TutorTube: Paired Samples in SPSS and Excel

Spring 2021

Introduction

Hello, welcome to another edition of TutorTube, where the Learning Center's Lead Tutors help you understand challenging course concepts with easy to understand videos. My name is Kelly Schmidt, Lead Tutor for statistics at the Learning Center. In today's video, we will go through the process conducting a paired samples (or dependent samples) t-test in SPSS. Let's get started!

Research Scenario

First let's look at a research scenario. In this situation, we are interested in testing the effectiveness of an after-school math program for 5th graders. We administer a pretest to 12 students before a month-long intervention and then a post-test immediately following the program's end.

Student	Pretest	Post-test
1	47	67
2	52	60
3	68	79
4	80	78
5	91	92
6	77	80
7	76	82
8	81	94
9	53	71
10	67	70
11	89	90
12	54	60

At first, you might be tempted to use an independent samples t-test to compare the means of these groups. Unfortunately, that kind of test doesn't work here. Our samples (the pretest and post-test groups) are NOT independent; the same students are measured in each group, making this a **dependent** test. As such, we need to do a paired-samples test instead.

Data Entry

First, start by opening up a blank data sheet in SPSS.

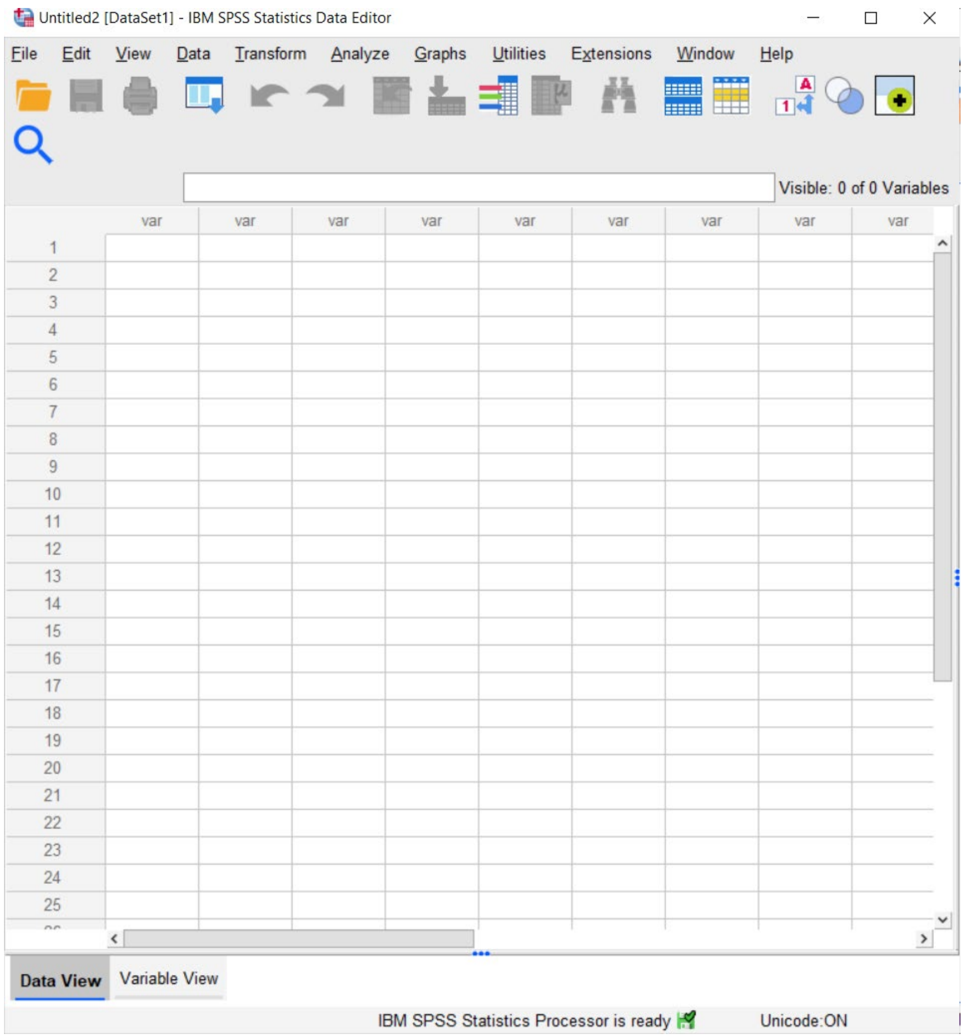


Figure 1. Initial Data View

Be sure you are in Data View, instead of Variable View. Data View is the place where we enter the raw data that we are given. So, let's copy and paste the scores into the sheet.

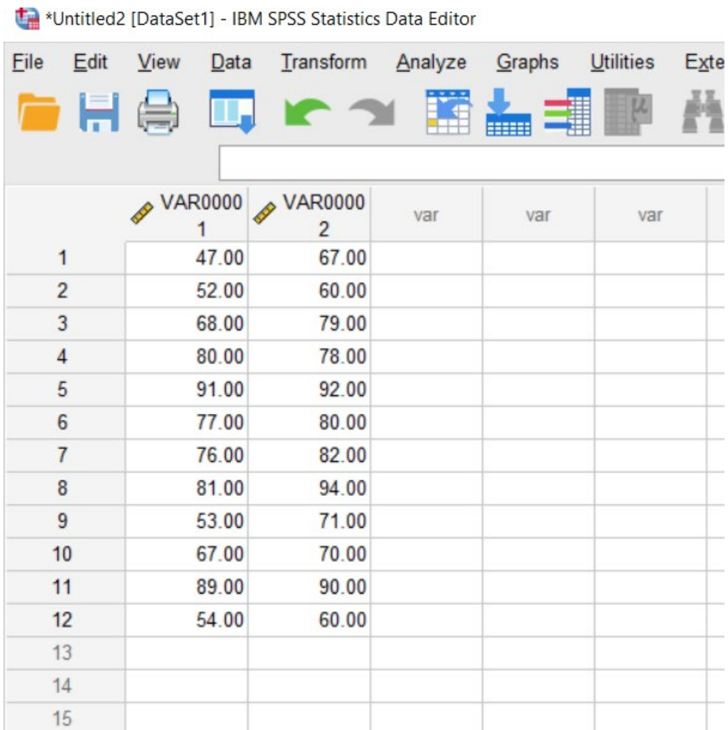


Figure 2. Data Entered in Data View

Notice that when we first copy them in, SPSS automatically labels the columns as scale variables called VAR00001 AND VAR00002. Remember that SPSS treats each row as a single person, so we can check to make sure we have 12 rows for 12 students.

Next, click on Variable View to switch to the variable coding window. This is the place where we name our variables and specify their measures and labels.

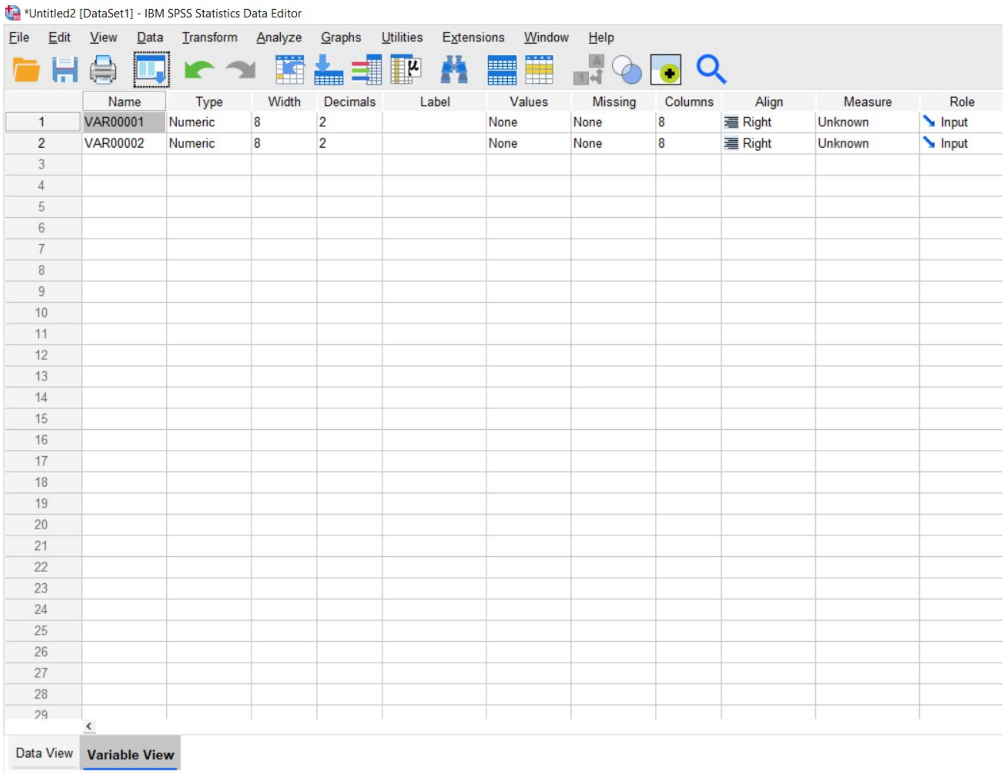


Figure 3. Variable View Window

We can start by clicking in the Name column and renaming VAR00001 as Pretest and VAR00002 as Posttest. Next, we can go to the Measure column and code each variable as scale. Unlike an independent-samples *t*-test, we don't need to specify labels for groups here. Also, be sure that each variable is coded as Numeric under Type.

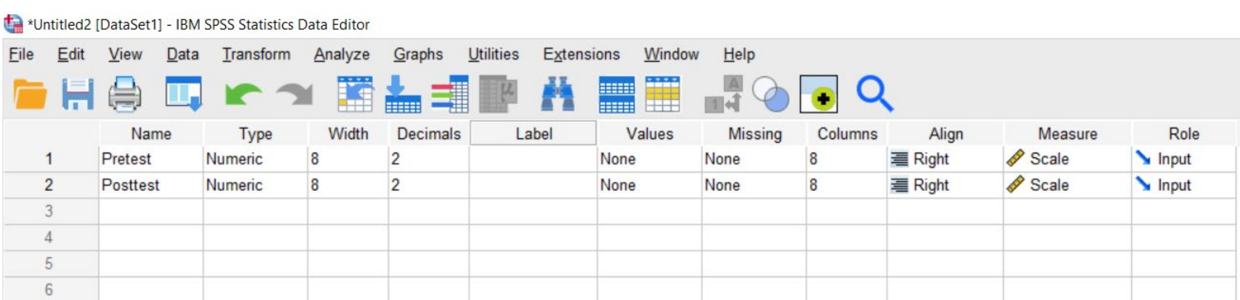


Figure 4. Variable View with Variables Coded

We can switch back to Data View to see how our data has changed. Notice that the columns are now labeled correctly. Now we are ready to do the analysis.

Analysis

Click on Analyze, Compare Means, Paired Samples T Test.

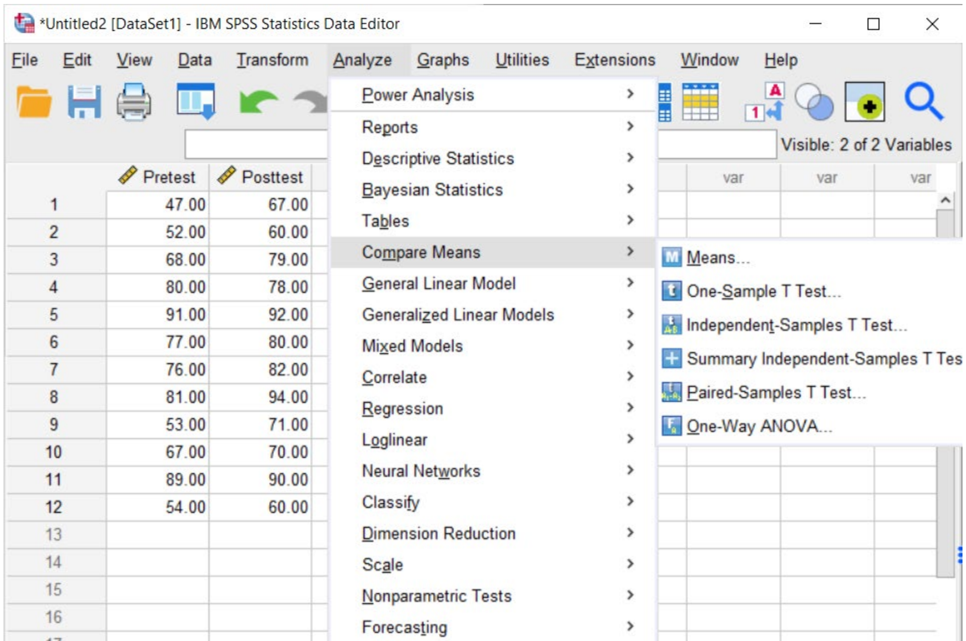


Figure 5. Analyze > Compare Means > Paired-Samples T Test

Once we get to this window, we need to choose the order that we want SPSS to use when comparing the Pretest and Posttest variables. Usually, we go with an After minus Before model for these types of tests because this lets us see more easily if scores have improved. So, start by moving Posttest into the Variable 1 box, and then move Pretest into the Variable 2 box.

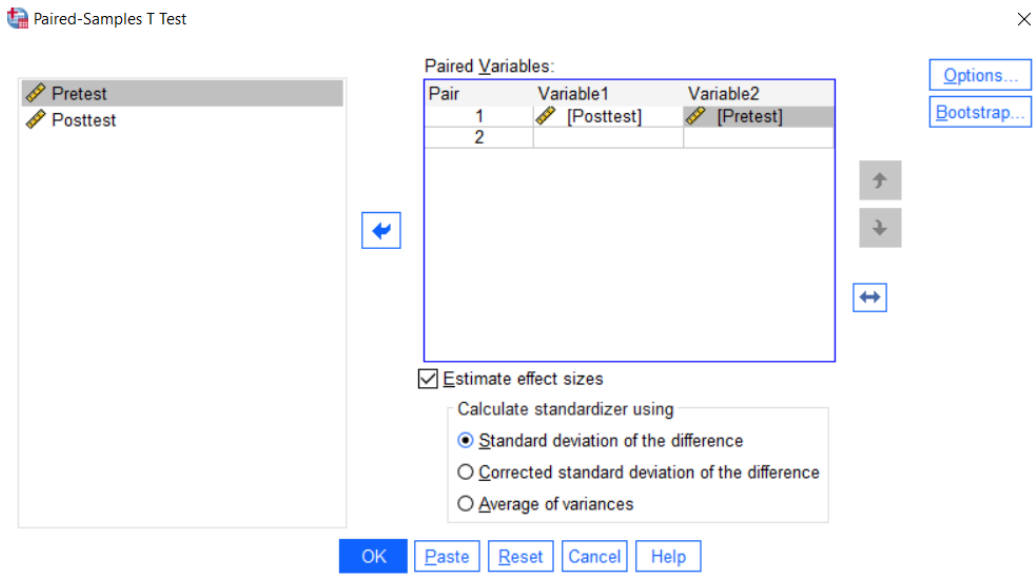


Figure 6. Paired Samples Window

From here, we don't need to change anything else, so we can just hit OK.

Output

This will bring up our output window. The first box shows the descriptive statistics for the test. We can see that the mean of the post-test was about 7 points higher than the mean of the pretest.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Posttest	76.9167	12	11.61862	3.35401
	Pretest	69.5833	12	15.12649	4.36664

Figure 7. Descriptive Statistics

The Paired Samples Test box gives us the information we need to answer our research question.

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	Posttest - Pretest	7.33333	6.93258	2.00126	2.92858	11.73808	3.664	11	.004

Figure 8. Paired-Samples Output

First, we can see that the mean of the difference scores was 7.33 with a standard deviation of 6.93. So, on average, there was a 7.33 increase per student. In order to determine if this increase was statistically significant, we need to look at the last box labeled Sig. (2-tailed). This gives us the p-value (the p-calc) for the test.

Because our p-value of .004 is less than the standard significance level of .05, we can say that there is a statistically significant difference in scores between the Pretest and Post-test. In other words, there is evidence here to suggest that our intervention program had an effect.

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	Posttest - Pretest	Cohen's d	6.93258	1.058	.328	1.758
		Hedges' correction	7.18066	1.021	.317	1.697

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Figure 9. Cohen's d Effect Size

Next, we can look at the effect size of the test to determine just how impactful the increase in scores actually was. Looking at the Cohen's *d* point estimate, we can see that this is a large effect size at 1.058. When interpreting Cohen's *d*, a general rule of thumb is that anything smaller than .2 is considered a small effect and anything larger than .8 is considered large. Since our value of 1.058 is bigger than .8, we can say that that difference in scores between the pretest and post test groups was noticeably different from zero, which is an indication that our math intervention program might have been really good!

Paired Samples in Excel

SPSS is a good tool for this analysis, but if you want to get an idea of the mechanics behind the calculations, you can also get the same results in Excel.

First, paste the data into an empty sheet.

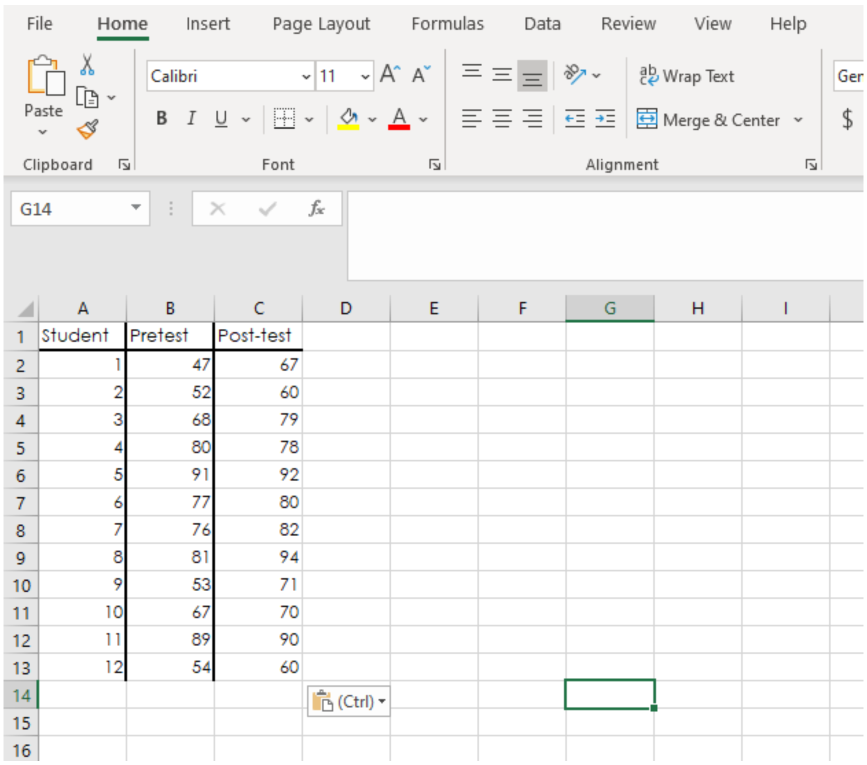
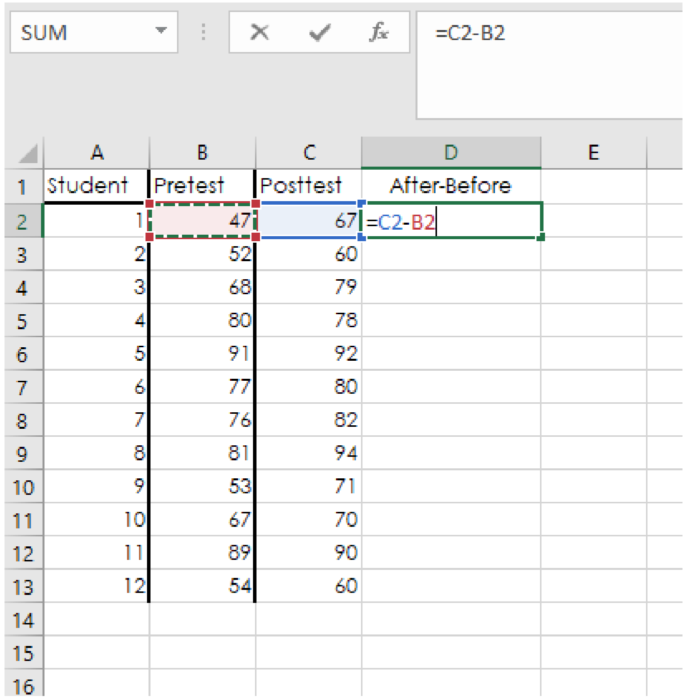


Figure 10. Excel Data

Next, we are going to create a new column containing the After-Before difference scores. I'll label this as Post-Pre. In the first empty cell, type an equal sign. Next click on the box containing the first Post-test score. Type a minus sign. Then click on the cell with the first score for the Pretest. Hit enter, and notice that the formula subtracted the two values to give us 20.

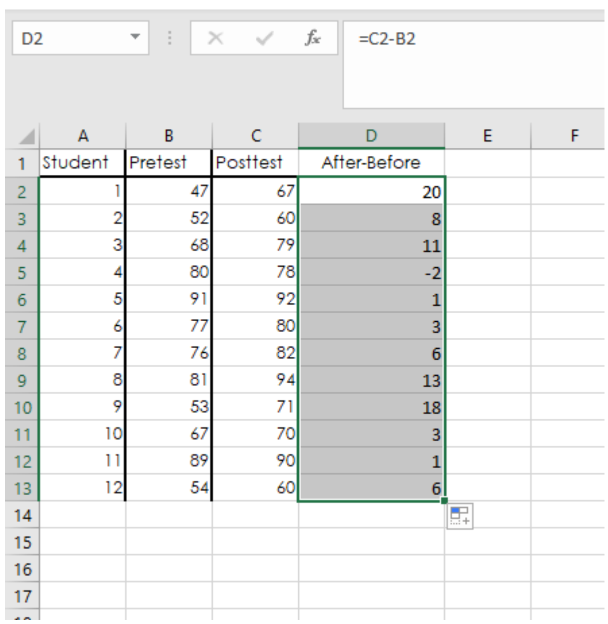


SUM | =C2-B2

	A	B	C	D	E
1	Student	Pretest	Posttest	After-Before	
2	1	47	67	=C2-B2	
3	2	52	60		
4	3	68	79		
5	4	80	78		
6	5	91	92		
7	6	77	80		
8	7	76	82		
9	8	81	94		
10	9	53	71		
11	10	67	70		
12	11	89	90		
13	12	54	60		
14					
15					
16					

Figure 11. Difference Scores

Next, click on the black arrow in the corner of the box and drag down to copy the formula to all the cells.



D2 | =C2-B2

	A	B	C	D	E	F
1	Student	Pretest	Posttest	After-Before		
2	1	47	67	20		
3	2	52	60	8		
4	3	68	79	11		
5	4	80	78	-2		
6	5	91	92	1		
7	6	77	80	3		
8	7	76	82	6		
9	8	81	94	13		
10	9	53	71	18		
11	10	67	70	3		
12	11	89	90	1		
13	12	54	60	6		
14						
15						
16						
17						

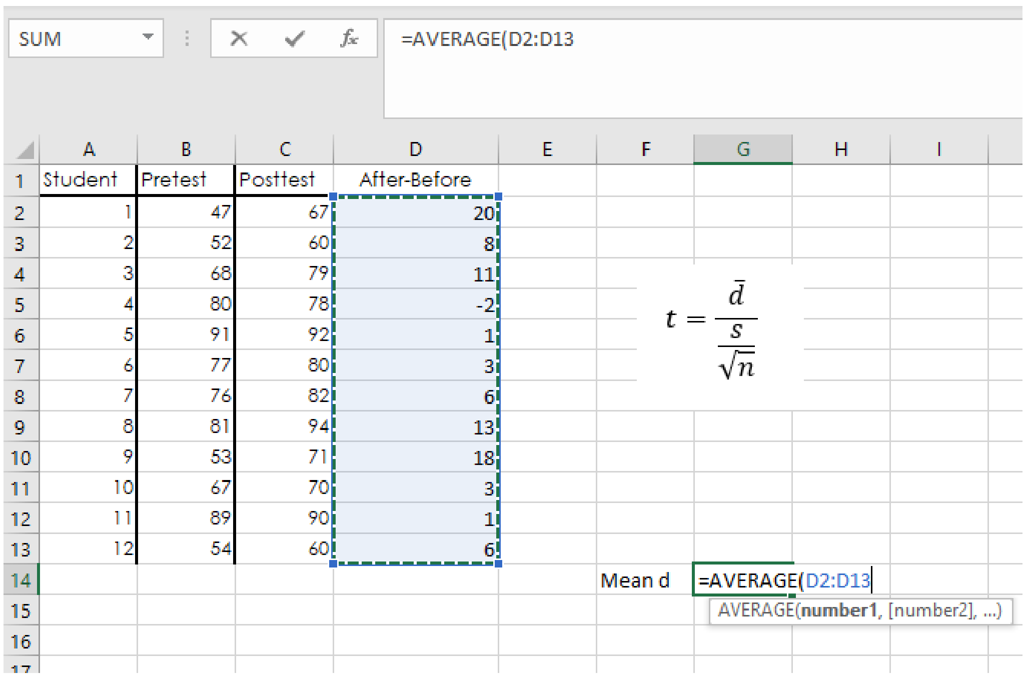
Figure 12. Difference Scores Column

Now we need to find out t-test statistic. The formula for this is test is:

$$t = \frac{\bar{d}}{\left(\frac{s}{\sqrt{n}}\right)}$$

where $\bar{d}$ is the mean of the paired difference scores that we just found, s is the sample standard deviation of the difference scores, and n is the sample size.

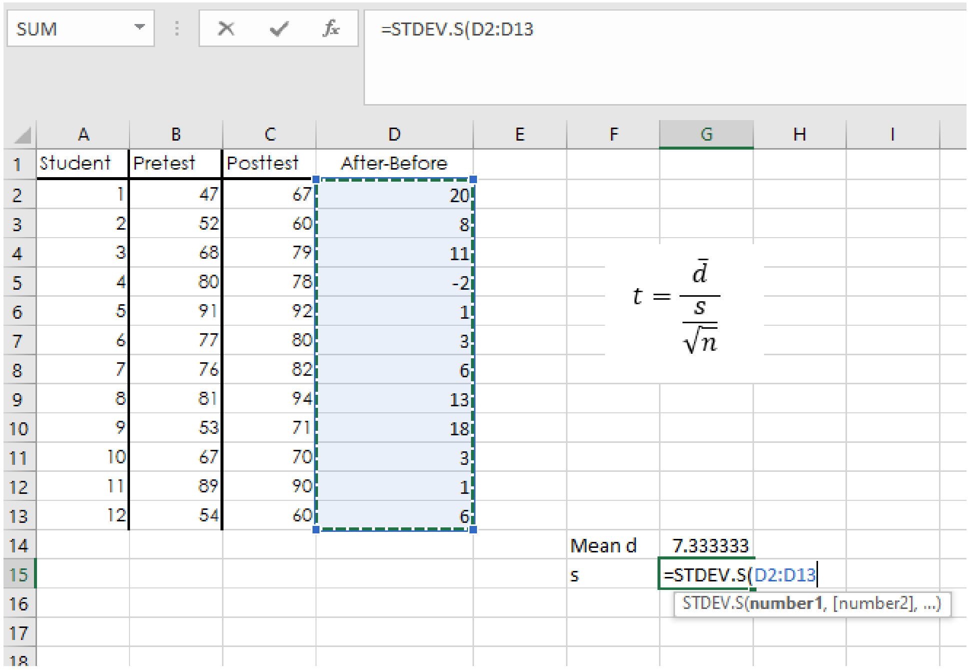
To find the mean of the scores, I'll pick an empty cell and type =AVERAGE(and then highlight my column and press enter. This will give us the mean difference, $\bar{d}$.



Student	Pretest	Posttest	After-Before
1	47	67	20
2	52	60	8
3	68	79	11
4	80	78	-2
5	91	92	1
6	77	80	3
7	76	82	6
8	81	94	13
9	53	71	18
10	67	70	3
11	89	90	1
12	54	60	6

Figure 13. Mean Difference

Next, I'll pick another cell and type =STDEV.S(and highlight the column again. This will give me my sample standard deviation.

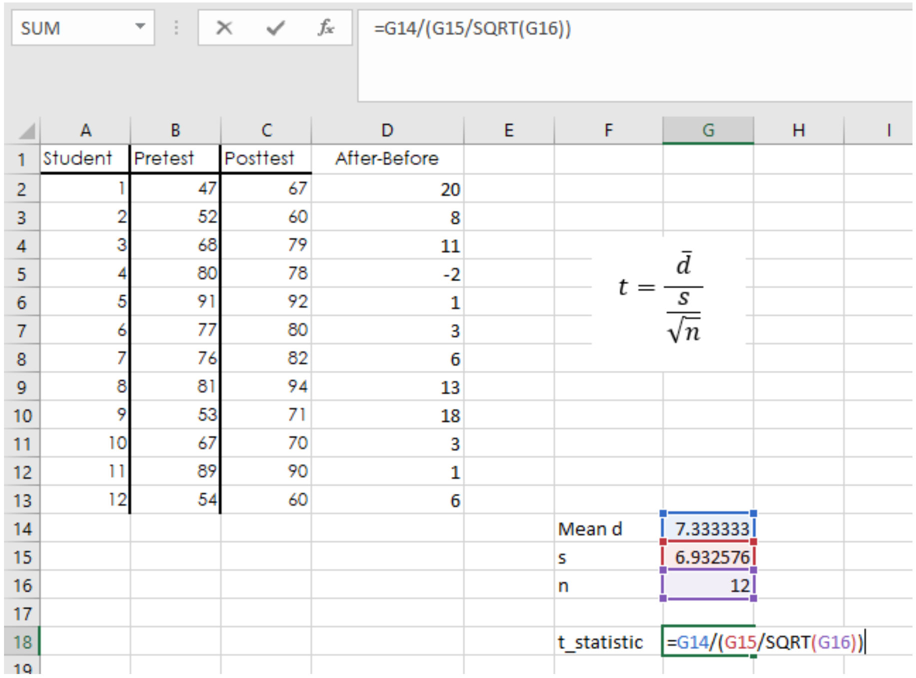


	A	B	C	D	E	F	G	H	I
1	Student	Pretest	Posttest	After-Before					
2	1	47	67	20					
3	2	52	60	8					
4	3	68	79	11					
5	4	80	78	-2					
6	5	91	92	1					
7	6	77	80	3					
8	7	76	82	6					
9	8	81	94	13					
10	9	53	71	18					
11	10	67	70	3					
12	11	89	90	1					
13	12	54	60	6					
14						Mean d	7.333333		
15						s	=STDEV.S(D2:D13		

Figure 14. Sample Standard Deviation

My sample size, n, is 12 because I have 12 students and 12 scores here.

Now I'll find my *t* statistic by building the formula. If you wanted to, you could plug these numbers into a standard calculator, but we can do it in Excel too. In another cell type an equal sign then click on the value of our mean. Type a division sign, then a parenthesis. Click on the value for standard deviation, type another division sign, and then type SQRT(, and click on the value of n. Finally, we type two closed parentheses and hit enter.



SUM | =G14/(G15/SQRT(G16))

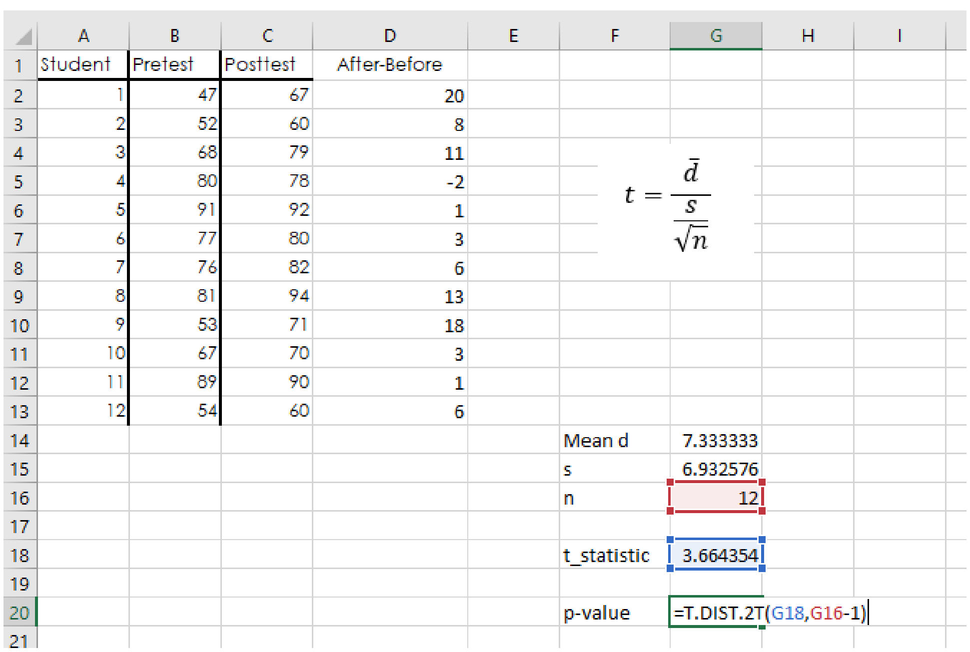
	A	B	C	D	E	F	G	H	I
1	Student	Pretest	Posttest	After-Before					
2	1	47	67	20					
3	2	52	60	8					
4	3	68	79	11					
5	4	80	78	-2					
6	5	91	92	1					
7	6	77	80	3					
8	7	76	82	6					
9	8	81	94	13					
10	9	53	71	18					
11	10	67	70	3					
12	11	89	90	1					
13	12	54	60	6					
14						Mean d	7.333333		
15						s	6.932576		
16						n	12		
17									
18						t_statistic	=G14/(G15/SQRT(G16))		

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}$$

Figure 15. t Statistic

This will give us our *t* statistic of 3.664, which is exactly the same number that SPSS calculated.

Finally, we can calculate our two-tailed *p*-value. In another cell, type =T.DIST.2T(for two-tailed. Then click on our *t* statistic, type a comma. For degrees of freedom, we need our n minus 1.



	A	B	C	D	E	F	G	H	I
1	Student	Pretest	Posttest	After-Before					
2	1	47	67	20					
3	2	52	60	8					
4	3	68	79	11					
5	4	80	78	-2					
6	5	91	92	1					
7	6	77	80	3					
8	7	76	82	6					
9	8	81	94	13					
10	9	53	71	18					
11	10	67	70	3					
12	11	89	90	1					
13	12	54	60	6					
14						Mean d	7.333333		
15						s	6.932576		
16						n	12		
17									
18						t_statistic	3.664354		
19									
20						p-value	=T.DIST.2T(G18,G16-1)		

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}$$

Figure 16. p-Value

Once we hit enter, we see that the p-value is .0037, which rounds to the same value that SPSS found of .004.

And that is all there is to it!

Outro

Thank you for watching this TutorTube presentation! I hope you enjoyed this video. Please subscribe to our channel for more exciting videos. Check out the links in the description below for more information about The Learning Center and follow us on social media. See you next time!