National College of Ireland

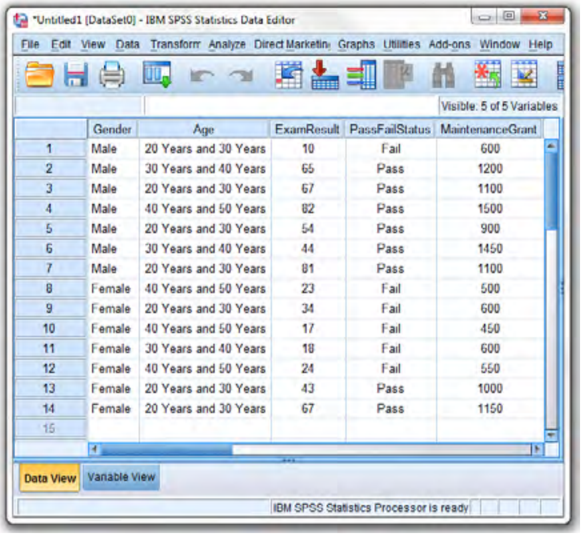
QUANTITATIVE GUIDE TO DISSERTATION CONSTRUCTION

Mr Jonathan Lambert and Dr Colette Darcy

Table of Contents

1 SPSS Statistical Analysis Package

The SPSS Statistical Analysis Package is a self-contained software package that allows the researcher to perform standard statistical evaluations of results, as well as allowing for advanced data modelling and exploration.



	Gender	Age	ExamResult	PassFailStatus	MaintenanceGrant
1	Male	20 Years and 30 Years	10	Fail	600
2	Male	30 Years and 40 Years	65	Pass	1200
3	Male	20 Years and 30 Years	67	Pass	1100
4	Male	40 Years and 50 Years	82	Pass	1500
5	Male	20 Years and 30 Years	54	Pass	900
6	Male	30 Years and 40 Years	44	Pass	1450
7	Male	20 Years and 30 Years	81	Pass	1100
8	Female	40 Years and 50 Years	23	Fail	500
9	Female	20 Years and 30 Years	34	Fail	600
10	Female	40 Years and 50 Years	17	Fail	450
11	Female	30 Years and 40 Years	18	Fail	600
12	Female	40 Years and 50 Years	24	Fail	550
13	Female	20 Years and 30 Years	43	Pass	1000
14	Female	20 Years and 30 Years	67	Pass	1150
15					

Figure 1: Case 1 - SPSS Data View

The SPSS Statistics engine has two views: the **Data View** and the **Variable View**. Within the Data View, **Figure 2**, each variable is represented as a column within the view. Each entry in a particular column represents a particular response value associated with the variable. More importantly, each row represents all responses captured for all variables, for a single respondent.

For example, the row labelled **6** in **Figure 2** captures the data for a Male, aged between 30 years and 40 years, with an exam result of 44%, achieving an overall Pass and that this respondent has indicated that they are in receipt of €1450 euro from a maintenance grant.

As can be seen in **Figure 2**, there are five variables being measured: Gender; Age; Exam Result; Pass-Fail status and Maintenance Grant receipts.

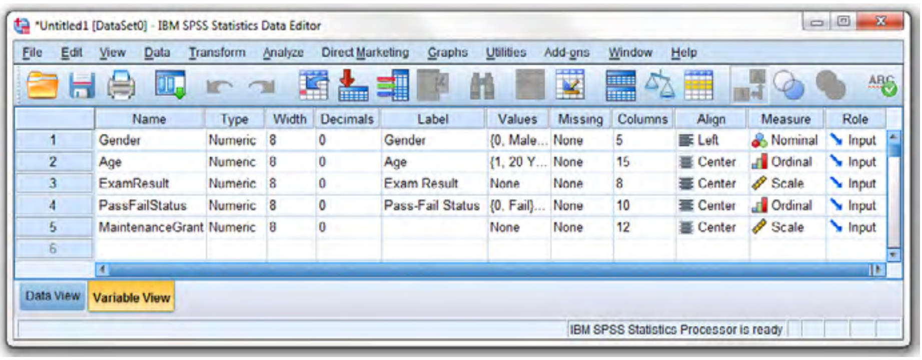


Figure 2: Case 1 - SPSS Variable View

Each variable within SPSS is coded through the **Variable View** window, **Figure 3**. This window informs SPSS of some fundamental properties associated with the variables being analysed. The important fields are: **Name**, **Values** and **Measure**.

All variables must be assigned a descriptive name, this name **cannot** contain spaces. All variables are measured at a particular scale of measurement, either: **Nominal**, **Ordinal**; or **Scale** for variables measured at the **Interval** or **Ratio** scale of measurement.

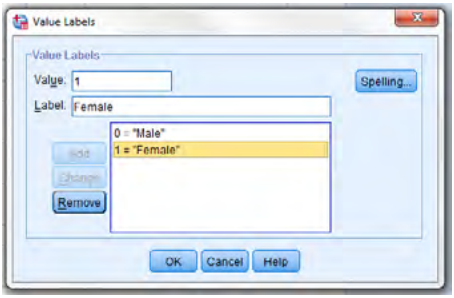


Figure 3: Case 1 - SPSS Value Labels

In the case of variables measured on a Nominal or Ordinal scale of measurement, SPSS allows you to code numerical values against possible variable responses.

For example, **Figure 4** shows the coding's for the Gender variable, with a value of '**0**' representing a '**Male**' and a value of '**1**' representing a '**Female**' response.

2 Case 1: Impact of Gender and Age on Examination Performance

The first case that we will consider is a hypothetical quantitative research question concerning the factors that influence student performance in terminal examinations. The hypothesis under consideration is presented in **Table 3** below.

Hypothesis
The academic performance of undergraduate first year students in college terminal examinations is influenced by a student's Gender and Age, as well as their financial stress while within college.

Table 1: Hypothesis being Tested

In order to verify a particular hypothesis it is essential that all variables be appropriately defined and understood. This is possibly one of the most important steps and ultimately will contribute to any proposed model. In the case above, the following variables have been identified, listed in **Table 4,** and thus will form questions within any survey being used for data capture.

Variables of Interest	Scale of Measurement
Pass Fail Status	Ordinal
Gender	Nominal
Age	Ordinal
Exam Result	Scale
Student Maintenance Grant	Scale

Table 2: Variables Identified from Hypothesis

2.1 SPSS Descriptive Statistics

The first step in any exploratory study concerns the summary of the data captured through sampling. This summary can be accomplished through tables, charts and graphs; and primarily involves the construction of: Bar Charts; Pie Charts; Histograms; Box and Whisker Plots; as well as the presentation of measures of centre and dispersion, such as: the mean, median and mode; as well as the range, the variance and standard deviation of our sample data.

There are many descriptive techniques that can be used to leverage trends and patterns from a data set, we only consider the most popular here in this section.

2.1.1 Frequencies

A frequency analysis allows the researcher to consider a variable, say Gender, and explore the distribution of data values associated with that variable. In particular, the total counts of observations that fall into either of the two categories: Male or Female. This output can be represented within tables and bar charts.

Figure 5 depicts the process for selecting a Frequency Analysis of a variable. Select '*Analyze'*, then from the drop down list select '*Descriptive Statistics'*, finally select '*Frequencies'* from the subsequent drop down list.

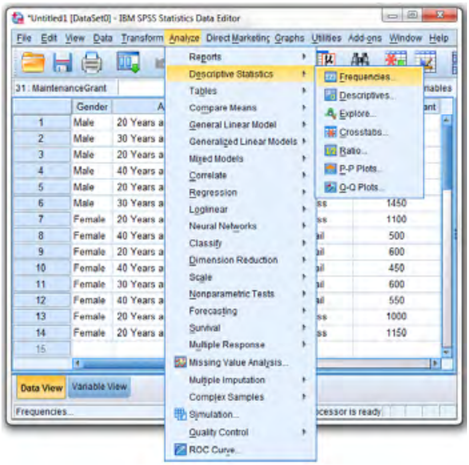


Figure 4: Selecting a Frequency Analysis

The Variable View window is depicted in **Figure 6**; the descriptive analysis can be undertaken on one or more variables. In this case we choose the single variable 'Gender' and ensure that this variable is listed under the 'Variable(s)' Pane.

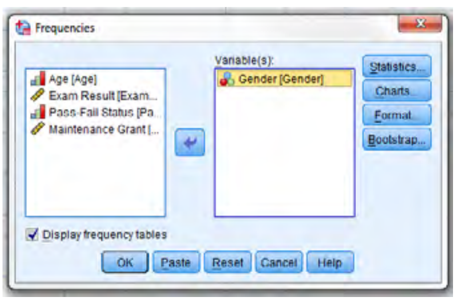


Figure 5: Defining the Variables

A number of additional options are available when undertaking a Frequencies Analysis. The Variable View window, **Figure 6**, shows a number of buttons: 'Statistics' and 'Charts'; which we will explore in this case.

Choosing the 'Statistics' button, an additional window is presented, shown in **Figure 7**. The statistics of interest for the 'Gender' variable are the Mode.

Choosing the 'Charts' button, a number of options for charting the results are presented, shown in **Figure 8**. We choose 'Bar Chart(s)' from the 'Charts' window, **Figure 8**.

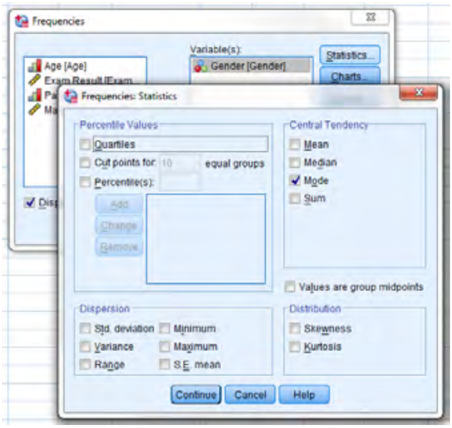


Figure 6: Choosing the Statistics

The typical output from this selection is shown in **Tables 5** and **6** and **Figure 8** below. **Table 5** identifies the statistic selected earlier: **Mode**, and in this case presents a mode of 2. This value is the encoding used within SPSS and signifies 'Female'. **Table 6** presents the actual frequencies accounted for

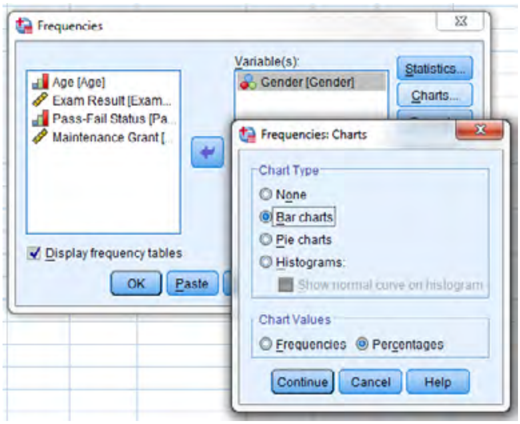


Figure 7: Choosing the Charts

at each level of the Gender variable, and their frequencies as percentages of the total number of cases examined.

Statistics

Gender

N	Valid	14
	Missing	0
Mode		2

Table 3: Descriptive Statistics

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	6	42.9	42.9	42.9
	Female	8	57.1	57.1	100.0
	Total	14	100.0	100.0	

Table 4: Summary Statistics

Figure 9 depicts a Bar Chart representation of the Gender distribution. The horizontal axis depicts the levels of measurement: Male and Female; with the vertical axis representing the actual number of observations falling within each of these classes.

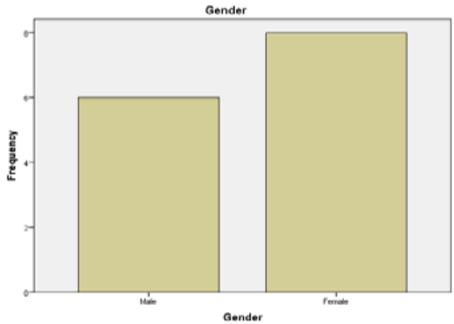


Figure 8: Bar Chart Gender Distribution

It is important to realise that the scale of measurement associated with your variable fully determines what statistics can be calculated for the variable. **Figure 7** depicts four options: '**Mean**'; '**Median**'; '**Mode**'; and '**Sum**'.

As we are undertaking a frequencies analysis of the variable 'Gender' which is measured on a nominal scale of measurement the only applicable statistic is 'Mode', as all others do not produce understandable results.

	Variable Scale of Measurement			
Statistical Measure	Nominal	Ordinal	Interval	Ratio
Mean			X	X
Median		X	X	X
Mode	X	X	X	X
Range			X	X
Variance + Standard Deviation			X	X

2.1.2 Descriptives

In this section we consider the procedures for listing relevant Descriptive Statistics through SPSS. In particular measures of centre, such as the: mean, median and mode; and measures of dispersion, such as the: range, variance, and standard deviation.

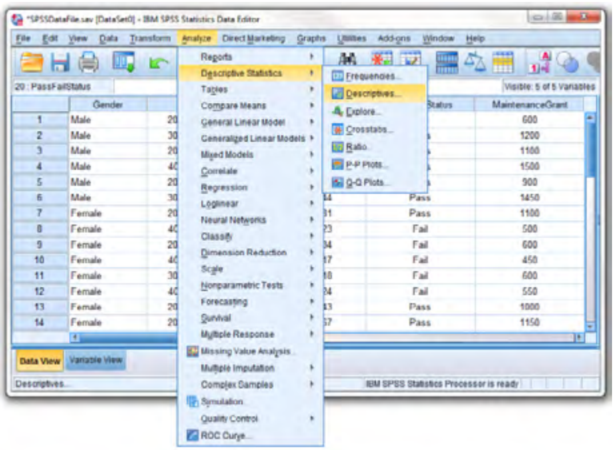


Figure 9: Selecting Descriptive Statistics

Figure 10 depicts the procedure for selecting Descriptive Statistics through SPSS. Select ‘*Analyze*’, then select ‘*Descriptives*’ from the drop down list.

Within the Descriptives window, **Figure 11**, we define the variables to be analysed. In this case we choose: Exam Results and Maintenance Grant.

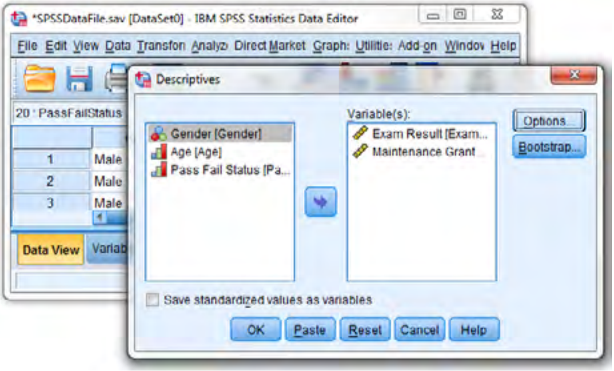


Figure 10: Selecting Variables to be Analysed

The 'Options' button, shown in **Figure 11**, allows for the selection of the appropriate statistics. A list of available statistics is shown in **Figure 12**.

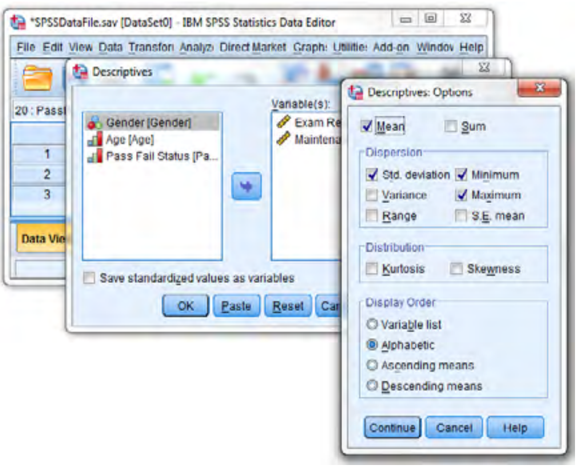


Figure 11: Selecting Appropriate Statistics

In **Figure 12** we have selected for consideration and analysis the: Mean, Standard Deviation, Minimum and Maximum statistics. The results of this selection are shown in **Table 8**.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Exam Result	14	10	82	44.93	24.656
Maintenance Grant Recieved	14	450	1500	907.14	357.264
Valid N (listwise)	14				

Table 6: SPSS Descriptive Statistics Results Output

Table 8 depicts six columns. The first column, not labelled, lists all the variables that have been included in the analysis.

The second column, labelled 'N', depicts the sample size associated with each variable. The third and fourth columns list the minimum and maximum values for each of the variables analysed.

Finally, the columns labelled 'Mean' and 'Standard Deviation' list the mean and standard deviation for each of the variables, respectively.

2.2 SPSS Inferential Statistics (Parametric)

A key challenge for the researcher when undertaking a quantitative research project, and as part of the Results and Analysis phase of their design, is whether their experimental data merits an application of Parametric Analysis procedures or a Non-Parametric Analysis approach.

The importance of this distinction **cannot and should not** be **underestimated** by the researcher. Ultimately both procedures are dependent on the scale of measurement associated with your variables.

In the case were a variable is measured on an **interval** or **ratio** scale – **parametric analysis is appropriate**.

In the case were a variable is measured on a **nominal** or **ordinal** scale of measurement, **non-parametric analysis is appropriate**.

With that said, an **additional key prerequisite** for using **parametric procedures** is the assumption that your variable data, sample data, has been drawn from a reasonably **normal distribution**.

This section assumes that our variables under analysis have met with the prerequisite conditions for the use of Parametric tests.

3.2.1 Single Sample t-Test

In this section we consider the procedure for performing a Single Sample t-Test through SPSS.

Purpose: A statistical procedure to test if the average (**mean**) of a single sample of observations significantly deviates from a predefined expected value.

Example: Prior evidence suggests that students, on average, score 75% in a National Mathematics Exam. Entry to a prestigious School requires that applicants achieve this level. Is it still the case that students achieve this grade on average or has it changed?

Null Hypothesis: There is no difference between a samples' mean value and the hypothesised value.

Alternative Hypothesis: There is a significant difference between the measures.

Decision: If **p-value** is less than **0.05** we reject the Null Hypothesis in favour of the Alternative Hypothesis

An example:

If we consider the data captured through **Case 1** for the variable 'Exam Result', the average exam grade for this sample, shown in **Table 8**, is **44.93%**. If we want to test if this sample provides evidence to suggest that the actual exam grade is different from **40%** we would perform a Single Sample t-Test.

Figure 13 depicts the selection process for choosing a Single Sample t-Test. Select '*Analyze'*, then select '*Compare Means'* from the drop down list, finally select '*One Sample t-Test'* from the subsequent drop down list.

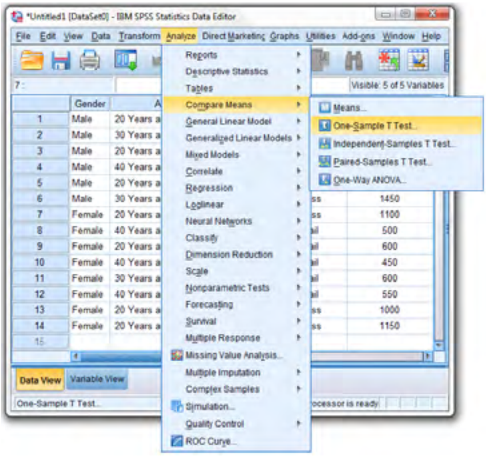


Figure 12: Selecting a Single Sample t-Test

Figure 14 depicts the Single Sample t-Test variable selection window. The 'Test Variable(s)' pane is where we define the variable to be tested, in this case Exam Result. In the field 'Test Value' we define the hypothesised test value for the population; in this case we are hypothesising that the average exam result for a class should be **40%**.

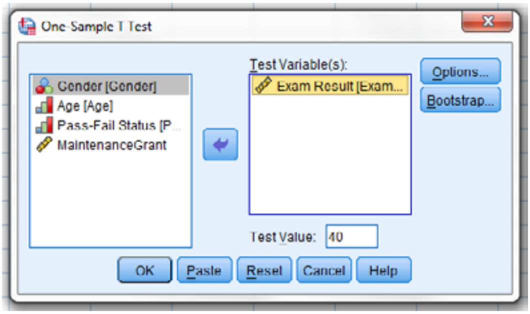


Figure 13: Defining the Variables

Table 9 and **10** depict the typical output from a Single Sample t-Test calculation. **Table 9** presents the usual statistics for our variable, in this case Exam Result. From this we can see that the average of our sample was recorded at **44.93%** with a Standard Deviation of **24.656**.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Exam Result	14	44.93	24.656	6.590

Table 7: Descriptive Statistics – Exam Result

Table 10 presents the results of the test. The important column, at this stage, is the column labelled '**Sig (2-tailed)**' and the entry in this column.

One-Sample Test

	Test Value = 40					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Exam Result	.748	13	.468	4.929	-9.31	19.16

Table 8: Single Sample t-Test Results Output

If the '**Sig (2-tailed)**' value is less than **0.05** we reject the null hypothesis of no difference between the average observed and our hypothesised average, in favour of the alternative that there is a significant difference between the averages.

In this case we do not reject the null hypothesis (**0.468 > 0.05**) and infer that there is insufficient evidence to suggest that the average Exam grade of this population is not **40%** based on the sample data observed.

Note: Like all tests that we have undertaken up to this point it is important to realise that each test has a number of preconditions that must be met in order for the result of the test to be reliable and robust.

In the case of a Single Sample t-Test there are two important preconditions. The first is normality; we must test to ensure that the population from which the sample has been captured is somewhat normally distributed. The second requires homogeneity of sample variances.

3.2.2 Independent Samples t-Test

In this section we consider the procedure for performing an Independent Samples t-Test through SPSS.

Purpose: A statistical procedure to test if the averages (**means**) of two samples of observations are significantly different from each other.

Example: The results of a group of students in a Mathematics examination were captured. The class is divided into two independent groups: males and females. Is there significant evidence to infer that the average examination results of males differs to that of females?

Null Hypothesis: There is no difference between the average Exam results for Males compared to that of Females.

Alternative Hypothesis: There is a significant difference between the measures.

Decision: If **p-value** is less than **0.05** we reject the Null Hypothesis in favour of the Alternative Hypothesis

An example:

Let us consider the data captured through **Case 1** for the variable 'Exam Result'. This set of results can be divided into two groups, the results for the Males and the results for the Females. Is there significant evidence to suggest that there is a difference between the Exam performance of Males compared to that of Females?

To perform this type of analysis we need two variables. The first is the test variable, in our case Exam Result; the second is a grouping variable, in our case Gender. The Gender variable effectively divides the Exam Result variable into two Independent Samples.

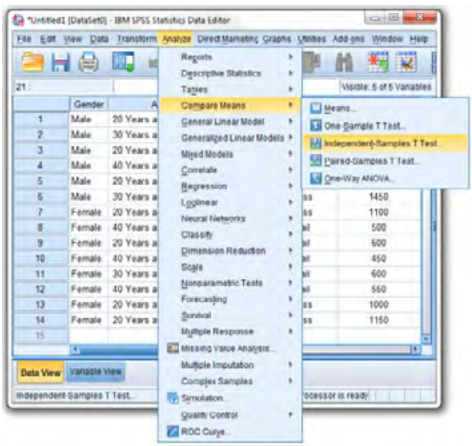


Figure 14: Selecting an Independent Samples t-Test

Figure 15 depicts the selection process for choosing an Independent Samples t-Test. Select '*Analyze*', then select '*Compare Means*' from the drop down list, finally select '*Independent Samples t-Test*' from the subsequent drop down list.

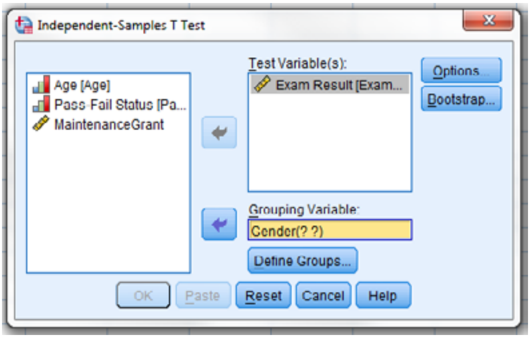


Figure 15: Defining the Variables

Figure 16 depicts the Independent Samples t-Test variable selection window. The 'Test Variable(s)' pane is where we define the variable to be tested, in this case Exam Result. In the field 'Grouping Variable' we define the variable that will partition our test variable, in this case Gender.

Once we have defined our variables it is important that we define the coding for our Grouping Variable.

This coding has previously been defined when we defined our variables through the SPSS Variable View, **Figure 3**. These values are the values entered in the Values field of **Figure 3**.

Figure 17 depicts how we define our groups. In this example we defined '**0**' for Males and '**1**' for Females.

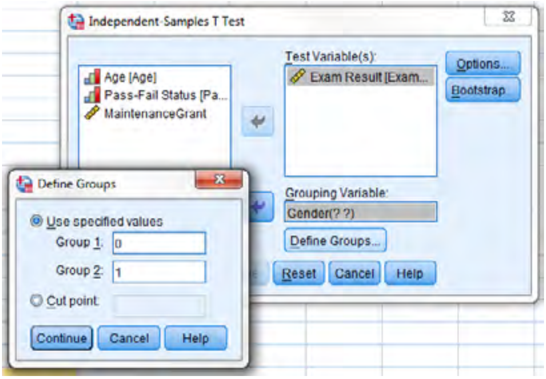


Figure 16: Defining the Grouping Variable

Table 11 and **12** depict the typical output from an Independent Samples t-Test. From **Table 11** we can see that the mean Exam Grade for Males is **53.67%** with a Standard Deviation of **24.937**; and the mean Exam Grade for Females is **38.38%** with a Standard Deviation of **23.880**.

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Exam Result	Male	6	53.67	24.937	10.181
	Female	8	38.38	23.880	8.443

Table 9: Descriptive Statistics – Exam Result by Gender

Table 12 presents a number of important statistics that need to be considered before an inference on differences can be made. First, the main row of column headings: **F**; **Sig.**; **t**; **df**; and **Sig. (2 tailed)** are what we are interested in.

When conducting an Independent Samples t-Test there are two variants of the test.

The first assumes that the two samples have been drawn from populations that have the same population variances; the second assumes that the two samples have been drawn from populations that have significantly different variances. **Table 12** depicts the results of two different types of Independent Samples t-Test, both listed across the rows: 'Equal Variances Assumed' and 'Equal Variances not Assumed'.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Exam Result	Equal variances assumed	.025	.877	1.164	12	.267	15.292	13.138	-13.333	43.916
	Equal variances not assumed			1.156	10.646	.273	15.292	13.226	-13.937	44.521

Table 10: Independent Samples t-Test Results Output

To decide on the appropriate variant of the t-Test, **Table 12** also lists two columns under the heading: "Levene's Test for Equality of Variances".

This test is associated with the null hypothesis of Equal Variances and in order to reject this hypothesis the significance value listed under the '**Sig**' column must be less than **0.05**. Clearly this is not the case (**0.877 > 0.05**) and as such we proceed to interpret the results assuming equal population variances.

To infer a difference between the mean Exam grades of Males versus Females we now consider the portion of **Table 12** under the heading: "t-Test for Equality of Means".

As the previous test indicated that we should proceed under the assumption of equal population variances, we will only consider entries across the first row, under those column headings.

The statistic of interest for this test is listed under the column heading: '**Sig. (2 tailed)**'. If this value is less than **0.05** there is significant evidence to suggest a difference.

In this case **.267** is not less than **0.05** and as such our analysis indicates that there is **no difference** between the average Exam grades of Males compared to Females.

Note: Like all tests that we have undertaken up to this point it is important to realise that each test has a number of preconditions that must be met in order for the result of the test to be reliable and robust.

In the case of an Independent Samples t-Test there are two important preconditions. The first is normality; we must test to ensure that the populations from which the samples have been captured are somewhat normally distributed. The second requires homogeneity of sample variances.

3.2.3 Single Factor ANOVA

In this section we consider the procedure for performing an Analysis of Variance through SPSS. The Analysis of Variance technique is appropriate when we have more than two independent groups and we wish to explore if there exists significant differences across those groups.

Purpose: A statistical procedure to test if the averages (**means**) of more than two samples of observations are significantly different from each other.

Example: The results of a group of students in a Mathematics examination were captured. The class is divided into three independent groups based on age: 20 to 30 years; 30 to 40 years and 40 to 50 years of age. Is there significant evidence to infer that the average examination results across age groups differ?

Null Hypothesis: There is no difference between the average results for any of the groups based on age.

Alternative Hypothesis: There is a significant difference between the measures.

Decision: If **p-value** is less than **0.05** we reject the Null Hypothesis in favour of the Alternative Hypothesis

It is important to note that we are looking at a single variable, for example: Exam Result; and we are segmenting the Exam Result observations into groups based on a factor that has a number of measurements, say Age.

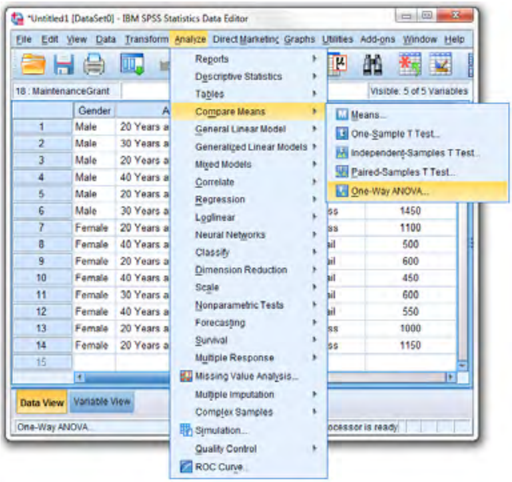


Figure 17: Selecting an ANOVA Test

Figure 18 depicts the selection process for choosing an ANOVA test. Select '*Analyze*', then from the drop down list select '*Compare Means*', finally select '*One-way ANOVA'* from the subsequent drop down list.

Figure 19 depicts the ANOVA variable selection window. The 'Dependent List' pane is where we define the variable to be segmented into groups. The 'Factor' field is where we define the variable that determines the groups.

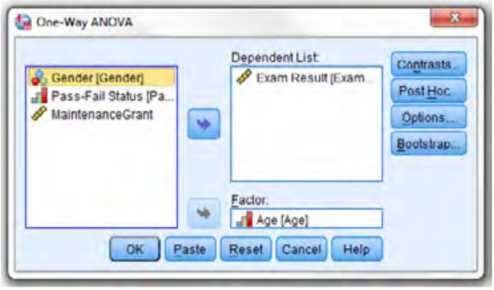


Figure 18: Defining the Variables

Homogeneity of Variances is a necessary condition in order to undertake an ANOVA calculation. To test for Homogeneity of Variance we can select this test through the 'Options' button shown in **Figure 19**.

Table 13 depicts the typical output for this test. The most important entry, at this stage, is the entry in the column labelled `**Sig.**'

Test of Homogeneity of Variances

Exam Result

Levene Statistic	df1	df2	Sig.
.211	2	11	.813

Table 11: Results for Homogeneity of Variances in Groups

The null hypothesis associated with the Leven's test is that homogeneity of variances is assumed. As the 'Sig' value (**0.813**) in **Table 13** is not less than **0.05**, we do not reject this assumption and thus proceed knowing that homogeneity of variances is assured.

ANOVA

Exam Result

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	550.405	2	275.202	.412	.672
Within Groups	7352.524	11	668.411		
Total	7902.929	13			

Table 12: ANOVA Results Output

The result of the ANOVA calculation is presented in **Table 14**. If the '**Sig.**' value is less than **0.05** we reject the null hypothesis of no difference between the groups, in favour of the alternative that there is a significant difference between the groups.

In this case the '**Sig.**' value is: **'.672**' indicating that we cannot reject the null hypothesis of no difference between the group means; and infer that there is insufficient evidence to assume that age is an important factor that influences Exam Results.

Note: Like all tests that we have undertaken it is important to realise that each test has a number of preconditions that must be met in order for the result of the test to be reliable and robust.

In the case of an Analysis of Variance there are two important preconditions. The first is normality; we must test to ensure that the populations from which the samples have been captured are somewhat normally distributed. The second requires homogeneity of sample variances.

3.2.4 Correlation (Pearson and Spearman)

Up-to this point of our analysis we have primarily been concentrating on statistics that test for the existence of differences between groups within a variable. This section deals with Correlation, a procedure that will allow us to test if changes in one variable (**independent variable**) result in changes in a second variable (**dependent variable**).

For example, we might be interested in exploring if changes in financial means will contribute to changes in exam score or alternatively whether a person's age impacts on exam score.

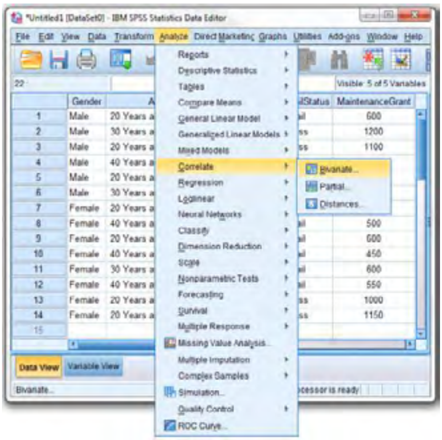


Figure 19: Selecting a Correlation Test

Figure 20 depicts the selection process for choosing a Correlation test. Select '*Analyze*', then select '*Correlate*' from the drop down list, finally select '*Bivariate*' from the subsequent drop down list.

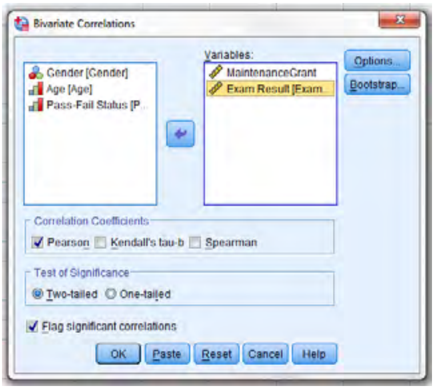


Figure 20: Defining the Variables (Pearson)

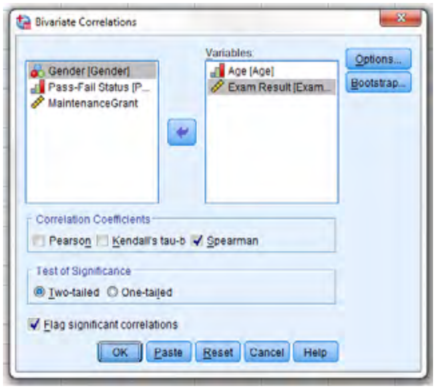


Figure 21: Defining the Variables (Spearman)

Figure 21 and **22** depict a number of options with respect to the type of correlation that we want to conduct.

Depending on the variable type there are three correlation techniques: Pearson; Kendall or Spearman. When both variables are measured on an interval or ratio scale of measurement we will choose a Pearson test. If at least one of the variables is measured on an ordinal scale of measurement we will choose a Spearman or alternatively a Kendall test.

Consider the case, as depicted in **Figure 21**. We are interested in whether the examination grade is influenced by the level of maintenance grant received. That is does being in receipt of a higher maintenance grant result in higher or lower exam grades.

Table 15 depicts the results of the correlation test. To interpret these results we identify a cell in the table where the 'Maintenance Grant' variable intersects with the 'Exam Result' variable. The first value: '**.584**' indicates the strength of association between Maintenance Grant and Exam Result, and the second value: '**.028**' indicates the significance of this result.

Correlations

		Maintenance Grant	Exam Result
Maintenance Grant	Pearson Correlation	1	.584*
	Sig. (2-tailed)		.028
	N	14	14
Exam Result	Pearson Correlation	.584*	1
	Sig. (2-tailed)	.028	
	N	14	14

*. Correlation is significant at the 0.05 level (2-tailed).

Table 13: Pearson Correlation Results

In this case the results suggest that there exists a moderate association between both variables '**.584'**, and the result is statistically significant '**.028**' meaning that it is unlikely that these results are due to chance.

Correlations

			Age	Exam Result
Spearman's rho	Age	Correlation Coefficient	1.000	.805**
		Sig. (2-tailed)	.	.001
		N	14	14
	Exam Result	Correlation Coefficient	.805**	1.000
		Sig. (2-tailed)	.001	.
		N	14	14

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

Table 14: Spearman Correlation Results

Considering the case as depicted in **Figure 22**, the ordinal variable Age is being correlated against the ratio variable Exam Grade. The results of this test are presented in **Table 16**. In this case the results suggest that there exists a strong association between both variables '**.805'**, and the result is statistically significant '**.001**' meaning that it is unlikely that these results are due to chance.

2.3 SPSS Inferential Statistics (Non-Parametric)

In this section we present a number of alternatives to the parametric tests considered in section 3.2 and in particular alternatives to the: Single Sample t-Test; Independent Samples t-Test; Analysis of Variance (ANOVA) and the Pearson correlation test.

Non-parametric analysis provides the researcher with two advantages over the parametric procedures considered in section 3.2.

The first is concerned with a relaxation on the preconditioned assumption that our sample data has been effectively drawn from normally distributed populations and the second allows us to analyse variables measured on an ordinal scale of measurement through the use of non-parametric procedures that are based on the ranking of data.

The alternative Non-Parametric tests in place of the Parametric versions are depicted in **Table 17**.

Parametric Test	Non-Parametric Equivalent
Single Sample t-Test	**Wilcoxin Signed-Rank Test**
Independent Samples t-Test	**Mann-Whitney U-Test**
Single Factor ANOVA	**Kruskal-Wallis H-Test**
Pearson Correlation	**Spearman Rank Correlation**

Table 15: Parametric Test Alternatives when Normality is Violated

2.3.1 Single Sample Test (Wilcoxin Signed-Rank Test)

The Wilcoxin Single Sample Signed-Rank Test is an alternative test to the Single Sample t-Test. It is used to compare the median rank of a sample against a hypothetical median value.

This test is appropriate when the normality assumption is violated and thus a Single Sample t-Test result cannot be statistically relied upon.

Figure 23 depicts the selection process for choosing a Wilcoxin Single Sample Signed-Rank test. Select *'Analyze'*, then select *'Nonparametric Tests'* from the drop down list, finally select *'One Sample'* from the subsequent drop down list.

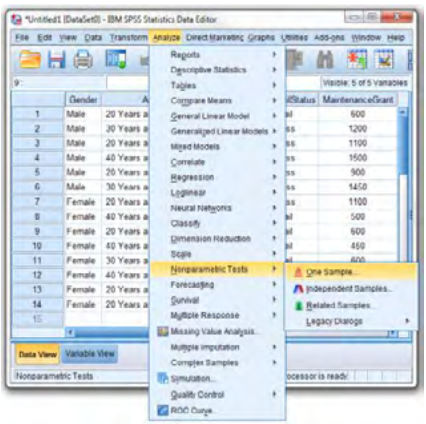


Figure 22: Selecting a Wilcoxin Single Sample Signed Rank Test

2.3.2 Independent Samples Test (Mann-Whitney U-test)

The Mann-Whitney U-Test is an alternative test to an Independent Samples t-Test. It is used to compare the median rank of two samples against each other so as to identify if they are significantly different.

This test is appropriate when the normality assumption is violated and thus an Independent Samples t-Test result cannot be statistically relied upon.

Figure 24 depicts the selection process for choosing a Mann-Whitney U-test. Select '*Analyze*', then select '*Nonparametric Tests*' from the drop down list, then select 'Legacy Dialogs' and final from the subsequent drop down list select 'Independent Samples'.

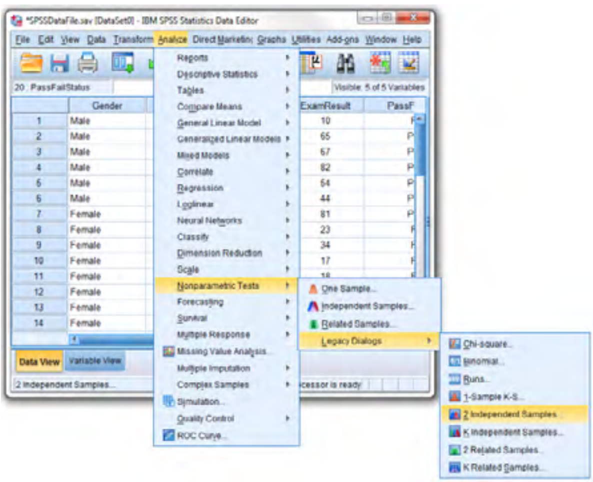


Figure 23: Selecting the Non-Parametric Independent Samples Test

2.3.3 K-Independent Samples Test (Kruskal-Wallis H)

The Kruskal-Wallis H-Test is an alternative to the Analysis of Variance test. It considers the relationship between the mean ranks of more than two groups so as to identify if they are significantly different.

This test is appropriate when the normality assumption is violated and thus an Independent Samples t-Test result cannot be statistically relied upon.

Figure 25 depicts the selection process for choosing a Mann-Whitney U-test. Select '*Analyze*', then select '*Nonparametric Tests*' from the drop down list, then select 'Legacy Dialogs' and final from the subsequent drop down list select 'Independent Samples'.

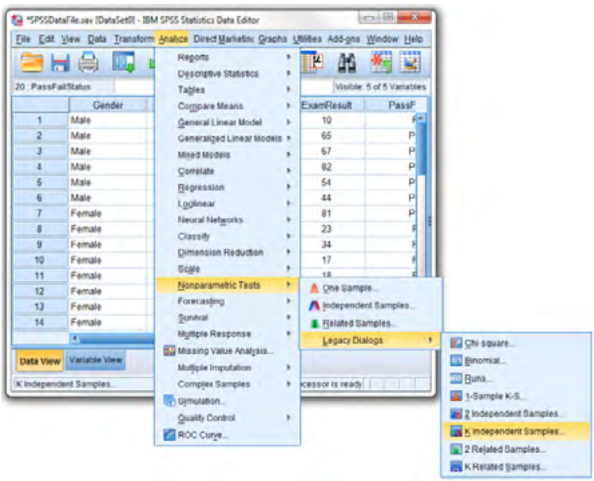


Figure 24: Selecting the Non-Parametric K Independent Samples Test

3 Survey Item (Consistency and Reliability – Cronbach's Alpha)

In this section we consider a simple example on how to test a Surveys Instrument Items to ensure that the items are measuring the same latent concept. This type of analysis is known as Internal Consistency and Reliability testing.

The statistic that we will rely upon to measure Survey Items internal consistency and reliability will be Cronbach's Alpha.

Item	Item Question
Item 1	I worry about my privacy and data security while using the internet.
Item 2	I worry that if I use my credit card to buy something on the internet my credit card number will be obtained / intercepted by someone else.
Item 3	I worry about people online not being who they say they are.
Item 4	I feel that identity theft could be real privacy risk.
Item 5	I worry that if I use internet with my mobile phone and someone steals it, he/she can find out some of my personal information or data.
Item 6	I'm familiar with data protection and securing while using the Internet in general.

Table 16: Items Intended to Measure Privacy and Data Security Concerns

As an example let us consider the six questions items listed in **Table 18**. These questions have been taken from a Survey Instrument developed by Tuunainen and Pitkanen (2012) and are a subset of 19 questions developed to measure different aspects of Awareness of Privacy on Online Social Networking Sites.

The six items listed in **Table 18** are designed to measure the same latent factor, namely: 'Privacy and Data Security Concerns in General'.

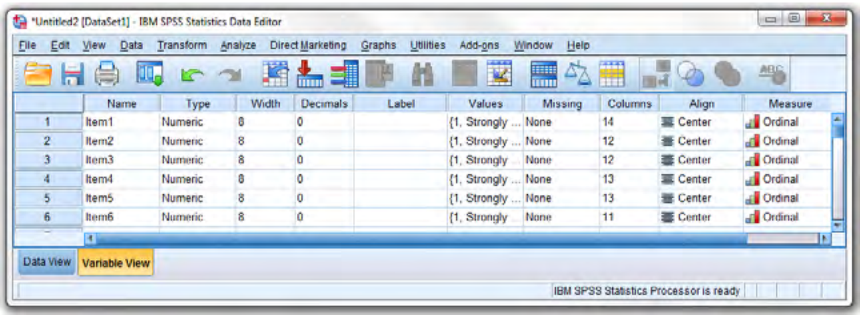


Figure 25: Defining the Variables

Figure 26 depicts the relevant encoding of these survey items within SPSS. Each survey item was measured on a 7 point Likert Scale and **Figure 27** depicts the encoding of each items' values in SPSS.

Figure 28 depicts a contrived set of responses for each of the six survey items, and in particular, represents the fictitious responses of 11 participants.

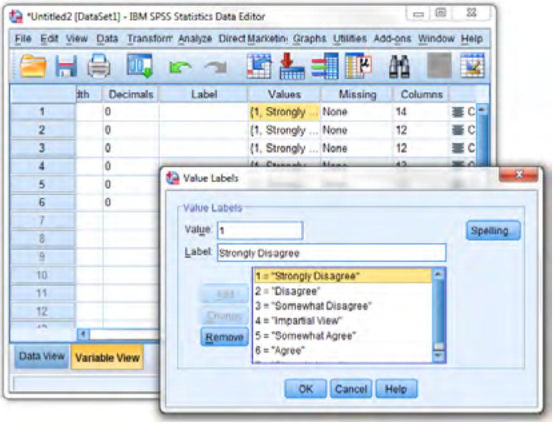


Figure 26: Defining the Likert Codings

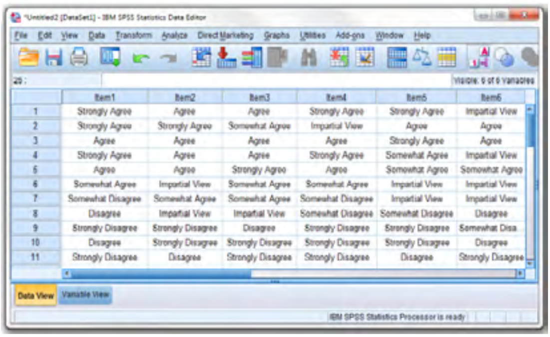


Figure 27: Sample Survey Response Data

Figure 29 depicts the selection process for choosing a Cronbach's Alpha Reliability Analysis. Select '*Analyze*', then select '*Scale*' and from the drop down list select 'Reliability Analysis'.

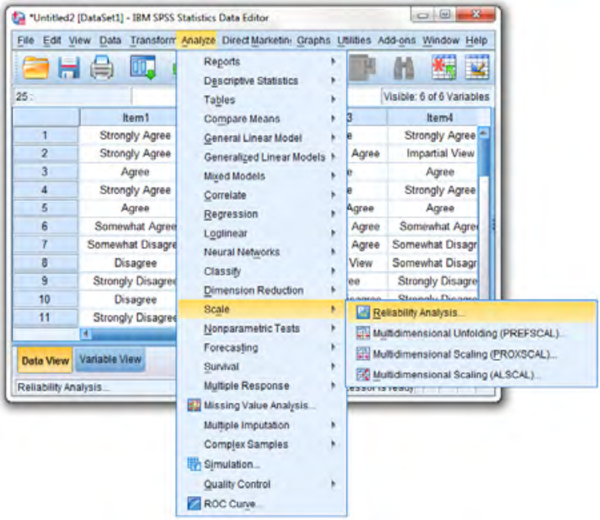


Figure 28: Selecting a Reliability Analysis Cronbach Alpha Test

A preliminary understanding of the directionality of the item scales and in particular the inter-item correlations can be achieved through the selection of 'Correlations' from the statistics option, presented when you choose the 'Reliability Analysis' option. These results are depicted in **Table 19**.

Inter-Item Correlation Matrix

	Item1	Item2	Item3	Item4	Item5	Item6
Item1	1.000	.891	.853	.915	.902	.813
Item2	.891	1.000	.912	.842	.934	.805
Item3	.853	.912	1.000	.921	.867	.793
Item4	.915	.842	.921	1.000	.878	.679
Item5	.902	.934	.867	.878	1.000	.807
Item6	.813	.805	.793	.679	.807	1.000

Table 17: Test Item Inter Item Correlations

It is important that all inter-item relationships exhibit moderately strong correlations in the same direction, that is, that all correlations are positive or all are negative. This can clearly be seen to be the case, as depicted in **Table 19**.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.970	.972	6

Table 18: Reliability Test Results Output

Table 20 depicts the result of the Cronbach Alpha procedure. It is generally accepted that a Cronbach Alpha value in excess of **0.70** is a sufficient value to infer internal consistency and reliability between survey items. Although, there is strong evidence to suggest that a value greater than **0.95** should be used as the appropriate benchmark.

In this example, our survey item results exhibit a Cronbach Alpha value of **0.97**, as shown in **Table 20**. This certainly exceeds both predefined levels; and as such, we can safely assume that the test items measured; together show significant evidence of internal consistency and reliability.

4 Dissertation Results

The Results section of your Dissertation should present all of the pertinent results that the author will use to defend the underlying hypothesis of the Thesis. Results can vary significantly in form, but generally are summarised in: tables, figures and graphs.

This Results section generally exploits the typical graphical descriptive techniques for presenting data, with all the typical descriptive measures summarised within tables and figures, as required.

There can be a tendency, at this stage, to present a commentary of the significance of a particular result; this should wait till the Analysis section.

In the Results section of a Dissertation we do comment on what a table, figure or graph is presenting. In the case of tables we would comment on the structure: the first row presents …; the first column presents the … etc. In the case of a graph: the horizontal axis presents …; the vertical axis presents … etc.

4.1 Tables and Figures

All tables, figures and graphs should be referenced and appropriately captioned, as well as introduced. The reader should not be surprised by seeing a table or figure and should be appropriately introduced so as to be able to interpret the results summarised within. For example:

Table 1 presents a summary of the usual descriptive statistics for the variables Examination Grade and Gender. The first row of **Table 1** lists the usual statistics; with row two depicting the levels of measurement associated with the variable Exam Result; and row three depicts the levels of measurement associated with the Gender variable.

For example, we can see that there were **14** cases considered valid with respect to the Exam Result variable and our results record a mean exam result of approximately **44.93** with a standard deviation of approximately **24.66**.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Exam Result	14	44.9286	24.65598	10.00	82.00
Gender	14	.50	.519	0	1

Table 19: Descriptive Statistics - Exam Result and Gender

An examination of the differences between examination performances based on gender is presented in **Figure 1**. The horizontal axis of **Figure 1** represents both Male and Female levels of measurement, with the vertical axis signifying the average Examination Grade observed at these levels of measurement.

For example, Males are recorded with an average Examination Grade of approximately **57%** compared to an average Examination Grade of **32%** for Females.

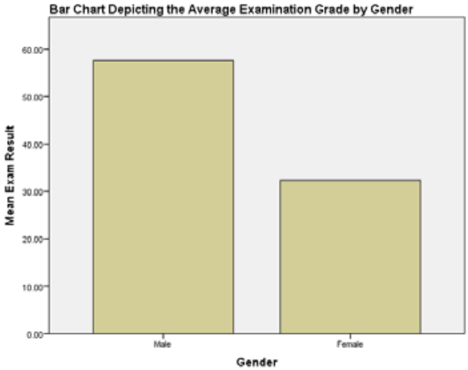


Figure 29: Average Examination Grade by Gender

A more granular examination of the impact of candidate Age on examination performances is presented in **Table 2** below. The first row of **Table 2** depicts the usual Analysis of Variance statistics. The second and third row depicting both between groups and within groups measures, respectively. For example, **Table 2** depicts an F-statistic of **0.412** with a significance value (**p-value**) of **0.672**.

ANOVA

Exam Result

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	550.405	2	275.202	.412	.672
Within Groups	7352.524	11	668.411		
Total	7902.929	13			

Table 20: One-way Analysis of Variance – Age

5 Dissertation Checklist

Introduction	Completed
Have you presented/stated the problem or issue which is to be addressed?	
Have you stated why your research is worthy of study?	
Have you given your research problem context? i.e. have you said what has been done previously or why this is an important area of study?	
Have you given some indication of the key literature identifying any gaps which your research hopes to address?	
Have you provided an overview of the structure of the research project?	
Are there references in your introduction? There should be!	

Literature Review	Completed
Is there evidence of up-to-date material pertaining to your area of study?	
Is the material mostly journal based rather than text book based?	
Have you provided a synthesis, not a summary, of previous studies/research?	

Literature Review Cont.	Completed
Is there evidence of up-to-date material pertaining to your area of study?	
Is the material mostly journal based rather than text book based?	
Have you provided a synthesis, not a summary, of previous studies/research?	
Are you guilty of summarising/describing what others have said? If so you need to address this!	
Does each paragraph describe simply what someone else has said? i.e. does it only contain one reference albeit multiple times to the same piece of work? If so you need to address this!	
Does your literature have a logical flow? Does it jump from one section to another without any link?	
Does your literature review have a conclusion?	

Research Question	Completed
Do you have a clearly stated research question or hypothesis?	
Have you identified and explained the aims and objectives of your study?	

Methodology	Completed
Have you provided summaries of each possible research method without ever linking it to your own work? If so please revisit.	
This section should describe how the problem was investigated and why particular methods and techniques were employed. Have you done this?	
Have you been able to link your methodological approach to other previous research in terms of adopting a similar approach?	
Have you stated what you did? Often referred to as the procedure adopted.	
Have you provided details of your sample? Who did you ask and why these and not others?	
What you asked them? Details about your research instrument.	
How you asked them? Details about your data collection	
What you did with what you collected? Data analysis – i.e. how you treated the data. This is NOT what they actually said – that comes in the next section! Here you want to be clear about how you treated the data not what you found.	
Did you pilot your data collection tool?	
Does your methodology section refer to ethical considerations?	
Have you included a limitations section?	

Findings/ Results**	Completed
Does your findings section simply list the answer to each question in your questionnaire/interview etc., one after the other? If so revisit!	
Is there a logical flow to your findings?	
Have you highlighted the key findings for the reader?	

Discussion	Completed
Have you linked your findings back to your literature?	
Have you highlighted for the reader what was important in your findings and how this relates back to previous studies/knowledge on your research topic?	
Have you included practical implications (if appropriate)?	
Have you considered the limitations of your study including your methodological approach?	

Conclusions	Completed
Have you provided a strong conclusion to your work?	
Have you provided a summary of what you have found out in relation to each research question posed?	
Is your conclusion section less than 1 page in length? If so please revisit.	
What are the next steps? Future research possibilities?	

Reference List	Completed
Have you included ALL references cited in your work?	
Have you adhered to the Harvard referencing system?	
Have you included all references in alphabetical order by surname?	
Have you separated books from journals etc.? If so please revisit. All reference material should appear in alphabetical order irrespective of whether it is a book or journal or working paper etc.	
Have you included material that is not directly referenced in your research report? If so please revisit. Only material directly cited/referenced in your report should be included in your reference list. All other material consulted but not directly cited should appear in your bibliography should you choose to include one.	

In-text referencing	Completed
Have you ensured that all in-text references appear in your reference list?	
Have you used the Harvard method? Surname/Date approach?	
Have you put the full stop after the close bracket for in-text references – example below; Students should pay attention to the advice they have been given by their lecturer (Darcy 2013).	

In-text referencing	Completed
Have you included an author's initial in your in-text referencing? If so please revisit. Only the surname and year should appear.	
Have you ensured that all in-text references appear in your reference list?	

Housekeeping Issues	Completed
Have you checked your spelling and grammar?	
Are there unexplained gaps between sections of work/ blank pages or pages where the work begins half way down the sheet? If so please revisit	
Is your work neat and tidy with a professional presentation?	

**** Some lecturers look to split out the section on Findings/Results. Your supervisor may have a preference for a 'findings or results section' which simply states your results and then a 'discussion section' where you set about discussing what these results mean in relation to your research question and previous research. If this is the case always be guided by your supervisor. It is simply a reflection of different traditions and no one approach is better than the other.**

This publication was produced by **Mr. Jonathan Lambert**, Mathematics Development and Support Officer at the National College of Ireland; and **Dr. Colette Darcy**, Vice Dean Postgraduate Studies & Research at the National College of Ireland.

NATIONAL COLLEGE OF IRELAND
Mayor Street, IFSC, Dublin 1
Telephone: 01 449 8624
Web: https://myncistudent@ncirl.ie