

**DIGITAL SKILLS****INTRODUCTION TO SPSS PART 2**

Last updated 13 SEPTEMBER 2021

This document takes you through some more of the procedures available in SPSS, with a focus on inferential statistics. Topics covered include testing for normality; confidence intervals; one, paired and independent samples t tests; one-way ANOVA; Pearson's correlation coefficient and the Chi-square test. The document uses the *Health Data.sav* file, which contains fictitious data for 80 people for the following questionnaire:

HEALTH QUESTIONNAIRE

Please complete this questionnaire by circling the number next to the response you want to give or by writing your answer on the line provided. Thank you for your co-operation in providing the information.

BACKGROUND INFORMATION

Q1	What gender do you identify as? (Circle one number only)	Male Female Other	1 2 3
Q2	What is your marital status? (Circle one number only)	Single Married Other	1 2 3
Q3	What is your age in years? (Write the number on the line)		_____
Q4	Do you have any children? (Circle one number only)	Yes No	1 2 (go to Q6)
Q5	If you answered "YES" to Q4 , how many children do you have? (Circle one number only)	One Two Three Four or more	1 2 3 4

WEIGHT

Q6	Did you participate in the diet plan? (Circle one number only)	Yes No	1 2
Q7	What was your weight in kilograms before going on the diet? (Write the numbers on the line, e.g. 199.2)		_____.
Q8	What was your weight in kilograms at the completion of the diet plan? (Write the numbers on the line, e.g. 92.2)		_____.
Q9	What type of exercise do you do?	Minimal exercise Low level exercise Medium level exercise Strenuous exercise	1 2 3 4



LIFE SATISFACTION AND HEALTH

Q10	How satisfied are you with your work (if employed) or home duties? (Circle one number only)	Very unsatisfied Unsatisfied Slightly satisfied Moderately satisfied Very satisfied	1 2 3 4 5
Q11	How satisfied are you with your personal life? (Circle one number only)	Very unsatisfied Unsatisfied Slightly satisfied Moderately satisfied Very satisfied	1 2 3 4 5
Q12	How would you currently describe your overall state of health and well being? (Circle one number only)	Ill Not well OK Well Very well	1 2 3 4 5
Q13	When you have personal problems or are very worried, how much understanding and help do you get from people you know? (Circle one number only)	Very little Some A fair amount A lot A great deal	1 2 3 4 5
Q14	Who is the main person in your family that you would turn to for help with a problem? Please specify their relationship to you, e.g. mother, uncle (Write your answer on the line)	_____	

EFFECTS OF THE SUN

Q15	Have you experienced any adverse health effects from the sun?	Yes No Unsure	1 2 3
	If you answered 'YES' to question 14, please list the three main adverse effects	_____ _____ _____	
Q16	Is it possible to get sun burnt:		
	On a cloudy day?	Yes No	1 2
	Through clothing?	Yes No	1 2
	Through closed car windows?	Yes No	1 2
	In the shade?	Yes No	1 2
	While swimming?	Yes No	1 2

This session works through some of the various ways you can use inferential statistics to draw conclusions about the population from which the sample of 80 questionnaire participants was drawn.



Before we begin, note that to view the variable names (i.e. **q7**) rather than labels (i.e. **Pre diet weight**) in the list of variables (as variable names are used in the examples), from the main menu select:

Edit

Options

Then use the toggle to switch to the **Display Names** option in the **Variable Lists** section of the Options window that appears.

1. Testing for normality

Before choosing an inferential statistical test to analyse your data, it is necessary to test the variable in question for normality. The outcome will determine whether you will be able to use a parametric test, or if you will need to use a non-parametric test (the latter are not covered in any detail here). Recall that it is helpful to consult a number of different measures in order to make a decision about normality - the Explore procedure will allow you to do this.

Example question – *Is the pre diet weight variable normally distributed?*

To answer the question above, from the menus choose:

Analyse

Descriptive Statistics

Explore

move the variable (**q7**) that you are checking for normality into the **Dependent List**

click on the **Plots** button

tick the **Histogram** option in the Descriptive section (you can also keep **Stem-and leaf** ticked)

tick the **Normality plots with tests** option

click on **Continue**

click on **OK**

The output should then be as shown on the following pages:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Pre diet weight	44	55.0%	36	45.0%	80	100.0%

Descriptives

		Statistic	Std. Error
Pre diet weight	Mean	64.523	1.7986
	95% Confidence Interval for		
	Lower Bound	60.895	
	Upper Bound	68.150	
	5% Trimmed Mean	64.303	
	Median	63.000	
	Variance	142.345	
	Std. Deviation	11.9308	
	Minimum	42.0	



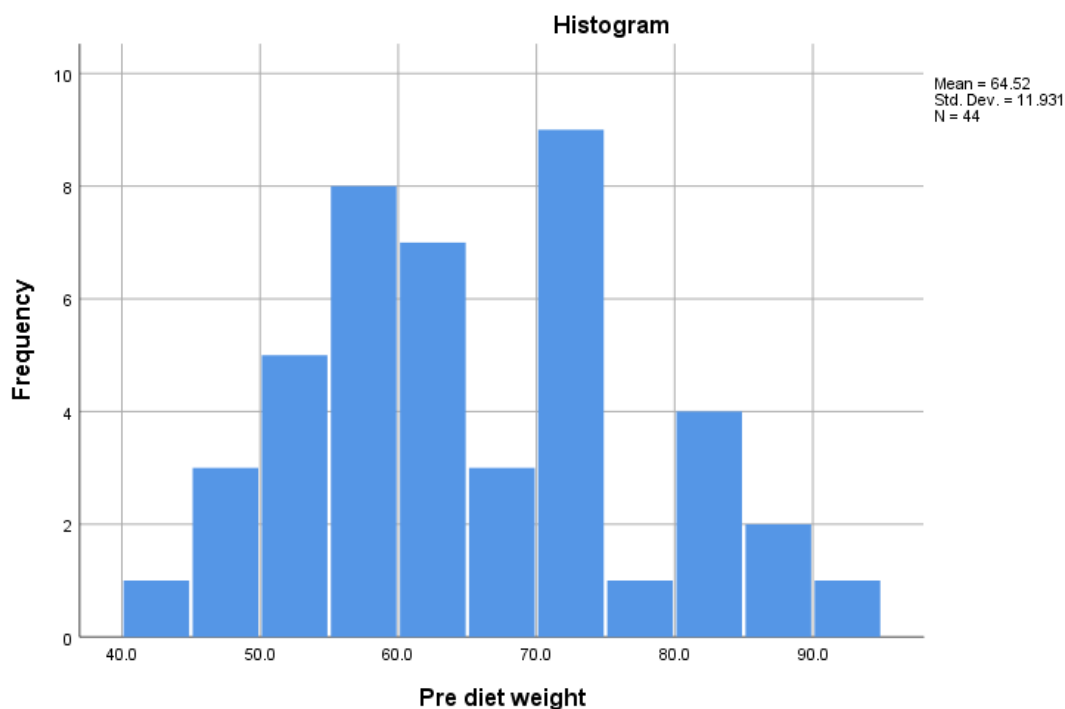
Maximum	92.0	
Range	50.0	
Interquartile Range	16.5	
Skewness	.333	.357
Kurtosis	-.573	.702

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre diet weight	.102	44	.200*	.973	44	.398

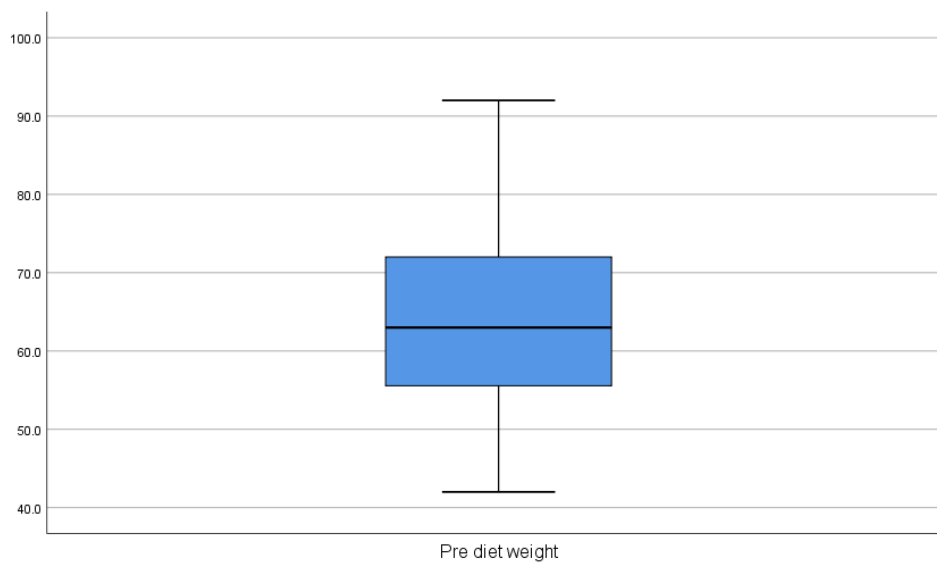
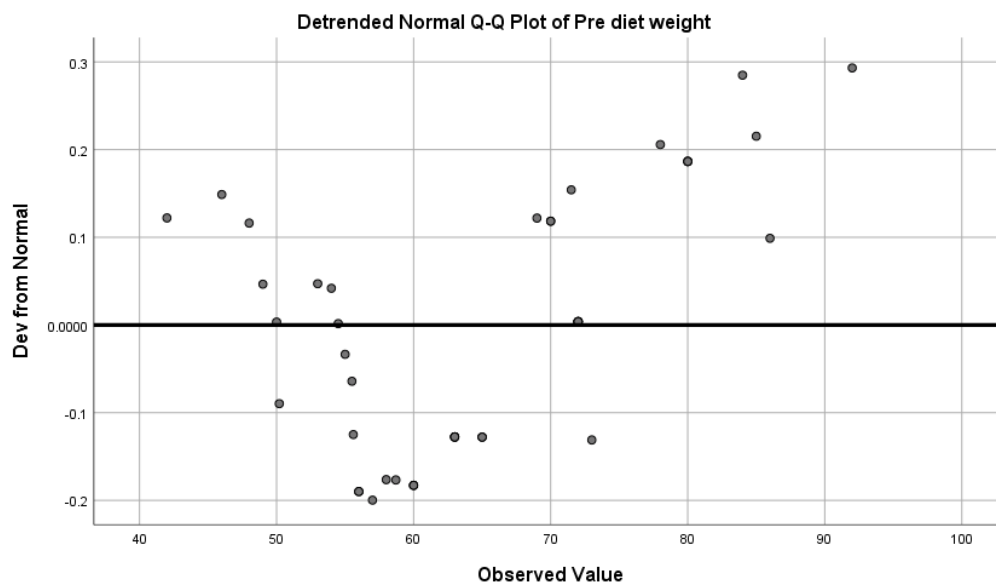
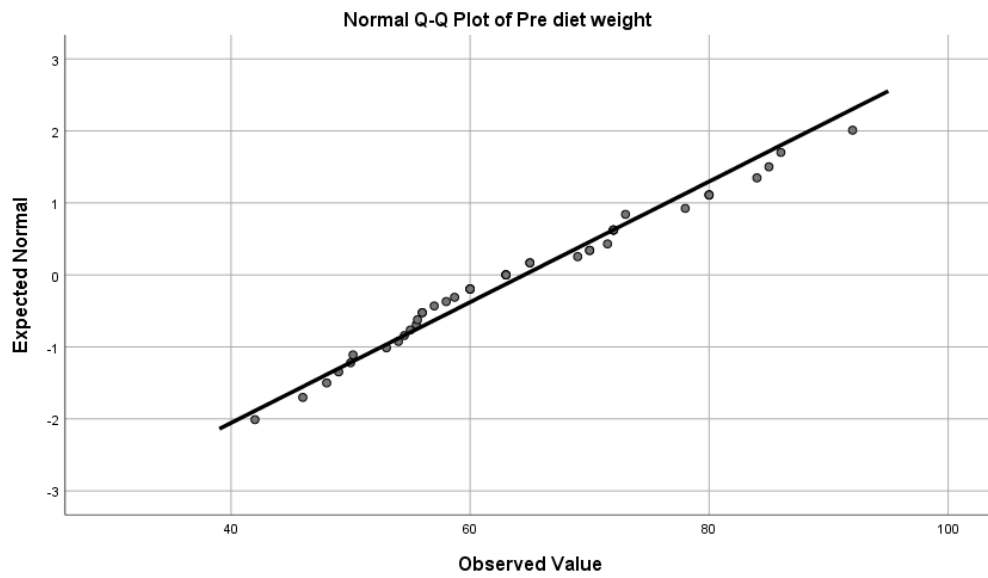
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Pre diet weight Stem-and-Leaf Plot

Frequency	Stem & Leaf
4.00	4 . 2689
13.00	5 . 0034455566788
10.00	6 . 0003333559
10.00	7 . 0012222238
6.00	8 . 000456
1.00	9 . 2
Stem width:	10.0
Each leaf:	1 case(s)



You then need to analyse this output as described in the following sections, to come to a conclusion regarding the normality (or otherwise) of the variable. This can be done in any particular order, but is detailed in the order of the output here:

1.1 Mean, median and mode

For a perfect normal distribution these three values should all be the same, so checking whether they are similar is a good (and simple) way to start. The mean and median are included in the Descriptives table above, and checking these is usually sufficient, however the mode can also be obtained through use of the **Frequencies** procedure (as detailed in the previous session) if required:

Statistics

Pre diet weight		
N	Valid	44
	Missing	36
Mean		64.523
Median		63.000
Mode		72.0

Note that in this particular example the mean and median are very similar, and while the mode is a bit higher this is less of a concern.

1.2 Skewness

Skewness indicates the direction of any tail of the histogram. If the data is normally distributed the skewness should be close to 0 (0 indicates perfect normal distribution), but at least in the range of -1 to 1 (negative values indicate negative skew with the tail to the left, i.e. skewed to the left; positive values indicate positive skew with the tail to the right, i.e. skewed to the right). A z-score can also be calculated for skewness by dividing the skewness by its standard error, and this should be within the range of -1.96 to 1.96.

Note that in this particular example the skewness is given as 0.333, meaning that it is inside of the required range and hence passes this first test for normality. The corresponding z-score is $0.333 \div 0.357 = 0.933$, which again is inside of the acceptable range for normality.

1.3 Kurtosis

If the data is normally distributed the kurtosis should be close to 0, but at least in the range of -1 to 1 (positive kurtosis indicates a high peak around the mean and fatter tails; negative kurtosis indicates a lower peak around the mean and thinner tails). A z-score can also be calculated for kurtosis by dividing the kurtosis by its standard error, and this should be within the range of -1.96 to 1.96.

Note that in this particular example the kurtosis is given as -0.573, meaning that it is inside of the required range and hence passes this second test for normality. The corresponding z-score is $-0.573 \div 0.702 = -0.816$, which again is inside of the acceptable range for normality.

1.4 Shapiro-Wilk and Kolmogorov-Smirnov tests

The Shapiro-Wilk (S-W) and Kolmogorov-Smirnov (K-S) tests check whether the distribution of scores is significantly different from a normal distribution; both tests are only recommended for sample sizes less than 100 though as they are too sensitive for larger samples. The tests both test the null hypothesis that the data is normally distributed (more on hypotheses later); if the significance value (p value) is greater than .05 we can assume that the null hypothesis is true and assume normal distribution but if the significance value (p value) is less than .05 we must reject the null hypothesis and cannot assume normal distribution.

Note that in this particular example we could use either test since we have a sample size of 80 (some people prefer one test over the other, and it is up to you which to interpret- or you can look at both). The p value ('Sig.') for these tests are given as .200 and .398 respectively, both of which are greater than .05 and hence we can assume normal distribution based on these tests.



1.5 Histogram

If the data is normally distributed the histogram should be symmetric and centred around its mean value. A histogram is not symmetric if there is a long tail to one side which is not reflected on the other side. A skewed distribution can be either positively / right skewed (long tail to the right) or negatively / left skewed (long tail to the left).

Note that, while far from perfect, the histogram in this example is reasonably symmetric about the mean (at least, it does not appear to have a long 'tail' on either side); so the variable passes this test of normality.

1.6 Stem and leaf plot

This plot is less frequently analysed but if you do choose to include it, it should be interpreted in the same way as a histogram only rotated on its side.

In this case the numbers in the plot do form a roughly symmetrical bell-shaped curve when the plot is viewed on its side, again indicating that the variable passes this test of normality.

1.7 Normal probability plots

If the data is normally distributed the data points will lie close to the straight diagonal line on the Normal Q-Q Plot, and there should be roughly equal number of points above and below the horizontal line (and no obvious trend) for the Detrended Normal Q-Q Plot.

Note that in this example the points on the Normal Q-Q Plot follow the diagonal line closely, and there are roughly equal numbers of points above and below the line and no obvious trend for the Detrended Normal Q-Q Plot; so the variable passes this test for normality.

1.8 Boxplot

The line at the bottom of the boxplot marks the minimum value for the variable and the line at the top marks the maximum value (excluding any outliers; these are marked separately), while the bottom of the box marks the 25th percentile and the top marks the 75th percentile. The thicker line positioned in the box is the median or 50th percentile. If the data is normally distributed the median should be positioned approximately in the centre of the box, and both left and right (or top and bottom) whiskers should have similar length. Ideally there should also be no outliers.

Note that there are no outliers in this example, and the boxplot is reasonably symmetrical, so the variable passes this final test for normality.

1.9 Conclusion

After analysing the data relating to these tests for normality, you should come to an overall conclusion based on what the majority of the tests indicate. In this case, all tests say the same thing- that the variable is normally distributed.

If the variable is not normally distributed, however, you can try transforming the variable to see if you can get it to conform more to a normal distribution. To transform a skewed continuous variable, you can apply one or all of the following transformations using the **Compute Variable** procedure described in the previous session :

- Natural logarithms (i.e. \ln) - to correct a positively skewed continuous variable (most commonly used)
- Square root - to correct a positively skewed continuous variable
- Reciprocal - to correct a positively skewed continuous variable
- Squares - to correct a negatively skewed continuous variable

Once the data has been transformed, it should be tested again for normality to check that the transformation has 'worked'. If it has, then any parametric inferential statistical tests you do should be conducted on the transformed data.



1.10 Testing for normality in two (or more) different groups

Sometimes rather than testing that all of a continuous variable's data is normally distributed, you need to check that the continuous variable's data is normally distributed for each category of a categorical variable.

Example question – *Is the pre diet weight variable normally distributed for both groups of the children variable? (i.e. is the pre diet weight normally distributed both for the group with children, and those without?)*

To do this using the Explore procedure, from the menus choose:

Analyse

Descriptive Statistics

Explore

move the variable (**q7**) that you are checking for normality into the **Dependent List**

move the grouping variable (**q4**) into the **Factor List**

To test for normality:

click on the **Plots** button

tick the **Histogram** option in the Descriptive section (you can also keep **Stem-and leaf** ticked)

tick the **Normality plots with tests** option

click on **Continue**

click on **OK**

Given that the majority of indicators suggest that the data is normally distributed for both groups, it is appropriate to go on and use a parametric test to compare the means as we will do later.

2. Determining a confidence interval

One type of inferential statistics is estimation, which involves estimating population parameters using sample statistics. A key example of this is a confidence interval, which is a range of probable values for an unknown population parameter based on the sample statistic; e.g. the mean. The percentage associated with the confidence interval (i.e. 95%) is termed the confidence coefficient; it is the level of confidence you have that the range actually includes the true value.

Example question – *Obtain and interpret the 95% confidence interval for the population mean weight at the end of the diet.*

The assumptions for a confidence interval (which we will assume to be valid) are that:

- The sample is a random sample that is representative of the population.
- The observations are independent, i.e. measurements for one subject have no bearing on any other subject's measurements.
- The variable is normally distributed, or the sample size is large enough to ensure normality of the sampling distribution.

To obtain a 95% confidence interval in SPSS, from the menus choose:

Analyse

Descriptive Statistics

Explore

move the variable (**q8**) into the **Dependent List** (remove other variables if required)

select **Statistics** in the **Display** section underneath

click on **OK**

The output should then include the following:

Descriptives

		Statistic	Std. Error	
Post diet weight	Mean	60.407	1.5530	
	95% Confidence Interval for Mean	Lower Bound	57.275	
		Upper Bound	63.539	
	5% Trimmed Mean	60.136		
	Median	59.000		
	Variance	106.124		
	Std. Deviation	10.3017		
	Minimum	40.1		
	Maximum	86.0		
	Range	45.9		
	Interquartile Range	16.5		
	Skewness	.490	.357	
	Kurtosis	-.249	.702	

This tells us that we can be 95% confident that the mean weight of the population after the diet lies between 57.275kg (lower bound) and 63.539kg (upper bound).

Note that to calculate a confidence interval with a confidence coefficient other than 95%, follow the previous instructions but also click on the **Statistics** button and change the **Confidence Interval for Mean** as required.

3. Hypothesis testing

The other type of inferential statistics is hypothesis testing, which involves testing hypothesised statements about the population using data collected in the sample. The particular test to use depends on the nature of the hypotheses, and as previously mentioned there are parametric (assume normal distribution and require at least one continuous variable) and non-parametric (don't assume normal distribution and can be used for ordinal variables) versions of each test. This section will cover some of the most frequently used tests, which are used in the following situations:

Situation	Parametric Test	Non-Parametric Version
You want to test whether there is a significant difference between a sample mean and a specific value.	One Sample t Test	
You want to test whether there is a significant difference between sample means for two related groups	Paired Samples t Test	Wilcoxon Signed Rank Test (compares mean ranks)
You want to test whether there is a significant difference between sample means for two independent groups	Independent Sample t Test	Mann-Whitney U Test (compares mean ranks)
You want to test whether there is a significant difference between sample means for more than two independent groups	One-way ANOVA	Kruskal-Wallis One-way ANOVA (compares mean ranks)
You want to test whether there is linear correlation between two continuous variables	Pearson's Correlation Coefficient	Spearman's Rho or Kendall's Tau-B (can test for non-linear correlation)
You want to test whether there is association between two categorical variables		Chi-Square Test

Whichever test you are using, you should follow these five steps to test a hypothesis using SPSS:

- 1) Write the hypotheses



- 2) Choose the appropriate test
 - 3) Check that all assumptions for the test are valid
 - 4) Use SPSS to conduct the relevant test and determine the p value and confidence interval (if applicable)
 - 5) Interpret the result
- An example of how to do this for each test is covered in this section.

3.1 Conducting a one-sample t-test

Example question – *Is there a statistically significant difference between the mean pre diet weight of females in the sample, and the mean weight of the population of Australian females (71.1kg)?*

- 1) Appropriate hypotheses for this question are:
 H_0 : The sample of females comes from a population with a mean pre-diet weight of 71.1kg
 H_A : The sample of females does not come from a population with a mean pre-diet weight of 71.1kg
- 2) Appropriate test is a one-sample t-test, as we are testing whether the sample comes from a population with a specific mean (71.1kg in this case).
- 3) The assumptions for a one-sample t-test (which we will assume to be valid) are that:
 - The sample is a random sample that is representative of the population.
 - The observations are independent, i.e. measurements for one subject have no bearing on any other subject's measurements.
 - The data comes from a normal distribution or the sample size is large enough to ensure normality of the sampling distribution.
- 4) To conduct this one-sample t-test in SPSS we first need to select only the female cases, using the method described in the previous session. Once this is done, we can run the one-sample t-test by choosing the following from the menus:

Analyse

Compare Means

One-Sample T Test

move the variable (**q7**) into the **Test Variable(s)** box (remove other variables if required)

enter the **Test Value** of 71.1

click on **OK**

The output should look like this:

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
Pre diet weight	38	62.184	10.7446	1.7430

One-Sample Test						
Test Value = 71.1						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Pre diet weight	-5.115	37	.000	-8.9158	-12.447	-5.384



- From the first table we can see that the mean pre diet weight of females in our sample is 62.18kg, which is 8.92kg less than our hypothesised value. To test whether this is a significant difference we refer to the second table, and in particular to the p value (listed as 'Sig. (2-tailed)'). If this value is less than .05 it tells us that the likelihood of obtaining our test statistic (i.e. the value for 't') if our null hypothesis is true is very low, in which case you would reject the null hypothesis in favour of the alternative hypothesis. Conversely, if this value is greater than .05 then you cannot reject the null hypothesis. In this case the p value is listed as .000 (i.e. it is less than .05), meaning we reject the null hypothesis and conclude that there is a statistically significant difference between the sample mean female pre diet weight and 71.1kg.

Furthermore, the table shows us that the 95% confidence interval for the difference between the population mean female pre diet weight and the hypothesised value has a lower bound of -12.447kg and an upper bound of -5.384kg. The fact that this confidence interval does not contain zero again indicates that there is a statistically significant difference between the sample mean female pre diet weight and 71.1kg.

NOTE: DON'T FORGET TO SELECT ALL CASES AGAIN BEFORE PROCEEDING TO THE NEXT SECTION.

3.2 Conducting a paired-samples t-test

Example question - *Is there a statistically significant difference between the mean weight of the respondents before going on the diet and the mean weight after the diet?*

- Appropriate hypotheses for this question are:
 H_0 : There is no significant difference between the mean pre diet and post diet weights
 H_A : There is a significant difference between the mean pre diet and post diet weights
- Appropriate test is a paired-samples t-test, as we have 'pairs' of data (i.e. before and after weights).
- The assumptions for a paired -samples t-test (which we will assume to be valid) are that:
 - The sample is a random sample that is representative of the population.
 - The observations are independent, i.e. measurements for one subject have no bearing on any other subject's measurements.
 - The data comes from a normal distribution or the sample size is large enough to ensure normality of the sampling distribution.

Note that if the last assumption is violated (and a transformation doesn't work to fix the problem), or if you are comparing ordinal data instead, you will need to use a Wilcoxon Signed Rank Test.

- To conduct this paired-samples t-test in SPSS, from the menus choose:

Analyse

Compare Means

Paired-samples T test

select the **two weight** variables (**q7** and **q8**) and move them into the **Paired Variables** box

click on **OK**

The output should look like this:

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre diet weight	64.523	44	11.9308	1.7986
	Post diet weight	60.407	44	10.3017	1.5530



Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pre diet weight & Post diet weight	44	.926	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pre diet weight - Post diet weight	4.1159	4.5719	.6892	2.7259	5.5059	5.972	43	.000

- 5) From the first table we can see that the mean pre diet weight of the sample is 64.52kg, while the mean post diet weight of the sample is 60.41kg. So there is a difference of 4.11kg, which is different to our hypothesised value of 0. To test whether this is a significance difference we refer to the third table, and in particular to the p value (listed as 'Sig. (2-tailed)') (the second table is showing us the correlation between the variables, which we don't need here). In this case the p value is listed as .000 (i.e. it is less than .05), meaning we reject the null hypothesis and conclude that, based on the sample provided, there is a statistically significant difference between the population mean weight pre diet and the population mean weight post diet.

Furthermore, the confidence interval shows us that we are 95% confident that the mean difference between pre and post diet weights in the population is not zero, and in fact is between 2.73kg and 5.51kg.

3.3 Conducting a two (independent) samples t-test

Example question - *Is there a statistically significant difference between the mean pre diet weight of those with children and those without children?*

- Appropriate hypotheses for this question are:
 H_0 : There is no significant difference between the mean pre diet weight of those with children and those without
 H_A : There is a significant difference between the mean pre diet weight of those with children and those without
- Appropriate test is a two (independent) samples t-test, as we are comparing two independent groups (i.e. those with children and those without).
- The assumptions for a two (independent) samples t-test (which we will assume to be valid) are that:
 - The sample is a random sample that is representative of the population.
 - The observations are independent, i.e. measurements for one subject have no bearing on any other subject's measurements.
 - The data comes from a normal distribution or the sample size is large enough to ensure normality of the sampling distribution.

Note that if the last assumption is violated (and a transformation doesn't work to fix the problem), or if you are comparing ordinal data instead, you will need to use a Mann-Whitney U Test.
- To conduct this two (independent) samples t-test in SPSS, from the menus choose:

Analyse



Compare Means

Independent Samples T test

move **pre diet weight (q7)** to the **test** variable (the one you want the mean of)

move **have children (q4)** to the **grouping** variable

click on the **Define Groups** button

Group 1 has the value **1** to represent those that have children

Group 2 has the value **2** to represent those that don't have children

click on **Continue**

click on **OK**

The output should look like this:

	Have children	N	Mean	Std. Deviation	Std. Error Mean
Pre diet weight	Yes	14	66.450	12.3323	3.2959
	No	30	63.623	11.8430	2.1622

		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	
									Lower	Upper
Pre diet weight	Equal variances assumed	.009	.924	.728	42	.471	2.8267	3.8829	-5.0094	10.6627
	Equal variances not assumed			.717	24.558	.480	2.8267	3.9419	-5.2992	10.9525

- 5) From the first table we can see that the mean pre diet weight of those with children is 66.45kg, while the mean pre diet weight of those without children is 63.62kg. So there is a difference of 2.83kg, which is different to our hypothesised value of 0. To test whether this is a significance difference we refer to the second table, and in particular to the p value (listed as 'Sig. (2-tailed)').

Notice however that this time there are two p value provided; which one to use depends on whether the two samples have equal or different variances. This has been tested by Levene's Test for Equality of Variances, which is another hypothesis test, this time with the null hypothesis being that the variances are equal. Hence a p value less than .05 indicates that you should reject this hypothesis, assume the variances are NOT equal, and read the information in the second row of the remainder of the table. Otherwise, you can assume the variances are equal and read the information in the first row.



In this case the p value for Levene's Test is .924 (i.e. it is greater than .05), which means that we can assume the variances are equal and read the information in the first row. You are interested in the p value for this row, which is .471, meaning we fail to reject the null hypothesis and conclude that, based on the sample provided, there is not a statistically significant difference between the population mean pre diet weight of those with children and those without. Furthermore, the confidence interval shows us that we are 95% confident that the mean difference between pre diet weights of those with children and those without in the population is between -5.01kg and 10.66kg.

3.4 Conducting a one-way ANOVA

Example question - *Is there a statistically significant difference between the mean pre diet weights for the four exercise types?*

- 1) Appropriate hypotheses for this question are:
 H_0 : There is no significant difference between the mean pre diet weight for any of the four exercise groups
 H_A : The mean pre diet weight of at least one of the exercise groups is significantly different from the others
- 2) Appropriate test is one-way ANOVA, as we are comparing more than two independent groups (i.e. those who do minimal, low level, medium and strenuous exercise).
- 3) The assumptions for ANOVA (first three of which we will assume to be valid) are that:
 - The sample is a random sample that is representative of the population
 - The observations are independent, i.e. measurements for one subject have no bearing on any other subject's measurements
 - The data comes from a normal distribution or the sample size is large enough to ensure normality of the sampling distribution
 - The populations being compared have equal variances (need to test this)

Note that if the second to last assumption is violated (and a transformation doesn't work to fix the problem), or if you are comparing ordinal data instead, you will need to use a Kruskal-Wallis One-Way ANOVA. If the last assumption is violated you will need to use a Welch or Brown-Forsyth statistic instead.

- 4) To run one-way ANOVA in SPSS, from the menus choose:

Analyse

Compare Means

One-way ANOVA

move the **pre diet weight (q7)** variable to the **Dependent List** (one you want mean of)

move the **exercise type (q9)** variable to the **Factor** box

click on the **Options** button

select **Descriptive** and **Homogeneity of variance test** in the new dialogue box

select click on **Continue**

click on **Post Hoc...**

select **Tukey**

click on **Continue**

click on **OK**

The output should look like this:



Descriptives

Pre diet weight

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Minimal exercise	8	79.500	7.7460	2.7386	73.024	85.976	70.0	92.0
Low level exercise	8	69.875	9.0780	3.2096	62.286	77.464	56.0	80.0
Medium level exercise	9	60.522	4.8700	1.6233	56.779	64.266	54.5	69.0
Strenuous exercise	8	52.475	6.9734	2.4655	46.645	58.305	42.0	63.0
Total	33	65.439	12.2834	2.1383	61.084	69.795	42.0	92.0

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Pre diet weight	Based on Mean	.738	3	29	.538
	Based on Median	.488	3	29	.693
	Based on Median and with adjusted df	.488	3	19.347	.695
	Based on trimmed mean	.701	3	29	.559

ANOVA

Pre diet weight

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3301.213	3	1100.404	20.898	.000
Within Groups	1527.006	29	52.655		
Total	4828.219	32			

Multiple Comparisons

Dependent Variable: Pre diet weight

Tukey HSD

(I) Exercise type	(J) Exercise type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Minimal exercise	Low level exercise	9.6250	3.6282	.059	-.260	19.510
	Medium level exercise	18.9778*	3.5260	.000	9.371	28.584
	Strenuous exercise	27.0250*	3.6282	.000	17.140	36.910
Low level exercise	Minimal exercise	-9.6250	3.6282	.059	-19.510	.260
	Medium level exercise	9.3528	3.5260	.059	-.254	18.959
	Strenuous exercise	17.4000*	3.6282	.000	7.515	27.285



Medium level exercise	Minimal exercise	-18.9778*	3.5260	.000	-28.584	-9.371
	Low level exercise	-9.3528	3.5260	.059	-18.959	.254
	Strenuous exercise	8.0472	3.5260	.126	-1.559	17.654
Strenuous exercise	Minimal exercise	-27.0250*	3.6282	.000	-36.910	-17.140
	Low level exercise	-17.4000*	3.6282	.000	-27.285	-7.515
	Medium level exercise	-8.0472	3.5260	.126	-17.654	1.559

*. The mean difference is significant at the 0.05 level.

Pre diet weight

Tukey HSD^{a,b}

Exercise type	N	Subset for alpha = 0.05		
		1	2	3
Strenuous exercise	8	52.475		
Medium level exercise	9	60.522	60.522	
Low level exercise	8		69.875	69.875
Minimal exercise	8			79.500
Sig.		.134	.064	.054

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 8.229.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

- 5) The first thing you need to observe from this output is the 'Test of Homogeneity of Variances' table. If the p value in this table (listed under 'Sig.')
- is greater than or equal to .05, then we can assume that the variances are equal, and hence that our equal variances assumption is valid. In this case the p value is .538, which is greater than .05, so our assumption is valid and the ANOVA table can be used.

Next observe the p value, listed as 'Sig.', in the ANOVA table. In this case the p value is listed as .000 (i.e. it is less than .05), meaning we reject the null hypothesis and conclude that, based on the sample provided, there is a statistically significant difference in population mean pre diet weight for at least one of the exercise types. We can get an idea of where the differences are by looking at the first (Descriptives) table, but to see whether each difference is statistically significant we need to refer to the results of our Tukey HSD post hoc test, in the Multiple Comparisons table and/or the Homogeneous Subsets table. From this we can see that there is a significant difference in mean pre diet weights for the following pairs; minimal and medium exercise, minimal and strenuous exercise, and low level and strenuous exercise (this can be determined by the significance values in the Multiple Comparisons table, and by the groupings in the Homogeneous Subsets table- groups in the same column are not significantly different).

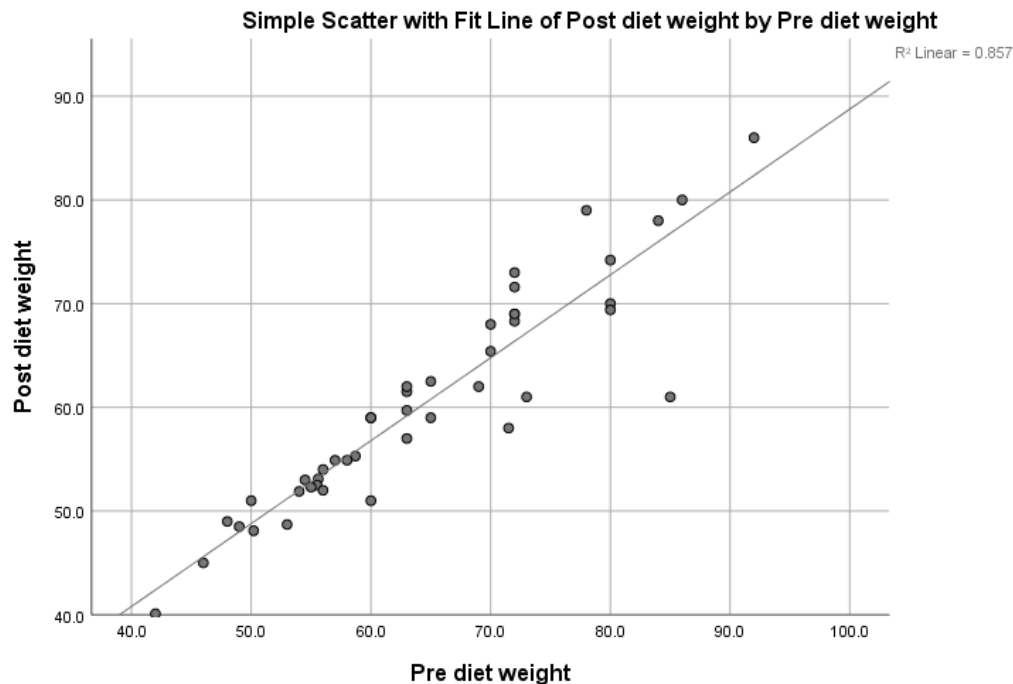
You can test for an association between two categorical variables in a population by conducting a Chi-square test, and for linear correlation between two continuous variables using Pearson's correlation coefficient. How to run each of these tests in SPSS is covered in the following sections.

3.5 Pearson's correlation coefficient

Example question - *Is there a statistically significant linear relationship (correlation) between pre diet and post diet weight in the population?*

Since this question concerns two continuous variables, you can test for a linear correlation between them using Pearson's correlation coefficient. Before doing this though, you should visualise the relationship

between the variables in a scatterplot, as detailed in the previous session. The resultant graph should look like this:



Doing this will also give you a value for the coefficient of determination (r^2), which we will cover shortly.

We can now quantify the linear correlation between the variables using Pearson's correlation coefficient, which is a number (denoted by r) between -1 and 1. The same procedure in SPSS is used both to calculate the correlation coefficient for a pair of continuous variables and to test whether linear correlation is significant in the population, the latter of which is another hypothesis test and therefore requires the five steps of hypothesis testing:

- 1) Appropriate hypotheses for this question are:
 - H_0 : There is no significant linear correlation between the variables
 - H_A : There is significant linear correlation between the variables
- 2) Appropriate test is Pearson's correlation coefficient, as explained previously.
- 3) The assumptions for Pearson's correlation coefficient (which are valid in this case) are that:
 - The observations are independent.
 - Both variables are normally distributed.
 - There is a linear relationship between the variables, as observed in the scatter plot (this is not strictly an assumption as Pearson's correlation coefficient is still valid without it, but if you already know the relationship is not linear further interpretation is not necessary)
 - There is a homoscedastic relationship between the variables (i.e. variability in one variable is similar across all values of the other variable), as observed in the scatter plot (dots should be similar distance from line of best fit all the way along)

If the variables are not normally distributed, or if the data is ordinal, you can use Spearman's rho or Kendall's tau-b instead.

- 4) To conduct the test in SPSS, from the menus choose:

Analyse
Correlate
Bivariate

select **q7 (Pre diet weight)** and **q8 (Post diet weight)** and move them into the **Variables** box
click on **OK**

Note: Even though the form of the correlation used here is ‘bivariate’, i.e. two variables, any number of variables can be moved into the **Variables** box specified above. By default, SPSS will produce a results table with ALL of the variables listed on both dimensions of the table. You then have to pick the cell on the table that gives the combination of variables needed. In this example, that cell will be the combination of Pre diet weight and Post diet weight.

The output should look like this:

		Pre diet weight	Post diet weight
Pre diet weight	Pearson Correlation	1	.926**
	Sig. (2-tailed)		.000
	N	44	44
Post diet weight	Pearson Correlation	.926**	1
	Sig. (2-tailed)	.000	
	N	44	44

** . Correlation is significant at the 0.01 level (2-tailed).

- To interpret this table first look at the Pearson correlation coefficient in the cell for the appropriate combination of variables (i.e. Pre diet weight and Post diet weight). This is listed as .926, which indicates that there is a strong positive linear relationship between the two variables in the sample. Note that the correlation coefficient is a measure of effect size, so we don't need to calculate a separate measure for it, although we can also square the value to obtain another measure known as the coefficient of determination (which is .857 in this case, as displayed previously when we added the line of best fit to our scatter plot). The coefficient of variation measures the proportion of variance in the dependent variable that can be accounted for by variance in the independent variable, so in this instance it indicates that 85.7% of variation in post diet weight can be accounted for by variance in post diet weight.

To test whether this linear relationship between the variables is statistically significant in the population, we interpret the *p* value (listed as 'Sig. (2-tailed)'). The fact that this is given as .000 indicates that we can reject the null hypothesis of no linear relationship, and instead conclude that there is a statistically significant linear relationship between the variables in the population.

3.6 Conducting a Chi-square test

Example question - *Is there a statistically significant association between marital status and having children in the population?*

Since this question concerns two categorical variables you need to use a Chi-square test to test for association in the population. This is a hypothesis test, so we will again use the five steps of hypothesis testing:

- Appropriate hypotheses for this question are:
 - H₀: There is no significant association between the variables
 - H_A: There is significant association between the variables
- Appropriate test is a Chi-square test, as explained previously.
- The assumptions for a Chi-square test (which are valid in this case) are that:
 - The cells are mutually exclusive, i.e. each subject can only belong to one category of each variable.



- The cells are exhaustive, i.e. it is possible to classify each subject into a cell.
- The expected frequencies are sufficiently large. Specifically, not more than 20% of the expected frequencies are less than 5 (note that if this last assumption is not satisfied, i.e. the percentage figure is 20 or greater and you have a 2x2 table, you can use the Fisher's Exact-test statistics instead. This is produced automatically by SPSS for 2x2 tables).

4) To conduct this Chi-square test in SPSS, from the menus choose:

Analyse

Descriptive Statistics

Crosstabs

transfer one variable (i.e. **q2**) to the **Row(s)** box

transfer the other variable (i.e. **q4**) to the **Column(s)** box

click on the **Statistics...** button, and check the **Chi-square** box and the **Phi and Cramer's V** box

select **Continue**

click on the **Cells...** button, and check both the **Observed** and **Expected** boxes (if required)

select **Continue**

check the **Display clustered bar charts** box

select **OK**

The output should look like this:

Marital status * Have children Crosstabulation

		Have children		Total	
		Yes	No		
Marital status	Single	Count	1	43	44
		Expected Count	11.0	33.0	44.0
	Married	Count	17	12	29
		Expected Count	7.3	21.8	29.0
	Other	Count	2	5	7
		Expected Count	1.8	5.3	7.0
Total		Count	20	60	80
		Expected Count	20.0	60.0	80.0

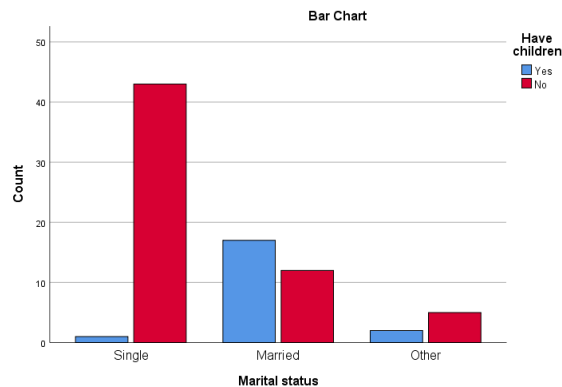
Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	29.652 ^a	2	.000
Likelihood Ratio	32.716	2	.000
Linear-by-Linear Association	16.328	1	.000
N of Valid Cases	80		

a. 1 cells (16.7%) have expected count less than 5. The minimum expected count is 1.75.

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	.609	.000
	Cramer's V	.609	.000
N of Valid Cases		80	



- 5) The first thing you need to observe from this output is the Crosstabulation table, which in this case shows observed and expected frequencies for the marital and child status of the subjects in the sample. The expected frequencies (listed in the 'Expected Count' rows of the table) are the frequencies we would expect if there is no association between the variables (i.e. if the proportion for each category is the same as the overall proportion- these are calculated by multiplying the column total by the row total and dividing by the overall total in each case), while the observed frequencies (listed in the 'Count' rows of the table) are those that have been calculated using the sample data. The greater the discrepancy between these two sets of values, the less likely it is that the null hypothesis is true (i.e. the less likely it is that there is no association between the variables). So you should be able to get a feel for this from looking at the table. For example, the expected frequency for single people who have children is 11, which is calculated by multiplying 20 (column total) by 44 (row total) and dividing by 80 (overall total) (i.e. $(20 \times 44) / 80 = 11$). This is quite different from the observed frequency (1), and in fact you should notice that most of the expected and observed frequencies are quite dissimilar. You can also visualise the relationship between the variables in the sample using the clustered bar chart, by comparing the heights of the columns for each category of the independent variable.

The second table is the results of our Chi-square test, which tests whether any observed association is significant in the population or not. So as with our previous tests on continuous data a test statistic is calculated, and our p value (listed as 'Asymp. Sig. (2-sided)') indicates how likely it is that this test statistic would have been obtained if our null hypothesis is true. Again, if our p value is less than .05 we reject the null hypothesis (and say that there IS an association between the variables), while if our p value is greater than or equal to .05 we fail to reject the null hypothesis (and instead say that there is NO association between the variables). In this case our p value is listed as .000 (i.e. is less than .05), so we reject the null hypothesis and conclude that there IS a statistically significant association between marital status and having children in the population.

Furthermore, the Symmetric Measures table gives us the effect size, which measures the strength of the association between the variables. There are various ways of measuring this but we asked for Phi and Cramer's V, which are two common measures. Phi should be interpreted in the case of a 2x2 table, while Cramer's V should be used in the case where one or both variables have more categories. While values for both statistics are the same here note that the correct measure to state is Cramer's V, and that a value of .609 indicates that the association between the variables is large (for Phi and Cramer's V where one variable has only two categories, the latter being true in this case, a value of .1 is considered a small effect size, .3 medium and .5 large).