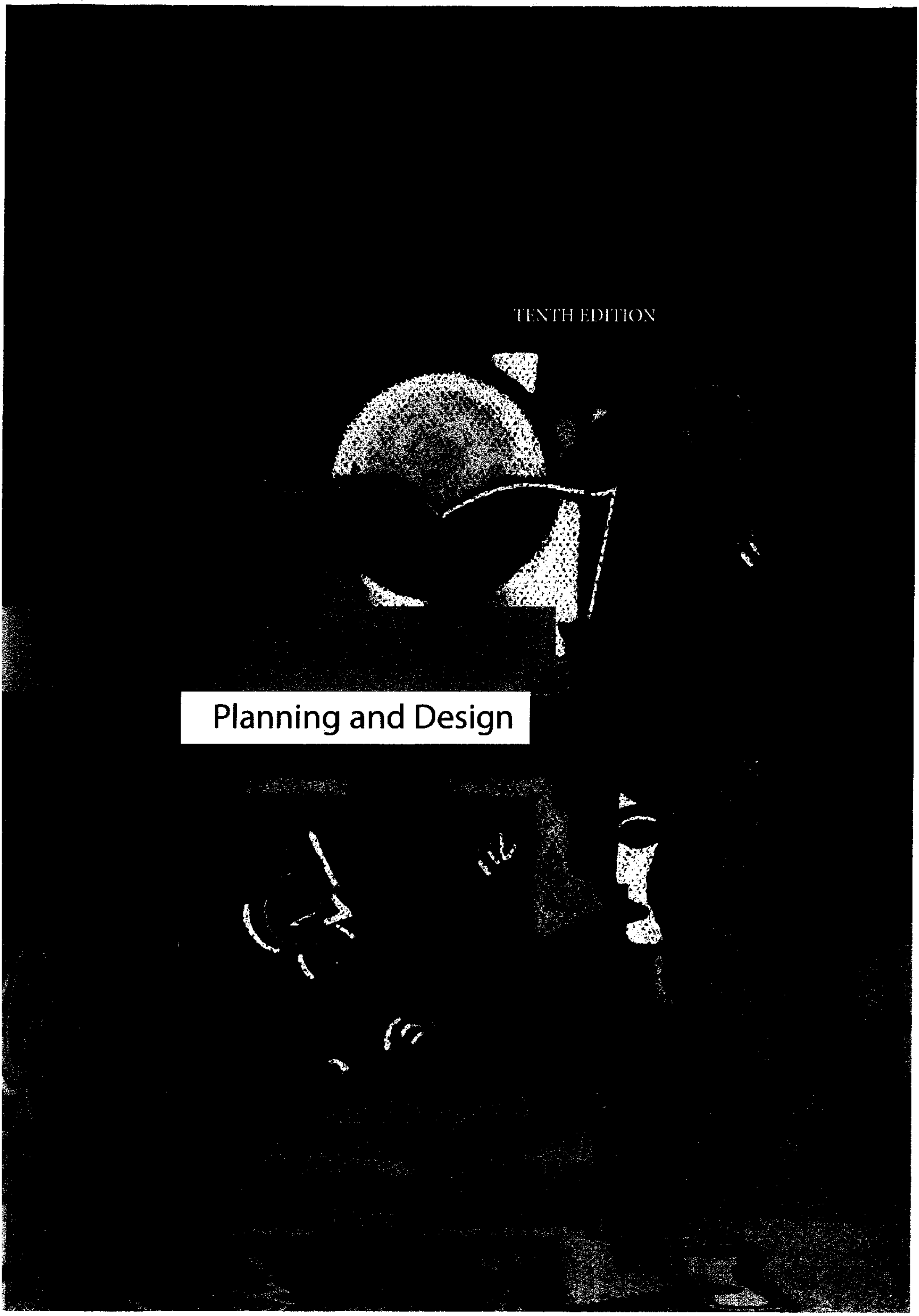


TENTH EDITION

Planning and Design

Pearson's MyLab™

PROVEN RESULTS

For over 10 years, instructors and students have reported better grades through increased engagement and real-time insights into progress.

ENGAGING EXPERIENCES

MyLab is designed to reach students in a personal way. Engaging learning and practice opportunities lead to assessments that create a personalized study plan.

A TRUSTED PARTNERSHIP

With millions of students registered annually, MyLab is the most effective and reliable learning solution available today.

To order this text with **MyEducationLab™**
use ISBN 0-13-289361-4

ISBN-13: 978-0-13-269324-0
ISBN-10: 0-13-269324-0
EAN
9 780132 693240 90000

www.pearsonhighered.com

PEARSON

ALWAYS LEARNING

Appendix B
Using SPSS

A complete explanation of how to use SPSS—short for Statistical Package for the Social Sciences—is well beyond the scope of a short appendix. However, a brief explanation of some of the basics can get you started. The version of SPSS we describe is PASW Statistics Student Version 18.0 for Macintosh.[1]

Creating a Data Set

USING TECHNOLOGY

Once you have loaded SPSS onto the hard drive of your computer (see the directions that come with the program for details), open the program. On your screen you will see a two-dimensional table that looks very much like a spreadsheet. Each row in the table designates a specific individual (human participant, animal subject, artifact, etc.) in your data set. Each column designates a specific variable in the data set. Once filled in, this table will provide the basis for your data analyses.

As an example, we use data from a pilot study that Dinah Jackson conducted in preparation for her dissertation study (1996; excerpts from her dissertation appear in Chapters 1 and 12). The data include the following information for 15 students in a college psychology class; (a) their gender; (b) their scores on three exams given during the semester; (c) the total of the three exam scores; (d) the quantity of class notes (i.e., number of pages) they took during the semester; and (e) the quality of their class notes. The last of these variables—quality of notes—is based on content analyses of students' notes; the numbers are proportions of notes that reflect an integration of two or more ideas rather than a single, isolated fact. In Jackson's study, better-integrated notes (reflected in higher numbers, such as .406 or .496) were theorized to facilitate better learning—and thus to be of better quality—than relatively non-integrated notes (reflected in lower numbers, such as .166 or .040). Jackson's pilot data are shown in Figure B.1

Notice that the seven columns in the table in Figure B.1 have short labels that tell us what each variable is. To insert such labels, we go down to the bottom of the screen, where there are two "buttons" called "Data View" and "Variable View." If we click on "Variable View," we get another table, which looks like Figure B.2. In this table, we have entered information about each of the variables in the data set. Here the variables are the rows (rather than the columns, as they are in the "Data" table), and the things we want to say about the variables are the columns. To keep our discussion simple, we describe only some of these columns:

- *Name:* Indicates the label that will appear for the variable in the "Data View" table. This label can include alphabet letters, numbers, and a few other meaningful symbols (e.g., "$").

[1] At the instructor's request, this book can be packaged with the Student Version of SPSS at a discount; the CD for the software provides versions for both Windows and Macintosh users. Please contact your local Pearson representative if you are an instructor who is interested in setting up such a package for your students.

FIGURE B.1
The "Data" table

	Gender	Exam1	Exam2	Exam3	TotalExam	NoteQuan	NoteQual	var
1	1	35.00	36.50	33.50	105.00	31	.315	
2	2	39.50	41.00	41.50	122.00	28	.384	
3	1	45.00	45.50	39.00	129.50	37	.381	
4	1	34.50	33.00	29.00	96.50	31	.251	
5	2	31.00	43.00	37.00	111.00	42	.305	
6	2	38.00	30.25	32.00	100.25	27	.190	
7	2	40.50	43.00	42.00	125.50	43	.350	
8	1	44.00	45.00	43.00	132.00	26	.166	
9	2	43.00	38.00	40.00	121.00	52	.406	
10	1	38.00	32.00	30.00	100.00	33	.208	
11	1	43.00	47.00	44.00	134.00	43	.496	
12	2	43.50	34.75	41.00	119.25	24	.201	
13	2	45.00	43.00	42.00	130.00	50	.321	
14	1	39.50	44.00	43.00	126.50	23	.179	
15	2	40.00	36.00	33.00	109.00	14	.040	
16								

Data View | Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Gender	Numeric	8	0	Gender	{1, Male}...	None	8	Right	Nominal	Input
2	Exam1	Numeric	8	2	Exam 1 Score	None	None	8	Right	Scale	Input
3	Exam2	Numeric	8	2	Exam 2 Score	None	None	8	Right	Scale	Input
4	Exam3	Numeric	8	2	Exam 3 Score	None	None	8	Right	Scale	Input
5	TotalExam	Numeric	8	2	Exam Score Total	None	None	8	Right	Scale	Input
6	NoteQuan	Numeric	8	0	Quantity of No...	None	None	8	Right	Scale	Input
7	NoteQual	Numeric	8	3	Quality of Notes	None	None	8	Right	Scale	Input
8											
9											

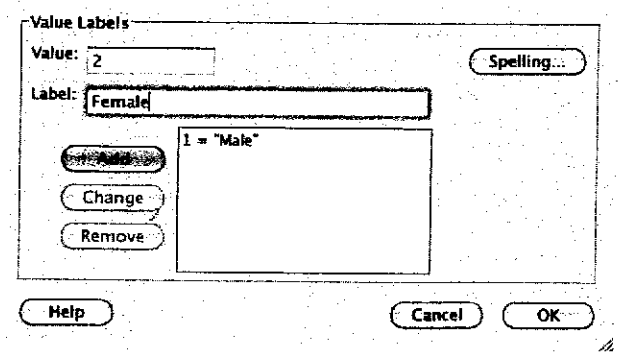
Data View | Variable View

FIGURE B.2
The "Variables" table

- *Type:* Indicates the type of data the variable represents, perhaps a number (numeric data), a letter string, a dollar amount, a date, or something else altogether.
- *Decimals:* Indicates an upper limit on the number of digits that will appear to the right of a decimal point.
- *Label:* Indicates the labels the variables will have when we create a table or graph—perhaps one to be included in a dissertation or research report.
- *Values:* Indicates labels that might be attached to particular values of a variable. For example, one of our variables is gender, a nominal scale. If we click on this "values" cell in the "Gender" row, a little button appears at the right side of the cell. We click on the button, and a box appears that allows us to tell the computer that a value of 1 means "male" and a value of 2 means "female." In Figure B.3, we show this box midway through the process: We've already told the software that a value of 1 means "Male," and we're in the process of telling it that 2 means "Female"; at this point, we click on "Add" and then on "OK" to say that we have labeled all possible values of the Gender variable.

FIGURE B.3

The "Value Labels" box for the "Variable View" table



- *Measure:* Indicates whether the variable reflects a nominal scale or an ordinal scale; the category "ordinal scale" also encompasses interval and ratio scales (Chapters 4 and 11 describe the four kinds of scales). As you can see in Figure B.2, our sample data set consists of one variable (Gender) on a nominal scale and six variables that are on interval or ratio scales—hence also on an ordinal scale, which in the Variables table is simply called "scale."

Computing Basic Descriptive Statistics

Now that we have our data set, let's conduct some simple analyses. First, let's compute basic descriptive statistics for six of the seven variables (computing a mean and standard deviation for the "gender" variable would, of course, be meaningless). We move the cursor to the word "Analyze" at the top of the screen and click on the mouse. A pull-down menu appears, and we move the mouse down until the term "Descriptive Statistics" is highlighted, at which point another menu appears to its right. We click on "Descriptives" in the right-hand box. A new box appears in front of our data set. This box contains two smaller boxes, with all seven of our variables listed in the left box. To calculate descriptive statistics for the last six variables, we want to move them into the right box. We do this by highlighting each one and then clicking the right-arrow button between the two boxes. After we've moved the six variables, we click on the "OK" button (see Figure B.4). At this point, a table appears that lists the number of observations (*N*), minimum and maximum values, mean, and standard deviation for each variable. The final row in the table, "Valid N (listwise)," simply means that SPSS found all 15 numbers for each variable to be appropriate ones; in other words, it didn't omit any scores in doing the calculations.

Now let's suppose that we want to see how overall exam performance (Exam Score Total), quantity of notes (Quantity of Notes), and quality of notes (Quality of Notes) are intercorrelated. To do this, we can calculate Pearson *r* correlation coefficients for each possible pairing of these three variables. Once again, we go up to "Analyze" at the top of the screen and click on the mouse. When the pull-down menu appears, we move the cursor down until the word "Correlate" is highlighted, then move the cursor to the right to highlight "Bivariate," and then click on the mouse. Once again, the two-box box appears, and we must move the three variables we want to analyze to the right box and then click on "OK." We now have a table that gives us the intercorrelations among these variables, which we can print out by going to the "File" pull-down menu and then to "Print" (see Figure B.5). The first number in each cell of the table tells us the Pearson *r* for a particular pair of variables (this number is 1 when a variable is correlated with itself), and the third number tells us the number of people for whom the *r* has been calculated. The middle number tells us the probability (*p*) that we would obtain an *r* that high if the two variables were *not* correlated in the overall population from which the sample has been drawn.

The table in Figure B.5 marks with two asterisks (**) all *r*s that are significant at an α level of .01. But we don't necessarily have to use that alpha level. Imagine, instead, that we decide to

FIGURE B.4

Identifying variables for which we want basic descriptive statistics to be calculated

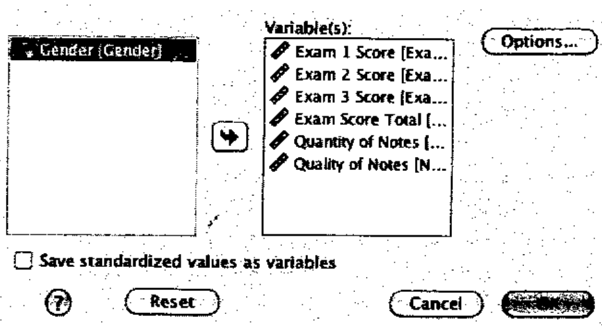


FIGURE B.5

Correlations among exam score total, quantity of notes, and quality of notes

Correlations

		Exam Score Total	Quantity of Notes	Quality of Notes
Exam Score Total	Pearson Correlation	1	.323	.425
	Sig. (2-tailed)		.241	.114
	N	15	15	15
Quantity of Notes	Pearson Correlation	.323	1	.777**
	Sig. (2-tailed)	.241		.001
	N	15	15	15
Quality of Notes	Pearson Correlation	.425	.777**	1
	Sig. (2-tailed)	.114	.001	
	N	15	15	15

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

use a significance level (α) of 0.05 for all of our analyses. Any p value in the table that is *smaller* than 0.05 indicates that the variables probably *are* correlated in the population from which our sample has been drawn. For example, the correlations between Exam Score Total and the Quantity and Quality of Notes are .323 and .425, respectively. Although these correlations are in the low-to-moderate range, the p values associated with them (.241 and .114) tell us that we might get correlations this high *simply by chance* when the two variables are actually unrelated in the overall population. (With a much larger sample size, such correlations would be statistically significant. Our small sample size may be leading us to make Type I errors here.) Now let's look at the correlation between Quantity of Notes and Quality of Notes. This correlation is .777, which has an associated probability of 0.001. This r is statistically significant: Students who take more notes also take better notes. We must be careful, however, that we don't conclude that there is a causal relationship here: Taking more notes does not necessarily cause a student to take better ones, nor does taking better ones cause a student to take more of them. Correlational data alone *never* allows us to draw clear-cut conclusions about cause-and-effect relationships.

Computing Inferential Statistics

In the preceding section we already ventured into inferential statistics a bit. When we looked at the probabilities that our correlation coefficients occurred by chance for a set of possibly unrelated variables, we were drawing inferences. But now let's do so intentionally. Let's see if there are any gender differences in the test performance of males and females. To find out, we need to perform a t-test between the two groups. Once again, we go up to "Analyze," and this time we highlight "Compare Means" and then "Independent Samples T Test."[2] A box similar to that shown in Figure B.4 appears, but this one has three boxes within it. We move our dependent variable (Exam Score Total) into the "Test Variable(s)" box and our independent variable (Gender) into the "Grouping Variable" box. Next, we click on the "Define Groups" button and tell the computer that a value of "1" puts a person in Group 1 (the males) and a value of "2" puts a person in Group 2 (the females). We click on the "Continue" button and then click on "OK." We get tables that provide descriptive statistics for the two groups, information about whether the variances of the two groups are equivalent, and results of t-tests. We can, of course, print out these tables (see Figure B.6). The program has calculated two ts, one based on the assumption of equal variances and another based on the assumption of unequal variances. Given the unequal variances for the two groups (the F value for Levene's test has a probability of .013), we'll look at the second t, which is .055. This value indicates that the two groups are probably not different in their overall exam performance (the p value is .957). (You can find explanations for the other numbers in this table in many statistics textbooks or through an Internet search.)

We have room for one final statistical analysis. Let's say we want to know whether the students performed differently on the three exams they took during the semester. To compare three means for the same group of students, we would ideally want to conduct a repeated-measures analysis of variance. Unfortunately, the version of SPSS we are using here performs only between-subjects ANOVAs, so we will have to settle for three paired-samples t-tests.

To conduct our t-tests, we go back up to "Analyze," move the mouse down to highlight "Compare Means," and then move it to the right to highlight "Paired-Samples T Tests." We release the mouse. Once again, we see a two-box box, but in this one the second box includes three columns labeled "Pair," "Variable 1," and "Variable 2." When we click on Exam 1 in the left box and then click on the arrow, and then subsequently do the same thing for Exam 2, we get an Exam 1–Exam 2 pair in the right box. In a similar manner, we can form Exam1–Exam 3 and Exam 2–Exam 3 pairs. We now have three pairs of variables in the right-hand box. We click on "OK" and print out the three tables that the analysis generates (Figure B.7). The first table gives us descriptive statistics; we've seen most of these before, but the column for standard error

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Exam Score Total	Male	7	117.6429	16.38524	6.19304
	Female	8	117.2500	9.76418	3.45216

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Exam Score Total	Equal variances assumed	8.335	.013	.057	13	.955	.39286	6.85139	-14.40868	15.19439
	Equal variances not assumed			.055	9.520	.957	.39286	7.09022	-15.51373	16.29945

FIGURE B.6

Computing t to determine if males and females have different total exam scores

[2]As noted in Table 11.5 in Chapter 11, a t-test can take either of two basic forms. An *independent-samples t*-test enables a comparision of means for two separate, independent groups. For instance, an independent-samples t-test enables a comparison of males versus females, as in the example presented here. In contrast, a *dependent-samples t*-test—also known as a *paired samples t*-test—enables a comparison of means for a single group of individuals or, instead, for two related groups. For example, a researcher might obtain measures of two characteristics of a single group of students or, alternatively, might obtain measures of one particular characteristic both for a group of fathers and for their first-born sons.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Exam 1 Score	39.9667	15	4.15102	1.07179
	Exam 2 Score	39.4667	15	5.42388	1.40044
Pair 2	Exam 1 Score	39.9667	15	4.15102	1.07179
	Exam 3 Score	38.0000	15	5.15128	1.33006
Pair 3	Exam 2 Score	39.4667	15	5.42388	1.40044
	Exam 3 Score	38.0000	15	5.15128	1.33006

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Exam 1 Score & Exam 2 Score	15	.388	.153
Pair 2	Exam 1 Score & Exam 3 Score	15	.622	.013
Pair 3	Exam 2 Score & Exam 3 Score	15	.814	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Exam 1 Score - Exam 2 Score	.50000	5.40089	1.39450	-2.49091	3.49091	.359	14	.725
Pair 2	Exam 1 Score - Exam 3 Score	1.96667	4.14241	1.06956	-.32732	4.26065	1.839	14	.087
Pair 3	Exam 2 Score - Exam 3 Score	1.46667	3.23329	.83483	-.32387	3.25720	1.757	14	.101

FIGURE B.7

Computing *t*s to determine if students performed differently on the three exams

of the mean is new. We also see Pearson *r*s for the three pairs. We are most interested in the *t* values for three pairs of exam scores, which are shown in the seventh column in the bottom table. None of these *t*s is statistically significant at our significance level of .05 (see the rightmost column), although the Exam 1–Exam 3 pair comes close, with a *p* value of .087.

We have merely scratched the surface of what SPSS can offer. We have ignored some of the values in the statistical tables we've presented. And we haven't even touched on SPSS's graphing capabilities. We urge you to explore SPSS for yourself to discover the many analyses it can perform and the many graphical displays it can create.