

72160 Statistical Analysis

Project formatting instructions

Here's an example of what your report might look like. A covering memo and a Terms of Reference are not required, but a Title page is. Including an Executive Summary or Abstract is optional. You may prefer to use alternative terms (for example Procedure instead of Methodology) in line with reports you have written for other subjects.

Title Page	<i>(This will include your full name, student number and the title of your report.)</i>
Table of Contents	<i>(This lists the sections below.)</i>
1. Introduction to topic and aims	<i>(Give a general discussion about your topic, discuss why the questions you are asking are important and include two or more clearly stated aims.)</i>
<i>(Title page, Table of Contents, Introduction: 20 marks)</i>	
2. Data	
2.1 Data description	<i>(Give a general description of the population, data, its source, how it was collected and name and describe the variables you will be using.)</i>
2.2 Method of sampling	<i>(Describe sampling method you have used and state the sample size. You need to actually take a sample from your population.)</i>
<i>(Data description, method of sampling: 10 marks)</i>	
3. Data analysis	
3.1 Methodology	<i>(Give a brief description of the analysis which will follow.)</i>
3.2 Analysis	<i>(Show your analysis here – an example is shown on page 3 of this document)</i>
<i>(Methodology, data analysis: 40 marks)</i>	
4. Conclusions	<i>(Summarise your findings in the context of your topic/aims and discuss the implications).</i>
5. Recommendations	<i>(Not all reports will have recommendations.)</i>
<i>(Conclusion, recommendations: 20 marks)</i>	
6. List of references	<i>(Data and any publications used should be referenced in correct APA format. You do not need to reference information obtained from the course material.)</i>
7. Appendices	<i>(Include complete sample of your data and course software output.)</i>
<i>(Presentation, referencing: 10 marks)</i>	

Headings: Use bold headings, with larger font for main headings. Imagine this is a business report and you are preparing it for publication.

Tables: Tables should be labelled Table 1, Table 2 in the order they occur. This should go above the table and on the following line there should be a title in italics according to APA style requirements.

For example:

Table 1

Means and standard deviations of cooking times

	Mean (mins)	Standard deviation (mins)
Men	42	23
Women	39	6

The table must be clear and easily read. Data and headings should be centred. If you are copying summary statistic tables R/RStudio or Shiny applications, delete any unnecessary values, round to an appropriate number of decimal places and put in the units of measurement.

Graphs. Graphs should be labelled *Figure 1*, *Figure 2* etc below the graph, with a title. For example

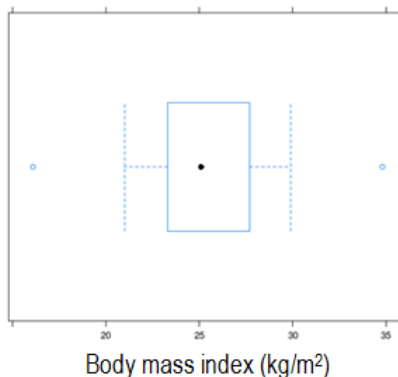


Figure 1. Distribution of BMI variable for NZ athletes

Appendices. This should start on a new page and be labelled Appendix. If there are more than one, start each on a new page and label as Appendix A, Appendix B etc.

Page breaks. As we will be marking online, you need to use page breaks to keep the correct pagination. Use the keyboard shortcut: `Ctrl + Enter` in all versions of Word. Put a page break after your title page, after your Table of Contents and between appendices. Keep all tables on one page, rather than splitting them in the middle.

Footer. Insert a footer and include your full name, student number, 72160 Assignment 3 and page numbers in this footer. You should find this on the *Insert* ribbon.

APA referencing style. If you refer to any publications, you must cite and reference these correctly.

Here's an example of what the analysis section might look like.

3. Data Analysis

3.1 Methodology

(Give a brief description of the analysis which will follow)

3.2 Analysis

3.2.1 Comparison of mean cooking times taken by men and women

To compare the mean times taken by men and women to complete the task of cooking dinner, the side-by-side boxplots were prepared (Figure 1).

(Insert your graph here)

Figure 1. Graph caption goes here

As can be seen, there appears to be little difference in median times, which are around 40 minutes. However, the men have one observation at 240 minutes, which appears to be an outlier. This observation was checked for authenticity, found to be a legitimate value and retained for further analysis. The minimum times are similar for both men and women at around 20 minutes.

The means and standard deviations for cooking time are given in Table 1.

Table 1

Table title goes here

(Insert your table here)

Although the means are similar, the men's times have more variation.

Using a one-sided alternative hypothesis that men take longer than women, on average, to cook dinner, the hypotheses for this were:

$$H_0 : \mu_w = \mu_m$$

$$H_a : \mu_w < \mu_m$$

where μ_w and μ_m are population mean cooking time for women and men respectively.

The test is valid as the sum of the sample sizes exceeds 40.

A two-sample t test $t(79) = 1.01$, $P = .17$ (see software output in Appendix A) confirms that there is no statistically significant difference in mean cooking times taken by men and women.

3.2.2 Comparison of average number of dishes used by men and women

.....

Reporting results of descriptive and inferential statistics in APA format

In reporting the results of statistical tests, report the descriptive statistics (means and standard deviations), the test statistic, degrees of freedom, obtained value of the test, and the p -value. Test statistics and p -values should be rounded to two decimal places. All statistical symbols that are not Greek letters should be italicised (M , SD , n , t , p , etc.)

Tables are useful if you find that a paragraph has almost as many numbers as words. If you do use a table, do not also report the same information in the text. It's either one or the other.

Use a zero before the decimal point with numbers that are less than 1 when the statistic can exceed 1 (e.g. 0.75 cm). Do not use a zero before a decimal fraction when the statistic cannot be greater than 1 (e.g. correlation, proportion and levels of statistical significance).

Mean and Standard Deviation are most clearly presented in parentheses:

In the sample of 100 *Statistical Analysis* students the average Assignment 1 mark was relatively high ($M = 72.5\%$, $SD = 9.4\%$).

In the *Statistical Analysis* course the average Assignment 1 mark for the sample of 100 students was 72.5% ($SD = 9.4\%$).

Percentages are also most clearly displayed in parentheses with no decimal places:

Nearly half (49%) of participants in the sample were working full time.

Correlation:

The two variables were strongly positively correlated, $R = .91$.

Correlation could be presented in the table together with the mean value and standard deviation:

Table 2.

Means, Standard Deviations, and Correlations in the Crop and Area Dataset

Variable	M	SD	R
<i>Area</i>	1173.12	431.69	
<i>Crop</i>	186.77	55.43	.69

Note. M and SD are used to represent mean and standard deviation, respectively.

Regression results are often best presented in a table. APA doesn't say much about how to report regression results in the text. You may use the following format:

$$\text{Crop} = 83.13 + 0.09 \text{ Area}, R = .69, R^2 = 47\%$$

Table 3.

Regression Results Using Crop Variable as the Response Variable

Variable	<i>b</i>	<i>SE</i>	<i>beta</i>	<i>R</i>	Fit
Model 1					
(Intercept)	83.13**	23.72			
Area	0.09**	0.02	0.69	.69**	
					$R^2 = .473$
					$F(1, 24) = 21.57$

Note. * indicates $p < .05$; ** indicates $p < .01$.

A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant.

b represents unstandardized regression weights; *SE* represents the standard error of the unstandardized regression weights; *beta* indicates the beta-weights or standardized regression weights, *R* is correlation and R^2 measures goodness-of-fit of the regression model; *SE*, *beta* and *F*-test are not covered in this course.

Confidence interval should be presented in squared parentheses with confidence level:

Mean: $M = \$300.5$, 95% CI [$\255.75, $\$344.25$]

Proportion: $p = .55$, 95% CI [.45, .65]

***p*-values**

p = value (no leading zeros); e.g. $p = .07$

$p < .01$ (if the value is less than .01)

***t*-test** should include test statistic, degrees of freedom, value of the test and p -value, i.e. $t(df) = \text{value}$, p -value:

one sample *t*-test: One sample *t*-test showed that the difference in quiz scores between the current sample ($n = 6$, $M = 3.45$, $SD = 2.11$) and the hypothesized value ($\mu = 3.00$) was not statistically significant, $t(5) = 1.25$, $p = .26$, 95% CI [1.24, 5.66].

two sample *t*-test: Two sample *t*-test showed that the difference in Assignment 1 marks between *Statistical Analysis* students in Trimester 1 and Trimester 2 was *marginally* significant, $t(5) = 1.25$, $p = .08$.

match pairs *t*-test: The 25 participants had an average difference from pre-test anxiety scores of -4.8 ($SD = 5.5$), indicating the anxiety treatment resulted in significant decrease in anxiety levels, $t(24) = -4.36$, $p < .01$ (one-tailed).

Results of the *t*-tests could be also presented in a tabular format:

Table 4.

Two Sample t-tests for Variables in the GymSurvey Dataset by Gender

Variable	Females		Males		<i>t</i> (20)	<i>p</i>	95% CI
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Age	35.2	3.5	35.9	3.9	-.5	.6	[-3.9, 2.4]
Exercise	2.2	1.3	3.3	2.1	-1.0	.2	[-2.6, 0.4]
BMI	24.7	4.5	26.1	2.8	-.9	.4	[-4.6, 1.8]

Chi-square statistics are reported with degrees of freedom and sample size in parentheses, and *p*-value:

As can be seen by the counts in the two-way table in Table 5, there is a significant relationship between Active Marketing and Growth Rate, $\chi^2(4, n = 356) = 18.51, p < .01$.

Or

Based on the result of the χ^2 -test: $\chi^2(4, n = 356) = 18.51, p < .01$ we reject the null hypothesis and conclude that there is an association between application of the active marketing strategy and the growth rate of a company.

Table 5.

Crosstabulation of Active Marketing and Growth Rate

Active Marketing	Growth Rate			χ^2
	High	Middle	Low	
Current	51	43	22	18.51**
Former	92	28	21	
Never	68	22	9	

Note. ** $p < .01$.

Based on:

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.