

ANALYSIS OF VARIANCE (ANOVA): STEP-BY-STEP INSTRUCTIONS

Please note that the analysis tools included in the MAX toolkit require that you have Microsoft Excel 2010 or later, and that you have also installed the **Analysis ToolPak Pack** and the **Real Statistics Resource Pack** add-in programmes. Both add-ins are free to use and guidance on how to install these add-ins can be accessed [here](#).



An overview of the relevant statistical concepts used in these instructions is provided in the [Getting Started with Statistics](#) guide. An overview of ANOVA is provided in the [exploring the differences between respondent groups guide](#) and [presentation](#),¹ along with an overview of Independent t-tests, which can be used to explore the means of two groups. Additional guidance is also provided in the **FURTHER ANALYSIS GUIDES**, which can be downloaded [here](#).

INTRODUCTION

Analysis of variance (ANOVA) is a form of **'hypothesis testing'** and can be used to determine whether there is a statistically significant difference between the means of more than two unrelated groups (e.g. characterised by rating of satisfaction with services) on the same continuous dependent variable (e.g. SCRQoL). The statistical test generates a **p-value** which enables you to determine whether or not to reject the null hypothesis. This p-value is the probability that the null hypothesis is true. If the p-value is less than a set **significance level** (usually 5% [0.05] or 1% [0.001]), it is said that the pattern is 'statistically significant'. This means that there is a low probability that the patterns observed in your data arose by chance.

¹ Please note that ANOVAs and t-tests can also be used to explore the relationship between survey variables and, conversely, that cross-tabulations and chi-square tests can be used to explore the differences between respondent groups. The format of these guides and presentations were used to introduce increasingly complex concepts to MAX toolkit users and attendees of the 2016 webinars.

ANOVA overcomes many of the issues associated with conducting multiple t-tests (e.g. resource-intensive, prone to generating **Type I errors**) and can be used to explore a range of variables in the Adult Social Care Survey (ASCS) and Carers Survey (PSS SACE).

This guide provides you with a step-by-step overview of how to conduct ANOVA in Excel. These instructions use the following question ***Do overall social care-related quality of life (SCRQoL) scores (ASCOF 1A) differ by rating of satisfaction with services?*** but are also relevant to the PSS SACE.

In this example, the six responses for satisfaction are collapsed into four categories: [1] extremely/very satisfied; [2] quite satisfied; [3] neither satisfied or dissatisfied; and [4] quite/very/extremely dissatisfied. Depending on your data, you may want to run ANOVAs using the 6 individual response categories or collapse them further (e.g. into 3 categories).

GETTING STARTED

Before running analysis of variance (ANOVA), ensure that your data fulfils the requirements of the statistical test. In particular, that:

1. **The dependent variable can be measured on a continuous scale:** in this instance, SCRQoL scores are measured on a scale of 0 to 24 (✓).
2. **Independence of observations:** each respondent must provide data for one group only and cannot be in both groups. Survey respondents can only provide one answer so this assumption is fulfilled (✓). If, however, this assumption is violated, a different type of ANOVA will need to be conducted. Please note that this test is not covered in the MAX toolkit.

The two other requirements – that the dependent variable should be normally distributed for each group (the **normal distribution** assumption) and that the variance of the dependent variable is also approximately equal for each group (the **homogeneity of variances** assumption) – are tested during the analysis.

STEP-BY-STEP INSTRUCTIONS

STEP 1: EXTRACT DATA

Select the relevant data from your NHS Digital data return and copy and paste it into a blank Excel sheet. Ensure the data is arranged in columns and has suitable headings.

	Extremely/ very satisfied	Quite satisfied	Neither	Quite, very or extremely dissatisfied
	10	7	8	8
	11	8	8	9
	12	9	9	9
	12	9	10	9
	12	9	10	10
	12	9	11	11
	13	9	11	12
	13	10	12	12
	13	10	12	12
	13	10	13	12
	13	10	13	12
	14	11	13	13

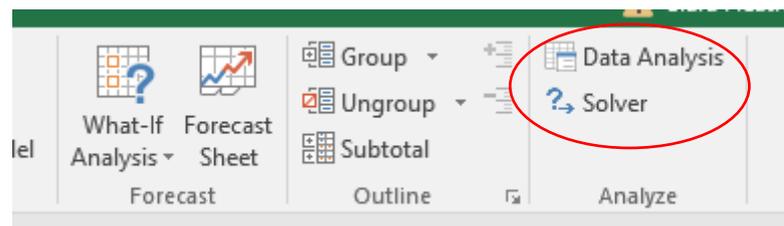
Screen shot from 'dummy' data

STEP 2: CHECK WHETHER DATA IS NORMALLY DISTRIBUTED

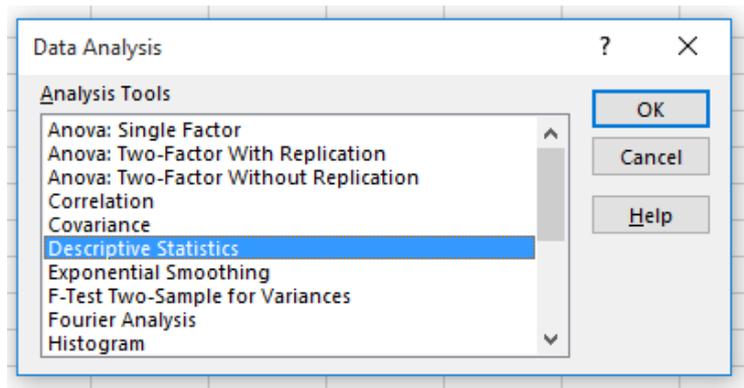
Before you conduct an ANOVA, you need to establish whether the data is normally distributed.

This assumption can be determined using the **Descriptives Statistics** function in Excel.

- I. Open data tab of Excel, select **Data Analysis** from Analyze section.



- II. In the **Data Analysis window**, select **Descriptive Statistics** and press **OK**.



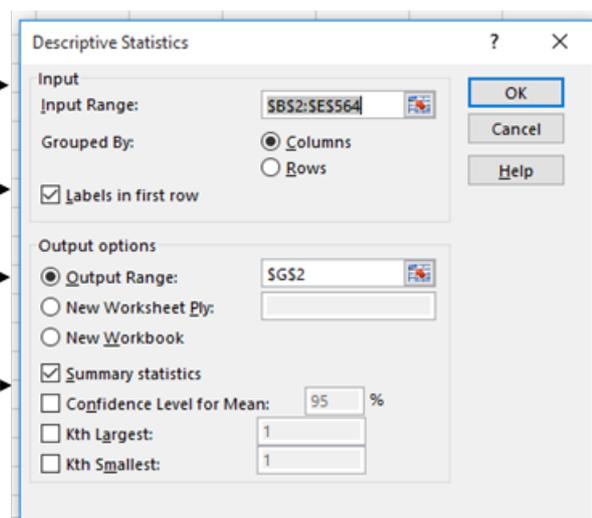
- III. In the **Descriptive Statistics window**, select the **Input Range** by clicking on . Select the entire dataset, taking care to ensure you include the row headings. Go back to Descriptive Statistics window, select **Labels in first row** [in Input section] and **Output range**. Click on **summary statistics** [in Output section] and then press **OK**.

Input range: select entire data set (i.e. data for all four groups) and include row labels

Click **Labels in first row**

Output range: where in Excel you would like analysis results to be posted.

Click **Summary statistics**



Note: you can also click the **Confidence level for Mean box** and amend the significance level (optional).

The descriptives statistics summary will be posted into your datasheet. **Stretch the columns** so you can see the data. You may also want to format the numbers to 2 decimal places.

Extremely/very satisfied	Quite satisfied	Neither	Quite, very or extremely dissatisfied
Mean	20.19	Mean	17.46
Standard Error	0.12	Standard Error	0.23
Median	21.00	Median	18.00
Mode	24.00	Mode	18.00
Standard Deviation	2.96	Standard Deviation	3.70
Sample Variance	8.73	Sample Variance	13.68
Kurtosis	-0.16	Kurtosis	-0.36
Skewness	-0.62	Skewness	-0.34
Range	14	Range	17
Minimum	10	Minimum	7
Maximum	24	Maximum	24
Sum	11348	Sum	4523
Count	562	Count	259

Screen shot from 'dummy' data

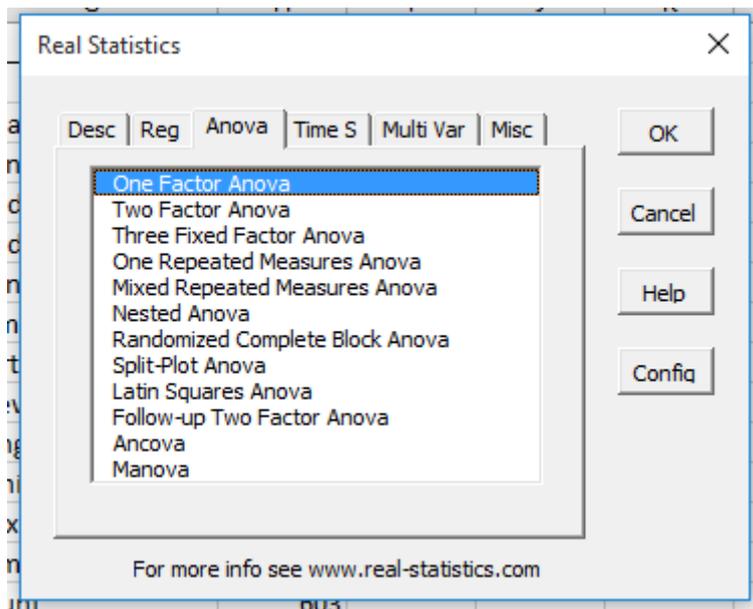
The extent to which a dataset is normally distributed is established using the measure of **skewness**. A frequency distribution with perfect symmetry has a skew of 0. Most distributions, however, are not perfect, and the most frequently occurring scores cluster at the lower end/left side of the distribution (**positively skewed distributions**) or at the higher end/right side of the distribution (**negatively skewed distributions**).

The data in this example is not normally distributed. This is probably the case for most SCRQoL data, as the majority of respondents report higher levels of quality of life, and means that the assumption of normal distribution is violated. In this instance, you should run a **Kruskal Wallis Test**, the non-parametric equivalent of the ANOVA. Instructions on how to do this are provided in step 4.

STEP 3: CHECK WHETHER VARIANCES IN DATASET ARE APPROXIMATELY EQUAL.

Before you conduct an ANOVA, you also need to establish whether the variance of the dependent variables (in this instance, SCRQoL) are approximately equal in each group. This assumption is known as the **homogeneity of variances** assumption and can be determined using the **Real Statistics** function in Excel.

- I. Click **CTRL + M** to open up the **Real Statistics window**. Click on the **Anova** tab and select **One Factor ANOVA**. Press **OK**.



- II. In the **ANOVA window**, select **Input Range** (i.e. the entire data set), ensuring you include the column headings. Select **Input Format** [Excel format with column headings]. In the Options section, unselect ANOVA and select **Levene's Test**. Specify **Output Range** and press **OK**.

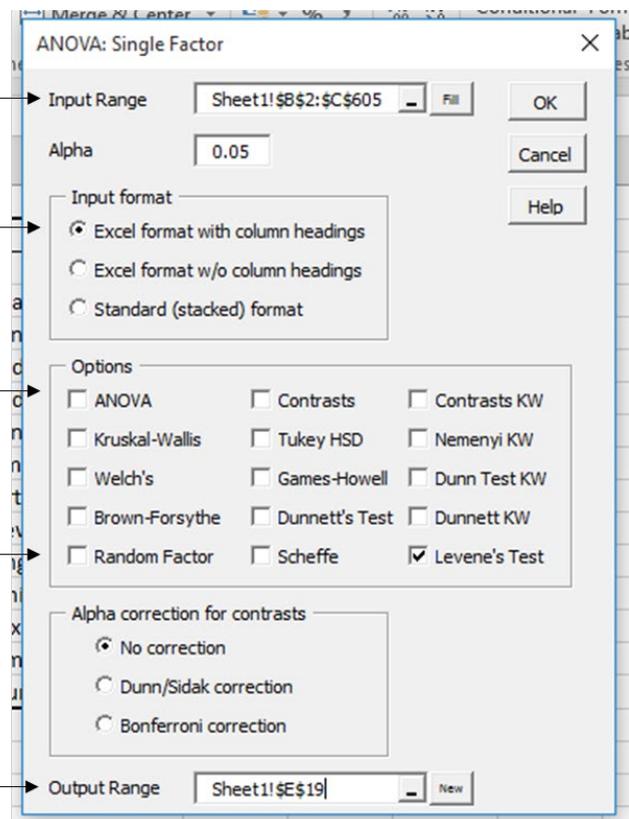
Input range: select entire data set (i.e. data for all four groups) and include row labels

Click **Excel format with column headings**

Unselect **ANOVA**

Select **Levene's Test**

Output range: where in Excel you would like analysis results to be posted.



The Levene's test output will be posted into your datasheet. **Stretch the columns** so you can see the data and format the numbers to decimal places. **If the p-value is less than**

0.05: the assumption of homogeneity of variances is violated (i.e. there is a difference between the variances for each group) and you will need to run a **Kruskal-Wallis test**. **If the p-value is greater than 0.05:** the assumption of homogeneity of variances is not violated (i.e. there is no – or marginal – differences between the variances for each group) and you can run a **one-way ANOVA**.

Levene's Tests	
type	p-value
means	0.00
medians	0.00
trimmed	0.00

In this example, the p-value is less than 0.05 [0.00]. The assumption of homogeneity of variances has therefore been violated and you will need to run a Kruskal-Wallis test.

Please note: as well as determining the type of ANOVA you will run [see Step 4], the results of your Levene's test will also determine the type of t-tests you will use during post-hoc analyses. If the null hypothesis is not rejected (i.e. the p-value is greater than 0.05 or the adjusted p-value used for multiple pairwise comparisons²), the t-test will be run with the option 'assuming equal variances'. If, on the other hand, the null hypothesis is rejected (i.e. the p-value is less than 0.05 or the adjusted p-value used for multiple pairwise comparisons), the t-test will be run using the option 'assuming unequal variances'

STEP 4: RUN ANOVA

There are two variations of an independent measures ANOVA and the most appropriate test will be determined by your results from Steps 2 and 3. The options are summarised in the table below.

² See the guidance on the **Bonferroni correction** on page 11-12.

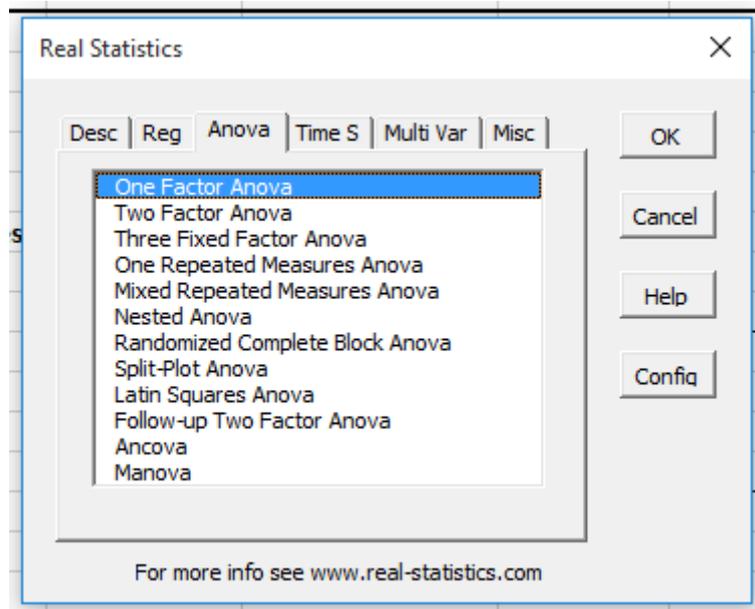
Test	Normal Distribution		Homogeneity of Variance	
	Yes	No	Yes	No
One way ANOVA	✓		✓	
Kruskal Wallis test ³		✓		✓

Note: in this example, you would run the **Kruskal Wallis Test**. This is therefore outlined first.

KRUSKAL WALLIS TEST

The Kruskal-Wallis test is a **non-parametric statistical test**. You may want to run this test (instead of a one-way ANOVA) if the SCRQoL score is highly skewed (i.e. not normally distributed) and/or the assumption of homogeneity of variance is violated. This analysis can be conducted in Excel using the **Real Statistics Resource Pack**.

- I. Click **CTRL + M** to open up the **Real Statistics window**. Click on the **Anova** tab and select **One Factor Anova**. Press **OK**.



³ **The Kruskal-Wallis test** is the non-parametric equivalent of the one-way (or independent measures) ANOVA so does not make any assumptions about the dataset.

- II. In the **ANOVA: Single Factor** window, select **Input Range**, tick **Excel format with column headings** [in Input format] and select **Kruskal-Wallis Test** [in Options]. Define **Output Range** and then press **OK**.

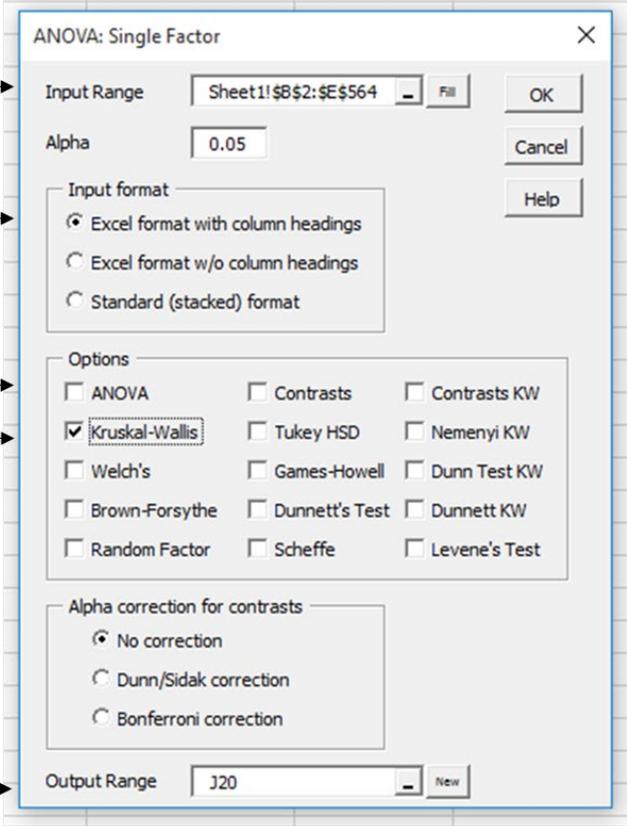
Input range: select entire data set (i.e. data for all four groups) and include row labels

Click **Excel format with column headings**

Unselect **ANOVA**

Select **Kruskal Wallis Test**

Output range: where in Excel you would like analysis results to be posted.

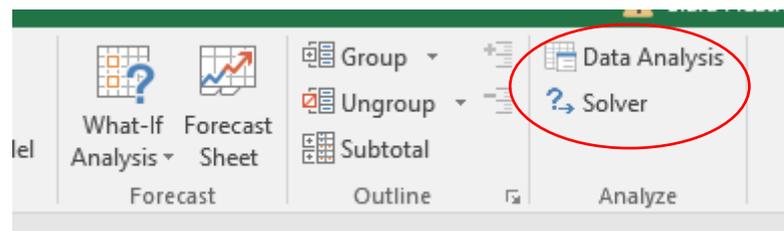


- III. Review output data and look at the **p-value**. In this instance, the p-value is 0.00. This means the results are statistically significant (i.e. there is a difference between the groups) and that the null hypothesis can be rejected. **Note:** The Kruskal-Wallis test confirms whether or not the result is significant in the last row.

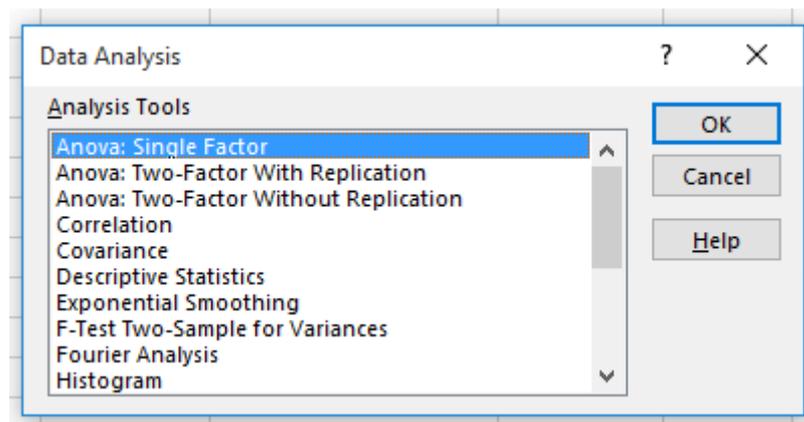
Kruskal-Wallis Test				
	Extremely/very	Quite satisfied	Neither	Quite, very or extremely dissatisfied
median	21	18	17	13.5
rank sum	305066.5	90581	17997.5	9095
count	562	259	56	42
r ² /n	165597098.6	31679218.38	5784107.254	1969500.595
H-stat				150.01
H-ties				151.21
df				3.00
p-value				0.00
alpha				0.05
sig				yes

SINGLE FACTOR ANOVA

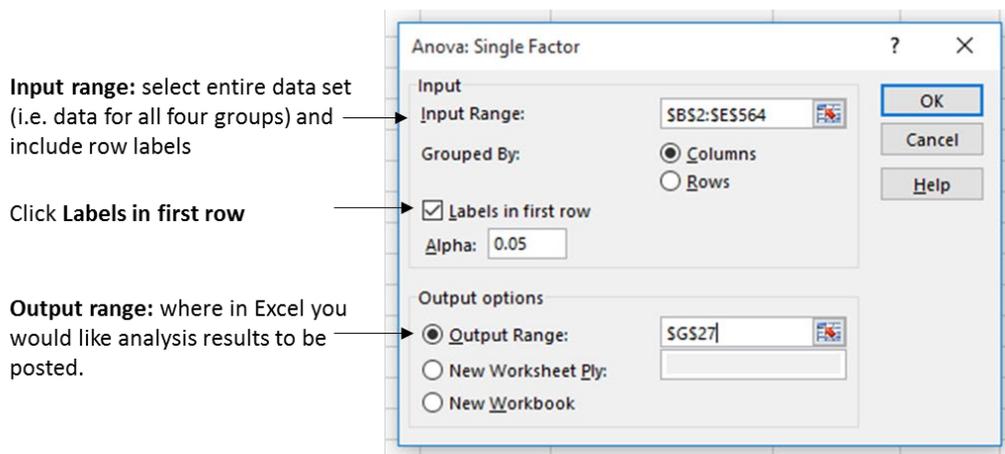
- I. Open data tab of Excel, select **Data Analysis** from Analyze section.



- II. In the **Data Analysis window**, select **Anova: Single Factor**. Press **OK**.



- III. Select **Input Range** (including row headings) and **Grouped by: columns**. Put a tick beside **Labels in first row** and define **Output Range**. Press OK.



The test output will be posted into your datasheet. **Stretch the columns** so you can see the data. You may also want to format the numbers to decimal places.

The first table [SUMMARY] provides an overview of the data, including the number of responses in each group [count], the average [in this instance, of SCRQoL rating] and the variance within each group. The second table [ANOVA] shows the p-value. **If the p-value is less than 0.05:** you can reject the null hypothesis and assert that there is a statistically significant difference between the groups. **If the p-value is greater than 0.05:** you cannot reject the null hypothesis.

SUMMARY						
Groups	Count	Sum	Average	Variance		
Extremely/very satis	562	11348	20.19	8.73		
Quite satisfied	259	4523	17.46	13.68		
Neither	56	939	16.77	17.60		
Quite, very or extre	42	614	14.62	21.46		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2493.98	3	831.33	74.01	0.00	2.61
Within Groups	10277.53	915	11.23			
Total	12771.51	918				

Screen shot from 'dummy' data

In this instance, the p-value is 0.00 so you can reject the null hypothesis and assert that there is a significant variation in SCRQoL between the four groups. Based on this, we may run further

analysis to compare the groups pairwise using ('post-hoc') T-tests. These are described further in section Running Post Hoc Analysis.

RUNNING POST HOC ANALYSIS

If you find a statistically significant difference between your groups, you may want to run further analysis to compare the groups pairwise using ('post-hoc') T-tests. In this example, there are six possible comparisons that you could run. These are summarised in the table below.

Group 1	Group 2
Extremely/Very satisfied	Quite satisfied
Extremely/Very satisfied	Neither satisfied or dissatisfied
Extremely/Very satisfied	Quite, very or extremely dissatisfied
Quite satisfied	Neither satisfied or dissatisfied
Quite satisfied	Quite, very or extremely dissatisfied
Neither satisfied or dissatisfied	Quite, very or extremely dissatisfied

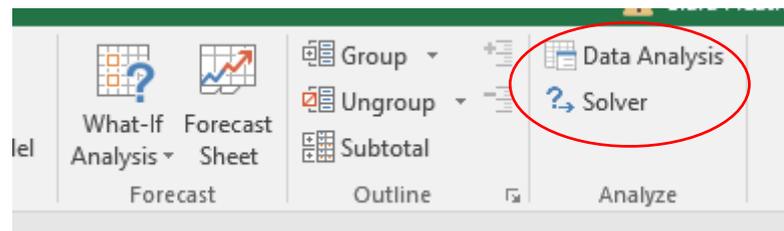
The type of t-test that you will run is determined by your results from Step 3. In this step, you ran a Levene's test to establish whether the variance of the dependent variables (in this instance, SCRQoL) is approximately equal in each group. Your results showed that the variance was not equal so you will therefore run a **t-test with equal variances not assumed** [this is shown as t-Test: Two-Sample Assuming Unequal Variances in the Analysis Toolpak]. If, on the other hand, your results showed that the variance was approximately equal, you would run a **t-test with equal variances assumed** [this is shown as t-Test: Two-Sample Assuming Equal Variances in the Analysis Toolpak].

Warning: there is a risk of generating a **Type I error** when running multiple pair-wise comparisons. A Type I error occur when the null hypothesis is rejected when, in fact, it is true (i.e. you conclude your data shows there is an effect or difference when there is none). The risk of generating a type I error can be reduced by using a lower significance value. The **Bonferroni correction** is an adjustment made to p-values when several dependent or independent statistical tests are being performed simultaneously on a single data set. A Bonferroni correction

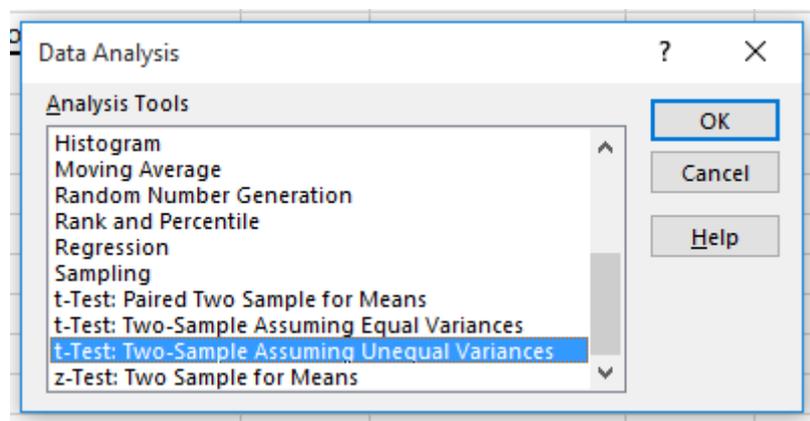
can be calculated by simply dividing the critical P -value (α), usually 0.05, by the number of comparisons being made. For example, in this example 6 pair-wise comparisons are being conducted. **The new critical P -value would therefore be 0.008.** The statistical power of the study is then calculated based on this modified P -value.

T-TEST WITH EQUAL VARIANCES NOT ASSUMED

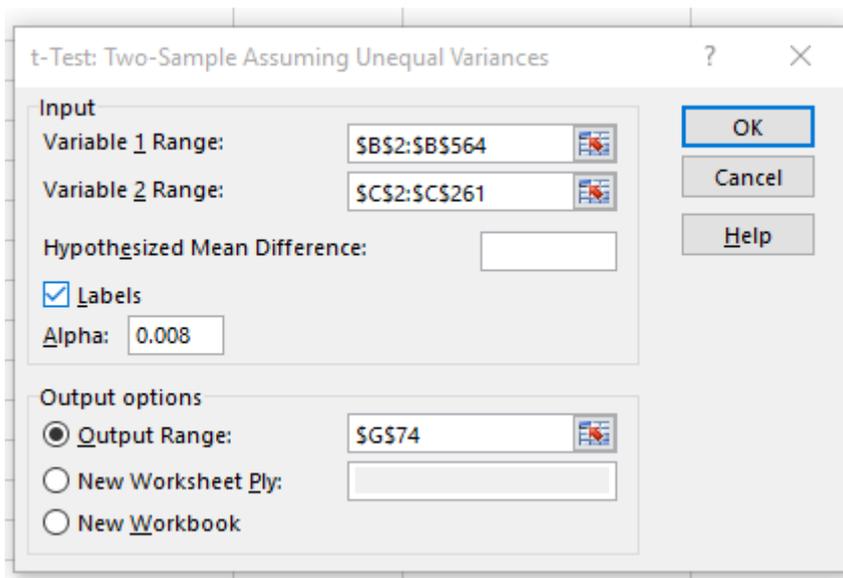
- I. Open data tab of excel, select **Data Analysis** from Analyze section.



- II. In the Data Analysis window, **scroll down** and select **t- test: Two Sample Assuming Unequal Variances**. Press **OK**.



- III. Select **Input Range** for both variables, put a tick beside **Labels** and, if using a Bonferroni correction, amend **Alpha** to your adjusted p-value. Define **Output Range** and press OK.



The test output will be posted into your datasheet. **If the p-value is less than your revised p-value (in this case, 0.008):** you can reject the null hypothesis and assert that there is a statistically significant difference between the groups. **If the p-value is greater than your revised p-value (in this case, 0.008):** you cannot reject the null hypothesis.

t-Test: Two-Sample Assuming Unequal Variances		
	<i>Extremely/very satis</i>	<i>Quite satisfied</i>
Mean	20.19	17.46
Variance	8.73	13.68
Observations	562	259
Hypothesized Mean	0	
df	416	
t Stat	10.44	
P(T<=t) one-tail	0.00	
t Critical one-tail	2.42	
P(T<=t) two-tail	0.00	
t Critical two-tail	2.66	

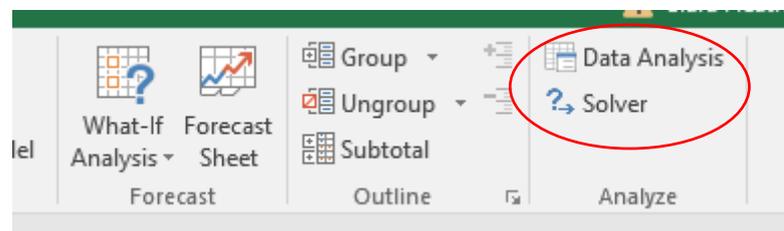
Screen shot from 'dummy' data

In this example, you are exploring the differences in SCRQoL scores between respondents who were extremely or very satisfied with services and respondents who were quite satisfied but have not specified the direction of the relationship. You will therefore need to look at the **p-value** for the **two-tailed hypothesis**. This is 0.00 and is therefore statistically significant.

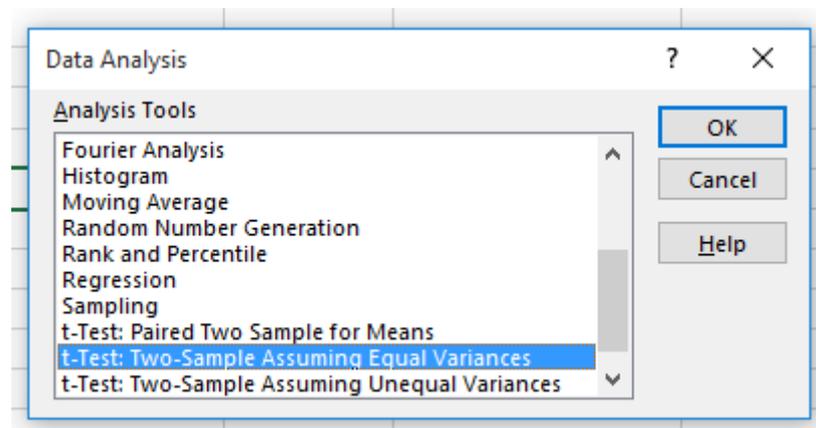
The outputs from the remaining t-tests are shown in the data provided with these instructions and also in the accompanying dummy data and [SCRQoL Tool](#) [see the ANOVA example tab]. Overall, these results indicate that there is a statistically significant relationship between satisfaction with services and social care-related quality of life. In particular, that people with lower satisfaction with services also report a lower quality of life. This raises the question of whether and how services could do more to improve satisfaction - and quality of life.

T-TEST WITH EQUAL VARIANCES ASSUMED

- I. Open data tab of Excel, select **Data Analysis** from Analyze section.



- II. In the Data Analysis window, scroll down and select **t-test: Two Sample Assuming Unequal Variances**. Press **OK**.



- III. Select **Input Range** for both variables, put a tick beside **Labels** and, if using a Bonferroni correction, amend **Alpha** to your adjusted p-value. Define **Output Range** and press OK.

Input range 1: select data for 1st group (e.g. extremely/v. satisfied) + include row labels

Input range 2: select data for 2nd group (e.g. quite satisfied) + include row labels

Click Labels

Remember to change **Alpha** if using a Bonferroni correction

Output range: where in Excel you would like analysis results to be posted.

t-Test: Two-Sample Assuming Unequal Variances

Input

Variable 1 Range: \$B\$2:\$B\$564

Variable 2 Range: \$C\$2:\$C\$261

Hypothesized Mean Difference:

Labels

Alpha: 0.008

Output options

Output Range: \$G\$74

New Worksheet Ply:

New Workbook

OK

Cancel

Help

The test output will be posted into your datasheet. **If the p-value is less than your revised p-value (in this case, 0.008):** you can reject the null hypothesis and assert that there is a statistically significant difference between the groups. **If the p-value is greater than your revised p-value (in this case, 0.008):** you cannot reject the null hypothesis.

INTERPRETING THE TEST STATISTICS

The tests outlined in this summary all generate a **p-value** which enables you to determine whether or not to reject the null hypothesis. This p-value is the probability that the null hypothesis is true. If the p-value is less than a set **significance level** (usually 5% [0.05] or 1% [0.001]) it is said that the pattern is 'statistically significant'. This means that there is a low probability that the observed pattern arose by chance.

The p-value in this example is 0.00 (see Kruskal-Wallis Test in Step 4) and means that the **null hypothesis can be rejected** – in other words, ***there is a statistically significant difference between SCRQoL and satisfaction.***

With the exception of one test (which compared SCRQoL for quite satisfied and neither satisfied or dissatisfied group), the p-values in the post-hoc analyses (see previous section) are all significant too.

REPORTING THE TEST STATISTICS

SINGLE FACTOR ANOVA

The usual format for reporting the results of a single factor ANOVA is shown below:

$$F(\underline{df [BG]}, \underline{df [WG]}) = f \text{ value}, p = \text{significance value}$$

↑
↑
↑

Degrees of freedom
For between groups [BG] and within-groups [WG]. Taken from t-test output (step 4)

F statistic value
Taken from output (step 4)

F statistic p value
Either one or two tailed. Taken from output (Step 4)

To ensure report recipients understand what the results mean, you should report your findings in more everyday terms. For example:

Social care related quality of life (SCRQoL) [ASCOF 1A] is significantly affected by ratings of satisfaction with services, $F(3,915) = 74, p = 0.00$.

SUMMARY						
Groups	Count	Sum	Average	Variance		
Extremely/very satisfied	562	11348	20.19	8.73		
Quite satisfied	259	4523	17.46	13.68		
Neither	56	939	16.77	17.60		
Quite, very or extremely dissatisfied	42	614	14.62	21.46		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2493.98	3	831.33	74.01	0.00	2.61
Within Groups	10277.53	915	11.23			
Total	12771.51	918				

KRUSKAL-WALLIS-TEST

The usual format for reporting the results of a Kruskal-Wallis test is shown below:

$$H(\underline{df}) = h \text{ value}, p = \text{significance value}$$

↑
↑
↑

Degrees of freedom
Taken from t-test output (step 4)

H statistic value
Taken from output (step 4)

t statistic p value
Either one or two tailed. Taken from output (Step 4)

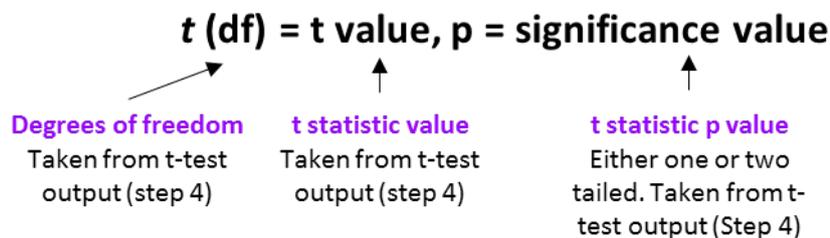
To ensure report recipients understand what the results mean, you should report your findings in more everyday terms. For example:

Social care related quality of life (SCRQoL) [ASCOF 1A] is significantly affected by satisfaction with services, $H(3) = 150.01, p = 0.00$.

Kruskal-Wallis Test					
	Extremely/very	Quite satisfied	Neither	Quite, very or extremely	dissatisfied
median	21	18	17	13.5	
rank sum	305066.5	90581	17997.5	9095	
count	562	259	56	42	919
r ² /n	165597098.6	31679218.38	5784107.254	1969500.595	205029924.8
H-stat					150.01
H-ties					151.21
df					3.00
p-value					0.00
alpha					0.05
sig					yes

POST HOC T-TESTS

The usual format for reporting the results of an independent t-test is shown below:



You can report the findings from your post-hoc analysis, if conducted, after the results from your ANOVA or Kruskal-Wallis test. To ensure report recipients understand what the results mean, you should report your findings in more every day terms. For example:

Post- hoc t-tests showed that, with the exception of ASCS respondents who are ‘quite’ or ‘neither satisfied or dissatisfied’ with services [$t(75) = 1.15, p = 0.25$], all between group differences were significant.

t-Test: Two-Sample Assuming Unequal Variances		
	<i>Quite satisfied</i>	<i>Neither</i>
Mean	17.46	16.77
Variance	13.68	17.60
Observations	259	56
Hypothesized Mean	0	
df	75	
t Stat	1.15	
P(T<=t) one-tail	0.13	
t Critical one-tail	2.46	
P(T<=t) two-tail	0.25	
t Critical two-tail	2.72	

A word of caution: a significant association does not imply causation but does identify potential important relationships that should be further explored using multi-variate and/or themed analysis. Additional guidance is provided in the **FURTHER ANALYSIS GUIDES** included in the MAX toolkit and can be downloaded [here](#).

DISCLAIMER

The MAX toolkit and website are based on independent research commissioned and funded by the NIHR Policy Research Programme (Maximising the value of survey data in adult social care (MAX) project and the MAX toolkit implementation and impact project). The views expressed on the website and in publications are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or its arm’s length bodies or other government departments.