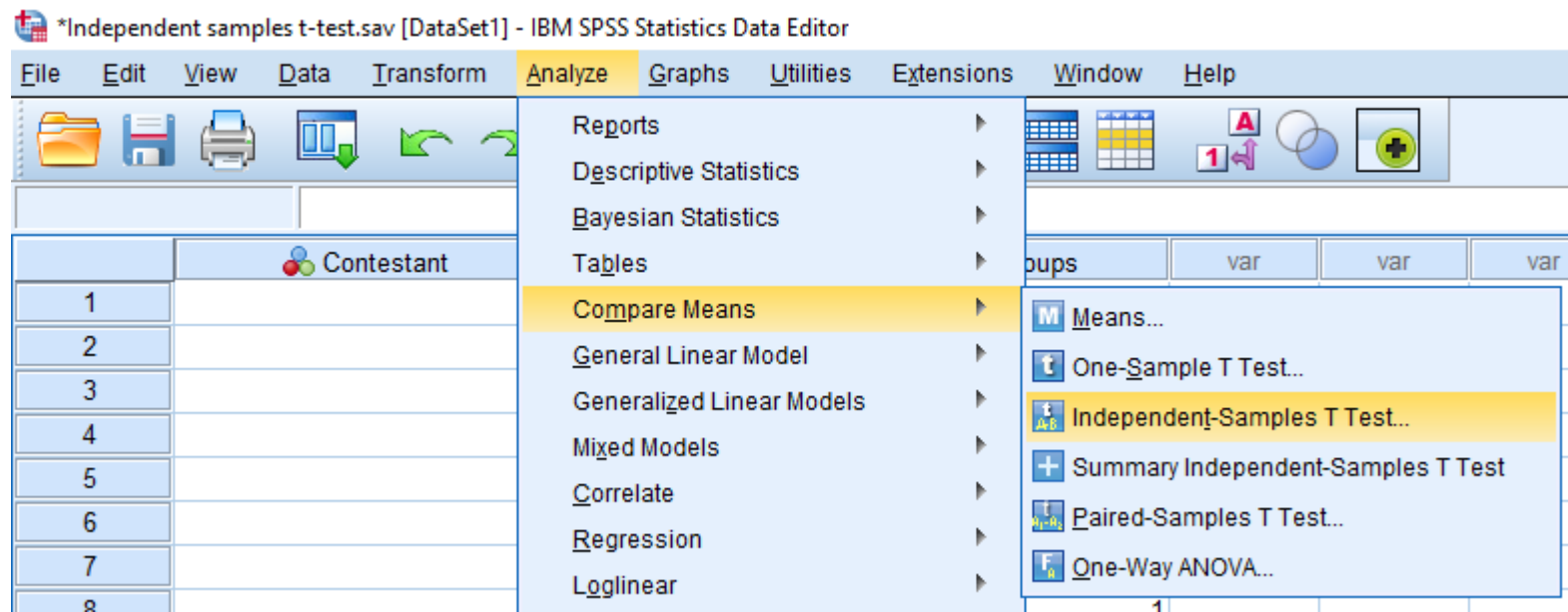
There is just one more assumption to be satisfied. Homogeneity of variance. A common error that people make about this assumption is that the variance value of one group must exactly equal to the variance value of the second group. This is not the case, instead what this assumption tests is that the variability of one group is approximately similar to the variability in the second group. This assumption can be checked using what's called the Levene's test.

The Levene's test is automatically applied when running an independent samples t-test in SPSS.

## 3.1.2 Independent samples t-test

Assumptions of the independent samples t-test

Click <**Analyze**>, <**Compare Means**>, <**Independent-Samples T Test**>





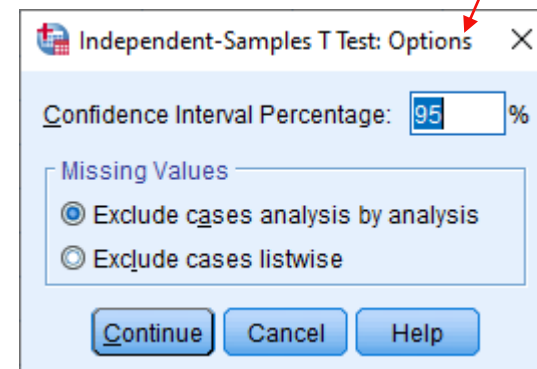
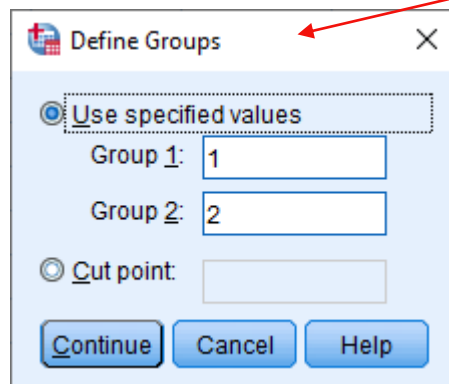
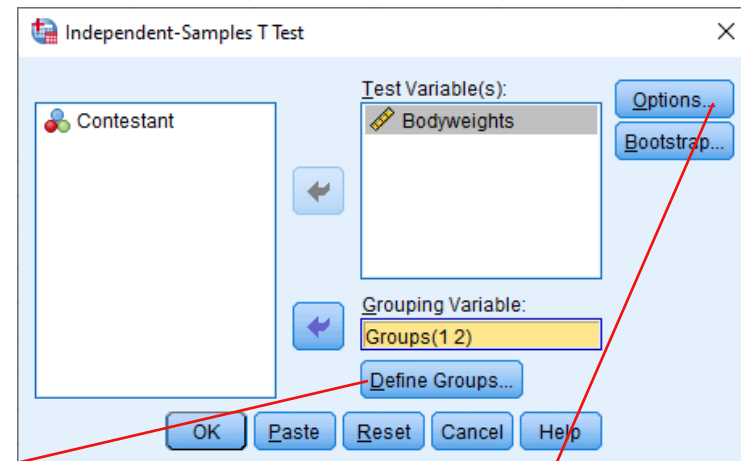
## 3.1.2 Independent samples t-test

Assumptions of the independent samples t-test

Remember:

Group 1 = 2013 Cohort

Group 2 = 2014 Cohort



## 3.1.2 Independent samples t-test

Assumptions of the independent samples t-test

Levene's test

Independent Samples Test										
Levene's Test for Equality of Variances				t-test for Equality of Means						
								95% Confidence Interval of the Difference		
				t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Bodyweights	Equal variances assumed	.838	.367	-.149	29	.883	-1.3752	9.2572	-20.3083	17.5579
	Equal variances not assumed			-.150	28.738	.882	-1.3752	9.1603	-20.1175	17.3670

The Levene's test is similar to the Kolmogorov-Smirnov test in that stating the  $H_0$  suggests that the variances are equal:

$H_0$ : there are equal variances between the 2013 Cohort and 2014 Cohort.

$H_1$ : there are not equal variances between the 2013 Cohort and 2014 Cohort.

Therefore, we can see that due to the Sig. value being 0.367 this is  $> 0.05$  so we accept the  $H_0$  that variances are equal and assumption 4 is satisfied.

Note that it is possible to have non-equal variances for an independent samples t-test. This is known as a heteroscedastic t-test. A t-test that has equal variances assumed is called a homoscedastic t-test.

## 3.1.2 Independent samples t-test

Calculating the test statistic for an independent samples t-test

Formula for an independent samples t-test:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left[ \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} \right] \left[ \frac{(n_1 + n_2)}{(n_1 n_2)} \right]}}$$

Difference in mean values  
between the two groups

Overall variance across  
the two groups

We will use an  $\alpha = .05$  two-tailed test. For an independent t-test, the degrees of freedom =  $(n_1 + n_2) - 2 = 29$ , this is less than 30 so our critical t value will be slightly different.

$s_1$  and  $s_2$  are the standard deviations of group 1 (2013 cohort) and group 2 (2014 cohort) respectively. But remember SD squared is the variance!

## 3.1.2 Independent samples t-test

Calculating the test statistic for an independent samples t-test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left[ \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} \right] \left[ \frac{(n_1 + n_2)}{n_1 n_2} \right]}}$$

$$t = \frac{132.1 - 133.5}{\sqrt{\left[ \frac{(17 - 1)719.3 + (14 - 1)582.4}{(17 + 14 - 2)} \right] \left[ \frac{(17 + 14)}{17 \times 14} \right]}}$$

$$t = -0.149$$

Calculated t = -0.149

Critical t = 2.043

Since we have a negative calculated t, we need to consider the negative range. In other words, our critical t becomes -2.043 and if our calculated t > critical t, which in this case it is not then we accept the  $H_0$  and say that the mean bodyweight between 2013 and 2014 cohorts is not statistically different. The best way to deal with negative test statistics is to simply take the absolute value and just treat them as positive. ie.  $0.149 < 2.043$

Degrees of Freedom	Alpha Level for one-tailed tests			
	0.05	0.025	0.01	0.005
	Alpha Level for two-tailed tests			
	0.1	0.05	0.02	0.01
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.269
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1.701	2.048	2.467	2.763
29	1.699	2.043	2.462	2.756
30	1.697	2.042	2.457	2.750

## 3.1.2 Independent samples t-test

### Independent samples t-test in SPSS

By checking assumption 4 (homogeneity of variance), we have already generated the SPSS output for the independent t-test.

**Group Statistics**

	Groups	N	Mean	Std. Deviation	Std. Error Mean
Bodyweights	2013 Cohort	17	132.118	26.8196	6.5047
	2014 Cohort	14	133.493	24.1327	6.4497

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Bodyweights	Equal variances assumed	.838	.367	-.149	29	.883	-1.3752	9.2572	-20.3083	17.5579
	Equal variances not assumed			-.150	28.738	.882	-1.3752	9.1603	-20.1175	17.3670

We see that our test statistic is the same as what we calculated. The Sig. (2-tailed) value shows that  $p > 0.05$  and therefore we do not have a statistical difference in the mean bodyweights of the 2013 and 2014 cohorts.

## 3.1.2 Independent samples t-test

### Calculating effect size

In most cases, p-value can only tell us so much. It indicates whether there was a statistical significance in the data that was obtained but it fails to tell us how significant this really is!

This is where effect sizes come in. Effect sizes tell us how much of an effect or how important the findings of our analysis are. This allows us “to appreciate the magnitude of the work that has been done to analyse the data” (Allan et al., 2019).

Effect sizes typically come in two families. The d family (looking at the difference between groups) and the r family (measuring the strength of a relationship).

## 3.1.2 Independent samples t-test

### Calculating effect size

For an independent samples t-test, we will use Cohen's d to determine the effect size.

$$d = \frac{\mu_1 - \mu_2}{s_p}$$

Where  $\mu_1$  and  $\mu_2$  are means of groups 1 and 2 respectively and  $s_p$  is the pooled standard deviation given by:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Where  $n_1$  and  $n_2$  are the sample sizes of groups 1 and 2 respectively and  $s_1^2$  and  $s_2^2$  are the variances of groups 1 and 2 respectively.

## 3.1.2 Independent samples t-test

Calculating effect size

Hence:

$$s_p = \sqrt{\frac{(17 - 1)26.8^2 + (14 - 1)24.1^2}{17 + 14 - 2}}$$

$$s_p = 25.62$$

Therefore:

$$d = \frac{132.1 - 133.5}{25.62} = -0.05$$

This indicates a small effect!



## 3.1.2 Independent samples t-test

Complete reporting of the independent samples t-test:

An independent samples *t* test was used to compare the mean bodyweights in The Biggest Loser 2013 cohort ( $n = 17$ ) to the mean bodyweights of the 2014 cohort ( $n = 14$ ).

Kolmogorov-Smirnov and Shapiro-Wilk tests showed that each group followed a normal distribution. A Levene's test was used to confirm that both groups have equal variances.

The *t* test was not statistically significant, with the 2013 cohort ( $M = 132.11$ ,  $SD = 26.82$ ) reporting only 1.38 kg less, 95% CI [-20.31, 17.56], than the 2014 cohort ( $M = 133.49$ ,  $SD = 24.13$ ),  $t(29) = -0.15$ ,  $p = .88$ .

When comparing means, it is always important to state which group had the higher mean value and the mean difference.

Note as there was no statistical difference, we have not included an effect size.

## 3.1.3 Paired samples t-test

A paired samples t-test is used to compare the means between two dependent (related) groups.

The comparison that is done is used to see if there is statistical significance between the means of the two related groups.

Eg. Patients that experience cardiac arrest might find themselves undergoing a 5-week rehabilitation program. A 6MWT (6 minute walk test) is done once before the rehabilitation program and once after the rehabilitation program and the distance walked is recorded in metres.

Let's see what this data set might look like in SPSS.

## 3.1.3 Paired samples t-test

Doing a paired t-test in SPSS is a bit more straightforward. Each group has a column of its own.

Next step is to check assumptions of paired t-test:

1. Data are measured on ratio or interval scale.
2. The differences between pairs are approximately normally distributed.

\*Paired samples t-test.sav [DataSet2] - IBM SPSS Statistics

	Before6MWT	After6MWT
1	300.00	334.00
2	295.00	388.00
3	300.00	416.00
4	254.00	302.00
5	345.00	311.00
6	349.00	338.00
7	297.00	447.00
8	354.00	406.00
9	341.00	447.00
10	306.00	340.00
11	302.00	370.00
12	371.00	387.00
13	240.00	334.00
14	306.00	321.00
15	275.00	302.00
16	349.00	421.00
17	274.00	450.00
18	278.00	321.00
19	255.00	383.00
20	368.00	427.00
21	380.00	395.00

## 3.1.3 Paired samples t-test

Assumptions of the paired t-test

Assumption 1 is satisfied as we can see that we have continuous data (distance walked in metres).

Assumption 2 requires using the transform menu in SPSS.

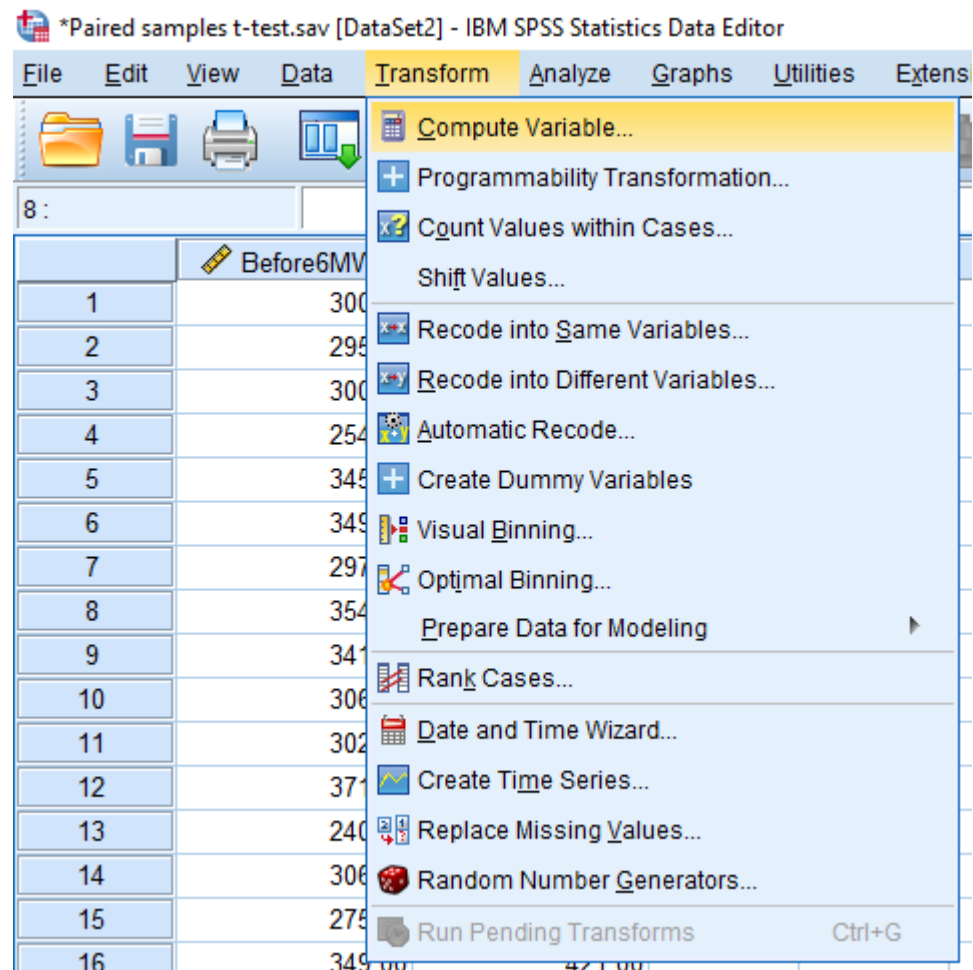
## 3.1.3 Paired samples t-test

Assumptions of the paired t-test

<Transform>

<Compute Variable>

This will allow us to calculate the differences in values between the two related groups.

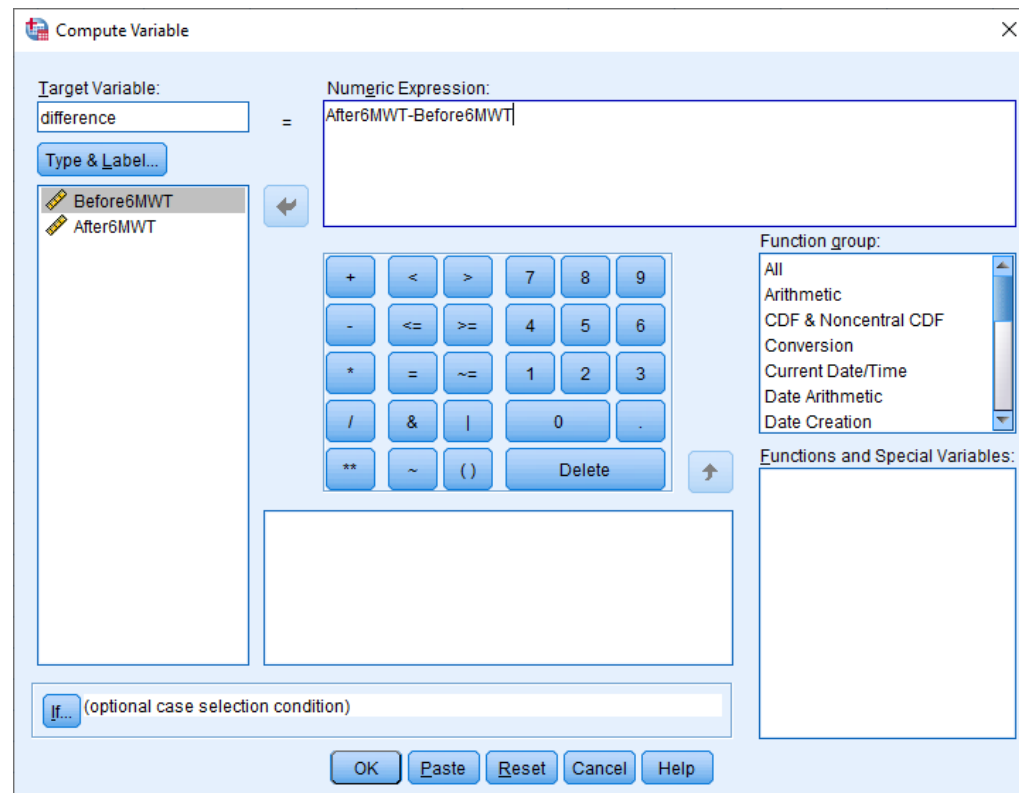


## 3.1.3 Paired samples t-test

### Assumptions of the paired t-test

Here the differences between Before6MWT and After6MWT have been computed.

Click OK.



## 3.1.3 Paired samples t-test

### Assumptions of the paired t-test

We see that a new column has been created “difference”, which is the result of our computed variable.

We can now test this column for normality as it represents the differences.

Note: it doesn't matter which direction of difference you use as long as you are consistent. ie. You could have computed Before6MWT – After6MWT instead.

\*Paired samples t-test.sav [DataSet2] - IBM SPSS Statistics Data Editor

	Before6MWT	After6MWT	difference
1	300.00	334.00	34.00
2	295.00	388.00	93.00
3	300.00	416.00	116.00
4	254.00	302.00	48.00
5	345.00	311.00	-34.00
6	349.00	338.00	-11.00
7	297.00	447.00	150.00
8	354.00	406.00	52.00
9	341.00	447.00	106.00
10	306.00	340.00	34.00
11	302.00	370.00	68.00
12	371.00	387.00	16.00
13	240.00	334.00	94.00
14	306.00	321.00	15.00
15	275.00	302.00	27.00
16	349.00	421.00	72.00
17	274.00	450.00	176.00
18	278.00	321.00	43.00
19	255.00	383.00	128.00
20	368.00	427.00	59.00
21	380.00	395.00	15.00

## 3.1.3 Paired samples t-test

### Assumptions of the paired t-test

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of difference is normal with mean 61.95 and standard deviation 53.20101.	One-Sample Kolmogorov-Smirnov Test	.200 <sup>a,b</sup>	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .050.

a. Lilliefors Corrected

b. This is a lower bound of the true significance.

Assumption 2 satisfied!

Now to establish the null hypothesis:

$H_0$ : there is no difference in the mean distance between Before6MWT and After6MWT

$H_1$ : there is a difference in the mean distance between Before6MWT and After6MWT



## 3.1.3 Paired samples t-test

Calculating the test statistic for a paired samples t-test

Formula:

$$t = \frac{\sum d}{\sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n - 1}}}$$

Sum of differences for each matched pair

Degrees of freedom =  $n - 1 = 20$

Using the formula above,  $t = 5.336$

Critical  $t = 2.086$

Therefore, reject the  $H_0$  and we have statistical significance ( $p < 0.05$ ).

Are you starting to see a pattern/recipe with these hypothesis tests?

## 3.1.3 Paired samples t-test

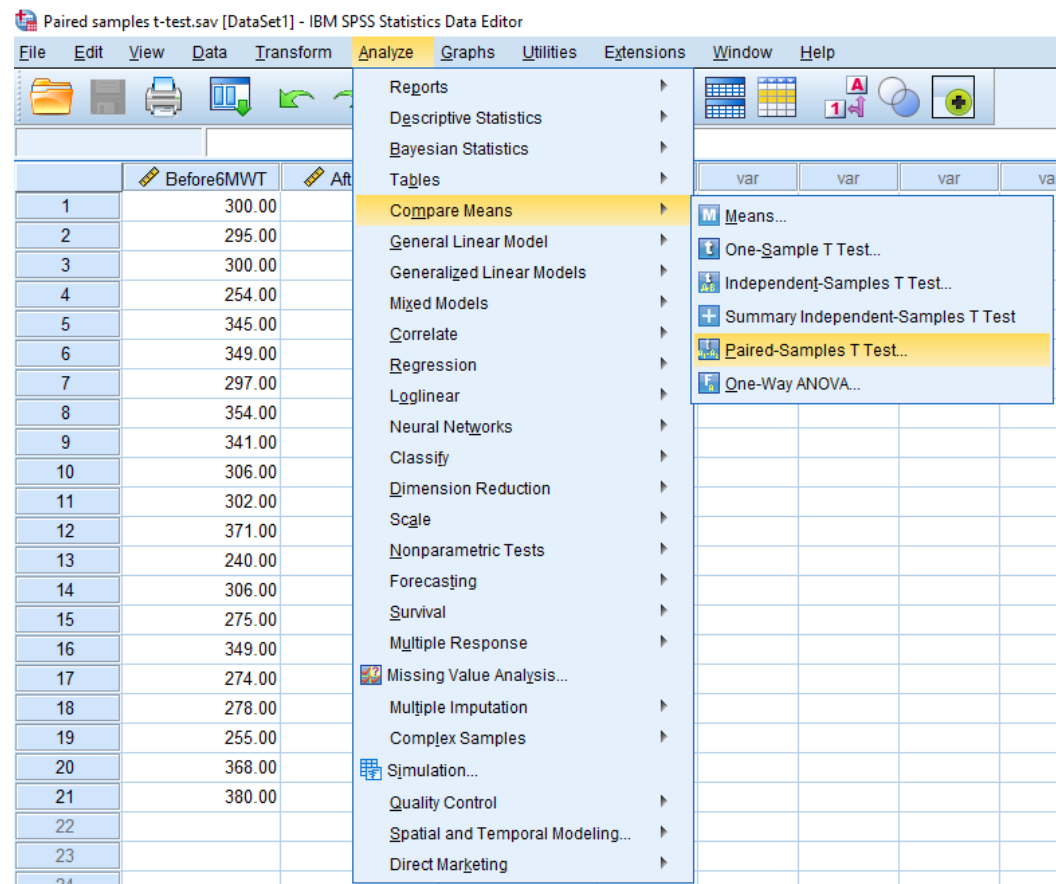
Paired samples t-test in SPSS

<Analyze>

<Compare Means>

<Paired-Samples T Test>

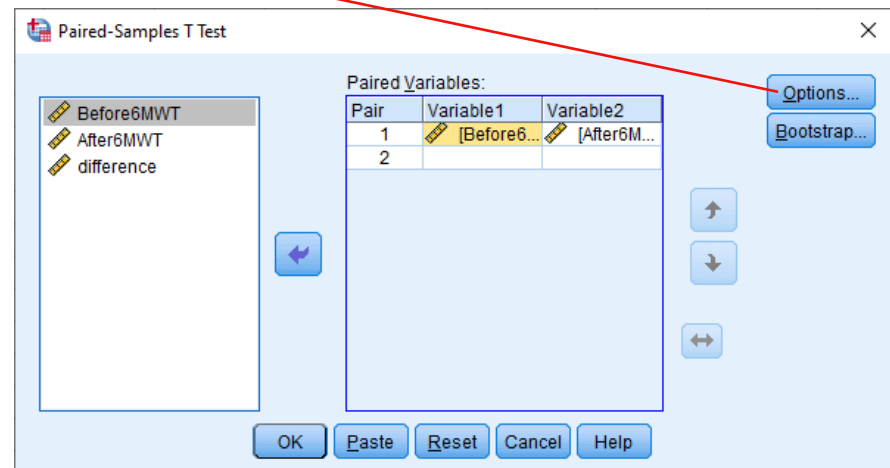
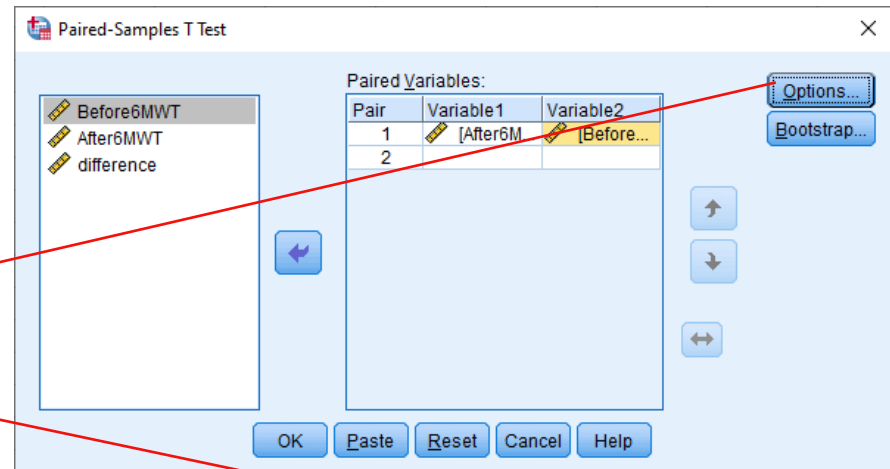
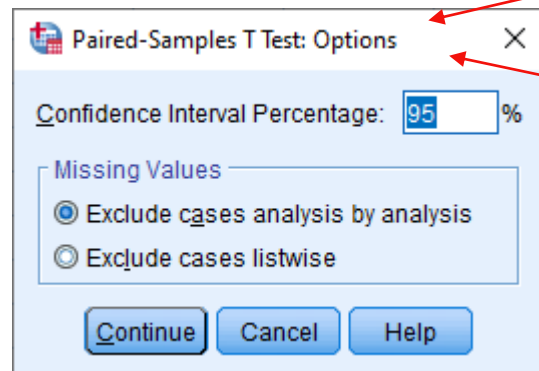
As usual, you will be led to a new dialogue box...



## 3.1.3 Paired samples t-test

Paired samples t-test in SPSS

Both options will work →



To get the SPSS output of the paired samples t-test.

## 3.1.3 Paired samples t-test

Paired samples t-test in SPSS

Basic descriptives →

This table reports a Pearson's correlation coefficient (Week 6). →

**Paired Samples Statistics**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	After6MWT	373.3333	21	50.36897	10.99141
	Before6MWT	311.3810	21	41.47587	9.05078

**Paired Samples Correlations**

		N	Correlation	Sig.
Pair 1	After6MWT & Before6MWT	21	.342	.130

**Paired Samples Test**

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	After6MWT - Before6MWT	61.95238	53.20101	11.60941	37.73557	86.16919	5.336	20	.000

At this stage, we should know what to look for.  $t = 5.336$ ,  $df = 20$  and  $p < 0.001$  therefore we have statistical significance and reject the null.

## 3.1.3 Paired samples t-test

Calculating effect size

For a paired samples t-test:

$$d = \frac{\mu_1 - \mu_2}{s_p}$$

Where  $\mu_1$  and  $\mu_2$  are means of groups 1 and 2 respectively and  $s_p$  is the pooled standard deviation given by:

$$s_p = \frac{s_1 + s_2}{2}$$

Where  $s_1$  and  $s_2$  are the standard deviations of groups 1 and 2 respectively.

## 3.1.3 Paired samples t-test

Calculating effect size

Hence:

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Before6MWT	21	240.00	380.00	311.3810	41.47587
After6MWT	21	302.00	450.00	373.3333	50.36897
Valid N (listwise)	21				

$$s_p = \frac{41.5 - 50.4}{2}$$

$$s_p = -4.45$$

Therefore:

$$d = \frac{311.1 - 373.3}{-4.45} = 13.98$$

This indicates a very large effect!

## 3.1.3 Paired samples t-test

Complete reporting of a paired samples t-test:

A paired samples *t* test was used to compare mean walking distances (in metres) for cardiac arrest patients employed in a rehabilitation program. Patients completed a 6MWT before the program ( $M = 311.38$ ,  $SD = 41.48$ ) and after the program ( $M = 373.33$ ,  $SD = 50.37$ ).

A Kolmogorov-Smirnov test was used to check the normality of difference assumption and these were found to be non-significant.

The average patient's walking distance was 61.95 metres, 95% CI [37.74, 86.17], longer after completing the program than prior to commencing it. This difference was statistically significant,  $t(20) = 5.34$ ,  $p < .001$ .

## 3.1.3 Paired samples t-test

Interpreting the results of a paired samples t-test (APA format)

A paired samples *t* test was used to compare mean walking distances (in metres) for cardiac arrest patients employed in a rehabilitation program. Patients completed a 6MWT before the program ( $M = 311.38$ ,  $SD = 41.48$ ) and after the program ( $M = 373.33$ ,  $SD = 50.37$ ).

A Kolmogorov-Smirnov test was used to check the normality of difference assumption and these were found to be non-significant.

The average patient's walking distance was 61.95 metres, 95% CI [37.74, 86.17], longer after completing the program than prior to commencing it. This difference was statistically significant,  $t(20) = 5.34$ ,  $p < .001$  with a large effect,  $d = 13.98$ .



## 3.2 Chi-square tests

- A chi-square test is a commonly used test to compare frequencies or proportions of categorised groups.
- We will learn three types of chi-square tests in this unit: (1) chi-square test for goodness of fit, (2) chi-square test of association and (3) chi-square test of association with Yate's correction.
- Because we are comparing frequencies or proportions, data must be categorical and is measured on a nominal scale.

Once again, while we will do some calculations by hand, it is advisable that you focus more on the procedures in SPSS as this will save you time and be more relevant to you in the future.

## 3.2.1 Chi-squared test for goodness of fit

A chi-square test for goodness of fit is used to compare whether observed frequencies (of categorical data) differ from expected frequencies (of categorical data).

Chi-square test statistic is denoted by  $\chi^2$ .

A chi-square test is known as a non-parametric test and doesn't require the assumption of normality unlike the t-tests.

Eg. The distribution of blood types A, B, AB, O is expected to follow a ratio of 10 : 4 : 1 : 15. A random sample of 150 people are surveyed in Melbourne CBD and a frequency table is generated:

Blood Type	A	B	AB	O
Frequency	37	28	2	83

## 3.2.1 Chi-squared test for goodness of fit

Now that we have a table of observed frequencies, we need to calculate the table of expected frequencies but before we do that, let's establish the null hypothesis and check the assumption of the chi-square test.

$H_0$ : there is no difference between observed and expected frequencies.

$H_1$ : there is a difference between observed and expected frequencies.

Like every hypothesis test we do, we need to check the assumptions of the test. The chi-square test only has one assumption.

1.  $\geq 80\%$  of expected frequencies  $\geq$  the value of 5

Again, we use SPSS to check this assumption. By checking the assumption of the chi-square test, we will conduct the chi-square analysis so let's try doing the calculation by hand first.

## 3.2.1 Chi-squared test for goodness of fit

To calculate the expected value, use the formula below:

$$\text{Expected value} = \text{sample size} \times \frac{\text{ratio}}{\text{total ratios}}$$

Therefore:

Blood Type	A	B	AB	O
Observed (O)	37	28	2	83

Blood type A Expected value =  $150 \times 10/30 = 50$

Blood type B Expected value =  $150 \times 4/30 = 20$

Blood type AB Expected value =  $150 \times 1/30 = 5$

Blood type O Expected value =  $150 \times 15/30 = 75$

Expected values are 50:20:5:75

Blood Type	A	B	AB	O
Expected (E)	50	20	5	75

## 3.2.1 Chi-squared test for goodness of fit

To calculate the  $\chi^2$  test statistic, use the formula below:

$$\text{Calculated } \chi^2 = \sum \frac{(O - E)^2}{E}$$

$$\begin{aligned} &= \frac{(37-50)^2}{50} + \frac{(28-20)^2}{20} + \frac{(2-5)^2}{5} + \frac{(83-75)^2}{75} \\ &= 3.38 + 3.2 + 1.8 + 0.853 \\ &= 9.233 \end{aligned}$$

This time we compare the calculated  $\chi^2$  value to a critical  $\chi^2$  value. Hence, a new statistics table.

Degrees of freedom =  $n - 1$  where  $n$  = number of categories (in this case, 4 blood types  $- 1 = 3$ )

If calculated  $\chi^2$  value > critical  $\chi^2$  value, then we reject null and  $p < 0.05$ .

## 3.2.1 Chi-squared test for goodness of fit

$df = 3$  at  $\alpha = 0.05$ , critical  $\chi^2$  value is 7.81.

Calculated  $\chi^2$  is 9.233 therefore we reject the null hypothesis and have a statistical difference between observed and expected frequencies (values) and  $p < 0.05$ .

We can confirm this using SPSS.

Let's go back to checking that one assumption...

Degrees of Freedom	Alpha level	
	0.05	0.01
1	3.84	6.63
2	5.99	9.21
3	7.81	11.34
4	9.49	13.28
5	11.07	15.09
6	12.59	16.81
7	14.07	18.48
8	15.51	20.09
9	16.92	21.67
10	18.31	23.21
11	19.68	24.72
12	21.03	26.22
13	22.36	27.69
14	23.68	29.14
15	25.00	30.58
16	26.30	32.00
17	27.59	33.41
18	28.87	34.81
19	30.14	36.19
20	31.41	37.57

## 3.2.1 Chi-squared test for goodness of fit

Chi-square test for goodness of fit in SPSS

One column for blood type and one column for observed frequencies.

Remember to use labels for categories as shown below:

Chi-square test for goodness of fit.sav [DataSet2] - IBM SPSS Statistics

	BloodType	Frequency
1	1.00	37.00
2	2.00	28.00
3	3.00	2.00
4	4.00	83.00

Chi-square test for goodness of fit.sav [DataSet2] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	BloodType	Numeric	8	2		{1.00, A}...	None	13	Right	Nominal	Input
2	Frequency	Numeric	8	2		None	None	14	Right	Scale	Input
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											

Value Labels

Value	Label
1.00	"A"
2.00	"B"
3.00	"AB"
4.00	"O"

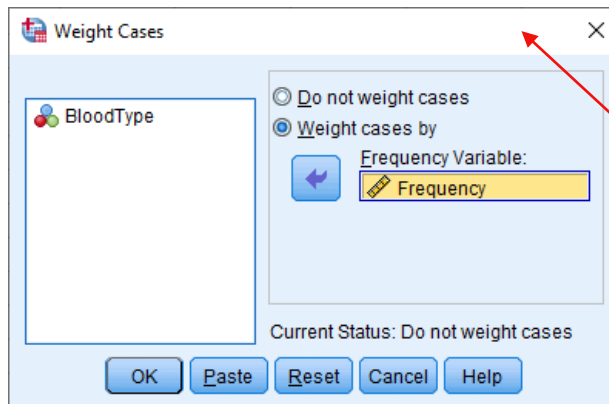
## 3.2.1 Chi-squared test for goodness of fit

Chi-square test for goodness of fit in SPSS

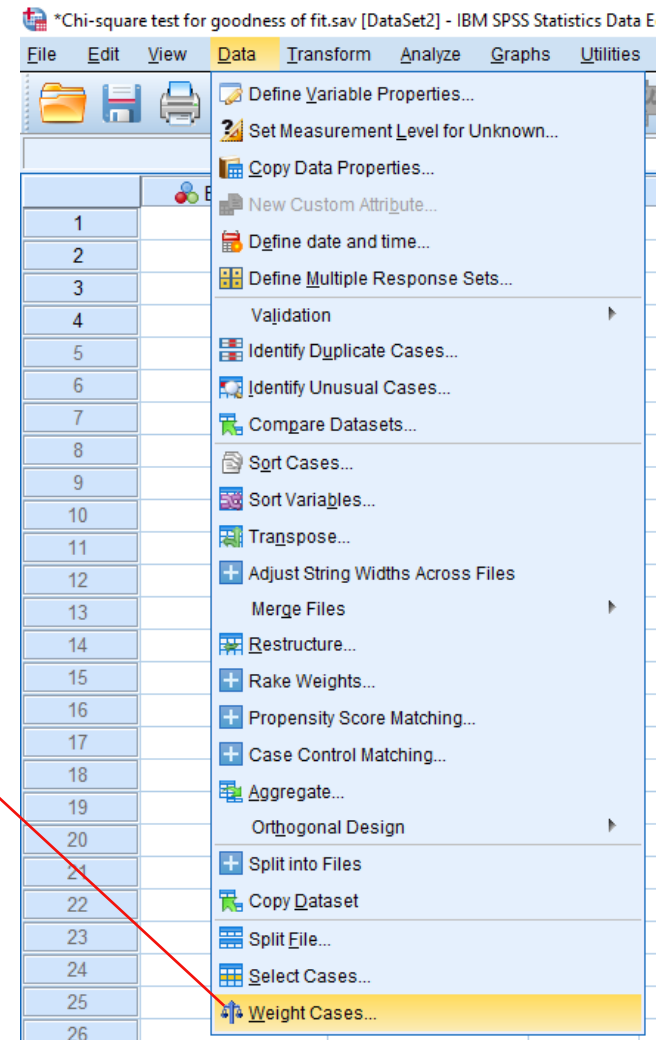
<Data>

<Weight Cases>

As we have uneven ratios in our observed frequencies, we need to weight cases by our frequency variable as shown below:



Click OK





## 3.2.1 Chi-squared test for goodness of fit

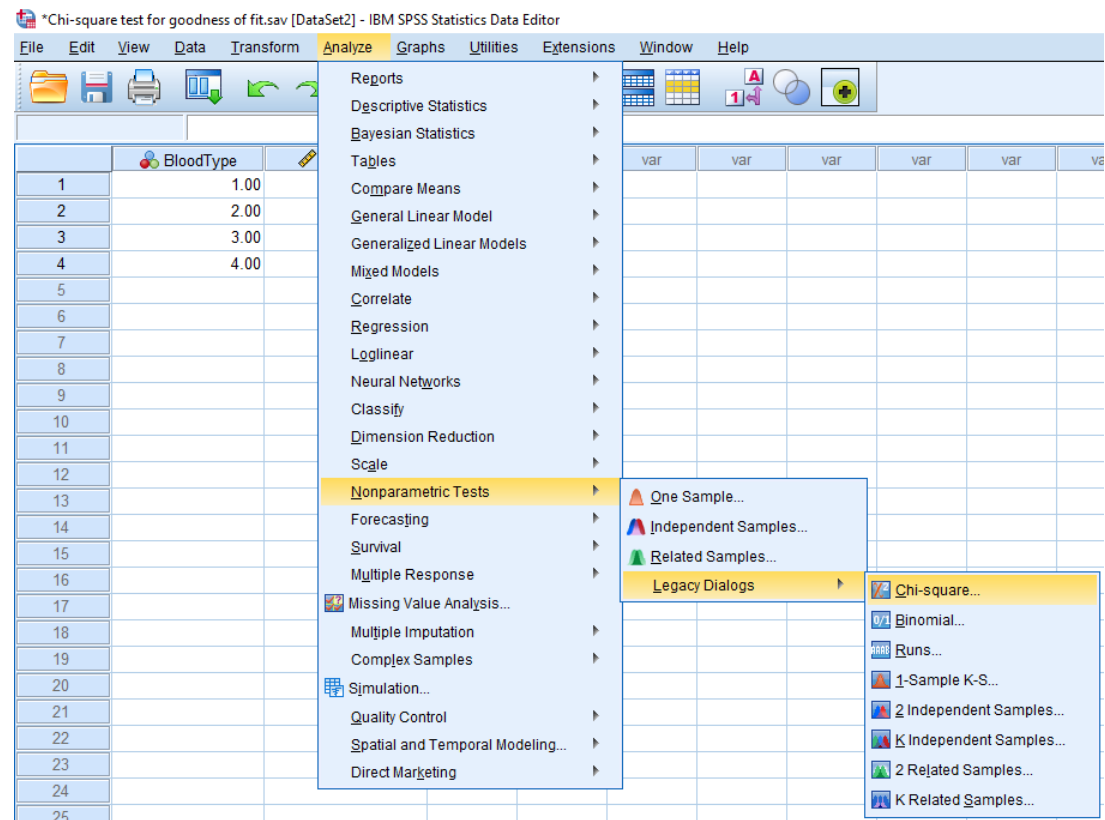
Chi-square test for goodness of fit in SPSS

<Analyze>

<Nonparametric Tests>

<Legacy Dialogs>

<Chi-square>

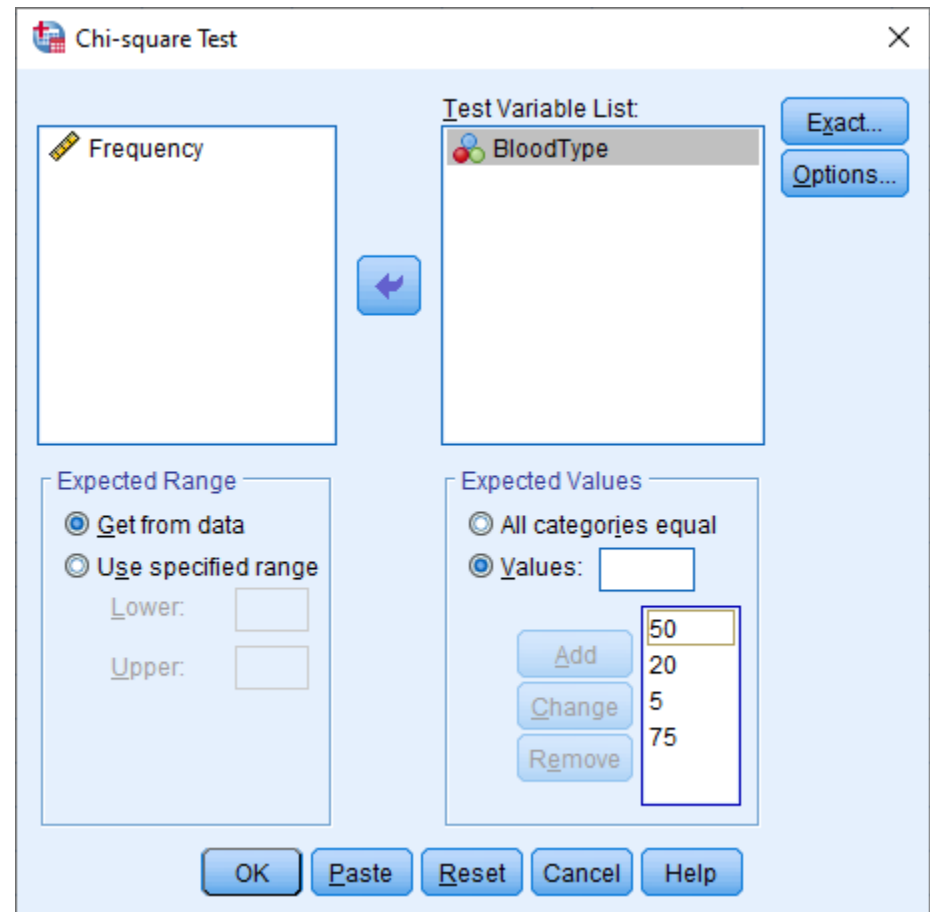


## 3.2.1 Chi-square test for goodness of fit

Chi-square test for goodness of fit in SPSS

Put BloodType in the Test Variable and enter in the expected values as shown →

For Options, we can just leave this as default.



## 3.2.1 Chi-square test for goodness of fit

### Chi-square test for goodness of fit in SPSS

The first table BloodType shows both observed and expected frequencies (values) as well as a residual column which is the difference between O and E.

BloodType			
	Observed N	Expected N	Residual
A	37	50.0	-13.0
B	28	20.0	8.0
AB	2	5.0	-3.0
O	83	75.0	8.0
Total	150		

The second table shows us the  $\chi^2$  test statistic, which is 9.233 (just as we calculated), the  $df = 3$  and  $p$ -value = 0.026, which is  $< 0.05$  and hence we reject the null and have statistical significance between the observed and expected values in sampling blood types in Melbourne CBD.

Test Statistics	
	BloodType
Chi-Square	9.233 <sup>a</sup>
df	3
Asymp. Sig.	.026

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 5.0.

The highlighted section is our chi-square test assumption, as it states 0 cells (0.0%) have expected frequencies less than 5 then the assumption is satisfied. If this was 20% or more then the assumption is not satisfied.

Finally, as always we report the results for the chi-square test for goodness of fit.

## 3.2.1 Chi-square test for goodness of fit

Reporting the results of a chi-square test for goodness of fit (APA format)

### Results

A chi-square test for goodness of fit was used to determine the commonality of blood types in the Melbourne CBD.

The chi-square test was statistically significant,  $\chi^2(3) = 9.233$ ,  $p = .03$ , indicating that some blood types were reported to have a greater frequency than others.

In our case, we see that blood type O was the most common.

You may also wish to include a summary table:

Table 1

Summary of blood type frequencies of a sample ( $N = 150$ ) of people in Melbourne CBD.

Blood Type	Percentage (%) of sample
A	24.67
B	18.67
AB	1.33
O	55.33
<b>Total</b>	<b>100</b>

## 3.2.2 Chi-square test of association

A chi-square test of association is used to determine if there is a relationship between two categorical (nominal) variables.

Chi-square of association test statistic is denoted by  $\chi^2$  and is also known as the Pearson's Chi-square.

Eg. Is there an association between sex and attitudes to smoking? Of the 15247 women in the sample, 4422 were smokers, 2592 were ex-smokers, and 8233 had never smoked regularly. Of the 14108 men in the sample, 4797 were smokers, 3950 were ex-smokers, and 5361 had never smoked.

Sex	Smoker	Ex-smoker	Never smoked
Male	4797	3950	5361
Female	4422	2592	8233

(Clarke, 1993)

## 3.2.2 Chi-square test of association

Note what we see here are the observed values.

Sex	Smoker	Ex-smoker	Never smoked
Male	4797	3950	5361
Female	4422	2592	8233

To calculate the expected values, we need to create a new contingency table:

$$Expected\ value = \frac{column\ total \times row\ total}{grand\ total}$$

But first, let's establish the hypotheses:

$H_0$ : there is no association between sex and smoking status

$H_1$ : there is an association between sex and smoking status

## 3.2.2 Chi-square test of association

Adding totals to the table...we can create the new contingency table of expected values.

Observed values	Sex	Smoker	Ex-smoker	Never smoked	TOTAL
	Male	4797	3950	5361	14108
	Female	4422	2592	8233	15247
	<b>TOTAL</b>	<b>9219</b>	<b>6542</b>	<b>13594</b>	<b>29355</b>

$$\text{Expected value} = \frac{\text{column total} \times \text{row total}}{\text{grand total}}$$

Expected values	Sex	Smoker	Ex-smoker	Never smoked	TOTAL
	Male	4430.65	3144.08	6533.27	14108
	Female	4788.35	3397.92	7060.73	15247
	<b>TOTAL</b>	<b>9219</b>	<b>6542</b>	<b>13594</b>	<b>29355</b>

## 3.2.2 Chi-square test of association

Calculating the chi-square of association test statistic:

$$\begin{aligned}\text{Calculated } \chi^2 &= \sum \frac{(O - E)^2}{E} \\&= \frac{(4797 - 4430.65)^2}{4430.65} + \frac{(3950 - 3144.08)^2}{3144.08} + \frac{(5361 - 6533.27)^2}{6533.27} + \\&\quad \frac{(4422 - 4788.35)^2}{4788.35} + \frac{(2592 - 3397.92)^2}{3397.92} + \frac{(8233 - 7060.73)^2}{7060.73} = \mathbf{861.02}\end{aligned}$$



## 3.2.2 Chi-square test of association

Calculating the chi-square of association test statistic:

Degrees of freedom for a chi-square test is:

$$df = (\text{number of rows} - 1) \times (\text{number of columns} - 1) = 2$$

Using the critical chi-square table we get a value of 5.99 at  $\alpha = 0.05$ .

Calculated  $\chi^2 >$  Critical  $\chi^2$ , therefore we reject the null and can say there is an association between sex and smoking status and  $p < 0.05$ .

Degrees of Freedom	Alpha level	
	0.05	0.01
1	3.84	6.63
2	5.99	9.21
3	7.81	11.34
4	9.49	13.28
5	11.07	15.09
6	12.59	16.81
7	14.07	18.48
8	15.51	20.09
9	16.92	21.67
10	18.31	23.21
11	19.68	24.72
12	21.03	26.22
13	22.36	27.69
14	23.68	29.14
15	25.00	30.58
16	26.30	32.00
17	27.59	33.41
18	28.87	34.81
19	30.14	36.19
20	31.41	37.57

## 3.2.2 Chi-square test of association

### Chi-square test of association in SPSS

Setup the data like so...

\*Chi-square test of association.sav [DataSet3] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Sex	Numeric	8	2		{1.00, Male}...	None	11	Right	Nominal	Input
2	SmokingStatus	Numeric	8	2		{1.00, Smok...	None	16	Right	Nominal	Input
3	Frequency	Numeric	8	2		None	None	14	Right	Scale	Input

\*Chi-square test of association.sav [DataSet3] - IBM SPSS Statistics Data Editor

	Sex	SmokingStatus	Frequency
1	1.00	1.00	4797.00
2	1.00	2.00	3950.00
3	1.00	3.00	5361.00
4	2.00	1.00	4422.00
5	2.00	2.00	2592.00
6	2.00	3.00	8233.00

Data View

Variable View

Value Labels

Value:

Label:

1.00 = "Smoker"  
2.00 = "Ex-smoker"  
3.00 = "Never smoked"

Add Change Remove

OK Cancel Help

Value Labels

Value:

Label:

1.00 = "Male"  
2.00 = "Female"

Add Change Remove

OK Cancel Help

## 3.2.2 Chi-square test of association

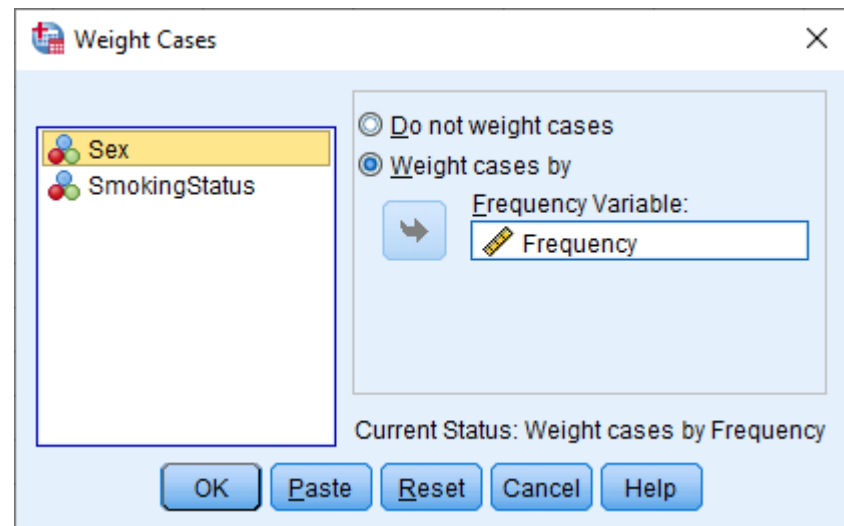
Chi-square test of association in SPSS

<Data>

<Weight Cases>

Input Frequency under the Frequency Variable.

Click OK



## 3.2.2 Chi-square test of association

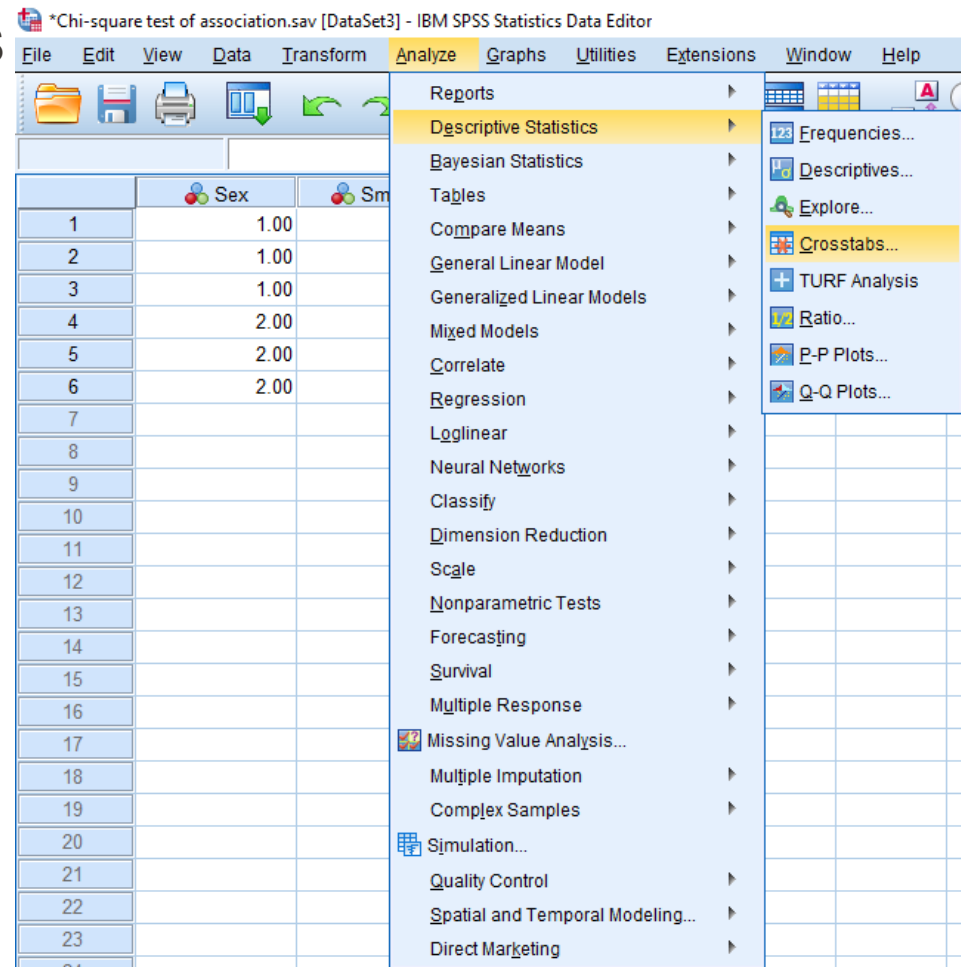
Chi-square test of association in SPSS

<Analyze>

<Descriptive Statistics>

<Crosstabs>

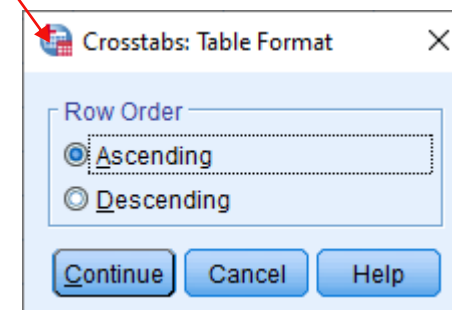
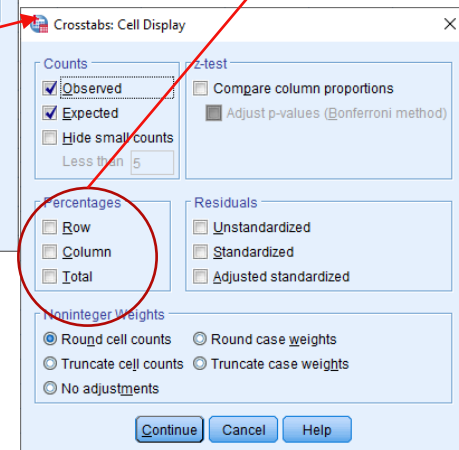
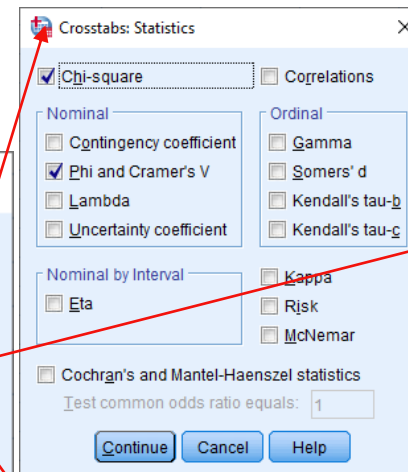
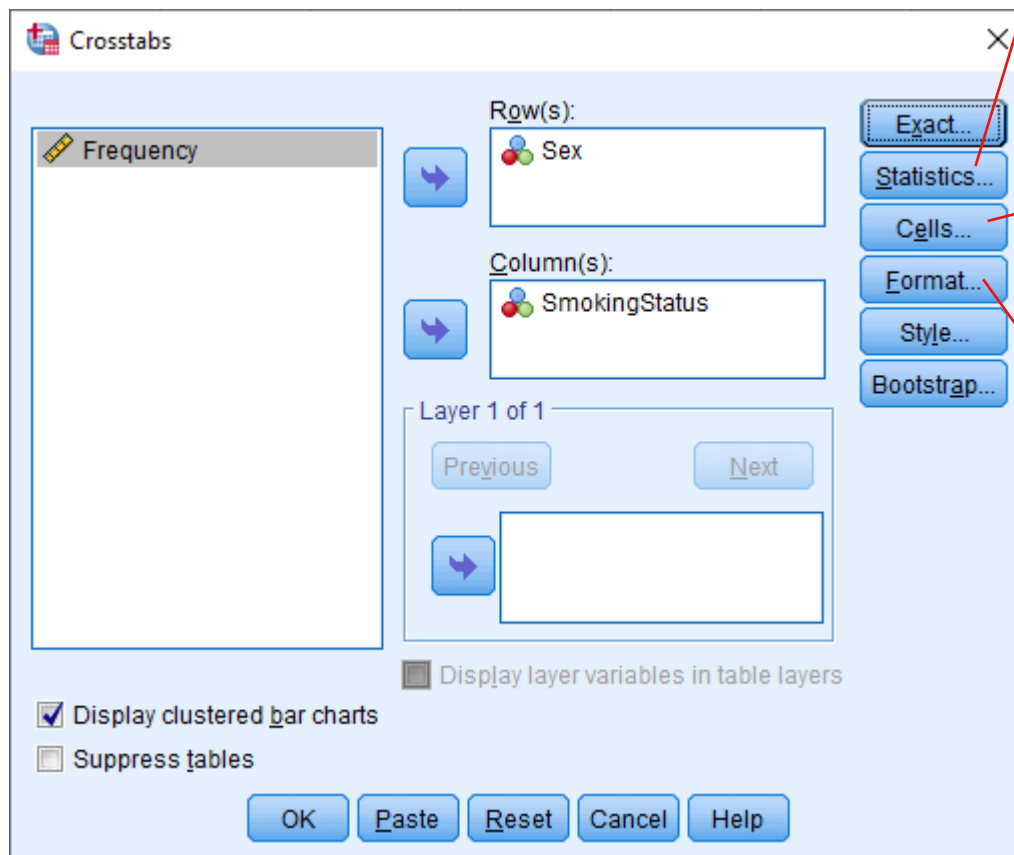
A bit different to how we've usually done a hypothesis test!



## 3.2.2 Chi-square test of association

### Chi-square test of association in SPSS

You may choose to use percentages (%).



## 3.2.2 Chi-square test of association

Chi-square test of association in SPSS

SPSS will give you the observed and expected values in a contingency table.

**Sex \* SmokingStatus Crosstabulation**

				SmokingStatus			
				Smoker	Ex-smoker	Never smoked	Total
Sex	Male	Count	4797	3950	5361	14108	
		Expected Count	4430.6	3144.1	6533.3	14108.0	
		% within Sex	34.0%	28.0%	38.0%	100.0%	
		% within SmokingStatus	52.0%	60.4%	39.4%	48.1%	
		% of Total	16.3%	13.5%	18.3%	48.1%	
	Female	Count	4422	2592	8233	15247	
		Expected Count	4788.4	3397.9	7060.7	15247.0	
		% within Sex	29.0%	17.0%	54.0%	100.0%	
		% within SmokingStatus	48.0%	39.6%	60.6%	51.9%	
		% of Total	15.1%	8.8%	28.0%	51.9%	
Total	Count	9219	6542	13594	29355		
	Expected Count	9219.0	6542.0	13594.0	29355.0		
	% within Sex	31.4%	22.3%	46.3%	100.0%		
	% within SmokingStatus	100.0%	100.0%	100.0%	100.0%		
	% of Total	31.4%	22.3%	46.3%	100.0%		

## 3.2.2 Chi-square test of association

Chi-square test of association in SPSS

As well as the Chi-square test statistic:

**Chi-Square Tests**

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	861.018 <sup>a</sup>	2	.000
Likelihood Ratio	866.372	2	.000
Linear-by-Linear Association	427.933	1	.000
N of Valid Cases	29355		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 3144.08.

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	.171	.000
	Cramer's V	.171	.000
N of Valid Cases		29355	

The Pearson Chi-square matches what we calculated with  $df = 2$  and  $p < 0.001$ . We also see under the chi-square test table that our assumption has been satisfied.

Phi ( $\phi$ ) and Cramer's V ( $V$ ) give us an effect size (a measure of association for nominal data). A Phi or Cramer's V of 0.1 is considered to be small, 0.3 medium and 0.5 large.

## 3.2.2 Chi-square test of association

Interpreting the results of a chi-square test of association (APA format)

### Results

A Pearson's chi-square test of association was used to evaluate whether there was an association between sex and smoking status.

The chi-square test was statistically significant,  $\chi^2(2) = 861.02$ ,  $p < .001$  with a Phi and Cramer's V of 0.17, which indicates a small association.

Figure 1 shows that there were significantly more females who had never smoked.

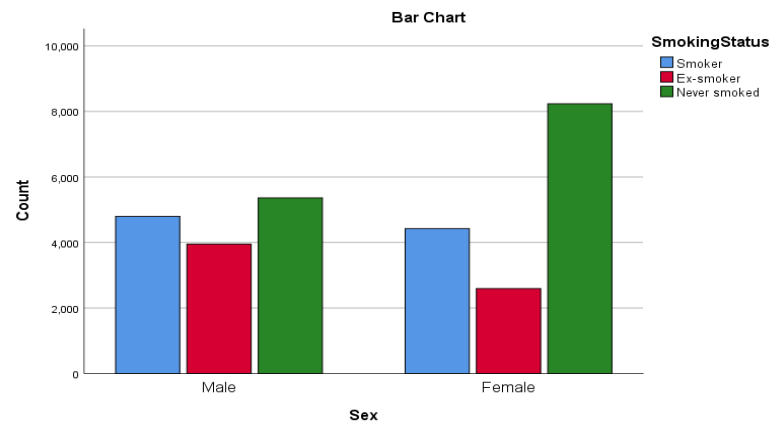


Figure 1: Clustered bar chart illustrating the frequency of males and females and their smoking statuses.



### 3.2.3 Chi-square test of association with Yate's correction method

Chi-square test with Yate's correction method

A chi-square test of association with Yate's correction is only used for a 2 x 2 contingency table ie.  $df = 1$ .

A 2 x 2 table results in the calculated  $\chi^2$  value being too high.

To correct this the general  $\chi^2$  equation is slightly modified:

$$\chi^2 = \sum \frac{(|O - E| - 0.5)^2}{E}$$

**Absolute value**

Eg. incidence of a disease while on "Treatment A"

### 3.2.3 Chi-square test of association with Yate's correction method

Categories: diseased or not diseased; administered Treatment A or not administered Treatment A - 2x2 table

Disease	Treatment A administered	Treatment A not administered	Row totals
Developed	15	22	37
Did not develop	100	93	193
Column totals	115	115	230

We can get the expected values quite easily using SPSS...

### 3.2.3 Chi-square test of association with Yate's correction method

**DiseasedStatus \* TreatmentA Crosstabulation**

			TreatmentA		
			Treatment A administered	Treatment A not administered	Total
DiseasedStatus	Developed	Count	15	22	37
		Expected Count	18.5	18.5	37.0
	Did not develop	Count	100	93	193
		Expected Count	96.5	96.5	193.0
Total		Count	115	115	230
		Expected Count	115.0	115.0	230.0

Expected counts are our expected values.

### 3.2.3 Chi-square test of association with Yate's correction method

$$\chi^2 = \sum \frac{(|O_i - E_i| - 0.5)^2}{E_i}$$

$$\chi^2 = \frac{(|O_1 - E_1| - 0.5)^2}{E_1} + \frac{(|O_2 - E_2| - 0.5)^2}{E_2} + \dots + \frac{(|O_n - E_n| - 0.5)^2}{E_n}$$

$$\begin{aligned} \chi^2 &= \frac{(|15 - 18.5| - 0.5)^2}{18.5} + \frac{(|100 - 96.5| - 0.5)^2}{96.5} + \frac{(|22 - 18.5| - 0.5)^2}{18.5} \\ &\quad + \frac{(|93 - 96.5| - 0.5)^2}{96.5} \end{aligned}$$

$$\chi^2 = 0.486 + 0.093 + 0.486 + 0.093$$

$$\chi^2 = 1.16$$

- Null hypothesis: that proportions in two groups are not different
- Significance: 0.05; d.f. = (number of columns - 1) x (number of rows - 1) = (2 - 1) X (2 - 1) = 1
- Critical value = 3.841
- Observed value less than critical value; null hypothesis accepted
- No difference in proportion of people with the disease across treatment groups

### 3.2.3 Chi-square test of association with Yate's correction method

When using Yate's correction method, we look at the Continuity Correction value in SPSS for the chi-square test statistic.

**Chi-Square Tests**

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	1.578 <sup>a</sup>	1	.209		
Continuity Correction <sup>b</sup>	1.160	1	.282		
Likelihood Ratio	1.586	1	.208		
Fisher's Exact Test				.281	.141
Linear-by-Linear Association	1.571	1	.210		
N of Valid Cases	230				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 18.50.

b. Computed only for a 2x2 table

Here, the continuity correction value is  $\chi^2 = 1.16$ , which is what we calculated!

### 3.2.3 Chi-square test of association with Yate's correction method

When reporting Yate's correction method, you simply say that you used a Pearson's chi-square test with Yate's correction and you quote the continuity correction value you obtained in SPSS.

#### Results

A Pearson's chi-square test with Yate's correction was used to evaluate incidence of disease while on Treatment A. The chi-square test was not statistically significant,  $\chi^2(1) = 1.16$ ,  $p = .282$ .

You may wish to include the Phi and Cramer's V and bar chart here too.

### 3.2.3 Chi-square test of association with Yate's correction method

When reporting Yate's correction method, you simply say that you used a Pearson's chi-square test with Yate's correction and you quote the continuity correction value you obtained in SPSS.

#### Results

A Pearson's chi-square test with Yate's correction was used to evaluate incidence of disease while on Treatment A. The chi-square test was not statistically significant,  $\chi^2(1) = 1.16$ ,  $p = .28$ .

You may wish to include the Phi and Cramer's V and bar chart here too.

## 3.3 APA statistical reporting

Important rules to follow:

1. **Italics** – All statistical parameters should be in italics HOWEVER, Greek letters should not be in italics e.g.  $M = 3.4$ ,  $SD = 1.1$  but  $\beta = 0.8$
2. **Decimal places** – When reporting test statistics, this is typically done to two decimal places. E.g.  $t(23) = 2.34$  as opposed to  $t(23) = 2.337$ . Note however that some p-values can have three decimal places e.g.  $p < .001$
3. **Leading zeroes** – If the value can be greater than 1 or smaller than -1 then include a zero before the decimal place. E.g.  $d = 1.08$  or  $0.72$  (Cohen's  $d$  can have values greater than 1). If the value is within -1 and 1 do not add a zero before the decimal e.g.  $p = .036$
4. **P-values** – p-values must be expressed exactly unless  $p < 0.001$ . e.g.  $p = .04$  as opposed to  $p < .05$ . P-values between .001 and .01 can be expressed to three decimal places. P-values larger than .01 should be expressed to two decimal places.
5. **Parenthesis** – round brackets can be used for degrees of freedom e.g.  $t(39)$ . Square brackets can be used for confidence intervals e.g. 95% CI [-1.3, 3.2]
6. **Capitalisations** – Typically “ $n$ ” should be used for sample size and not “ $N$ ”. Means and Standard Deviations should use capitalised letters i.e. “ $M$ ” and “ $SD$ ” respectively.
7. **Spacing** – Do not have unwanted spaces e.g.  $t(32)$  as opposed to  $t (32)$ .