Dependent Samples t-Test

The paired t-Test tests the difference between Mean in pairs of measurements to a Mean of zero

The Paired t-test compares the differences between pairs of *measurements taken from the same test subjects* at different times or under different conditions

	Paire	ed <i>t</i> -test	
	Before Training	After Training	Difference $n = 5$
K. Albert	74	78	+4
P. Jacobs	76	83	+7
T. Smith	73	81	+8
R. Wang	81	84	+3
D. Young	78	86	+8

Paired *t*-test: $\frac{\bar{d} - 0}{s_d / \sqrt{n}}$



is the Mean of the differences in the two values comprising each pair

Degree of freedom

n – 1(*n* is the number of pairs)

There are two sets of measurements taken –
 before and after

 The test treats it as one Sample which consists of the differences calculated between the two measurements taken from each individual test subject

Conditions for use

 Check for normality of the measurements after the experiment takes place (difference between means)

• The most common use for the Paired t-test is for a "before and after" analysis

EXAMPLE

Does a training program make a Statistically Significant difference in the production output of individual workers?

The 2-Sample t-test would not work as well for answering that question.

There are any number of **Factors** which can affect worker production output. Experience is one. If one Sample has workers with more experience than the other Sample, then that could give us misleading results.

But if we measure the same workers before and after training, we can eliminate other Factors, like experience, so that we can focus on only one Factor – **the training**

Before







Compared to the 2-Samples t-test the Paired ttest does a much better job of <u>Blocking</u> out the <u>effect of other Factors</u> (Experimental Design)

Null Hypothesis

The experiment (training) has had **NO** (Statistically Significant) **effect**

H0= Mean difference between the two paired measurements is zero

HA= Mean difference between the two paired measurements is not equal to zero





Paired Samples Statistics

		Mean	Ν	Std. Deviation	Std. Error Mean
Pair 1	pre	1073.00	20	165.119	36.922
	post	1088.80	20	164.560	36.797

Paired Samples Test

				Paired Differen	ces				
				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	-15.800	2.648	.592	-17.039	-14.561	-26.687	19	.000

Does the Biostatistics course add to the students' knowledge in statistics??









Student	Before mark	After mark	Diff
1	18	22	4
2	21	25	4
3	16	17	1
4	22	24	2
5	19	16	-3
6	24	29	5
7	17	20	3
8	21	23	2
9	23	19	-4
10	18	20	2
11	14	15	1
12	16	15	-1
13	16	18	2
14	19	26	7
15	18	18	0
16	20	24	4
17	12	18	6
18	22	25	3
19	15	19	4
20	17	16	-1





H0: There is no difference in mean pre- and post-marks

• HA: There is a difference in mean pre- and post-marks

P	aired	Sam	ples	Stat	tistics	
						_

				Std.	Std. Error
		Mean	N	Deviation	Mean
	Mark after	20.45	20	4.058	.907
	Mark				.705
Pair 1	before	18.40	20	3.152	

p-value

/

Paired Samples Test												
		Paired	Differences									
	Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference		95% Confidence Interval of the Difference		Std. Error Mean Std. Difference		t	df	Sig. (2- tailed)
			wean	Lower	Upper							
Mark after - Mark before	2.050	2.837	.634	.722	3.378	3.231	19	.004				

Conclusion:

There is strong evidence (t = 3.23, p = 0.004) that the teaching intervention improves marks

In an experiment to compare <u>anxiety levels</u> induced between looking at real spiders and pictures of spiders, the following data was collected from 12 people with a fear of spiders (arachnophobia)

	Anxiety	score	
Participant	Picture	Real	Diff
1	30	40	10
2	35	35	0
3	45	50	5
4	40	55	15
5	50	65	15
6	35	55	20
7	55	50	-5
8	25	35	10
9	30	30	0
10	45	50	5
11	40	60	20
12	50	39	11

Paired Samples Statistics

		Mean	Ν	Std. Deviation	Std. Error Mean
Pair 1	Anxiety score when looking at a spider picture	40.00	12	9.293	2.683
	Anxiety score when looking at a real spider	47.00	12	11.029	3.184

Paired Samples Test

				Paired Diff	erences				
				Std Error	95% CI of the Difference				Sig (2-
		Mean	SD	Mean	Lower	Upper	t	df	tailed)
Pair 1	Anxiety score when looking at a spider picture - Anxiety score when looking at a real spider	-7.000	9.807	2.831	-13.231	769	- 2.473	11	.031

There is evidence to suggest that participants experienced statistically significantly greater anxiety (p = 0.031) when exposed to real spiders (mean = 47.0 units, SD = 9.3) than to pictures of spiders (mean = 40.0 units, SD = 11.0). The 95% confidence interval for the difference is (-13.2,-0.77)

To test for normality of the quantitative variable: 1- Q-Q Plot

2- Kolmogorov-Smirnov Test

3- Shapiro-Wilk Test

Q-Q Plot



Null Hypothesis for Kolmogorov-Smirnov Test & Shapiro-Wilk Test

The DATA is normally distributed

•Ur	ntitled	1 [DataSet0] - F	ASW Statist	ics Data E	ditor						1 🛪
File	Edit	<u>View</u> Data	Transform	Analyze	Direct	Marketing	Graphs	Utilities	Add-ons	Window	Help
				2			i ii		¥		20
1:Da	ta	20	0.1333						Visit	ble: 1 of 1 V	ariables
1		Data	Var	va	ŕ	var	var		/ar	var	
	1	20.13	8								-
	2	14.90)								
1 1	3	14.14	L .								
1	4	12.66	5								
	5	18.57	7					_			
	6	17.50)								
1 87	7	16.56	5								
	8	22.54	1								
	9	22.68	3								
1	0	18.35	5								
1	1	23.80)								
1	2	18.33	3								
1	3	18.58	3								
1	4	23.07	7								
1	5	20.62	2								
1	6	23.55	5								
1	7	14.93	3								
1	8	18.45	5								
1	9	16.27	7								
2	20	13.82	2								
2	1										
2	2										-
-	_					-			1.11		E
Data	View	Variable View									
						PASW	Statistics	Processo	r is ready		

Select "Analyze -> Descriptive Statistics -> Explore

T *Untitled	1 [DataSet0] - PAS	SW Statist	ics Data E	ditor				
<u>File Edit</u>	<u>View D</u> ata <u>T</u> i	ransform	<u>A</u> nalyze	Direct <u>M</u> arketing	<u>G</u> raphs	Utilities Add	d- <u>o</u> ns	<u>W</u> indow <u>H</u> elp
		10	Rep	ports				
1 Data	20.1	333	D <u>e</u> s	criptive Statistics	+	Erequence	cies	f 1 Variables
T. Data	Data	var	Con	npare Means	р. 	Descripti	ves	
1	20.13		<u>G</u> en	 neral Linear Model		Crosstab		
2	14.90		Gen	neralized Linear Mo	dels 🕨		15	
3	14.14		Mi <u>x</u> e	ed Models	*	Ratio		
4	12.66		Cor	relate	*	P-P Plots	1.1 .2	
5	18.57		Reg	pression	ь.	🛃 Q-Q Plots	S	

Data 😽	Dependent List:	Statistics Plots
•	Eactor List:	Options Bootstrap
□Display ● <u>B</u> oth ◎ St <u>a</u> tistics ◎ P <u>l</u> ots		
OK Paste	Reset Cancel Help)

From the list on the left, select the variable "Data" to the "Dependent List"



Boxplots © Eactor levels together © Dependents together	Descriptive Stem-and-leaf <u>H</u> istogram
Normality plots with tests Spread vs Level with Lever	ne Test
100/10/10	
© Power estimation © Transformed Power:	Vatural log 🔹 🔻

Output1 (Document1) - PASW St.	atistics View	er.								Call	×
Elle Edit View Data Transfor	m insert	Format 2	malyze Dire	ect Marketin	g <u>G</u> raphs	Utilities	Add-ons y	Mindow	Help		_
😑 H 🖨 🗟 🤌		-	1 🎽			0					1
+ + + -											
Cutput	/мі	SSING LIS	STWISE	11				_			-
Explore	/N0	TOTAL.									
Title	Evol	ore									
Active Dataset	-Apr	ore									
Case Processing Descriptives	• [a Case Processing [DataSet0]										
Tests of Normality	- Le Tests of Normality										
Data	_	ř –	Case P	rocessing 1	Summary						
Normal Q-Q F	+	Va	lid	Miss	ting	Tot	al.				
Log	24	N	Percent	N	Percent	N	Percent				
	Data	20	100.0%	0	.0%	20	100.0%				
				Bassadat							
	Descriptives						Std. Erro	or l			
	Data	Mean				18,4715	.7712	6			
	95% Confidence Interval Lower Bound for Mean					16.8577					
1		Upper Bound 5% Trimmed Mean Median Variance									
		Std. Devia	tion			3.44918					
		Minimum				12.66					
		Range				11.14					
		Interquarti	le Range			6.80					
		Skewness					51	2			
		Kurtosis				+1.080		2			
	Tests of Normality										
		Kolmogorov-Smirnov* St				hapiro-Wilk					
		Statistic	df	Sig.	Statistic	đf	Sig.				
	Data a Li	,137 Illiefors Sign	20 ificance Corre	200	946	20	.316				
	*. Tr	tis is a lower	bound of the	true signifi	cance.						
[*]				0.000			In a new new				-
						_	PASW St	austics Pri	ocessoris	ready.)	

Case Processing Summary

	Cases								
	Valid		Miss	sing	Total				
1	N	Percent	N	Percent	N	Percent			
Data	20	100.0%	0	.0%	20	100.0%			

			Statistic	Std. Error
Data	Mean	18.4719	.77126	
	95% Confidence Interval for Mean	Lower Bound	16.8577	
		Upper Bound	20.0862	
	5% Trimmed Mean	18.4986		
	Median	18.3962		
	Variance	11.897		
	Std. Deviation	3.44918		
	Minimum	12.66		
	Maximum	23.80		
	Range	11.14		
	Interquartile Range	6.80		
	Skewness	.070	.512	
	Kurtosis	-1.060	.992	

Descriptives

Tests of Normality

-	Kolmo	gorov-Smirr	nov ^a	Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
Data	.137	20	.200	.946	20	.316	

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.