Chapter 7. Using SPSS to Conduct Hypothesis Tests for Two Samples

Introduction to Two-Sample t tests

In the lecture portion of the course, you learned about hypothesis testing. In its simplest form, a sample mean is compared to a known population mean. Under these circumstances, we can conduct either a *z* test or a single-sample *t* test.

In many if not most research applications, we do not have population parameters available to us. As a result, we need to compare two (or more) samples of data to determine if the means are significantly different. Such comparisons allow us to determine if the samples were drawn from the same population or if they were drawn from different populations. These types of analyses also allow us to describe cause-and-effect relationships. We can use the results from such analyses to determine if our independent variable caused a change in our measure, our dependent variable.

In this chapter, you will learn how to conduct two-sample *t* tests. These *t* tests compare the means from two samples of data, and allow us to determine if a significant difference exists between the two means.

We have two different forms of this *t* test. The first is the **independent-groups *t* test**. When we use this test, we take measures from two different groups of individuals, and each group is measured under a single level of our independent variable. We work from the assumption that there is absolutely no relation between the scores across the two samples; that is, they are **"independent."**

The second version is the **related-samples *t* test**. When we use this variation of the *t* test, we acknowledge that the scores in the two samples are somehow related to one another. We have three different scenarios in which we use the related-samples *t* test.

The most common version of the **related-samples *t* test** is the **repeated-measures *t* test**. In this case, we measure each individual under both levels of our independent variable; that is, we obtain a pair of scores from each individual, gathering one measure under each of the two experimental conditions.

We also use the **related-samples *t* test** when scores occur in "natural" pairs, for instance, when one sample contains the measures from parents, and the other sample results from taking the same measure from their children. Parents and children are a naturally occurring pair. They have much interaction, and therefore, are likely to exert some type of influence over one another's behavior. As a result, these scores are not likely to be independent, and are more likely to be related. This variation of the related-samples *t* test is referred to as the **"paired-samples" *t* test**.

In some instances, we gather measures from different groups of individuals; however, prior to testing, the researcher matches them up on some variable that is likely to have a great impact on our measure. In doing so, the researcher forces or "creates" a relation in the scores across the two groups. This third variation of the **related-samples *t* test** is referred to as the **"matched-samples" *t* test**.

In this chapter, you will learn how to conduct the independent *t* test, as well as the related-samples *t* test.

Data Set 7.1

The owner of a large appliance chain is seeing a drop off in sales. He wonders if there is some way to boost sales. Currently, all of his employees receive a fixed salary regardless of the number of sales they make. He wonders if sales might increase if he pays his employees a “commission” based on the number of sales they make.

To test to see if there is a difference in sales dependent upon payment method, he randomly selects a sample of 20 people from the television (TV) departments of his stores. These individuals are randomly assigned to be paid on a fixed salary (n = 10) or on a commission (n = 10) based on the number of TVs sold. He records the number of TVs each person sells across a 2-month period. The data are presented below.

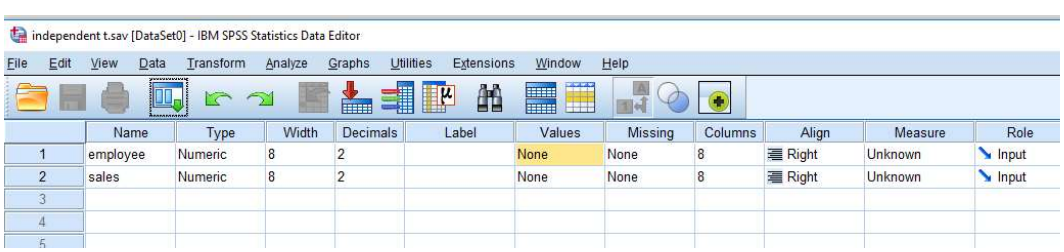
salary	commission
15	17
16	15
14	18
13	21
17	22
12	16
15	12
9	18
16	13
14	19

The Independent t test

1. Open SPSS, and start with a new data screen.

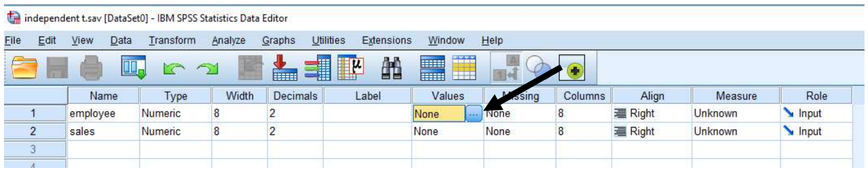
2. Switch to **“Variable View”** to begin to code your variables.

3. In the **“Name”** column, enter the word **“employee”** in the first row, and then enter **“sales”** in the next row.

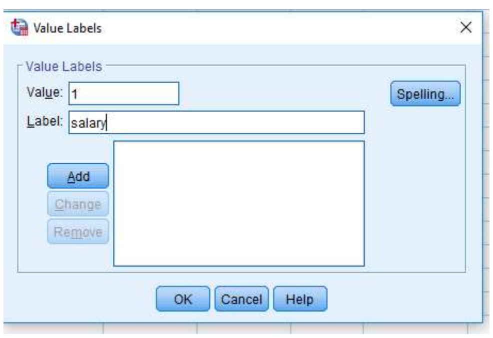


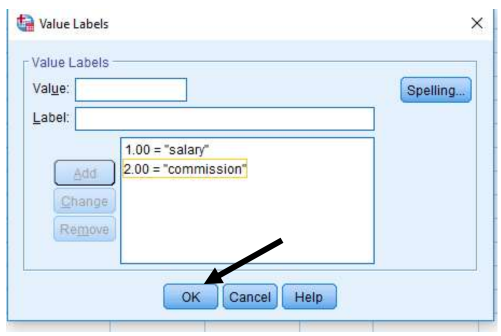
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	employee	Numeric	8	2		None	None	8	Right	Unknown	Input
2	sales	Numeric	8	2		None	None	8	Right	Unknown	Input
3											
4											
5											

4. The next step in coding the variables is to define the groups. Click in the box in the cell under the **"Values"** column in the **"employees"** row. Click the button with the dots. A pop up will appear.

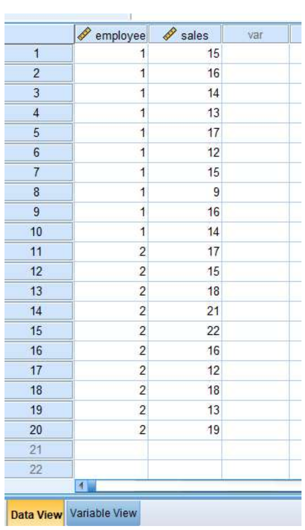


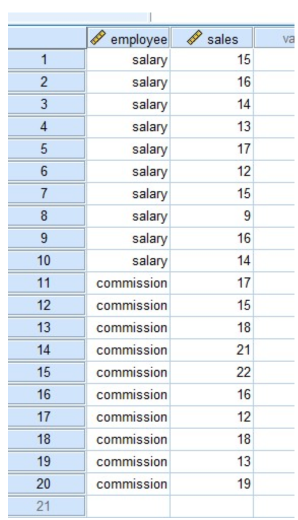
5. Now begin to code the values for the type of pay for the employees. Enter **"1"** in the **"Value"** box, and enter **"salary"** in the **"Label"** box, and then click **"Add."** Now code in the second type of employee, those on commission. Enter **"2"** in the **"Value"** box, and enter **"commission"** in the **"Label"** box, and then click **"Add."** Click **"OK"** to continue.



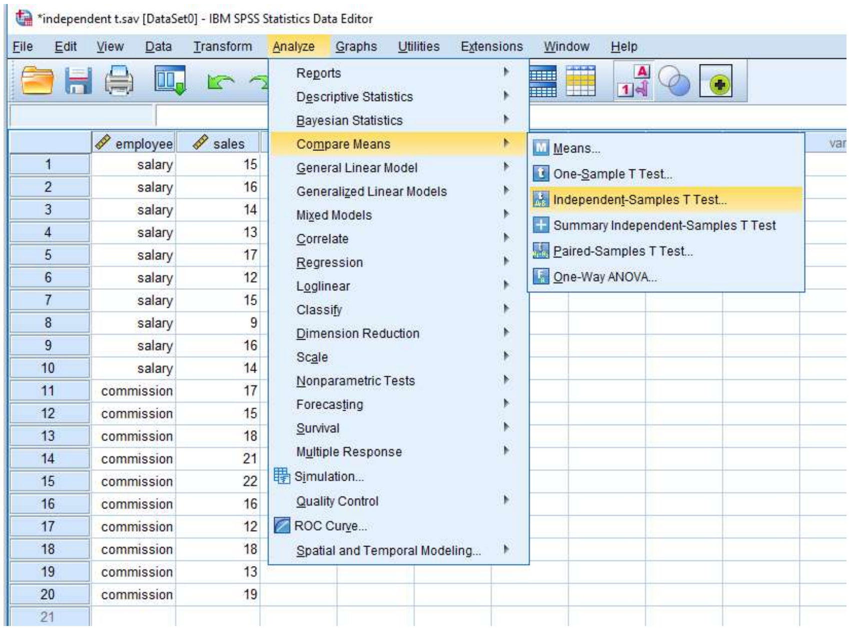


6. Switch to the **"Data View"** tab, and begin entering the sales for each employee. You may find it easier to first hit the **A→1** button  so that the labels appear before entering the data.

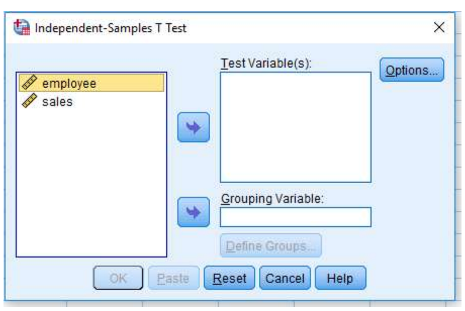


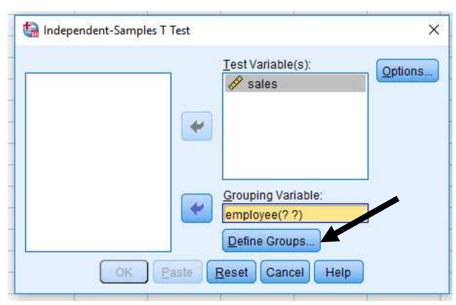


7. Now begin to analyze the data using the Independent *t* test. From the menu bar at the top of the **"Data View"** screen, select **"Analyze"** then **"Compare Means"** followed by **"Independent-Samples T Test."**

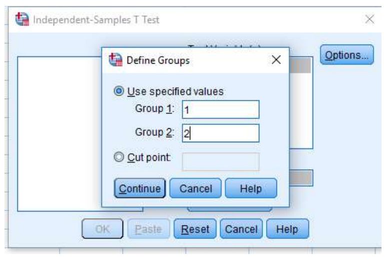


8. Once the selections are made, a pop up will appear. The next step is to move the independent and dependent variables to the correct places. First, select your dependent variable, **"sales,"** and move it to the **"Test Variable"** box using the arrow button. Then select your independent variable, **"employee,"** and move it to the **"Grouping Variable"** box using the arrow button. Then click on **"Define Groups."** (See arrow below.)





9. Now define the groups. Enter **"1"** in the **"Group 1"** box, and **"2"** in the **"Group 2"** box. Click **"Continue,"** and the pop up will disappear. Then click **"OK."** The results of the analysis will appear in the SPSS output window.



10. The results of the analysis are presented below. The **"Group Statistics"** table gives us our descriptive statistics including our group means, standard deviations and standard error of the mean. We need the means and the standard error to graph the results. (See boxes below.)

The next table gives us our *t* statistic, and our test for equality of variance. One of the assumptions of the independent *t* test is that the variances of the two samples are <u>not</u> significantly different. The *p* value for the Levene's test is .299, and this is greater than .05. That tells us the ***sample variances*** <u>do not</u> significantly differ.

Our *t* statistic is also presented in this table, and its value is –2.390. The *p* value for our *t* test is in the next table, and it is .028. As our *p* value is **less than .05**, this tells us we have <u>a significant difference</u> between our means.

Group Statistics

	employee	N	Mean	Std. Deviation	Std. Error Mean
sales	salary	10	14.10	2.331	.737
	commission	10	17.10	3.213	1.016

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means	
		F	Sig.	t	df
sales	Equal variances assumed	1.143	.299	-2.390	18
	Equal variances not assumed			-2.390	16.419

The *t* statistic

The *p* value for Levene's test

Degrees of freedom

Independent Samples Test

		t-test for Equality of Means			
		Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the ... Lower
sales	Equal variances assumed	.028	-3.000	1.255	-5.637
	Equal variances not assumed	.029	-3.000	1.255	-5.655

The *p* value for the *t* test

11. Once we have all of these values, we can interpret our results. While there is no "exact" way in which to write up an interpretation of the results, the fill-in-the-blank template (below) can be used as a guide in writing an interpretation.

If your result is **significant**, use the following guide:

- "The analysis showed significant effect of IV name on DV name, (report statistic)"
  - $t(df)=t_{obt}$, $p<\alpha$ (or use *p* = and your "*p*" value)
- Then compare means
  - "Mean measure name was significantly… pick an adjective – higher, greater, lower, etc. for those exposed to/given/treated with **IV level name** (*M* =) than those exposed to/given/treated with **IV level name** (*M* =)."
- Then interpret in simple English
  - "This finding suggests…"

If your result is **not significant**, you may use the following guide:

- "Even though mean measure (DV) name was… pick an adjective – higher, greater, lower, etc. for those exposed to/given/treated with **IV level name** (*M* =) than those exposed to/given/treated with **IV level name** (*M* =), this difference was not significant, report statistic."
  - $t(df)=t_{obt}$, $p>\alpha$
- Then interpret in simple English
  - "This finding suggests…"

Since we have a significant difference between our means, use the first model. A sample interpretation is presented below.

The analysis showed a significant effect of salary method on mean number of TVs sold, *t*(18) = –2.39, *p* = .028. Those employees paid on commission sold significantly more TVs (*M* = 17.10) than those paid on a fixed salary (*M* = 14.10). This suggests that paying employees a commission is more effective at promoting sales than paying them a fixed salary.

12. While the interpretation above provides a good description of the findings, many people find it helpful to view a graph to better understand the data.

    In Chapter 4, you learned how to use Excel to graph means and standard error of the mean using both bar graphs and line graphs. Follow the same steps to graph the data from the *t* test. It may be helpful to refer back to the detailed steps in Chapter 4 as you construct the graph.

13. First, open an Excel workbook. Create a table with the means and standard error of the mean (*SEM*). Next, select the words **"salary"** and **"commission"** and also select their corresponding means (See below.)

	Salary	**Commission**
mean	14.10	17.10
SEM	0.737	1.016

14. Next, choose the correct graph from the insert tab on the menu bar. For these data, we need a bar graph because our independent variable – salary method – is a nominal (categorical) variable.

15. Once the graph appears, the next step is to add the error bars using SEM. Click once on one of the bars to highlight the bars. Next, from the **"Chart Tools"** tab, select the **"Design"** tab, and then click **"Add Chart Element."** Then select **"Error Bars"** followed by **"More Error Bar Options."**

16. A formatting window will appear on the right. At the bottom of the window, select **"Custom,"** and then click **"Specify Value."** A pop up will appear. Highlight the values for the SEMs for both the positive and negative values of the error bars.

See next page

17. Once the error bars appear, work on fixing up the graph. Remove the extra gridlines, add axes and axes labels. You may consider removing the fill from the bars, leaving only the outline. Doing so makes it easier to see the negative error bar (end of bar that extends down below the mean). Below is a sample of a completed graph.

See Next Page for the Related-Samples t test

Data Set 7.2

A local college basketball team is having a tough time making free throws. Regardless of the amount of practice and number of drills, the coach is seeing no improvement in the players' performance. He consults with a Sports Psychologist, who agrees to come in and try to help. She notices that the timing and mechanics of many of the players' technique is a bit off. She decides to do a video intervention in which she breaks the technique down into small steps, and then teaches each step, providing corrective feedback at each step. She works with 10 players on the team. She measures the number of free throws each player makes (out of 30) before she starts working with them, and then again following 5 sessions of her intervention. The data follow.

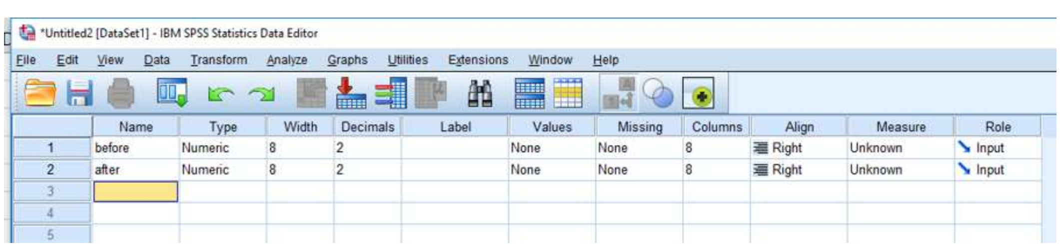
Player	Before	After
PD	8	9
KM	13	15
JV	17	19
ED	8	10
DL	6	15
CP	11	12
AL	10	14
EG	15	22
TF	19	23
SE	15	17

The Related-Samples t test

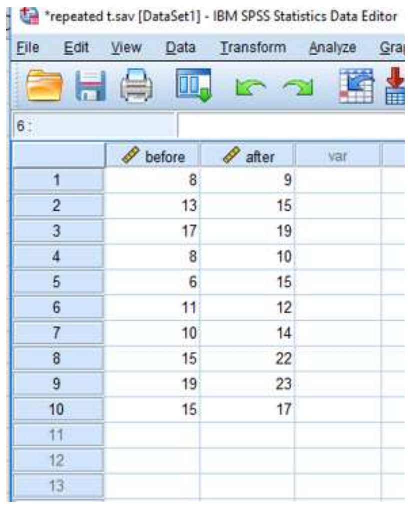
1. Open SPSS, and start with a new data screen.

2. Switch to **"Variable View"** to begin to code your variables. <u>Note</u>: some of the steps for coding the related-samples *t* test are different from those used for the independent *t* test.

3. In the **"Name"** column, enter the two times of measurement – **"before"** and **"after"** the intervention.

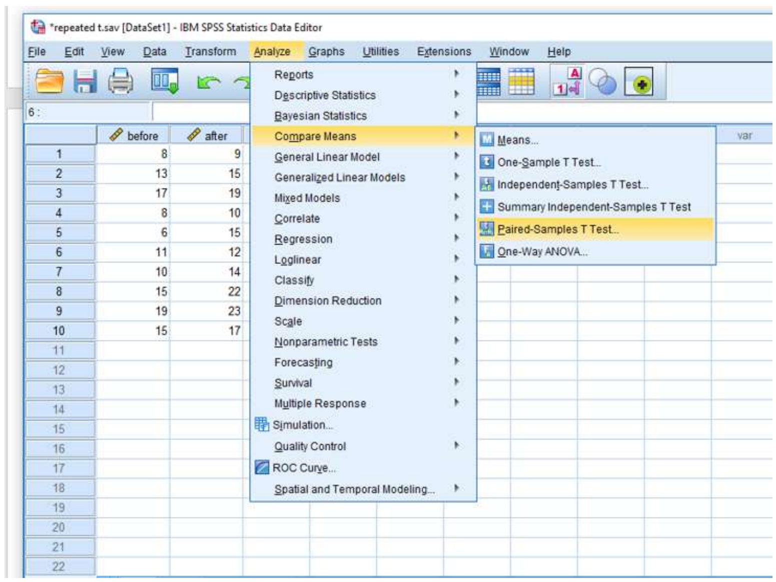


4. Next, switch to the **"Data View"** tab, and enter your data. Enter each player's **"before"** score in the first column, and then the corresponding **"after"** scores in the second column. In doing this, each **"row"** represents a pair of scores for a given player.

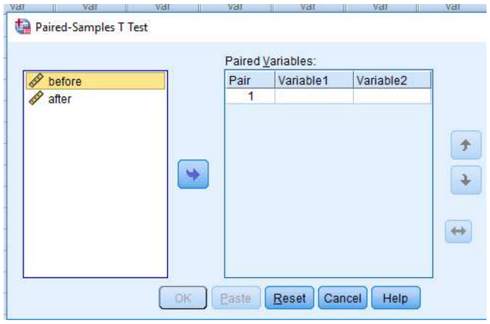


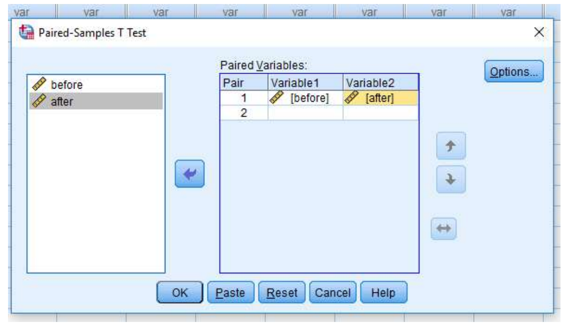
5. Once the data are entered, you can start to complete the steps to analyze the data.

6. From the **"Data View"** menu bar, select **"Compare Means,"** and then **"Paired-Samples T test."**



7. A pop up will appear. Use the arrow button to move **"before"** to the **"Variable 1"** box, and do this again to move **"after"** to the **"Variable 2"** box. Then click **"OK."**





8. The results of the analysis will appear in the SPSS output window. (See below.) The **"Paired Samples Statistics"** table provides our means, standard deviations, and standard errors. The next table provides the value of the underlying correlation in the score. Think about it; these are pairs of scores, and we might expect them to be correlated. The *p* value is **less than .05**, so we know there is a significant correlation present. The bottom tables, the **"Paired Samples Test"** tables, give us the difference between our means, as well as our *t* statistic. Our obtained *t* statistic is –4.019, the degrees of freedom are 9, and our *p* value is .003. Since the p value is **less than .05**, our *t* test **is significant**.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	12.20	10	4.290	1.356
	after	15.60	10	4.719	1.492

Means and *SEM*

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	before & after	10	.828	.003

Correlation and *p* value

Paired Samples Test

		Paired Differences				
					95% Confidence Interval of the Difference	
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper
Pair 1	before - after	-3.400	2.675	.846	-5.314	-1.486

Paired Samples Test

		t	df	Sig. (2-tailed)
Pair 1	before - after	-4.019	9	.003

The *t* statistic, degrees of freedom and *p* value for the *t* test

9. Once we have all of these values, we can interpret our results. Since the related-samples *t* tests compares pairs of scores, and not scores from different groups (as in the Independent *t* test), the write up of the interpretation is a bit different.

   If your result is **significant**, use the following guide:

   - "The analysis showed significant effect of IV name on DV name, (report statistic)"
     - $t(df)=t_{obt}$, $p<\alpha$ (or use *p* = and your "*p*" value)
       - "Mean measure name was significantly… pick an adjective – higher, greater, lower, etc. when participants were exposed to/given/treated with **IV level name** (*M* =)than when they were exposed to/given/treated with **IV level name** (*M* =)"
   - Then interpret in simple English
     - "This finding suggests…"

- **If not significant**
  - "Even though mean measure (DV) name was… … pick an adjective – higher, greater, lower, etc. when participants were exposed to/given/treated with **IV level name** (*M* =)than when they were exposed to/given/treated with **IV level name** (*M* =), this difference was not significant, report statistic."
    - $t(df)=t_{obt}$, $p>\alpha$
  - Then interpret in simple English
    - "This finding suggests…"

Since we have a significant difference between our means, use the first model. A sample interpretation is presented below.

The analysis showed a significant effect of the intervention on mean number of free throws made, $t(9) = -4.02$, $p = .003$. Mean number of successful free throws was significantly higher after the intervention ($M = 15.50$) than before training ($M = 12.20$). These findings suggest that the video and feedback intervention improved the free-throw completion rate of the team members.

See next page

10. Now graph the data in Excel. Try to do this without looking back at the graphing instructions given for the results of the independent *t* test. Trying to do this on your own will help you become more fluent in completing the steps needed to construct graphs. The data needed are organized in the table below. You will need a bar graph here because the independent variable in this scenario (time of intervention – before vs. after) is categorical (nominal).

	Before	After
mean	12.20	15.60
SEM	1.356	1.492

Data Collection Exercise

Your instructor will now conduct a simple experiment to collect data from the class. Once the data are gathered and put on screen, perform the correct *t* test to analyze the data, write an interpretation, and create a graph.

<u>Steps</u>

1. Decide if your data are from independent groups or related samples.
2. Conduct your *t* test.
3. Write an interpretation.
4. Graph the results.