# **SPSS** for Windows

# Versions 20.0 for (Windows 7) A Survival Guide For EPRS 8530 and EPRS 8540

Edition 5 (Version 1.0)

December 2014

Tianna C. S. Floyd

Department of Educational Policy Studies

Keith D. Wright

**Department of Educational Policy Studies** 

H. A. Russell III Department of Educational Psychology

J. Randy Beggs Department of Counseling & Psychological Services

Gary L. May Department of Public Administration & Urban Studies

> T. Chris Oshima Department of Educational Policy Studies

# TABLE OF CONTENTS

Foreword	4
About This Guidebook	5
Getting Started with SPSS for Windows, Version 20.0	6
Descriptive Statistics	12
SPSS Procedures for Frequency Distribution, Measures of	
Central Tendency, Measures of Dispersion, and Correlation	16
Selected Results for Correlations	
SPSS Procedures for Graphing	19
Example Reporting of Results	22
Inferential Statistics	
Procedural Model: Overview	27
Independent T-Test	
SPSS Procedures	29
Selected Results From SPSS Execution	
Example Results and Discussion	
Dependent T-Test	
SPSS Procedures	
Selected Results From SPSS Execution	
Example Results and Discussion	
Analysis of Variance	27
Analysis of variance	
Delenand One Way ANOVA	
SPSS Procedures	
Selected Results From SPSS Execution	
Example Results and Discussion	45
Unbalanced One-Way ANOVA	46
SPSS Procedures	
Selected Results From SPSS Execution	
Example Results and Discussion	51
Factorial Analysis of Variance	53
Procedural Model: Overview	
Factorial ANOVA Procedure Model	55
Balanced Two-Way ANOVA	
SPSS Procedures	
Selected Results From SPSS Execution	60
Example Results and Discussion	63
Unbalanced Two-Way ANOVA	65
SPSS Procedures, Part I	66

Selected Results From SPSS Execution	. 69
SPSS Procedures, Part II	. 75
Selected Results From SPSS Execution	. 77
Example Results and Discussion	. 80
Analysis of Covariance	82
SPSS Procedures, Part I	83
Selected Results From SPSS Execution	85
SPSS Procedures, Part II	88
Selected Results From SPSS Execution	89
SPSS Procedures, Part III	89
Selected Results From SPSS Execution	94
Example Results and Discussion	96
Repeated Measures Analysis	97
Procedural Model: Overview	
SPSS Procedures for One-Within Repeated Measures	.100
Selected Results From SPSS Execution	.103
Example Results and Discussion	.105
SPSS Procedures for One-Within One-Between Repeated Measures	.108
Selected Results From SPSS Execution	.111
Example Results and Discussion	.116
lvanced Tips and Tricks	.119
Default Operations of SPSS	.120
The SPSS Syntax Editor	.123
Importing and Exporting SPSS Data Files	125
Recall Dialog Box Icon	.127
The "Select Cases" Function	.128
Producing Graphs	.130
Technical Support from SPSS, Inc.	.130
oncluding Remarks	.131
eferences	.132

# Foreword

The purpose of this guidebook is to help you achieve the course objectives for either Quantitative Methods I and/or Quantitative Methods II: (1) to gain understanding of the various statistical techniques for analysis with descriptive statistics, (2) to gain conceptual understanding of the various techniques for determining the significance of group differences, (3) to use statistical software to run the analyses, and (4) to interpret and report the results according to the style used in research journals.

With these objectives in mind, you will be assigned a variety of computer homework problems to complete for course credit. This guidebook provides a working example for each type of statistical analysis to help you get started with each assignment. The section of this guidebook that addresses statistical analysis problems contains:

- 1. Procedures decision model (algorithm) for inferential statistical examples
- 2. An example of research problems with necessary data
- 3. A guide to SPSS for Windows Versions 20.0, with explanatory notes and tips
- 4. Selected portions of the SPSS printout with commentary
- 5. An example of how to report results and provide discussion

These example problems are basic and intended only as a guide. The steps you take to do an analysis on your own particular problem may vary slightly based upon your data and the outcome of your analysis. You will also be expected to go beyond the examples in your own work. To accomplish this, you will need to attend class regularly, study your text carefully, and complete the recommended non-graded problems assigned from the text.

Do *NOT* ignore the many notes explaining many of the finer points in the example to follow. Unlike many explanatory notes you may have encountered in the past, these notes are vitally important to your understanding of how to correctly use SPSS. *Ignore them at your own peril.* 

When you have completed the course, this guidebook should also serve as a ready procedure reference for your future work in your chosen field.

It is our hope this guidebook will help you complete your assignments with confidence, increase your learning, and make your statistics course more enjoyable.

Tianna C S Floyd Department of Educational Policy Studies, Georgia State University

Keith D Wright Former student of Department of Educational Policy Studies, Georgia State University

H. Arthur Russell III, M.S. Former student of Department of Educational Psychology, Georgia State University Gary L. May, M.S Former student of Dept. of Public Administration & Urban Studies, Georgia State University

J. Randy Beggs, M.B.A, M.S., Ed.S. Former student of Department of Counseling and Psychological Services, Georgia State University

T. C. Oshima, Ph.D. Professor Department of Educational Policy Studies, Georgia State University

# ABOUT THIS GUIDEBOOK

This guidebook was developed specifically for students taking statistical courses in the Educational Policy Studies department of the College of Education at Georgia State University. Anyone conducting research analysis and new to using SPSS for Windows 7 may find this guide helpful.

This guide is organized into three sections. First, the basic operation of SPSS in the Windows 7 environment is addressed followed by examples of various statistical analyses using the package. The inferential statistical example problems used are exactly the same as those presented in Student Guidebook for SAS<sup>TM</sup> users developed by Gary L. May and T. Chris Oshima. Some students find a comparison of the several software products helpful as multiple terms are used by different vendor products to refer to inferential data. The final section addresses more advanced tips and tricks that can be used to take full advantage of SPSS for Windows, Versions 20.0. The guide also contains references to common questions and trouble areas which students have encountered in the past.

Every effort has been made to ensure this guidebook is free from errors. If you should find any mistakes within this guide, please contact either Tianna Floyd (tfloyd6@student.gsu.edu) or Dr. Chris Oshima (oshima@gsu.edu) so that corrections can be made to the material.

Certain assumptions were made in the writing for this guidebook. For example, it is assumed you accepted the installation program default values when placing SPSS for Windows Version 20.0 on your computer. If this is not the case, you will need to consider how you modified the default installation when reading this guidebook. Secondly, this guide was composed specifically with the Windows 7 operating systems in mind. You should already possess a basic knowledge of Windows navigation techniques in order to successfully use the SPSS product. You may find that a more advanced knowledge of Windows navigation is required to perform some of the tasks discussed in the "Tips and Tricks" section of this guide.

As the guidebook progresses from one problem type to another, certain levels of detail, such as each window concerning data entry, are omitted due to space limitations. A cumulative working knowledge of SPSS is assumed as the guidebook progresses.

Finally, please note that SPSS<sup>TM</sup>, Microsoft Windows<sup>TM</sup>, Statistica<sup>TM</sup> and SAS<sup>TM</sup> are registered trademarks of SPSS, Inc., Microsoft Corporation, Statsoft, Inc., and the SAS Institute respectively. Users of all personal computer software should remember copyright laws of the United States protect these products. Use of software products is governed by the license granted to the purchaser of the product. Infringement of copyrights held by software publishers may result in legal action on the part of the copyright owner.

# Getting Started with SPSS<sup>TM</sup> for Windows 7 Versions 20.0

# Getting Started With SPSS for Windows Versions 20.0

## Introduction

SPSS for Windows, Version 20.0, is a powerful Microsoft Windows compliant package that allows easy statistical analysis of data. Additionally, Version 20.0 addresses many of the shortcomings noted in earlier versions of SPSS to provide many improved functions including enhanced paper management and charting capabilities. Becoming familiar with the package and its default operation prior to starting actual analysis can save the student many hours of time and a great degree of frustration. As a minimum, it is recommended that the user complete the SPSS "Tutorial" and completely familiarize him or herself with SPSS 's capabilities. The purpose of the *Getting Started* section of this manual is to explain the basic operation of this software package and to make suggestions that will ease the burden of becoming familiar with the operation of SPSS.

This section is divided into the following areas:

- 1. A few words on the use of Microsoft Windows
- 2. Initial software interface presentation
- 3. Explanation of a matrix concept of entering data for analysis
- 4. A more in-depth look at several of the data entry / results windows used by SPSS
- 5. Concluding remarks

### Use of Microsoft Windows

This manual is written with the assumption the user has a basic understanding of how to navigate in a Microsoft Windows 7 environment. It is strongly recommended that you become familiar with Windows 7 prior to use of SPSS, Version 20.0. Many one or two day courses regarding Windows 7 are available through commercial vendors or through local universities. Use of SPSS, Version 20.0 without a basic understanding of the Windows 7 environment will prove more difficult, time consuming, and may prove very frustrating

Untit	fed1 [D	Index Could				_									<u></u>	1
le E		atasetu) -	IBM SPSS	5 Statisti	cs Data Ed	itor										ê 🛃
- <u>-</u>	dit <u>V</u>	iew <u>D</u> a	ta <u>⊤</u> rar	nsform	Analyze	Gra	phs Utilitie	s Add-on:	: <u>W</u> indow	Help						
			00.	10			± =	A	* 2		5			AC.		
		-	-	_		100	-				9	14 🔍				
															Visible: 0 of	Variable:
		Var	1	/ar	var		var	var	Var	Var	var	Var	Var	var	Var	
1																1
2																
3						_										
4			_			_										
5			_			_										
6			_			_									_	
7	_															
8	_					-										
9	_		-			-									_	_
10	-		-			+									_	
11	-		-			+									_	
12	-		-												_	
1.0	-															
14	-		-			-									_	
16			-			+									-	+
17	-		-			+										+
18						-										
19	-															
20			-													+
21						-										+
22						-										
23																
		1										_		_	_	•
		ariabla 16			-											

SPSS starts with an Untitled -SPSS Data Editor window. The Data Editor provides most of the controls (menu and button selections) you will use to enter and manipulate data as you perform statistical analyses.

Notice also that across the top of the window is a row of 'cells' labeled 'var' in grayed out lettering while down the left side of the window is a series of numbers beginning with '1' and subsequently incremented by '1'. The row across the top is used to name data variables while the column on the left hand side of the window represents the 'case' number. Case is the term used in SPSS to identify the observation number of the record within the data set to be analyzed.

A second window called 'Output' will be displayed once you begin performing operations with SPSS. This is where you will be able to see the results of your analyses and to determine if you wish to perform additional analyses or otherwise manipulate your data. SPSS automatically brings the 'Output' window to the fore-ground when all execution of an analysis is complete.

A third window, known as the 'SPSS Syntax' window, may also be present when requested or during certain steps of your analyses. The syntax window is only mentioned at this point and will be covered in more depth during specific example problems and in the Advanced Tips and Tricks section of the guidebook. The utility of this method of data entry, while possibly not apparent at this time, will become more visible as you progress in the analysis of your various data sets.

## A More In-Depth Look at SPSS Data Entry and Results Windows

Now that you have a concept of how to initialize SPSS for use and a conceptual framework for the entry of data to be analyzed, a more in-depth look at variable definition will prove useful

### Data Entry Window

The data variables represented in the problem analysis must be defined for SPSS. This is done as follows:

le <u>E</u> dit	View Data	Transform	Analyze <u>G</u> ra	phs <u>U</u> tilities	Add-ons Wind	dow <u>H</u> elp					
26				1	H #	¥ =	₫.	19		j.	
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Rok
1	price	Numeric	8	2		None	None	8	Right	Unknown	> Input
2	location	Numeric	8	2		None	None	8	🗃 Right	Unknown	🔪 Input
3	sales	Numeric	8	2		None	None	8	🗃 Right	Unknown	S Input
4											
5											
6											
7	1										
8											
9											
10											
11	-										
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
ne	(4)										12
ata View	Variable View										
-									Saliatica Processo	or is ready	

Data variables are defined by clicking on the 'Variable View' tab below. This view can also be activated by double clicking on the grayed 'var' column heading illustrated earlier.

From the 'Variable View' screen, SPSS allow you to create and edit all of the variables in your data file. Each column represents some property of a variable, and each row represents a variable. All variables must be given a name. Just click on the first empty cell in the Name column and type in a name. Perform these same steps for the other variables.

-											
SPSS 20(1)	- 131.96.133.153	- Remote Desk	top Connectio	on			_	_	and the second		
*Untitled1	[DataSet0] - IBM	SPSS Statistics I	Jata Editor								
<u>File Edit</u>	<u>V</u> iew <u>D</u> ata	Transform <u>A</u> r	nalyze <u>G</u> rap	hs <u>U</u> tilities	Add-	ow <u>H</u> elp					
				▙						ð	
	Name	Туре	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	price	Numeric	8	2		None	None	8	🗮 Right	Unknown	ゝ Input 📥
2	location	Numeric	8	2		None	None	8	🗮 Right	Unknown	🔪 Input
3	sales	Numeric	8	2		None	None	8	🗮 Right	Unknown	🔪 Input
4											
5								L			
6			ta Va	lue Labels				×			
7				lue Labels—							
8			10	lue: lo							
9			Va	i <u>u</u> e. 3			Sb	elling			
10			La	bel: Checko	ut						
11					1.00 = "Entrance"						
12				Add	2.00 = "Sporting Go	ods Aisle"					
13				Change							
14				Remove							
15				Trennove							
16				l							
17					ок	ancel Heln					
18											
19											
20											
21											
22											
23											
24											
าย	1										
Data View	Variable View										

Many users find it imperative to have assigned labels for the various values of a data variable. This allows you the ability to associate a description with variables. Furthermore, these descriptions can describe the variables themselves or the values of the variables. Click the cell you want to assign values to on the

'Values' column. You will see a small gray box Click this box and this will bring up the 'Value Labels' box. In this box you will type in a value, such as 1 with a label of Entrance, and click add. Continue this process for as many values required for your variable. In the measure column, make sure you select the appropriate measurement type for each variable either scale (interval or ratio), ordinal or nominal.

5 SPSS 20(1) - 131.96.133.153 - Remote Desktop Connection											
👘 *Output1 [Document1] - IBM SPSS	Statistics Viewer										- 6 -
Elle <u>E</u> dit <u>V</u> iew <u>D</u> ata <u>T</u> ransform Insert Format <u>A</u> nalyze <u>O</u> raphs <u>U</u> tilities Add- <u>o</u> ns <u>W</u> indow <u>H</u> elp											
😂 H 🖨 🖄 🤌 💷 🖛 🛥 🧝 🕌 🏥 🖆 🖉 🕒 🗮											
<ul> <li>■ Goutput</li> <li>● Parameter</li> <li>● Summarize</li> <li>● Title</li> <li>● Title</li> <li>● Active Dataset</li> <li>■ Case Processing</li> </ul>	■     ■     ●     Summarize     ●       ■     ■     ■     ■     ■       ■     ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■       ■     ■     ■     ■								-		
🖵 🚡 Case Summaries				Cas	ses			1			
		Inclu	ded	Exclu	ıded	Total					
		Ν	Perce	nt N	Percent	N	Percent				
	price	6	100.0	0% 0	0.0%	6	100.0%				
	location	6	100.0	0% 0	0.0%	6	100.0%				
	sales	6	100.0	0% 0	0.0%	6	100.0%				
	a. Limite	ed to first 10	)0 case	s.							
		Cas	se Sumi	naries <sup>a</sup>							
		pri	ce	location	sales						
	1		1.00	Entrance	28.00						
	2		1.00	Entrance	21.00						
	3		1.00	Entrance	17.00						
	4		1.00	Sporting Goods Aisle	23.00						
	5		1.00	Sporting Goods Aisle	18.00						
	6		1.00	Sporting Goods Aisle	15.00						
	Total N		6	6	6						
	a. Limite	ed to first 10	00 case	S.							
4											-
									IBM SPSS Sta	atistics Proc	essor is ready

Output Window

All results generated as results of running analyses are automatically sent to the 'Output' SPSS Output Navigator' window.

## <u>In Conclusion</u>

Always remember to save your data using the 'SAVE' option listed under the 'FILE' menu selection on the SPSS Menu Bar. If your data is particularly important, you should also consider maintaining a second copy of the data and your SPSS syntax using the 'SAVE AS' option. While not necessary for the problems you will be doing in this course, keep in mind the American Psychological Association suggests research data be saved a minimum of five years. Considering the lead time regarding research article submission, editing, acceptance, and publication, many professionals recommend saving your data and program for seven years. Additional information regarding the use of the various SPSS windows is presented in the Advanced Tips and Tricks section at the end of this guidebook.

# **Descriptive Statistics**

Measures of Central Tendency

Measures of Variation

Correlation

# Descriptive Statistics Example Problem

A counselor working with a group of caregivers of patients living with a terminal illness is interested in forming a support group to share experiences and therefore reduce the sense of isolation often associated with catastrophic illness. The counselor, working with hospital staff, administers a depression and anxiety inventory to each caregiver that has volunteered for the program. The counselor feels that knowing the levels of depression and anxiety within the group will help in the design of an effective intervention program. The scores obtained from the administration of the two inventories are given below.

Individual	Anxiety Score	Depression Score
1	22	16
2	12	8
3	68	33
4	10	6
5	5	5
6	53	24
7	44	18
8	37	17
9	0	2
10	21	14
11	64	31
12	33	17
13	55	30
14	18	13
15	3	3
16	4	4
17	11	7
18	13	9
19	7	5

First we need to enter the data. Please refer to the introduction section to see how to enter the numbers. You should always verify the accuracy of your data entry. The most obvious method to do so is to review your data in the 'Data Editor' window. You can also obtain a list of all of your data by clicking on the command sequence 'Analyze', 'Reports', and 'Case Summaries...' in the main SPSS menu. Click on the right pointing arrow located between the two text boxes to move the variables to the 'Variables' text box Click 'OK' to instruct SPSS to create a summary list of cases.





The results are automatically displayed in the 'Output 1 - SPSS Output Navigator' window.

Now we will instruct SPSS to calculate frequency distribution(s), measures of central tendency, and measures of dispersion.

Using the mouse pointer, highlight 'Analyze', 'Descriptive Statistics', and 'Frequencies' starting from the menu selection on the SPSS main menu.

Using the same techniques previously discussed to highlight and move variables from one text box to another, move the variables to the 'Variable(s)' text box. Ensure the 'Display frequency tables' check box is checked. Using the mouse pointer, click on the 'Statistics' button.

The 'Frequencies: Statistics' window will be displayed. Click on the box to the left each measure of central tendency and each measure of dispersion to be calculated. Click on 'Continue' to proceed. The 'Frequencies' window will be redisplayed. Click on 'OK' to continue. The results of the analyses will be automatically displayed in the 'SPSS Output Navigator' window.

### SPSS 20.0 A Survival Guide

Frequencies: Statistics  Percentile Values  Quartiles  Cut points for: 10 equal groups	Central Tendency ✓ Mean ✓ Median
Percentile(s):	Mode
<u>A</u> dd <u>C</u> hange <u>R</u> emove	Values are group midpoints
Dispersion	Distribution
Std. deviation 🗹 Minimum	Ske <u>w</u> ness
✓ Variance ✓ Maximum	🕅 <u>K</u> urtosis
🗖 Ra <u>n</u> ge 📄 S. <u>E</u> . mean	
Continue	Help



J.B. Landers © www.CAUSEweb.org

We can also have SPSS calculate the correlation between two variables. From the SPSS for Windows main dialog, choose 'Analyze', 'Correlate' and 'Bivariate'. The 'Bivariate Correlations' window will be displayed. Using the mouse pointer, highlight both variables, 'Anxiety' and 'Depression', and click on the right pointing arrow to move the variable to the 'Variables' text box. Click on the type of correlation coefficient you wish to produce. The default correlation coefficient is the Pearson Correlation Coefficient. After moving the variables to the 'Variables' text box, click on 'OK' to continue. The results of the correlation analysis will be automatically displayed in the 'SPSS Output Navigator' window.

ta *Output1 [Document1] - IBM SPSS Statistics Viewer								
<u>File Edit View Data Transfo</u>	orm Insert Format Analyze Graphs Utilities Add-ons Window Help							
😑 🔚 🖨 🙇 🥏	😑 🖶 🖨 🖉 💷 🖛 🛥 🧱 🖆 📥 🗐 🙆 🍉 🚑 🛔							
<b>* * * -</b>								
Case Summarize  Case Summariz	CORRELATIONS /VARIABLES=Anxiety Depression /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE. Correlations [DataSet0] Correlations							
🖻 Title	Anxiety Depression							
🛱 Anxiety	Anxiety Pearson Correlation 1 .982							
Depression	Sig. (2-tailed) .000							
	N 19 19							
→ P Title	Depression Pearson Correlation .982 1							
Notes	Sig. (2-tailed) .000							
Active Dataset	N 19 19							
Correlations	**. Correlation is significant at the 0.01 level (2-tailed).							
	. Contenation is significant at the olor level (2-tailed).							
1								
	IBM SPSS Statistics Processor is ready							

The output that is given is in the form of a data matrix. For every possible combination of variables (anxiety, anxiety; anxiety, depression; depression, anxiety; and depression, depression) the correlation coefficient, the number of cases used to perform the correlation, and the 'p' value are given. As you might expect, a variable is perfectly correlated with itself which is indicated by the correlation coefficient of 1. This output shows the correlation between anxiety and depression is r = .98.

### SPSS Procedures for Graphing

Making graphs is a bit easier in the newer versions of SPSS. Procedures for producing a histogram and scatterplot, are outlined below.



To make a Histogram, from the SPSS for Windows main dialog, choose 'Graphs', 'Chart Builder'. SPSS might give you a warning to remind you that you need to set the correct measurement level for each variable. You can select 'Define Variable Properties' to make sure variables are identified correctly or you can select 'OK" to proceed. Now the chart builder window will appear. You will need to select histogram from the list of available graph types in the left hand side. You will click on the histogram that is on the far left and drag and drop it to the chart preview area.



Now we need to format the graph. The 'Element Properties' menu will appear. We will exit out of that for now. We would like our histogram to show the frequency information for anxiety levels. We will drag and drop anxiety from the variables window to the x axis. Select 'OK' and the histogram will be shown in the output window

Now double click on the graph in the output window. This will open up the "Chart Editor'. In this window we can modify the graph. In order to change the X-axis, go to 'Edit', "Select X axis". A new window will open and you can click on the 'Scale' tab. Here we can change the scale of the X axis. We want to change the major increments to 10. If we want to change the bin size, first click on one of the bins in the chart editor window, go to 'Edit', 'Properties' and click on the 'Binning' tab. This will give options to customize the bins by either specifying the number of intervals or the bin size

			Properties Properties	<b>—</b> ×-
5.0			Labels & Ticks Number Format Variables Chart Size Fill & Border Binning	Variables
			Chart Size Scale Lines  O X axis only	
			Range © Z axis only	
			Auto Custom Data	
4.0			Minimum 🔽 0 0 Automatic	
			Maximum 🔽 70 70 O Custom	
ž			Major Increment 🔲 10 💿 Number of intervals: 0	
<b>9</b> 3.0			Origin 🔽 0 O Interval width:	
nba			Display line at origin	
2.0 <sup>.</sup>			Type © Linear © Logarithmic	
			Base: 10 Safe        Base: 10 Safe	
1.0 <sup>.</sup>			Lower margin (%): 5 Upper margin (%): 5	
0.01	.0	00 20.00 40.00 60.00	Apply Close Help Apply	Close Help
		Anxiety		



Creating a scatter plot is very similar. Go to "Graphs", "Chart Builder". The same histogram might be in the chart preview menu which is OK because we will replace it. On the left hand side select 'Scatter/Dot' and drag and drop the first scatter/dot plot into the window. We will place Anxiety on the X- axis and Depression in the Y-Axis. Select 'OK' and the graph will come out in the output window.



J.B. Landers © www.CAUSEweb.org

## Example Reporting of Results Results

Table 1 presents the mean, median, mode, minimum and maximum scores, and standard deviations for anxiety and depression scores administered to caregivers of terminally ill individuals. Comparison of the mean, median, and mode indicates the distribution of anxiety and depression are both somewhat positively skewed. The mean for anxiety and depression are both affected by a small number of high scores suggesting the median may be a more reflective measure of central tendency. The standard deviation and range of values for both variables indicate a wide dispersion among scores of individuals.

A histogram of the anxiety frequency distribution is depicted in Figure 1. This identifies the number of individuals falling into the levels of anxiety.

A Person's coefficient of correlation was calculated between the variables of anxiety and depression. The high positive correlation, r(17) = .98, p < .001, indicates the greater an individual's level of depression, the greater the expected level of anxiety and vice versa. Figure 2 presents a scatterplot of the correlation.

Table 1

# Mean, Median, Mode, Minimum and Maximum Scores, and Standard Deviations for Anxiety and Depression Scores

Variable	Sample <u>Size</u>	<u>Minimum</u>	Maximum	Mean	Median	Mode	Standard Deviation
Anxiety	19	0.00	68.00	25.26	18.00	N/A*	21.99
Depression	19	2.00	33.0	13.79	13.00	5,17	9.85

\* Not Applicable



Figure 1. Histogram of the anxiety score frequency distribution



Figure 2. Scatterplot of correlation between anxiety and depression

SPSS 20.0 A Survival Guide

# Inferential

**Statistics** 



# **PROCEDURAL MODEL: OVERVIEW**

26



# Independent T-Test Dependent T-Test

# **INDEPENDENT T-TEST**

### **Example Problem**

A training manager believes that a new interactive computer based training package will help improve the production rate of order assemblers. She arranges for a production area of 21 experienced employees to complete the new training package over a six week period. Another group from the production area of 23 employees received no additional training. Following are the average production rates per person per hour based on a 12 week period following the training:

Table 1

Production Rates: Completed Assemblies Per Person Per Hour

Training Group	Control Group
26	44
45	45
60	57
73	28
45	64
51	39
63	35
46	43
69	21
51	56
55	22
58	87
61	48
54	12
64	19
56	62
59	55
35	44
48	39
45	44
59	57 BREAKING
	30
	30



WWW.BITSTRIPS.COM

File Edit	[DataSet0] - IBN View Data	SPSS Statistics I Transform A	Data Editor nalyze Gra	phs Utilities	Add-ons	Window
					h	*
28 :						
	rate	group	var	var	var	var
7	63.00	1.00				
8	46.00	1.00				
9	69.00	1.00				
10	51.00	1.00				
11	55.00	1.00				
12	58.00	1.00				
13	61.00	1.00				
14	54.00	1.00				
15	64.00	1.00				
16	56.00	1.00				
17	59.00	1.00				
18	35.00	1.00				
19	48.00	1.00				
20	45.00	1.00				
21	59.00	1.00				
22	44.00	2.00				
23	45.00	2.00				
24	57.00	2.00				
25	28.00	2.00				
26	64.00	2.00				
27	39.00	2.00				
28	35.00	2.00				
	4					
Data View	Variable View					

In order to do the t-test, we will have to enter the data. In this case we will add a new variable called "Group" and code 1 for "Training and 2 for "Control". Then we will enter each production rate as a new entry for a total of 44 entries. Do not forget to define the values for the group variable.

Now we can enter the data (only a sample of the data entry is shown). The 'SPSS Data Editor' window is shown with a sample of data. Remember, a total of 44 'Cases' or observations are entered. Notice the complete listing of cases is not shown in the example window at the left.

Independent-Samples T	Test	×
	Test Variable(s):	Options
	group(1 2)	
ОК	Paste Reset Cancel Help	
ta Define Group	os 💌	
© <u>U</u> se specif Group <u>1</u> :	fied values	
Group <u>2</u> :	2	
© <u>C</u> ut point:		
Continue	Cancel Help	

Now we will conduct the t test. Select 'Analyze', 'Compare Means' and "independent sample T test'. We want 'rate' to be in the test variable box and 'group' to be in the grouping variable. Now select 'Define Groups'. This will allow us to distinguish what two groups we want to compare. The group one value is '1' and the group 2 value is '2'.

At this point, SPSS will perform the statistical analysis directing the results to the 'Output' SPSS Output Navigator' window which automatically displays the results of the analysis.

### T-Test

[DataSet0]

	Group Statistics							
+		group	Ν	Mean	Std. Deviation	Std. Error Mean		
	rate	training	21	53.4762	11.00736	2.40200		
		control	23	42.6522	17.31161	3.60972		

The SPSS results provide the mean values of the scores for the training and control groups, along with the standard deviations and standard errors.

#### Independent Samples Test

	Levene's Test Varia	t-test for Equality of Means							
						Mean	Std. Error	95% Confidenc Differ	e Interval of the ence
	F	Sig.	t	df	Sig. (2-tailed)	Difference	Difference	Lower	Upper
rate Equal variances assumed	2.872	.098	2.448	42	.019	10.82402	4.42228	1.89950	19.74854
Equal variances not assumed			2.496	37.672	.017	10.82402	4.33586	2.04401	19.60402

Levene's Test of equality of variances performs hypothesis testing for equality of variances. If the probability (significance) is less than .05 (p < .05), then the hypothesis that the variances are equal is rejected which is not the case in this example so we can continue with interpreting the results of the t test.

In interpretation of the results, use the t-value, df, and probability (Sig, 2-tailed), labeled **UNEQUAL variances assumed** if the probability from the Levene's test is less than .05. If the probability is greater than .05, use the EQUAL variances assumed line. The Equal variance assumes line is what we will use.

### **INDEPENDENT T-TEST**

### **Example Results Write Up**

### Results

Table 1 presents the means and standard deviations by group. The results show that order assemblers in the training group produced on average about 10 more assemblies per person per hour (M = 53.48, SD = 11.01) than the assemblers in the control group (M = 42.65, SD = 17.31) and they also had more homogeneous production rates. This mean difference resulted in the effect size of .75 (Cohen's *d*), which is considered to be medium.

Since the measurements in the training group and control group were unrelated to each other, an independent ttest was performed. The assumption of homogeneous variances was satisfied using (Levene's test, F(1,42) = 2.87, p = .098). The mean score for the training group was significantly higher than the mean score for the control group, t(42) = 2.45, p = .019.

# Table 1Mean Assemblies Per Person Per Hours and Standard Deviations

Group	М	SD
Control	42.65	17.31
Training	53.48	11.01

The new computer-based training package used with the training group had a significant and positive effect on production rates. Therefore, it is appropriate for management to consider providing the new training to employees in all assembly areas.

# **DEPENDENT T-TEST**

### **Example Problem**

A sports psychologist was interested in testing the effect of a simple relaxation technique on college basket-ball players' free throw shooting accuracy. Each player was asked to shoot 20 consecutive free throws and the number of successful attempts was recorded. The players were then trained to use a simple 5 second relaxation technique while preparing to shoot a free throw. The players then returned to the court and shot 20 consecutive free throws again. The resulting data are given below:

Table 1

Number of Free Throws Completed Per Twenty Attempts

	Post
Pre-Training	Training
12	13
15	15
9	11
16	15
12	15
15	18
17	17
10	12
12	13
14	17



J.B. Landers © www.CAUSEweb.org

-				
ta *Untitled1	[DataSet0] - IBM	SPSS Statistics	Data Editor	
<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Transform A	nalyze <u>G</u> raj	ohs <u>U</u> tilitie
			¥ 🎬	▙
14 :				
	Pre	Post	var	var
1	12.00	13.00		
2	15.00	15.00		
3	9.00	11.00		
4	16.00	15.00		
5	12.00	15.00		
6	15.00	18.00		
7	17.00	17.00		
8	10.00	12.00		
9	12.00	13.00		
10	14.00	17.00		
11				
12				
13				
14				
15				

As usual, the first thing we need to do is enter the data. Make sure that you label the variable "pre" and "post"

Now go to 'Analyze', "Compare Means', and 'Paired Sample T Test'.

Paired-Samples T Test		Paired V	ariables:			×
A n	1	Paireu v	Variable1	VariableQ	1	Options
V Pre		raii 1	Variable I	Variablez		
I Post			[Fie]	∢r (rosij		
		2				
					7	
					÷	
					$\leftrightarrow$	
(			)(-			

You want to add 'Pre' to the variable 1 column and 'Post' to the variable 2 column. Now we are ready to run the test. Select 'OK'.

### **Paired Samples Statistics**

		Mean	Ν	Std. Deviation	Std. Error Mean
Doir 1	Pre	13.2000	10	2.61619	.82731
Part	Post	14.6000	10	2.31900	.73333

**Paired Samples Correlations** 

		N	Correlation	Sig.
Pair 1	Pre & Post	10	.839	.002

#### Paired Samples Test

			Paired Differences						
				Std. Error	95% Confidenc Differ	e Interval of the ence			
		Mean	Std. Deviation	Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pre - Post	-1.40000	1.42984	.45216	-2.42285	37715	-3.096	9	.013

The Mean in the analysis is the mean of the difference between the 'Pre' and 'Post' training scores.

🔄 Summarize Cases
Variables: Pre Post Options
Grouping Variable(s):
☑ Disp <u>l</u> ay cases
✓ Limit cases to first 100
Show only valid cases
Show <u>c</u> ase numbers
OK Paste Reset Cancel Help

A listing of data can be obtained by choosing 'Analyze', 'Reports', and 'Case Summaries' from the SPSS main menu.

As before, using the mouse pointer, highlight the variables 'Pre' and 'Post'. Click on the right hand pointing arrow to move the variables to the 'Variables' text box. Click on 'OK' to list the data.

As before, the results will be displayed in the 'Output 1 SPSS Output Navigator' window.

Case Processing Summary <sup>a</sup>								
	-	Cases						
		Included		Excluded		Total		
		N	Percent	Ν	Percent	N	Percent	
	Pre	10	100.0%	0	0.0%	10	100.0%	
	Post	10	100.0%	0	0.0%	10	100.0%	

a. Limited to first 100 cases.

Case Summaries"						
		Pre	Post			
1		12.00	13.00			
2		15.00	15.00			
3		9.00	11.00			
4		16.00	15.00			
5		12.00	15.00			
6		15.00	18.00			
7		17.00	17.00			
8		10.00	12.00			
9		12.00	13.00			
10		14.00	17.00			
Total	Ν	10	10			

The 'Case Processing Summary' and 'Case Summaries' listings provide a quick and convenient synopsis of the cases comprising your study.

a. Limited to first 100 cases.

# **DEPENDENT T-TEST** Example Reporting of Results

Results

Because the data from the pre-training observations were logically tied to the data from the post-training observations, a dependent t-test was performed. The mean difference between the pre-and post-training scores was 1.40, which was statistically significant, t(9) = 3.10, p = .013.

Discussion

The results showed the post-training scores to be significantly higher after exposure to the relaxation technique. This means that the relaxation technique appears to have a positive effect on the player's free throw shooting accuracy.
# ANALYSIS OF VARIANCE

**One-Way ANOVA** 



# **ANALYSIS OF VARIANCE**





### **Example Problem**

### **Balanced** groups

A direct marketer of insurance wanted to evaluate the effect of age on the response rate to a new insurance product. Below are the response rates per 1,000 mailings by age group from 12 different metropolitan areas.

Table 1

### Response Rate per One Thousand Mailings

Young	Middle	Elderly
25	30	25
27	29	22
23	29	27
24	31	23
23	28	24
24	31	25
22	29	23
25	32	22
21	30	21
24	29	22
21	28	24
23	31	23

# **BALANCED ONE-WAY ANOVA**

ta i					
<u>F</u> ile	<u>E</u> dit	⊻iew	<u>D</u> ata	Transform	<u>A</u> nalyze
2	• 6				
26:					
		а	ge	rate	va
	1		1.00	25.	00
	2		1.00	27.	00
	3		1.00	23.	00
	4		1.00	24.	00
	5		1.00	23.	00
	6		1.00	24.	00
	7		1.00	22.	00
	8		1.00	25.	00
	9		1.00	21.	00
	10		1.00	24.	00
	11		1.00	21.	00
	12		1.00	23.	00
	13		2.00	30.	00
	14		2.00	29.	00
	15		2.00	29.	00
· ·	16		2.00	31.	00
	·				

# **SPSS Procedures**

Define variables 'Age' and 'Rate' using the procedures described in the Independent T-Test. Thirty-six cases are entered for this example.

Go to ' Analyze", ' Compare Means', 'One-Way ANOVA'

				*Untitled'i [DataS
а	<u>T</u> ransform	<u>Analyze</u> <u>G</u> raphs <u>U</u> tilitie	es Add- <u>e</u>	ons <u>W</u> indow <u>H</u> elp
TT		Reports	*	👬 🗰 📅 🥅 👬 🗰 🖉
	•	Descriptive Statistics	۲.	
		Co <u>m</u> pare Means	- F	Means
	rate	<u>G</u> eneral Linear Model		Cone-Sample T Test
.00	25.0	Generalized Linear Mo	odels 🕨	Independent-Samples T Test
.00	27.0	Mi <u>x</u> ed Models	۰.	Paired-Samples T Test
.00	23.0	<u>C</u> orrelate	۰.	
.00	24.0	<u>R</u> egression		
.00	23.0	L <u>o</u> glinear	*	
.00	24.0	Classify	*	
.00	22.0	Dimension Reduction	n 🕨	
.00	25.0	Sc <u>a</u> le	•	
.00	21.0	<u>N</u> onparametric Tests	•	
.00	24.0	Forecasting	•	
.00	21.0	<u>S</u> urvival	*	
.00	23.0	Multiple Response	*	
00	29.0	🟭 Missing Value Analysi	s	
00	29.0	Multiple Imputation	•	
00	31.0	Simulation		
00	28.0	Quality Control	•	
.00	31.0	ROC Curve		
.00	29.0	0		
.00	32.0	00	_	
00	30 (	າດ		

### SPSS 20.0 A Survival Guide

		OK	One-Wa	ay ANOVA Dependent Lis rate	t.	Contrasts Post <u>H</u> oc Options	
--	--	----	--------	-----------------------------------	----	--	--

Place 'rate' in the box for dependent variable and 'age' in as the factor.

	One-Way ANOVA
Statistics	he-Way ANOVA: Post Hoc Multiple Comparisons
Descriptive	Equal Variances Assumed
Eixed and random effects Homogeneity of variance test	Bonferroni Jukey Type I/Type II Error Ratio: 100
Brown-Forsythe	Scheffe Duncan Control Category: Last
Welch	R-E-G-WF Hothberg's GT2 Test
🥅 <u>M</u> eans plot	
Missing Values	Equal Vanances Not Assumed Tamhane's T2 Dunnett's T3 Games-Howell Dunnett's C
© Exclude cases listwise	Significance level: 0.05
Continue Cancel Help	Continue Cancel Help

In the options box, make sure "Descriptives" and 'Homogeneity of Variance' are selected and select continue. In the Post Hoc box make sure 'Tukey" is selected and select continue again. Finally when back to the One Way ANOVA screen select OK.

## The following is the SPSS output from SPSS



This shows the means and standard deviation for each group as well as the grand mean (Total)

#### Descriptives

rate								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Low er Bound	Upper Bound		
young	12	23.5000	1.73205	.50000	22.3995	24.6005	21.00	27.00
middle	12	29.7500	1.28806	.37183	28.9316	30.5684	28.00	32.00
elderly	12	23.4167	1.67649	.48396	22.3515	24.4819	21.00	27.00
Total	36	25.5556	3.37592	.56265	24.4133	26.6978	21.00	32.00

Test of Homogeneity of Variances

rate			
Levene Statistic	df 1	df2	Sig.
.302	2	33	.741



Levene's test results

ANOVA

rate					
	Sum of df Mean Square		Mean Square	F	Sig.
	Squares				
Betw een Groups	316.722	2	158.361	63.601	.000
Within Groups	82.167	33	2.490		
Total	398.889	35			

One Way ANOVA results. Use this to calculate  $\eta^2$ .

$$\eta 2 = \frac{SS Between}{SS \ total}$$

# Tukey's Post Hoc Results

#### Multiple Comparisons

Dependent Variable: rate

Tukey HSD

ukey HSD									
(I) age	(J) age	Mean	Std. Error	Sig.	95% Confidence Interval				
		Difference (I-J)			Low er Bound	Upper Bound			
	middle	-6.25000*	.64419	.000	-7.8307	-4.6693			
young	elderly	.08333	.64419	.991	-1.4974	1.6640			
middle	young	6.25000*	.64419	.000	4.6693	7.8307			
midale	elderly	6.33333 <sup>*</sup>	.64419	.000	4.7526	7.9140			
a baba sabas	young	08333	.64419	.991	-1.6640	1.4974			
elderly	middle	-6.33333 <sup>*</sup>	.64419	.000	-7.9140	-4.7526			

### **ONE-WAY ANOVA**

### **Example Results Write Up**

Table 1 presents the means and standard deviations by age group. A one-way analysis of variance yielded a significant difference among the age group means, F(2,33) = 63.60, p < .001,  $\eta^2 = .794$ . Levene's test showed that the assumption of equal variances was tenable (F(2, 33) = .302, p = .741). Table 2 presents the analysis of variance summary. A post hoc analysis using Tukey's procedure ( $\alpha = .05$ ) revealed the mean for the middle age group was significantly higher than the young and elderly age groups. There was no significant difference between the means of the young and elderly groups.

 Table 1

 Response Rate per One Thousand Mailings by Age Groups and Standard Deviations

Group	N	М	SD	
Young	12	23.50	1.73	
Middle	12	29.75	1.29	
Elderly	12	23.42	1.68	

Table 2Analysis of Variance Summary

63.60*

\*p < .001

#### Discussion

The results indicate a significantly higher response rate from the middle age group for this particular insurance product. Therefore, it is recommended that the direct marketer targets the middle age group for selling the new insurance product.

# **ONE-WAY ANOVA Unbalanced Groups**

### **Example** problem

Locus of Control is a widely used concept that concerns individuals' perceptions about who or what influences important outcomes in their lives. A nursing administrator was interested in examining the relationship between locus of control and open heart surgery patients' length of stay in the hospital. Working with Counseling Services, the administrator arranged for the Multidimensional Health Locus of Control (MHLC) to be administered to 24 patients scheduled for heart valve replacement. Based on the MHLC scores and norms, the patients were then classified into one of three groups: high, average, or low. High locus of control indicates a sense of personal control. Low locus of control indicates a sense of control by outside factors. The number of days required in the hospital by patient by classification is presented below.

Table 1

#### Number of Hospital Days by Locus of Control Group

Group	Hospital days by Patient									
Low	14	27	23	25	28	25	15	27		
Average	15	20	24	13	16	16	14	20	14	18
High	11	17	6	5	8	7				



# UNBALANCED ONE-WAY ANOVA

# **SPSS Procedures**

<b>t</b>								
<u>F</u> ile	<u>E</u> dit	View	<u>D</u> ata	<u>T</u> ransform	An	nalyze	G	
			) 🗖	J 🗠	C			
			ays)	Group		va	r	
	1		14.00	1.	00			
:	2		27.00	1.	00			
:	3		23.00	1.	00			
4	4		25.00	1.	00			
	5		28.00	1.	00			
	6		25.00	1.	00			
	7		15.00	1.	00			
	8		27.00	1.	00	) (		
	9		15.00	2	00			
1	0		20.00	2	00			
1	1		24.00	2	00			
1	2		13.00	2	00			
1	3		16.00	2	00			
1	4		16.00	2	00			
1	5		14.00	2	00			
1	6		20.00	2	00			

Define variables 'Days' and 'Group'.

									*spssg	uide u	nbalan	cedA	NOVA.
v <u>D</u> ata	<u>T</u> ransform	<u>Analyze</u> G	raphs	<u>U</u> tilities	Add-	ons	<u>W</u> indow	Н	elp				
		Re <u>p</u> ort: D <u>e</u> scrij	s otive Sta	tistics	۲ ۲		1	5					/
		Co <u>m</u> pa	re Mear	IS	•	M	Means						
Days	Group	Genera	l Linear	Model	•	t	 One-San	əlar	T Test				var
14.00	1.0	Genera	lized Lir	near Mod	els ▶	3	- Independ	ient-	Sample	s T Tes	t		
27.00	1.0	Mixed M	lodels		•		Deired C			o 1 100			
23.00	1.0	<u>C</u> orrela	te		•		raneu-o	amp		SL			
25.00	1.0	<u>R</u> egres	sion		•		<u>O</u> ne-way	AN	JVA				
28.00	1.0	L <u>o</u> gline	ar		•								
25.00	1.0	Classif	y		•								
15.00	1.0	Dimen	sion Re	duction									
27.00	1.0	Scale			•								
15.00	2.0	Nonpa	ametric	Tests	*								
20.00	2.0	Eoreca	stina										
24.00	2.0	Surviva	1										
13.00	2.0	Multiple	Reeno	neo	, k								
16.00	2.0	Minoin	Volue	Analysia	r								
16.00	2.0	Wissing Wissing	, value /	niaiysis									
14.00	2.0	Multiple	Imputa	tion									
20.00	2.0	🎼 Simula	tion										
14.00	2.0	<u>Q</u> uality	Control		•								
18.00	2.0	🖉 ROC C	urve										
11.00	3.0	0				·							
17.00	3.0	0											
6.00	3.0	0											

Go to ' Analyze", ' Compare Means', 'One-Way ANOVA'



the One Way ANOVA screen select OK.

This shows the means and standard deviation for each group as well as the grand mean (Total)

Levene's test results.

#### Descriptives

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Low er Bound	Upper Bound		
low	8	23.0000	5.47723	1.93649	18.4209	27.5791	14.00	28.00
average	10	17.0000	3.46410	1.09545	14.5219	19.4781	13.00	24.00
high	6	9.0000	4.42719	1.80739	4.3540	13.6460	5.00	17.00
Total	24	17.0000	6.87782	1.40393	14.0958	19.9042	5.00	28.00

### Test of Homogeneity of Variances

Days

Days

Levene Statistic	df 1	df2	Sig.
.781	2	21	.471



Multiple Comparisons

Tukey's Post Hoc Results

Dependent Variable: Days

Tukey HSD

(I) group	(J) group	Mean	Std. Error	Sig.	95% Confide	ence Interval	
		Difference (I-J)			Low er Bound	Upper Bound	
low	average	6.00000*	2.11119	.025	.6786	11.3214	
IOW	high	14.00000*	2.40370	.000	7.9413	20.0587	
	low	-6.00000*	2.11119	.025	-11.3214	6786	
average	high	8.00000*	2.29838	.006	2.2068	13.7932	
h i e h	low	-14.00000*	2.40370	.000	-20.0587	-7.9413	
nign	average	-8.00000*	2.29838	.006	-13.7932	-2.2068	

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

### **UBALANCED ONE-WAY ANOVA**

#### **Example Reporting of Results**

Table 1 shows the mean number of hospital days and standard deviations by locus of control group. Levene's test showed that the assumption of equal variances was tenable (F(2, 21) = .781, p = .471). The one-way analysis of variance procedure yielded a significant difference among the locus of control group means  $F(2,21) = 16.96, p < .001, \eta^2 = .62$ . Table 2 presents the analysis of variance summary. A post hoc analysis using Tukey's procedure ( $\alpha = .05$ ) revealed the mean number of hospital days for all three groups were significantly different (see Figure 1).

Table 1

Mean Hospital days and Standard Deviations by Locus of Control Groups

Group	Ν	М	SD	
Low	8	23.00	5.48	
Average	10	17.00	3.46	
High	6	9.00	4.43	

Table 2Analysis of Variance Summary

Source	Sum of Squares	DF	Mean Square	F
Group	672.00	2	336.00	16.96*
Error	416.00	21	19.81	
Total	1088.00	23		
*p < .001				



Figure 1. Mean Number of Hospital Days

#### Discussion

The results of this study showed that patients who have a high locus of control, that is, feel they have a high degree of personal control over events, appear to be discharged much sooner than those patients with average or low locus of control. Furthermore, even patients with average locus of control can be expected to be discharged sooner than patients with a low locus of control. Figure 1 clearly shows the relationship of locus of control to the mean number of hospital days by group.

Comments about the Results and Discussion Section:

There are three ways you can report the results of the statistical treatments: (1) within the text of the section, (2) with one or more summary tables, and (3) graphically. As illustrated in this example, an appropriate graph (technically called a figure) sometimes can help the reader gain a clearer understanding of how the study turned out.

Note: Figure 1 was produced in SPSS. See the Descriptive statistics portion of this guide for instructions of how to produce the graph.

# FACTORIAL ANALYSIS OF VARIANCE

## Two-Way ANOVA

Two-Way ANOVA with Significant Interaction



# **PROCEDURAL MODEL: OVERVIEW**

One

Way

ANOVA

ANCOVA

Factorial

ANOVA

Factorial

ANCOVA

Independent

T-Test

Dependent

T-Test

Repeated

Measures

ANOVA



### **Factorial ANOVA Procedure Model**

(1) The MSw term from the main factorial ANOVA is more reflective of the population variance. The MSw term from the main factorial ANOVA should be used as the error term for the subsequent simple main effects analysis and in the post hoc analysis.

# **TWO-WAY ANOVA**

### **Example Problem**

A marketing manager for a supermarket chain was interested in the effect of both retail price and display location on a new promotional line of cookies. A group of 24 stores with matching store volume, layout, and customer demographics was split at random into 6 groups of 4 stores each. One group was assigned to each of the 6 combinations of retail price (regular retail vs. discounted retail) and display location (entrance aisle, cookie aisle, and checkout). Below are the average weekly unit sales by store over a 13 week period.

Table 1

Promotional Cookie: Average Weekly Unit Sales

		DISPLAY LOCATIONS				
		1 Entrance	2 Cookies	3 Checkout		
	1. Regular Retail	38	28	21		
		31	25	32		
		27	23	30		
		33	20	22		
Price						
	2. Discount Retail	35	22	19		
		21	24	15		
		39	16	25		
		30	17	20		

*Untitled1	[DataSet0] - IBN	1 SPSS Statistics	Data Editor
<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Transform A	nalyze <u>G</u> raphs
			> 🖹 📩
22 :			
	Price	Location	Sales
1	1.00	1.00	38.00
2	1.00	1.00	31.00
3	1.00	1.00	27.00
4	1.00	1.00	33.00
5	1.00	2.00	28.00
6	1.00	2.00	25.00
7	1.00	2.00	23.00
8	1.00	2.00	20.00
9	1.00	3.00	21.00
10	1.00	3.00	32.00
11	1.00	3.00	30.00
12	1.00	3.00	22.00
13	2.00	1.00	35.00
14	2.00	1.00	21.00
15	2.00	1.00	39.00
16	2.00	1.00	30.00
17	2.00	2.00	22.00

Define three variables as: 'Price', 'Location', and 'Sales'.

For the variable 'Price', assign the variable labels of 'Retail' to the value of '1' and 'Discount' to the value of '2'.

For the variable 'Location', assign the variable labels of 'Entrance' to the value of '1', 'Aisle' to the value '2', and 'Checkout' to the value of '3'. Note: The entire data set is not displayed.

AVO	sav [DataSet0	] - IBM SP	SS Statistic	s Data Edit	or								
ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Add-	ons	<u>W</u> indow	He	elp				
		Rep	orts		•	3				5			A (
	<ul> <li>▼</li> </ul>	D <u>e</u> s	criptive Sta	atistics	•							1.6	3
		Co <u>m</u>	Co <u>m</u> pare Means								_		
•	Location	<u>G</u> en	eral Linea	r Model	- Þ.	GLM	<u>U</u> nivariate	ə			var		
1.00	1.0	Gen	erali <u>z</u> ed Li	near Mode	els ►	GUM	Multivariat	te					
1.00	1.0	Mi <u>x</u> e	d Models		•	GLM	 Repeated	Mea	sures				
1.00	1.0	<u>C</u> orr	elate		•		Varianco (	Com	nonont		-		
1.00	1.0	<u>R</u> eg	ression				vanance v	Com	ponent	ə			
1.00	2.0	L <u>o</u> gl	inear		•								
1.00	2.0	Clas	sify		•								
1.00	2.0	<u>D</u> im	ension Re	duction	•								
1.00	2.0	Sc <u>a</u> l	e		•								
1.00	3.0	Non	parametrio	c Tests									
1.00	3.0	Fore	casting		•								
1.00	3.0	<u>S</u> urv	ival										
2.00	1.0	M <u>u</u> lti	iple Respo	onse	•								
2.00	1.0	ジ Miss	sing Value	Analysis									
2.00	10	Multi	iple Imputa	ation	•								
2.00	1.0	Qua	lity Control	I									
2.00	2.0	ROC	Curve										
2.00	2.0	0	24.00										
2.00	2.0	0	16.00										
2.00	2.0	0	17.00										
2.00	3.0	0	19.00										

Go to 'Analyze', 'General Linear Model', and select 'Univariate' since you have only one dependent variable.

ta Univariate		
Image: Constraint of the second s	Dependent Variable: ✓ Sales Fixed Factor(s): ✓ Location ✓ Price Random Factor(s): Covariate(s): <u>ULS Weight:</u> <u>Reset</u> Cancel Help	Model Co <u>n</u> trasts Plo <u>t</u> s Post <u>H</u> oc Save Options

Place 'Sales' in the Dependent variable box and 'Location' and 'Price' in the Fixed Factor(s) box.

늘 Univariate: Options	×				
Estimated Marginal Means					
Eactor(s) and Factor Interactions:	: Display <u>M</u> eans for:				
(OVERALL)	Location				
Price	Location*Price				
Location*Price					
	Compare main ellects				
	Confidence interval adjustment:				
	LSD(none)				
Display					
Descriptive statistics	✓ Homogeneity tests				
Estimates of effect size	Spread vs. level plot				
Observed power	🔲 <u>R</u> esidual plot				
Parameter estimates	Lack of fit				
Contrast coefficient matrix	General estimable function				
Significance level: .05 Confidence intervals are 95.0 %					
Continue Cancel Help					

In the options menu we want to make sure 'Descriptive statistics'', 'Estimates of Effect Size', 'Observed power', and 'Homogeneity test' are selected. We also want to display group means for 'Location', 'Price' and 'Location\*Price'. Select 'Continue'

👍 Univariate: Profile Plots	<b>—</b>						
<u>F</u> actors: Location Price	Horizontal Axis:         Location         Separate Lines:         Price         Separate Plots:						
Plo <u>t</u> s: <u>A</u> dd	Change Remove						
Continue Cancel Help							

We also would like to see plots so under the plots menu. Place 'Location' in the Horizontal Axis box and 'Price' in the Separate lines box. Select 'Add' to add this plot then 'Continue'. This will take you back to the Univariate box.

🔚 Univariate: Post Hoc Multiple Comparisons for Observed Means 🥢 💌							
Factor(s):	Post Hoc Tests for:						
Location	Location						
Price	•						
Equal Variances Assumed							
ESD S-N-K	Maller-Duncan						
🔲 Bonferroni 👿 Tukey	Type I/Type II Error Ratio: 100						
🔲 S <u>i</u> dak 📃 Tu <u>k</u> ey's-b	🕅 Dunn <u>e</u> tt						
🔲 Scheffe 🛛 🔲 Duncan	Control Category:						
🔲 <u>R</u> -E-G-W-F 📃 <u>H</u> ochberg's GT	2 Test						
🔲 R-E-G-W- <u>Q</u> 🕅 <u>G</u> abriel	O 2-sided $O$ < Control $O$ > Control						
Equal Variances Not Assumed Tamhane's T2 Dunnett's T3 Games-Howell Dunnett's C							
Continue Cancel Help							

Finally, let's go to the Post Hoc menu where we will place 'Location' in the box and select Tukey for the Post Hoc analysis. We do not need to do a Post Hoc analysis on Price since there are only two categories. Select 'Continue', then select 'OK' in the Univariate box.

# SELECTED SPSS OUTPUT

Dependent Variable: Sales							
Location	Price	Mean	Std. Deviation	Ν			
	Regular	32.2500	4.57347	4			
Entrance	Discount	31.2500	7.76209	4			
	Total	31.7500	5.92211	8			
	Regular	24.0000	3.36650	4			
Cookies	Discount	19.7500	3.86221	4			
	Total	21.8750	4.05101	8			
	Regular	26.2500	5.56028	4			
Checkout	Discount	19.7500	4.11299	4			
	Total	23.0000	5.70714	8			
	Regular	27.5000	5.51856	12			
Total	Discount	23.5833	7.56137	12			
	Total	25.5417	6.77578	24			

#### **Descriptive Statistics**

Descriptive statistics with cell means and Standard deviations

### Levene's Test of Equality of Error Variances<sup>a</sup>

Dependent Variable: Sales

F	df 1	df2	Sig.	
1.120	5	18	.385	

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

ine dependent variable is equal across groups.

a. Design: Intercept + Location + Price + Location \* Price





ANOVA Table

Dependent Variable: Sales									
Source	Type III Sum of	df	Mean Square	F	Sig.	Partial Eta	Noncent.	Observed	
	Squares					Squared	Parameter	Pow er <sup>b</sup>	
Corrected Model	590.208 <sup>a</sup>	5	118.042	4.562	.007	.559	22.810	.908	
Intercept	15657.042	1	15657.042	605.103	.000	.971	605.103	1.000	
Location	467.583	2	233.792	9.035	.002	.501	18.071	.947	
Price	92.042	1	92.042	3.557	.076	.165	3.557	.431	
Location * Price	30.583	2	15.292	.591	.564	.062	1.182	.133	
Error	465.750	18	25.875						
Total	16713.000	24							
Corrected Total	1055.958	23							

**Tests of Between-Subjects Effects** 

#### a. R Squared = .559 (Adjusted R Squared = .436)

b. Computed using alpha = .05

#### Notes:

[1] Initial interpretation.

[a] The main effects are significant. The question is 'which main effect(s)' is/are significant?

[b] The 'LOCATION' effect is significant, p = .002.

[c] The 'PRICE' effect is not significant, p = .076.

[d] The 'LOCATION PRICE' is called an interaction term. If the location differences were not the same for the two price factors, we would expect a significant interaction. For example, if the regular price did better in the entrance location, but the discount price did better at the checkout, we'd have an interaction, i.e., the effect of price depends on location. In this example, the interaction is not significant, p = .564

Tukey	HSD
-------	-----

(I) Location (J) Location		Mean	Std. Error	Sig.	95% Confide	ence Interval
		Difference (I-J)			Low er Bound	Upper Bound
	Cookies	9.8750 <sup>*</sup>	2.54337	.003	3.3839	16.3661
Entrance	Checkout	8.7500 <sup>*</sup>	2.54337	.008	2.2589	15.2411
Castrias	Entrance	-9.8750 <sup>*</sup>	2.54337	.003	-16.3661	-3.3839
Cookies	Checkout	-1.1250	2.54337	.898	-7.6161	5.3661
Checkout	Entrance	-8.7500 <sup>*</sup>	2.54337	.008	-15.2411	-2.2589
	Cookies	1.1250	2.54337	.898	-5.3661	7.6161

Post Hoc results



Based on observed means.

The error term is Mean Square(Error) = 25.875.

\*. The mean difference is significant at the .05 level.



Note: This graph shows that the graphs are basically parallel. You can see a large different between the means at the three locations however the difference between the prices is not as big.

# **TWO-WAY ANOVA**

### **Example Results and Discussion**

#### Results

A two-way analysis of variance yielded no significant sales difference due to price and no significant interaction between price and display location. However, there was a significant difference among the display locations, F(2,18) = 9.04, p = .002. A post hoc analysis using Tukey's procedure ( $\alpha = .05$ ) revealed the mean sales for the entrance location was significantly higher than both the cookie aisle and checkout location. There was no significant difference between the mean sales of cookie aisle and checkout location. Table 1 presents the means and standard deviations by group. Table 2 presents the analysis of variance summary.

#### Table 1

Mean Sales and Standard Deviations of Unit Sales as a Function of Price and Display Location

	Display Location								
Retail Price		Entrance	Entrance Cookies		Combined				
Regular									
	N	4	4	4	12				
	M	32.25	24.00	26.25	27.50				
	SD	4.57	3.37	5.56	5.52				
Discount									
	Ν	4	4	4	12				
	M	31.25	19.75	19.75	23.58				
	SD	7.76	3.86	4.11	7.56				
Combined									
	Ν	8	8	8	24				
	M	31.75	21.88	23.00	25.54				
	SD	5.92	4.05	5.71	6.78				

#### Table 2

#### Analysis of Variance Summary

Source	Sum of Squares	DF	Mean Square	F Value	p value
Price	92.04	1	92.04	3.56	.076
Location	467.58	2	233.79	9.04	.002
Price*Location	30.58	2	15.29	.591	.564
Error	465.75	18	25.86		

### Discussion

Locating the new promotional cookie display in the entrance aisle appears to produce the most unit sales. It is interesting to note that there was no significant difference in sales based on retail price. The company can enjoy higher profits by displaying the cookies at regular price, regardless of the location therefore it is recommended that the store keep regular prices and locate the cookies where customers can see them as they enter the store.

# TWO WAY ANOVA With Significant interaction

### **Example Problem**

Eighteen students were randomly assigned to one of three different classroom teaching methods (online, hybrid, and in person) to learn a new math concept. At the end of 12 weeks, students were given a test to assess their understanding of the concepts. A two-way ANOVA will be run to assess the difference between the teaching methods as well as between student's previous online experiences. The table below shows the scores by gender and method. Two students did not finish the 12 week course so the distribution between the teaching methods and the experience was not equal.

	1.In person	2. Hybrid	3. Online
1. Experienced(Has taken an online course in the past)	57.5 80 62.5	72.5 75	77.5 90 82.5
2. Not Experienced (Has not taken an online class in the past)	80 65 65	85 100 87.5	57.5 65

I CAN'T BELIEVE SCHOOLS ARE STILL TEACHING KIDS ABOUT THE NULL HYPOTHESIS. I REMEMBER READING A BIG STUDY THAT CONCLUSIVELY DISPROVED IT HEARS AGO.

<b>1</b>							
<u>F</u> ile	<u>E</u> dit	View	<u>D</u> ata	<u>T</u> ransform	<u>A</u> nalyze	<u>G</u> raph	ns <u>U</u> tilities .
		Exp	erience	Method	So	core	var
	1		1.00	1.0	0	57.50	
	2		1.00	1.0	0	80.00	
	3		1.00	1.0	0	62.50	
	4		1.00	2.0	0	72.50	
	5		1.00	2.0	0	75.00	
	6		1.00	3.0	0	77.50	
	7		1.00	3.0	0	90.00	
	8		1.00	3.0	0	82.50	
	9		2.00	1.0	0	80.00	
1	10		2.00	1.0	0	65.00	
1	11		2.00	1.0	0	65.00	
1	12		2.00	2.0	0	85.00	
1	13		2.00	2.0	0	100.00	

# **SPSS Procedure**

Define three variables for your data set: 'Experience', 'Method', and 'Score'.

Define variable labels for 'Experience' of '1' indicating experienced, and '2' indicating not experienced.

Define variable labels for 'Method' of '1' indicating in person and '2' indicating hybrid and '3' indicating online.

Go to 'Analyze', 'General Linear Model', and select 'Univariate' since you have only one dependent variable

								*unba	alanced	twoway/
n	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Add-	ons <u>V</u>	<u>V</u> indow	<u>H</u> e	lp		
	Repo	orts		•	- 64	*				🗠 🗄
	D <u>e</u> so	criptive Sta	atistics	•			Ĵ) ⊢			~⊖ ⊞
	Co <u>m</u>	pare Meai	ns	•						
ick	<u>G</u> ene	eral Linea	r Model	- •	翻 <u>U</u>	nivariate	e			var
2.0	Gene	erali <u>z</u> ed Li	near Mode	ls ►	GUM N	ultivaria	te			
2.0	Mixe	d Models		•		epeated	l Mea	sures		
2.0	Corre	elate		•	v	ariance (	Com	onent	s	
2.0	<u>R</u> egr	ession		►l	-					
2.0	L <u>o</u> gli	near		•						
1.0	Clas	sify		•						
1.0	<u>D</u> ime	ension Re	duction	•						
1.0	Sc <u>a</u> le	9		•						
1.0	Non	parametric	Tests	•						
1.0	Fore	casting		•						
1.0	<u>S</u> urvi	val		•						
1.0	Multi	ple Respo	onse	•						
1.0	ジ Miss	ing Value	Analysis							
1.0	Mul <u>t</u> i	ple Imputa	ation	•						
1.0	🖶 Simu	lation								
2.0	Qual	ity Control		•						
2.0	R0C	Curve								



Place 'Score' in the Dependent variable box and 'Experience' and 'Method' in the Fixed Factor(s) box.



In the options menu we want to make sure 'Descriptive statistics'', 'Estimates of Effect Size', 'Observed power', and 'Homogeneity test' are selected. We also want to display group means for 'Experience', 'Method' and 'Experience\*Method'. Select 'Continue'

		U Univaria	nivariate te: Profile	e Plots		×		
	actors: Experience Method Plo <u>t</u> s: Method*Experi	Add	H S Change	orizontal Axis: eparate Lines: eparate Plots: a) <u>Remova</u>	, , , , , , , , , , , , , , , , , , ,			
	(	Continue	Cancel	Help				
OK Paste Reset Cancel Help								

We also would like to see plots so under the plots menu. Place 'Method' in the Horizontal Axis box and 'Experience' in the Separate lines box. Select 'Add' to add this plot then 'Continue'

actor(s): Experience Method	Post Hoc Tests for:
Equal Variances Assumed LSD S-N-K Egenterron V Tukey Sidak Tukey's-b Scheffe Duncan R-E-G-W-F Hochberg's R-E-G-W-Q Gabriel	<u>Waller-Duncan</u> Type I/Type II Error Ratio: 100     Dunn <u>ett</u> Control Category: Last GT2 Test <u>@ 2-sided @ &lt; Control @ &gt; Cont</u>
Equal Variances Not Assumed	t s T3 III Games-Howell III Dunnett's C 3 Cancel Help

Since we have three groups we will need to perform a post hoc test if we see group significance. In the Post Hoc box add Method to the box for Post Hoc test and select Bonferoni and Tukey as the post hoc test to run. Select 'Continue'.

This will take you back to the Univariate box.

Select 'OK' in the Univariate box. To run the analysis.

# SELECTED SPSS OUTPUT

#### **Descriptive Statistics**

**Descriptive Statistics** 



Dependent Variable: Score						
Experience	Method	Mean	Std. Deviation	Ν		
	in person	61.6667	11.81454	3		
	hybrid	73.7500	1.76777	2		
experienced	online	83.3333	6.29153	3		
	Total	72.8125	12.35180	8		
	in person	65.0000	8.66025	3		
not experienced	hybrid	90.8333	8.03638	3		
not experienced	online	61.2500	5.30330	2		
	Total	73.7500	15.69804	8		
	in person	63.3333	9.44281	6		
Total	hybrid	84.0000	10.98294	5		
TOLAI	online	74.5000	13.15770	5		
	Total	73.2813	13.65402	16		

#### Levene's Test of Equality of Error Variances<sup>a</sup>

Dependent Variable: Score

F	df 1	df2	Sig.
1.779	5	10	.205

Tests the null hypothesis that the error variance of the dependent variable is equal across groups. a. Design: Intercept + Experience + Method + Experience \* Method Levene's Test





Dependent Variable: Score								
Source	Type III Sum	df	Mean	F	Sig.	Partial Eta	Noncent.	Observed
	of Squares		Square			Squared	Parameter	Pow er <sup>b</sup>
Corrected Model	2127.734 <sup>a</sup>	5	425.547	6.363	.007	.761	31.817	.934
Intercept	81407.440	1	81407.440	1217.308	.000	.992	1217.308	1.000
Experience	1.190	1	1.190	.018	.897	.002	.018	.052
Method	959.391	2	479.696	7.173	.012	.589	14.346	.833
Experience *	951.827	2	475.913	7.116	.012	.587	14.233	.830
Method								
Error	668.750	10	66.875					
Total	88718.750	16						
Corrected Total	2796.484	15						

#### Tests of Between-Subjects Effects

a. R Squared = .761 (Adjusted R Squared = .641)

b. Computed using alpha = .05

### Note:

# A significant interaction effect exists between the independent variables of 'Method' and 'Experience'.

Always check for an interaction effect first when performing analysis on a higher order design. If a significant interaction exists, do not interpret the main interaction. Instead, perform a 'simple main effects' analysis.

'Simple main effects' analysis is a One-Way ANOVA which is conducted to reduce the analysis to the levels of the independent variables in order to examine the cause of the interaction effects.

In this case, we wish to compare the cell means for the three teaching methods for each experience level horizontally in terms of the data chart (Figure 1.).

## Post Hoc analysis



#### **Multiple Comparisons**

Dependent Variable: Score							
	(I) Method (J) Method		Mean Difference Std. Error		Sig.	95% Confidence Interval	
			(I-J)			Low er Bound	Upper Bound
	in person	hybrid	-20.6667*	4.95185	.005	-34.2412	-7.0922
		online	-11.1667	4.95185	.110	-24.7412	2.4078
Tukey HSD	hybrid	in person	20.6667*	4.95185	.005	7.0922	34.2412
		online	9.5000	5.17204	.207	-4.6781	23.6781
	online	in person	11.1667	4.95185	.110	-2.4078	24.7412
		hybrid	-9.5000	5.17204	.207	-23.6781	4.6781
Bonferroni	in person	hybrid	-20.6667*	4.95185	.006	-34.8788	-6.4545
		online	-11.1667	4.95185	.143	-25.3788	3.0455
	hybrid	in person	20.6667*	4.95185	.006	6.4545	34.8788
		online	9.5000	5.17204	.288	-5.3441	24.3441
	online	in person	11.1667	4.95185	.143	-3.0455	25.3788
		hybrid	-9.5000	5.17204	.288	-24.3441	5.3441

Based on observed means.

The error term is Mean Square(Error) = 66.875.

\*. The mean difference is significant at the .05 level.



Interaction plot
# **Investigating Simple Main Effects**

There are two ways to analyze simple main effects. The figures below show both the horizontal and vertical way to analyze the effects.

Figure 1. Simple Main Effects in a Horizontal Direction.



Figure 2. Simple Main Effects in a Vertical Direction.



If one of the variables in the analysis is an attribute variable such as gender and another variable is an active variable such as method, performing a simple main effects analysis within each gender level is sufficient. Since gender cannot be randomly assigned, it makes sense to answer the following questions:

- (a) Given females, which method (1, 2, or 3) works better?
- (b) Given males, which method (1, 2, or 3) works better?

Instead of conducting a simple main effects analysis as described above, many researchers will use an alternative method to understand the nature of a significant interaction. In the alternative method, all pair wise comparisons among all cell means are made using a post hoc test such as the Scheffe' test.

In our case, experience will be the assigned (moderator) variable because it is intrinsic to the participant and method will be the active (focal) variable because it is determined by the researcher. We are assuming that the effect of the method depends on experience. For more information about simple main effects, please read the Interaction handout (Oshima & McCarty 2000) available at http://coeweb.gsu.edu/coshima/statistics\_2.htm

Simple main analysis can be done in SPSS using Syntax. If you are not familiar with how to use syntax, the SPSS Tips and Tricks will give you a short overview of what it is and how to use it. The example that follows just shows a little on how to use SPSS syntax.

Syntax for simple main effects

<b>t</b> a		taSet0
File Edit View Data Transform Analyze	Graphs Litilities Add-or	Elle Edit View Data Transform Analyze Graphs Utilities Add-gns Run Tools Window Help
New	🗊 <u>D</u> ata	Var
<u>O</u> pen ▶	🔁 Syntax	1.
Open Data <u>b</u> ase	Cutput	
Read Text Data	var	
	Script	
Read Cognos Data	80.00	
Close Ctrl+F4	80.00	
	62.50	
The Ctri+S	72.50	
Save As	75.00	
E Save All Data	05.00	
ave Ali Data	85.00	
🗞 Export to Database	77.50	
hark File Read Only	90.00	
	82.50	
🌃 Rena <u>m</u> e Dataset	80.00	
Display Data File Information	00.00	
Tache Data	65.00	
	65.00	
Stop Processor Ctrl+Period	85.00	
Cwitch Conjor		
		IBM SPSS Statistics Processor is ready In 1 Col 1
		🧀 🚞 🗷 P∄ 📲 ⊌ 🍕 😌 💊 🐼 🖉

With the data open, navigate to 'File', 'New', 'Syntax' and a new syntax editor will pop up

Copy and paste the following directly into the syntax editor as it is written.

UNIANOVA Score BY Method Experience /METHOD = SSTYPE(3) /INTERCEPT = INCLUDE /EMMEANS = TABLES (Method\*Experience) COMPARE (Method) ADJ(BONFERRONI) /CRITERIA = ALPHA(.05) /DESIGN =Method Experience Method\*Experience.





Finally select 'Run', 'All'

# **SPSS** Output

Between-Subjects Factors						
		Value Label	Ν			
	1.00	in person	6			
Method	2.00	hybrid	5			
	3.00	online	5			
Functionan	1.00	experienced	8			
Experience	2.00	not experienced	8			

Descriptive statistics

Regular ANOVA table

Tests of Between-Subjects Effects								
Dependent Variable: So Source	Type III Sum of Squares	df	Mean Square	F	Sig.			
Corrected Model	2127.734 <sup>a</sup>	5	425.547	6.363	.007			
Intercept	81407.440	1	81407.440	1217.308	.000			
Method	959.391	2	479.696	7.173	.012			
Experience	1.190	1	1.190	.018	.897			
Method * Experience	951.827	2	475.913	7.116	.012			
Error	668.750	10	66.875					
Total	88718.750	16						
Corrected Total	2796.484	15						

a. R Squared = .761 (Adjusted R Squared = .641)



#### Estimates

Dependent	Variable <sup>.</sup>	Score
Dependent	vanabie.	00010

Method	Experience	Mean	Std. Error	95% Confidence Interval	
				Low er Bound	Upper Bound
	experienced	61.667	4.721	51.147	72.187
in person	not experienced	65.000	4.721	54.480	75.520
ha da al al	experienced	73.750	5.783	60.866	86.634
nybrid	not experienced	90.833	4.721	80.313	101.353
	experienced	83.333	4.721	72.813	93.853
online	not experienced	61.250	5.783	48.366	74.134

## Bonferroni Post Hoc test



#### **Pairwise Comparisons**

Dependent Variable: Score

Experience	(I) Method	(J) Method	Mean Difference	Std. Error	Sig. <sup>b</sup>	95% Confiden	ce Interval for
			(I-J)			Differ	ence <sup>b</sup>
						Low er Bound	Upper Bound
		hybrid	-12.083	7.465	.410	-33.509	9.342
	in person	online	-21.667*	6.677	.026	-40.830	-2.503
	ha de sé d	in person	12.083	7.465	.410	-9.342	33.509
experienced	hybrid	online	-9.583	7.465	.685	-31.009	11.842
	P	in person	21.667 <sup>*</sup>	6.677	.026	2.503	40.830
	online	hybrid	9.583	7.465	.685	-11.842	31.009
	:	hybrid	-25.833 <sup>*</sup>	6.677	.009	-44.997	-6.670
	in person	online	3.750	7.465	1.000	-17.676	25.176
n at aumanian a al	las da ni al	in person	25.833 <sup>*</sup>	6.677	.009	6.670	44.997
not experiencea	nybrid	online	29.583 <sup>*</sup>	7.465	.008	8.158	51.009
	opling	in person	-3.750	7.465	1.000	-25.176	17.676
	online	hybrid	-29.583 <sup>*</sup>	7.465	.008	-51.009	-8.158

Based on estimated marginal means

 $^{\ast}\!.$  The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

# Simple Main Effects ANOVA table



#### Univariate Tests

Dependent Variable: Score

Experience		Sum of Squares	df	Mean Square	F	Sig.
	Contrast	706.510	2	353.255	5.282	.027
experienced	Error	668.750	10	66.875		
not experienced	Contrast	1417.708	2	708.854	10.600	.003
	Error	668.750	10	66.875		

Each F tests the simple effects of Method within each level combination of the other effects show n.

These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

### TWO WAY ANOVA with Significant Interaction

#### **Example Reporting of Results**

A two-way analysis of variance was conducted to investigate the effect of three teaching methods on math concept retention. The two independent variables in the analysis were method and experience. The interaction of the method factor by experience was significant, F(2,10) = 7.12, p = .012, suggesting that method is moderated through experience (see Figure 1).

Figure 1.



Estimated Marginal Means

A simple main effects analysis using method as the focal variable indicated the mean score for the methods were significantly different for those with experience, F(2,10) = 5.82, p = .027 as well as those without experience, F(2,10) = 10.60, p = .003. A follow up Bonferroni post hoc analysis shows that for students with previous online experience, the mean online score is significantly higher than the mean in person score. The post hoc results also show that the hybrid score is significantly higher than both the in person score and the online score for students who do not have previous online experience. Table 1 presents the number of subject, means, and standard deviations for all the groups. Table 2 presents the analysis of variance summary.

# Table 1

## **Descriptive** statistics

			Methods		
Experience		In person	Hybrid	Online	Combined
Experienced	Number of Subjects	3	2	3	8
-	Mean Scores	61.67	73.75	83.33	72.81
	Standard deviation	11.81	1.77	6.29	12.35
Not	Number of Subjects	3	3	2	8
Experienced	Mean Scores	65.00	90.83	61.25	73.75
-	Standard deviation	8.66	8.04	5.30	15.70
Combined	Number of Subjects	6	5	5	16
	Mean Scores	63.33	84.00	74.50	73.28
	Standard deviation	9.44	10.98	13.16	13.65

## Table 2.

#### Analysis of Variance Summary

Source	Sum of Squares	DF	Mean Squares	F value	p value
Method	959.39	2	479.70	7.17	.012
Experience	1.19	1	1.19	.018	.897
Method* Experience	951.83	2	475.91	7.12	.012
Error	668.75	10	66.86		

When those without experience are taught new math concepts using the hybrid method, they have significantly higher concept retention scores than the other methods. Those with experience had significantly higher online scores than in person scores. Due to these results, it is recommended that when teaching new math concepts, teachers should take students past online experiences into consideration when selecting a teaching method.

# ANALYSIS OF COVARIANCE

**One-Way ANCOVA** 

# **PROCEDURAL MODEL: OVERVIEW**

# **Analysis of Covariance**



# One Way ANCOVA

# **Example Problem**

A personnel manager wished to evaluate the effect of positive and negative reinforcement on tardiness in a large manufacturing plant. A group of 30 chronic late arrivers were identified based on their previous tardiness records. They were then randomly assigned in equal numbers to one of three study groups: (1) positive reinforcement when on time, (2) negative reinforcement when late, and (3) no reinforcement. After a 10 week treatment period, data was collected on the number of tardies over an additional 10 week period. The data below presents the number of tardies for 10 weeks prior to the intervention and 10 weeks following termination of the treatment.

Table 1.

Number of Tardies

1. Positive	Reinforcement	2. Negat	ive Reinforcement	3. No re	inforcement
Pre	Post	Pre	Post	Pre	Post
4	2	4	3	5	5
5	3	5	5	5	4
6	1	5	3	5	4
7	4	6	6	6	8
7	3	7	6	8	8
8	5	8	7	8	9
9	3	8	6	9	7
9	5	9	8	10	10
10	6	10	6	10	8
11	5	11	4	7	10

Type of Reinforcement

One	Way ANCOVA
SPS	SS Procedures

🔚 ANCOVA.sav [DataSet2] - IBM SPSS Statistics Data Editor						
<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Transform A	nalyze <u>G</u> raph	ıs <u>U</u>		
		📮 🗠 🦳	<b>&gt; i</b>	L.		
28 :						
	Group	Pre	Post	~		
1	1.00	4.00	2.00			
2	1.00	5.00	3.00			
3	1.00	6.00	1.00			
4	1.00	7.00	4.00			
5	1.00	7.00	3.00			
6	1.00	8.00	5.00			
7	1.00	9.00	3.00			
8	1.00	9.00	5.00			
9	1.00	10.00	6.00			
10	1.00	11.00	5.00			
11	2.00	4.00	3.00			
12	2.00	5.00	5.00			
13	2.00	5.00	3.00			
14	2.00	6.00	6.00			
15	2.00	7.00	6.00			
16	2.00	8.00	7.00			
17	2.00	8.00	6.00			

Define three variables for the ANCOVA: 'Group', 'Pre', and 'Post'. Assign appropriate variable labels to each level of the 'Group' factor.

Set2] -	IBM SPSS Stat	istics Data Editor								
<u>D</u> ata	Transform	Analyze Graph:	s <u>U</u> tilities	Add-	ons <u>W</u> ind	dow	<u>H</u> elp			
		Re <u>p</u> orts D <u>e</u> scriptive : Compare M	Statistics	* * *	*5	2		4		 1ର୍କ
bup	Pre	General Lin	ear Model	•	Univa	ariate			var	
1.00 1.00 1.00	4.0 5.0 6.0	Generalized Mixed Model Correlate	Linear Mode s	els ⊧ ⊧	Multiv Bepe	variate. eated N	 leasures.			
1.00	7.0	Regression			Varia	nce Co	omponent	S		
1.00	7.0	Loglinear								
1.00	8.0	Classify								
1.00	9.0	Dimension I	Reduction							
1.00	9.0	Scale								
1.00	10.0	Nonparame	tric Tests				_			
1.00	11.0	Forecasting								
2.00	4.0	Survival					_			
2.00	5.0	M <u>u</u> ltiple Res	ponse							
2.00	6.0	🌠 Missing Valu	ie Analysis							
2.00	7.0	Multiple Imp	utation							
2.00	8.0	Quality Cont	rol	- F						
2.00	8.0	🖉 ROC Curve.								
2.00	9.0	0 8.00								
2.00	10.0	0 6.00								
2.00	11.0	0 4.00								
3.00	5.0	0 5.00								
3.00	5.0	0 4.00								
3.00	5.0	0 4.00								

Navigate to 'Analyze', 'General Linear Model', 'Univariate'

ta Univariate			×
	*	Dependent Variable:	Model ontrasts Plots ost Hoc Save Options
ОК <u>Р</u> а	ste	Covariate(s): Pre <u>WLS Weight:</u> <u>Reset</u> Cancel Help	

🔄 Univariate: Model	
Specify Model Factors & Covariates: Pre Build Term(s) Tyge: Interaction	Model: Group*Pre Group Pre
Sum of squares: Type III  Continue	Include intercept in model
Contained	

"Post' is the dependent variable, 'Group' is the fixed factor and 'Pre' is the covariate.

In the model menu, we want to specify a custom model. Make sure that custom is chosen and then highlight both 'Group' and 'Pre' and place that into the model box, then place 'Group' and 'Pre' by themselves in the model box. Select 'Continue'

ta Univariate: Options	<b>—</b>					
Estimated Marginal Means						
<u>F</u> actor(s) and Factor Interactions: (OVERALL) Group	Display Means for:					
	Compare main effects Confidence interval adjustment LSD(none)					
Display-						
Descriptive statistics	Homogeneity tests					
Estimates of effect size	Spread vs. level plot					
Observed power	Residual plot					
Parameter estimates	Lack of fit					
Contrast coefficient matrix	General estimable function					
Significance level: 05 Confidence intervals are 95.0 %						

In the options menu make sure you check off 'Descriptive statistics', 'Estimates of effect size', and 'Observed Power'. Select 'Continue' and then 'OK' to run the model.

# One Way ANCOVA

## Selected SPSS Output

#### Source Type III Sum of df Mean Square F Partial Eta Noncent. Observed Sig. Squares Squared Pow er<sup>b</sup> Parameter Corrected Model 106.272<sup>a</sup> 5 21.254 9.964 .000 .675 49.820 .071 1.832 Intercept 3.907 3.907 1.832 .189 1 Group \* Pre .348 4.704 2.205 2 2.352 1.103 .084 2.048 Group 4.369 2 2.184 1.024 .374 .079 Pre 39.302 39.302 18.424 .000 .434 18.424 1 Error 51.195 24 2.133 1054.000 Total 30 Corrected Total 157.467 29

#### Tests of Between-Subjects Effects

a. R Squared = .675 (Adjusted R Squared = .607)

The interaction is not significant

1.000

.255

.221

.207

.984

b. Computed using alpha = .05

Dependent Variable: Post

The assumption of homogeneity of regression slopes (are the slopes parallel?) is tested in the interaction between 'GP' \* 'PRE'. If the interaction is NOT significant, the assumption of homogeneity of regression slopes is found tenable (the slopes are parallel).

In this example, the interaction is not significant; therefore the assumption of homogeneity of regression slopes is tenable.

# **One Way ANCOVA**

# **SPSS Procedures (Part 2)**

**Test of significance.** Once the assumption of homogeneity of regression slopes is found tenable, the ANCOVA is repeated using the 'Full Factorial' Model.

Set2] -	IBM SPSS Stat	istics Data Editor									
<u>D</u> ata	Transform	Analyze Graph	ns <u>U</u> tilities	Add-	ons	Window	He	lp			
	<b>1</b>	Re <u>p</u> orts D <u>e</u> scriptive Compare N	Statistics leans	* * *		*5 🖬			4		▲
oup	Pre	General Lir	ear Model	E.	1977	Univeriate				va	r
1.00	4.0	Generalize	d Linear Mode	els.≯							
1.00	5.0	Mixed Mode	Is		HOLD FREE		e				
1.00	6.0	Correlate				Repeated	Mea	sures		_	
1.00	7.0	Regression	1			Variance C	Com	ponent	S		
1.00	7.0	Loglinear									
1.00	8.0	Classify		- <u>-</u>							
1.00	9.0	Dimension	Reduction								
1.00	9.0	Scale	Reduction								
1.00	10.0	Scale	atria Talata								
1.00	11.0	Gooparam	euro resis								
2.00	4.0	Porecasiing	,								
2.00	5.0	Survival									
2.00	5.0	Multiple Re	sponse								
2.00	6.0	🏭 Missing Va	lue Anal <u>y</u> sis								
2.00	7.0	Multiple Imp	outation	•							
2.00	8.0	Quality Con	trol	•							
2.00	8.0	ROC Curve									
2.00	9.0	0 8.00									
2.00	10.0	0 6.00									
2.00	11.0	0 4.00									
3.00	5.0	0 5.00									
3.00	5.0	0 4.00									
3.00	5.0	0 4.00									

Navigate to 'Analyze', 'General Linear Model', 'Univariate'



"Post' is the dependent variable, 'Group' is the fixed factor and 'Pre' is the covariate.

Factors & Covariates:	Build Term(s) Type:	odel:	
	•		
Sum of sguares: Type II	T Inclu	de intercept in model	

This time we are going to run the full factorial model. Make sure this option is selected.

Eactor(s) and Factor Interactions:	Display Means for:
Group	
	Compare main effects
	Confidence interval adjustment: LSD(none)
Display	
Descriptive statistics	Homogeneity tests
Estimates of effect size	Spread vs. level plot
Observed power	Residual plot
Parameter estimates	Lack of fit
Contrast coefficient matrix	General estimable function
Significance level: .05 Confid	dence intervals are 95.0 %

In the options box we want to display the means for the groups so make sure 'Group' is in the display means box. Also make sure 'Descriptive statistics', 'Estimates of effect size', and 'Observed Power' are selected. Select 'Continue' and then 'OK' to run the model.

# **One-Way ANCOVA**

# Selected SPSS Output

#### **Descriptive Statistics**

Group	Mean	Std. Deviation	N
1.00	3.7000	1.56702	10
2.00	5.4000	1.64655	10
3.00	7.3000	2.26323	10
Total	5.4667	2.33021	30



#### Tests of Between-Subjects Effects

Dependent Variabl	e: Post							
Source	Type III Sum	df	Mean Square	F	Sig.	Partial Eta	Noncent.	Observed
	of Squares					Squared	Parameter	Pow er <sup>b</sup>
Corrected	101.567 <sup>a</sup>	3	33.856	15.747	.000	.645	47.241	1.000
Model								
Intercept	4.989	1	4.989	2.320	.140	.082	2.320	.311
Pre	36.701	1	36.701	17.070	.000	.396	17.070	.978
Group	70.456	2	35.228	16.385	.000	.558	32.771	.999
Error	55.899	26	2.150			K		
Total	1054.000	30						
Corrected Total	157.467	29						
Corrected Total	157.467	29						

a. R Squared = .645 (Adjusted R Squared = .604)

b. Computed using alpha = .05

Indicates at least one pair of means is significantly different.

Indication of a significant difference between *adjusted* means.

Dependen	t Variable:	Post		
Group	Mean	Std. Error	95% Confide	ence Interval
			Low er Bound	Upper Bound
1.00	3.593 <sup>a</sup>	.464	2.639	4.548
2.00	5.453 <sup>a</sup>	.464	4.500	6.407
3.00	7.353 <sup>a</sup>	.464	6.400	8.307

Group



a. Covariates appearing in the model are evaluated at the following values: Pre = 7.4000.

As there are more than two groups and this significance cannot be interpreted by visual inspection, it is necessary to perform a Bryant-Paulson post hoc analysis for randomized designs

92



# One Way ANCOVA Bryant-Paulson Post Hoc Procedure in SPSS

We need to conduct a One Way ANOVA for the covariate. Go to 'Analyze', 'Compare Means', 'One-Way ANOVA'

<b>t</b> a		One-Way	y ANOVA		×	
Post	OK	Paste Re	Pre Pre ctor: Group set Cancel	Help	Contrasts Post <u>H</u> oc Options	

Continue Cancel Heip

Move 'Pre' to the Dependent List box. Move the variable 'Group' to the Factor box.

In the options box, make sure you select 'Descriptives'. We will use this information in the write up. Click on 'Continue' then 'OK' to run the analysis.

# One Way ANCOVA Selected SPSS output for Bryant Paulson Test

# Pre treatment subgroup means. This will be used in the write up

Pre	Descriptives													
	Ν	Mean	Std. Deviation	Std. Error	95% Confidence	Interval for Mean	Minimum	Maximum						
					Low er Bound	Upper Bound								
1.00	10	7.6000	2.22111	.70238	6.0111	9.1889	4.00	11.00						
2.00	10	7.3000	2.31181	.73106	5.6462	8.9538	4.00	11.00						
3.00	10	7.3000	2.00278	.63333	5.8673	8.7327	5.00	10.00						
Total	30	7.4000	2.11073	.38536	6.6118	8.1882	4.00	11.00						

ANOVA

Pre					
	Sum of Squares	df	Mean Square	F	Sig.
Betw een Groups	.600	2	.300	.063	.939
Within Groups	128.600	27	4.763		
Total	129.200	29			

This is the between and within sums of squares that will be used in the Bryant-Paulson post hoc test

# **Bryant-Paulson equation:**

$$BP = \frac{\overline{Y_1^*} - \overline{Y_2^*}}{\sqrt{\frac{MSw^* \left[1 + \frac{MSbx}{MSwx}\right]}{n}}} \quad \text{Remember: } \frac{MSbx}{MSwx} = F$$

The adjusted means are Group 1=3.593, Group 2=5.453, and Group 3=7.353

Group 2 vs Group 1

$$BP = \frac{5.453 - 3.593}{\sqrt{\frac{2.150[1 + .063]}{10}}} = \frac{1.86}{.4781} = 3.89$$

Group 3 vs Group 2

$$BP = \frac{7.353 - 5.453}{\sqrt{\frac{2.150[1+.063]}{10}}} = \frac{1.90}{.4781} = 3.97$$

Group 3 vs Group 1

$$BP = \frac{7.353 - 3.593}{\sqrt{\frac{2.150[1 + .063]}{10}}} = \frac{3.76}{.47817} = 7.86$$

 $BP_{critical} = BP(.05,3,26)$  which is approximately equal to BP(.05,3,24) = 3.61. See Stevens, Appendix B5, for table values. Since all calculated results are greater than the critical value it implies the difference is significant.

# One Way ANCOVA Example Results and Discussion

An analysis of covariance was conducted to determine whether the post treatment tardies for the three groups differed after adjustments were made for pre treatment differences. The assumption of equal regression slopes was tested and found tenable, F(2, 24) = 1.10, p = .348. The ANCOVA indicated at least one pair of means was significantly different, F(2,26) = 16.385, p < .001

A post hoc analysis using the Bryant-Paulson technique (randomized design, a = .05) indicated the mean tardies for the positive reinforcement group was significantly lower than the mean tardies for the negative reinforcement group. Furthermore, the mean tardies for the negative reinforcement group was significantly lower than the mean tardies for the no reinforcement group.

Table 1 presents the means, standard deviations, and adjusted means for all of the groups. Table 2 presents the analysis of covariance summary.

Table 1.

# Mean Tardies, Standard Deviations, and Adjusted Means by Type of Reinforcement

	Pre Tre	atment	Post Tre	Post Treatment Obtained				
			Obtained					
Reinforcement Type	М	SD	М	SD	М			
Positive	7.60	2.22	3.70	1.57	3.59			
Negative	7.30	2.31	5.40	1.65	5.45			
None	7.30	2.00	7.30	2.26	7.35			

Table 2.

## Analysis of Variance Summary

Source	Sum of Squares	DF	Mean Square	F value	p value
Pre Treatment	36.70	1	36.70	17.07	<.001
Group	70.46	2	35.23	16.39	<.001
Error	55.90	26	2.15		

# Discussion

The method of positively reinforcing the employees when they are on time had the greatest impact on reducing tardiness. Negative reinforcement for lateness was not as effective as positive reinforcement for being on time, but was better than no reinforcement at all.



# **REPEATED MEASURES ANALYSIS**

# Example Problem One Within Factor Design

A high school math teacher studies the impact of paper color on mathematics test scores. The hypothesis was test scores would be higher on tests taken on pastel green paper than tests taken on bright yellow paper or traditional white paper because the cool color would have a calming effect and reduce test-taking anxiety. Weekly math tests for an Algebra I class were printed in equal quantities on the three colors of paper. The order of treatment was counterbalanced such that one third of the students were randomly assigned to a different color each week over a three week period. Below are the test scores by student by paper color.

### Table1

### Math Test Scores

	Paper Color								
Student	Yellow	Green	White						
1	80	76	77						
2	81	89	70						
3	39	64	55						
4	95	93	91						
5	71	90	87						
6	86	76	92						
7	98	94	83						
8	95	92	92						
9	73	77	53						
10	78	88	83						
11	54	64	57						
12	73	92	96						
13	82	75	69						
14	49	67	55						
15	83	91	79						
16	91	90	88						
17	94	97	91						
18	85	89	90						
19	62	69	45						

ſ	ta *Untitled1	[DataSet0] - IBM	SPSS Statistics	Data Editor		
	<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Transform A	nalyze <u>G</u> rapł	ns <u>U</u> tilities	Add- <u>o</u>
	😑 🔚			<b>1</b>	▙	<u>a</u> na
	9:					
		Student	yellow	green	white	v
	1	1.00	80.00	76.00	77.00	
	2	2.00	81.00	89.00	70.00	
	3	3.00	39.00	64.00	55.00	
	4	4.00	95.00	93.00	91.00	
	5	5.00	71.00	90.00	87.00	
	6	6.00	86.00	76.00	92.00	
	7	7.00	98.00	94.00	83.00	
	8	8.00	95.00	92.00	92.00	
	9	9.00	73.00	77.00	53.00	
	10	10.00	78.00	88.00	83.00	
	11	11.00	54.00	64.00	57.00	
	12	12.00	73.00	92.00	96.00	
	13	13.00	82.00	75.00	69.00	
	14	14.00	49.00	67.00	55.00	
	15	15.00	83.00	91.00	79.00	
	16	16.00	91.00	90.00	88.00	
	17	17.00	94.00	97.00	91.00	
	18	18.00	85.00	89.00	90.00	
	19	19.00	62.00	69.00	45.00	
	20					
	21					

Define and enter initial data. Define four variables: 'Student', 'yellow', 'green', and 'white'. Note that this method of data entry is different from that used in ANOVA and ANCOVA type problems. You previously used this method of data entry to perform the Dependent T-Test, which is a special case of this single group repeated measures design.

ta *	Untitled	1 (DataSet(	1 - IBN	SPSS Statistic	s Data Edit	tor										
File	Edit	View D	ata	Transform	Analyze	Graphs	<u>U</u> tilities	Add-	ons	Wind	ow	Help	р			
6			Ū,		Rep D <u>e</u> s	orts criptive S	tatistics	* *		*,	K			4		 1∜
21 :					Com	pare Me	ans									
		Stude	ent	Color1	Gen	eral Line	ar Model	•	SIM	Univa	riate				Vá	ar
	1		1.00	80.0	Gen	erali <u>z</u> ed l	Linear Model	s⊧	GUM	Multiv	ariate					
	2		2.00	81.0	Mixe	d Models	3		GLM	Rene	ated	1029	ures			
	3		3.00	39.0	<u>C</u> orr	elate		*	212	I IZehe	aleur	neaa	oures		_	
	4		4.00	95.0	Reg	ression			Ļ	Variar	nce C	omp	onents	S		
	5		5.00	71.0	Logi	inear										
	6		6.00	86.0	Clas	sify		•								
	7		7.00	98.0	Dim	ension R	eduction									
	8		8.00	95.0	Scal	e										
	9		9.00	73.0	Non	parametr	ric Tests									
	10		10.00	78.0	Eore	casting										
	11		11.00	54.0	Suni	ival										
	12		12.00	73.0	Multi	nia Door	00000	, ,								
	13		13.00	78.0		pie Resp	Julise	,								
	14		14.00	49.0	MISS	ing valu	e Analysis									
	15		15.00	83.0	Mul <u>t</u> i	ple Impu	itation	•								
	16		16.00	91.0	Qual	ity Contro	ol	•								
	17		17.00	94.0	ROC	Curve										
	18		18.00	85.00		89.00	90.00									
	19		19.00	62.00		69.00	45.00									
	20															
	21															

Navigate to 'Analyze', 'General Linear Model' and 'Repeated Measures'

Repeated Measures Define Factor(s)
Within-Subject Factor Name:
Number of Levels:
Add factor1(3)
Change
Re <u>m</u> ove
Measure <u>N</u> ame:
Add
Remove
Keniove
Define Reset Cancel Help

The 'Repeated Measures Define Factor(s)' dialog box is displayed. An entry of 'factor1' is displayed in the 'Within-Subject Factor Name:' text box. Enter the value '3' (corresponding to the number of levels in the study) in the 'Number of Levels' text box. Click on the 'Add' button.

As a result of clicking on the 'Add' button, the term 'factor1 (3)' will be inserted into the central text box.

Click on the 'Define' button located in the upper right hand corner of the dialog box.

ta Repeated Measures			<b>×</b>
Student	¢ +	Within-Subjects Variables (factor1): yellow(1) green(2) white(3)	Model Contrasts Plots Post <u>H</u> oc Save Options
	•	Between-Subjects Factor(s)	
ОК	Paste F	Reset Cancel Help	

The 'Repeated Measures' dialog box is displayed. Highlight the variable 'yellow' in the left side text box and then click on the right pointing arrow to move the variable into the 'Within-Subjects Variable (factor1):' text box. Move the variables 'green' and 'white' to the 'Within-Subjects Variable (factor1):' text box in the same manner. Once all three variables are in the text box, click on the 'Options' button at the bottom right corner of the dialog box.

ta Repeated Measures: Options	<b>—</b>
Estimated Marginal Means <u>Factor(s) and Factor Interactions:</u> (OVERALL) factor1	Display <u>M</u> eans for: factor1
	Compare main effects Confidence interval adjustment Bonferroni
-Display	
Descriptive statistics	Transformation matrix
Chserved power	Spread vs. level plot
Parameter estimates	Residual plot
SSCP matrices	
Residual SS <u>C</u> P matrix	General estimable function
Significance level: 05 Confid	lence intervals are 95.0 % Cancel Help

The 'Repeated Measures: Options' dialog box is displayed. Use the mouse pointer to check 'Descriptive statistics', 'Estimates of effect size' and 'Observed power'

In the 'Display' options section, we want to add 'factor1' to the box and select 'Compare main effects' and select 'Bonferroni' as the confidence interval adjustment.

Click on 'Continue' to return to the 'Repeated Measures' dialog box.

Click on 'OK' to perform the analysis. The results will be displayed in the SPSS Output Navigator window

# **ONE-WITHIN REPEATED MEASURES** Selected Results from SPSS Execution

	Descriptive Statistics									
	Mean	Std. Deviation	N							
yellow	77.3158	16.42171	19							
green	82.7895	11.07840	19							
w hite	76.4737	16.23565	19							

You can ignore this table of multivariate tests.



				Multivar	late lests				
Effect		Value	F	Hypothesis	Error df	Sig.	Partial Eta	Noncent.	Observed
				df			Squared	Parameter	Pow er <sup>c</sup>
	Pillai's Trace	.380	5.202 <sup>b</sup>	2.000	17.000	.017	.380	10.404	.755
	Wilks' Lambda	.620	5.202 <sup>b</sup>	2.000	17.000	.017	.380	10.404	.755
factor1	Hotelling's Trace	.612	5.202 <sup>b</sup>	2.000	17.000	.017	.380	10.404	.755
	Roy's Largest Root	.612	5.202 <sup>b</sup>	2.000	17.000	.017	.380	10.404	.755

.....

#### Mauchly's Test of Sphericity<sup>a</sup>

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-	df	Sig.	Epsilon <sup>b</sup>		
		Square			Greenhouse-	Huynh-Feldt	Low er-bound
					Geisser		
factor1	.937	1.110	2	.574	.941	1.000	.500



This shows that the sphericity assumption was met.

Using the sphericity assumed line, we see that factor 1 is significant.



Bonferroni Post Hoc Results

Measure: MEASURE_1										
Source		Type III Sum of	df	Mean Square	F	Sig.	Partial Eta			
		Squares					Squared			
	Sphericity Assumed	446.877	2	223.439	4.141	.024	.187			
factor1	Greenhouse-Geisser	446.877	1.881	237.566	4.141	.027	.187			
TACION	Huynh-Feldt	446.877	2.000	223.439	4.141	.024	.187			
	Low er-bound	446.877	1.000	446.877	4.141	.057	.187			
	Sphericity Assumed	1942.456	36	53.957						
Error(factor1)	Greenhouse-Geisser	1942.456	33.859	57.369						
	Huynh-Feldt	1942.456	36.000	53.957						
	Low er-bound	1942.456	18.000	107.914						

Tests of Within-Subjects Effects

## Estimates

Measure: MEASURE_1										
factor1	Mean	Std. Error	95% Confidence Interval							
			Low er Bound	Upper Bound						
1	77.316	3.767	69.401	85.231						
2	82.789	2.542	77.450	88.129						
3	76.474	3.725	68.648	84.299						

#### Pairwise Comparisons

Measure: MEASURE\_1

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Low er Bound	Upper Bound
	2	-5.474	2.239	.075	-11.383	.436
1	3	.842	2.666	1.000	-6.193	7.878
0	1	5.474	2.239	.075	436	11.383
2	3	6.316 <sup>*</sup>	2.218	.032	.463	12.169
	1	842	2.666	1.000	-7.878	6.193
3	2	-6.316 <sup>*</sup>	2.218	.032	-12.169	463

Based on estimated marginal means

# **Example Results and Discussion**

A univariate repeated measure analysis of variance indicated a significant difference between the means of the three paper color treatments, F(2,36) = 4.14, p = .024,  $\eta_p^2 = .187$ . The assumption of sphericity was found to be tenable according to Mauchly's test ( $\chi^2(2) = 1.110$ , p = .574). A Bonferroni post hoc analysis revealed the mean for the green paper treatment was significantly higher than the white treatment. There was no significant difference between the means of the green and yellow treatments and the yellow and white treatments. Table 1 presents the means and standard deviations for the three paper colors. Table 2 presents the repeated measure analysis of variance summary.

#### Table 1

Means and Standard Deviations of Math Test Scores as a Function of Paper Color

Paper Color	Ν	М	SD
Green	19	82. 79	11 .08
Yellow	19	77. 32	16. 42
White	19	76. 47	16. 24

Table 2

Univariate Repeated Measure Analysis of Variance Summary

Source	Sum of Squares	DF	Mean Square	F value	p value
Color	446.88	2	223.44	4.14	.024
Error	1942.46	36	53.96		

# **ONE WITHIN REPEATED MEASURES**

# **Example Results and Discussion**

The analysis indicates some support for the idea that paper color could make a difference in math test scores. The pastel green definitely resulted in higher test scores when compared to the traditional white paper. However, the pastel green did not show a large enough difference over the bright yellow to be significant, calling into question the theory on the effect of cool colors. Further study is suggested.

### **Example Problem** One Within Factor and One Between Factor Design

A researcher wanted to investigate the effect of anxiety on math performance for fourth graders in a variety of testing time constraints. She came up with a 20 question multiplication test and gave it to twelve participants. For the first trial, she allowed participants one hour. For the second trial, participants had 45 min to take the test, the time for the third trial was 30 min and the time for the final trial was 15 min. Prior to giving the math test, she assessed students test anxiety level and students with low test anxiety were in the group one and those with high test anxiety were in the second group. The tests, however, were the same between groups. The data can be seen below <sup>1</sup>.

Subject	Anxiety	Trial 1	Trial 2	Trial 3	Trial 4
1	1	18	14	12	6
2	1	19	12	8	4
3	1	14	10	6	2
4	1	16	12	10	4
5	1	12	8	6	2
6	1	18	10	5	1
7	2	16	10	8	4
8	2	18	8	4	1
9	2	16	12	6	2
10	2	19	16	10	8
11	2	16	14	10	9
12	2	16	12	8	8

1. Data was adapted from UC Denver College of Nursing. Their example can be found at http://www.ucdenver.edu/academics/colleges/nursing/Documents/PDF/RepeatedMeasuresANOVA.pdf

ta sp	litplot.si	av (Data	Set13] - I	BM SPSS Statis	tics Data Editor				
<u>F</u> ile	Edit	View	<u>D</u> ata	Transform	Analyze Grag	ohs <u>U</u> tilities	Add- <u>o</u> ns <u>W</u> ir	ndow <u>H</u> elp	
2			)		<b>~</b>	<b>*</b> =	H 🐮		4
		su	bject	anxiety	trial1	trial2	trial3	trial4	var
	1		1.00	1.0	0 18.0	0 14.00	12.00	6.00	
	2		2.00	1.0	0 19.0	0 12.00	8.00	4.00	
	3		3.00	1.0	0 14.0	0 10.00	6.00	2.00	
	4		4.00	1.0	0 16.0	0 12.00	10.00	4.00	
	5		5.00	1.0	0 12.0	0.800	6.00	2.00	
	6		6.00	1.0	0 18.0	0 10.00	5.00	1.00	
	7		7.00	2.0	0 16.0	0 10.00	8.00	4.00	
	8		8.00	2.0	0 18.0	0.800	4.00	1.00	
	9		9.00	2.0	0 16.0	0 12.00	6.00	2.00	
1	10		10.00	2.0	0 19.0	0 16.00	10.00	8.00	
1	11		11.00	2.0	0 16.0	0 14.00	10.00	9.00	
1	12		12.00	2.0	0 16.0	0 12.00	8.00	8.00	
1	13								
1	14								
1	15								
1	16								
1	17								
1	18								
1	19								
- 2	20								
2	21								
2	22								
1	23								

Define and enter initial data. Define four variables: 'Subject', 'Anxiety', 'Trial 1', 'Trial 2', 'Trial 3', and 'Trial 4'.

ta spl	litplot.si	av (DataS	et13] - I	IBM SPSS Stati	istics Data E	ditor						
<u>F</u> ile	Edit	View	<u>D</u> ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	Utilities	Add-g	ons <u>W</u> ir	dow <u>H</u> e	elp	
🖹 🖶 🖨 🛄 🖛				Rep D <u>e</u> s	orts criptive St	atistics	+ +		4			
23 :					Co <u>m</u>	ipare Mea	ans	•				
		sub	ject	anxiety	<u>G</u> en	eral Linea	ar Model	•	🔛 Univ	/ariate		
	1		1.00	1.(	Gen	erali <u>z</u> ed L	inear Mode	ls ►	Mult	ivariate		
	2		2.00	1.0	Mi <u>x</u> e	d Models		•	Ren	eated Mea	asures	
	3		3.00	1.(	<u>C</u> orr	elate		•	Vori			
	4		4.00	1.(	<u>R</u> egi	ression		•	van	ance Com	ponents	
	5		5.00	1.0	L <u>o</u> gi	inear		•	6.00		2.00	
	6		6.00	1.(	Clas	sify		•	5.00		1.00	
	7		7.00	2.0	Dim	ension Re	eduction	•	8.00		4.00	
	8		8.00	2.0	Scal	e		•	4.00		1.00	
	9		9.00	2.(	Non	parametri	c Tests	•	6.00		2.00	
	10		10.00	2.0	Fore	casting			10.00		8.00	
1	11		11.00	2.0	Surv	ival			10.00		9.00	
1	12		12.00	2.0	Multi	nia Raco	0050		8.00		8.00	
1	13					ing Value	Analysis					
1	14				MISS	ing value	Anaiysis					
1	15				Mul <u>t</u> i	ple Imput	ation	•				
1	16				Qual	ity Contro	il .	•				
1	17				ROC	Curve						
1	18											


ta Repeated Measures Define Factor(s)								
Within-Subject Factor Name:								
factor1								
Number of <u>L</u> evels:	4							
factor1	(4)							
<u>A</u> uu								
Change								
Remove								
Measure <u>N</u> ame:								
Add								
Change								
Cnange								
Remove								
Define Rese	et Cancel Help							

The 'Repeated Measures Define Factor(s)' dialog box is displayed. An entry of 'factor1' is displayed in the 'Within-Subject Factor Name:' text box. Enter the value '4' (corresponding to the number of levels in the study) in the 'Number of Levels' text box. Click on the 'Add' button.

As a result of clicking on the 'Add' button, the term 'factor1 (4)' will be inserted into the central text box.

Click on the 'Define' button located in the upper right hand corner of the dialog box.



The 'Repeated Measures' dialog box is displayed. Highlight the variable 'trial 1' in the left side text box and then click on the right pointing arrow to move the variable into the 'Within-Subjects Variable (factor1):' text box. Move the variables 'Trial 2', 'Trial 3' and 'Trail 4' to the 'Within-Subjects Variable (factor1):' text box in the same manner. Now move 'anxiety' to the between subject factor. Once all variables are in the text box, click on the 'Options' button at the bottom right corner of the dialog box.

â	anxiety trial1 trial2 trial3 trial4 var										
	1.0	0 18.0	0 14.00	12.00	6.00						
	<b>t</b> a -		Repeated Me	easures: Opti	ons	×					
ţ	Estimated Marginal Means Factor(s) and Factor Interactions: (OVERALL) anxiety factor1 anxiety*factor1  Compare main effects Confidence interval adjustment LSD(none)										
	Disp Disp Disp Disp Disp C Disp C Disp Disp Disp Disp Disp Disp Disp Disp	lay escriptive statis stimates of effe <u>b</u> served power arameter estim SCP matrices esidual SS <u>C</u> P icance level: [0	stics ct size nates matrix 5 Confider Continue C	Transformal Homogenei Spread vs. Ii Residual plu Lack of fit General esti nce intervals are ancel Hel	tion matrix ty tests evel plot ot mable function 9 95.0 %						

Repeated Measures: Profile <u>Factors:</u> anxiety factor1	Horizontal Axis: Factor1 Separate Lines: anxiety Separate Plots:
Plots: <u>A</u> dd	Change Remove

The 'Repeated Measures: Options' dialog box is displayed. Use the mouse pointer to check 'Descriptive statistics', 'Estimates of effect size', 'Observed power', and 'Homogeneity test'.

In the 'Display' options section, we want to add 'factor1' to the box and select 'Compare main effects' and select 'Bonferroni' as the confidence interval adjustment.

Click on 'Continue' to return to the 'Repeated Measures' dialog box.

The final thing we want to do is to add some plots to the output. On the Repeated Measures Dialog box, select 'Plots'

Next add 'Factor1' to the 'Horizontal Axis' box and 'anxiety' to the 'Separate Lines' box.

Select 'Add' to add the plot then select 'Continue'

Click on 'OK' to perform the analysis. The results will be displayed in the SPSS Output Navigator window

# Repeated Measures Selected SPSS Output

Descriptive Statistics									
	anxiety	Mean	Std. Deviation	Ν					
	1.00	16.1667	2.71416	6					
trial1	2.00	16.8333	1.32916	6					
	Total	16.5000	2.06706	12					
	1.00	11.0000	2.09762	6					
trial2	2.00	12.0000	2.82843	6					
	Total	11.5000	2.43086	12					
	1.00	7.8333	2.71416	6					
trial3	2.00	7.6667	2.33809	6					
	Total	7.7500	2.41680	12					
	1.00	3.1667	1.83485	6					
trial4	2.00	5.3333	3.44480	6					
	Total	4.2500	2.86436	12					

You can ignore this table of multivariate tests.



Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Pow er <sup>c</sup>
factor1	Pillai's Trace	.961	64.854 <sup>b</sup>	3.000	8.000	.000	.961	194.561	1.000
	Wilks' Lambda	.039	64.854 <sup>b</sup>	3.000	8.000	.000	.961	194.561	1.000
	Hotelling's Trace	24.320	64.854 <sup>b</sup>	3.000	8.000	.000	.961	194.561	1.000
	Roy's Largest Root	24.320	64.854 <sup>b</sup>	3.000	8.000	.000	.961	194.561	1.000
	Pillai's Trace	.479	2.451 <sup>b</sup>	3.000	8.000	.138	.479	7.354	.408
	Wilks' Lambda	.521	2.451 <sup>b</sup>	3.000	8.000	.138	.479	7.354	.408
factor1 * anxiety	Hotelling's Trace	.919	2.451 <sup>b</sup>	3.000	8.000	.138	.479	7.354	.408
	Roy's Largest Root	.919	2.451 <sup>b</sup>	3.000	8.000	.138	.479	7.354	.408

### Mauchly's Test of Sphericity<sup>a</sup>

Measure: MEASURE_1							
Within Subjects Effect	Mauchly's W	Approx. Chi-	df	Sig.	Epsilon <sup>b</sup>		
		Square			Greenhouse-	Huynh-Feldt	Low er-bound
					Geisser		
factor1	.283	11.011	5	.053	.544	.701	.333

Tests of Within-Subjects Effects

This shows that the sphericity assumption was met.

#### Measure: MEASURE\_1

Source		Type III Sum	df	Mean	F	Sig.	Partial Eta	Noncent.	Observed
		of Squares		Square			Squared	Parameter	Pow er <sup>a</sup>
	Sphericity Assumed	991.500	3	330.500	128.62	.000	.928	385.881	1.000
factor1	Greenhouse- Geisser	991.500	1.632	607.468	128.627	.000	.928	209.943	1.000
	Huynh-Feldt	991.500	2.102	471.773	128.627	.000	.928	270.329	1.000
	Low er-bound	991.500	1.000	991.500	128.627	000	.928	128.627	1.000
	Sphericity Assumed	8.417	3	2.806	1.091	.368	.098	3.276	.265
factor1 * anxiety	Greenhouse- Geisser	8.417	1.632	5.157	1.092	.346	.098	1.782	.194
	Huynh-Feldt	8.417	2.102	4.005	1.092	.357	.098	2.295	.220
	Low er-bound	8.417	1.000	8.417	1.092	.321	.098	1.092	.157
	Sphericity Assumed	77.083	30	2.569					
Error(factor1)	Greenhouse- Geisser	77.083	16.322	4.723		The f	etor variabl	a is significa	nt the
	Huynh-Feldt	77.083	21.016	3.668		interac	ction is not.		int, uit
	Low er-bound	77.083	10.000	7.708					

a. Computed using alpha = .05

	F	df1	df2	Sig.
trial1	3.312	1	10	.099
trial2	.156	1	10	.701
trial3	.266	1	10	.617
trial4	7.788	1	10	.019

### Levene's Test of Equality of Error Variances<sup>a</sup>

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + anxiety

Within Subjects Design: factor1

### Test of Equal Variance



ANOVA results for the between subject factor

Transformed Variable: Average

Measure: MEASURE\_1

Indiorenti	Tanoromod Valiable. Avolage										
Source	Type III Sum	df	Mean Square	F	Sig.	Partial Eta	Noncent.	Observed			
	of Squares					Squared	Parameter	Pow er <sup>a</sup>			
Intercept	4800.000	1	4800.000	280.839	.000	.966	280.839	1.000			
anxiety	10.083	1	10.083	.590	.460	.056	.590	.107			
Error	170.917	10	17.092								

Tests of Between-Subjects Effects

a. Computed using alpha = .05

		i un	milee oompe			
Measure: N	EASURE_1					
(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confiden Differ	ce Interval for ence <sup>b</sup>
					Low er Bound	Upper Bound
	2	5.000*	.693	.000	2.728	7.272
1	3	8.750 <sup>*</sup>	.827	.000	6.038	11.462
	4	12.250 <sup>*</sup>	.920	.000	9.236	15.264
	1	-5.000*	.693	.000	-7.272	-2.728
2	3	3.750 <sup>*</sup>	.410	.000	2.407	5.093
	4	7.250*	.484	.000	5.662	8.838
	1	-8.750 <sup>*</sup>	.827	.000	-11.462	-6.038
3	2	-3.750 <sup>*</sup>	.410	.000	-5.093	-2.407
	4	$3.500^{*}$	.394	.000	2.208	4.792
	1	-12.250*	.920	.000	-15.264	-9.236
4	2	-7.250*	.484	.000	-8.838	-5.662
	3	-3.500*	.394	.000	-4.792	-2.208

## Pairwise Comparisons

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

# Bonferroni Post Hoc results



### **Example Results**

A univariate repeated measure analysis of variance indicated a significant difference between the means of the four test trials, F(3,30) = 128.63,  $p < .001\eta_p^2$ , = .928. The interaction between anxiety and test time was not significant, F(3,30) = 2.81, p = .368,  $\eta_p^2 = .098$ . The assumption of sphericity was found to be tenable according to the Mauchly's Teest of Sphericity ( $\chi^2(5) = 11.011$ , p = .053). A Bonferroni post hoc analysis using revealed the mean difference was significant for all test times. Table 1 presents the means and standard deviations for the test times. Table 2 presents the repeated measure analysis of variance and the result of the between factor ANOVA. Figure 1 shows a graphical depiction of the mean test scores by anxiety level and test time.

### Table 1

Means and Standard Deviations of Math Test Scores by Trial and anxiety level

	Anxiety Level One			Anxie	Anxiety Level Two					-
	Ν	М	SD	Ν	М	SD	Ν	М	SD	
Trial 1	6	16.17	2.71	6	16.17	2.71	6	16.17	2.71	
Trial 2	6	11.00	2.10	6	11.00	2.10	6	11.00	2.10	
Trial 3	6	7.83	2.71	6	7.83	2.71	6	7.83	2.71	
Trial 4	6	3.17	1.83	6	3.17	1.83	6	3.17	1.83	

Table 2

Univariate Repeated Measure Analysis of Variance Summary

Source	Sum of Squares	DF	Mean Square	F value	p value
Anxiety	10.08	1	10.08	.590	.460
Error	170.92	10	17.09		
Trial	991.50	3	330.5	128.63	<.001
Trial*Anxiety	8.42	3	2.81	1.09	.368
Error	77.08	30	2.57		





# One Within and One Between REPEATED MEASURES

## **Example Discussion**

The analysis indicated that there was a significant difference in test scores between test time conditions, however, there is not a significant difference in test scores between anxiety levels. In addition, the effect of test time did not depend on the anxiety level. In other words, this study showed that while the amount of time a student has to take a test affected the performance; it failed to reveal any effect of anxiety on the performance. The small sample size could have masked the effect of anxiety on test scores. Further study is suggested to see if there is indeed an effect of anxiety on test scores under different time constraints.