SPSS DEMONSTRATIONS [GSS18SSDS-A]

Demonstration 1: Producing a One-Sample *t* Test

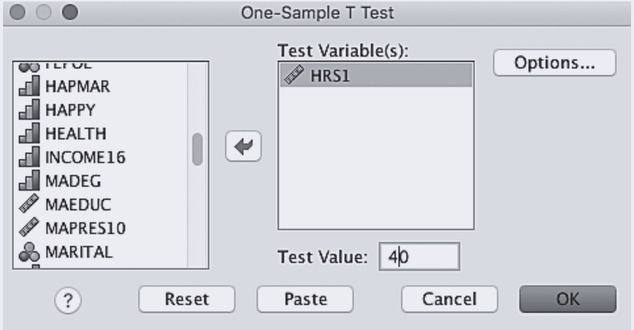
In this chapter, we discussed methods of testing differences in means between a sample and a population value. SPSS includes a One-Sample T Test procedure to do this test. SPSS does not compute the test with the *Z* statistic; instead, it uses the *t* statistic to test for all mean differences. The One-Sample T Test procedure can be found under the *Analyze* menu choice, then under *Compare Means*, where it is labeled *One-Sample T Test*. The opening dialog box (Figure 7.6) requires that you place at least one variable in the Test Variable(s) box. Then a test value must be specified.

We'll use the GSS18SSDS-A data set for this demonstration. The standard workweek is thought to be 40 hours, so let's test to see whether American adults work that many hours each week. In this example, place HRS1 in the Test Variable(s) box and "40" in the Test Value box. Then click on *OK* to run the procedure.

The output from the One-Sample T Test procedure is not very extensive (see Figure 7.7). A total of 875 people answered the question about number of hours worked per week. The mean number of hours worked is 41.32, with a standard deviation of 15.277. Below this, SPSS lists the test value, 40. It includes the two-tailed significance, or probability, for the one-sample test. This value is .011, given the calculated *t* statistic of 2.554, with 874 degrees of freedom. Thus, at the .01 significance level, we would reject the null hypothesis and conclude that American adults work more than 40 hours per week.

SPSS also supplies a 95% confidence interval for the mean difference between the test value and the sample mean. Here, the confidence interval runs from 0.31 to 2.33, providing estimates of how much more than 40 hours per week Americans work.

Figure 7.6 One-Sample T Test Dialog Box



Demonstration 2: Producing a Test of Mean Differences

In this chapter, we have also discussed methods of testing differences in means or proportions between two samples (or groups). The Two-Sample T Test procedure

Figure 7.7 One-Sample T Test Output

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Number of hours worked last week	875	41.32	15.277	.516

One-Sample Test

	Test Value = 40					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Number of hours worked last week	2.554	874	.011	1.319	.31	2.33

can be found under the *Analyze* menu choice, then under *Compare Means*, where it is labeled *Independent-Samples T Test*.

The opening dialog box requires that you specify various test variables (the dependent variable) and one independent or grouping variable (Figure 7.8). We'll test the null hypothesis that men and women work the same number of hours each week by using the variable HRS1. Place that variable in the Test Variable(s) box and SEX in the Grouping Variable box. When you do so, question marks appear next to SEX indicating that you must supply two values to define the two groups (independent samples). Click on *Define Groups*. Then put "1" in the first box and "2" in the second box (1 = male and 2 = female), as shown in Figure 7.9. Then click on *Continue* and *OK* to run the procedure.

The output from Independent-Samples T Test (Figure 7.10) is detailed and contains more information than we have reviewed in this chapter. The first part

Figure 7.8 Independent-Samples T Test Dialog Box

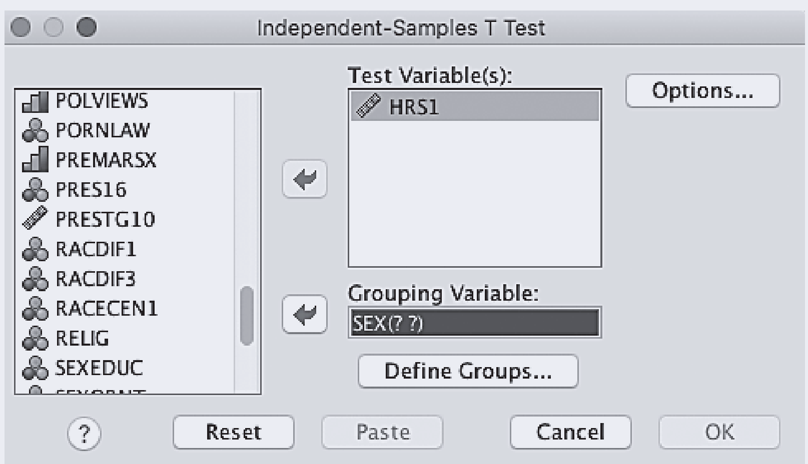


Figure 7.9 Define Groups Dialog Box

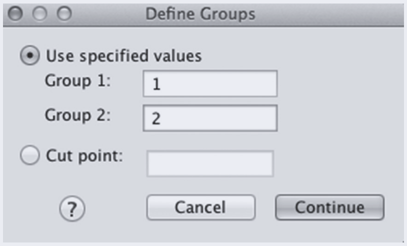


Figure 7.10 Independent-Samples T Test Output

Group Statistics

	Respondents sex	N	Mean	Std. Deviation	Std. Error Mean
Number of hours worked last week	Male	435	45.02	15.754	.755
	Female	440	37.66	13.870	.661

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
Number of hours worked last week	Equal variances assumed	5.764	.017	7.329	873	.000	7.352	1.003	5.384	9.321
	Equal variances not assumed			7.324	856.670	.000	7.352	1.004	5.382	9.323

of the output displays the mean number of hours worked for males and females, the number of respondents in each group, the standard deviation, and the standard error of the mean. We see that males worked 7.36 hours more per week than females (45.02 – 37.66 = 7.36).

Earlier in the chapter, we reviewed Levene's test and how to determine whether the variances of the two groups are equal. In this case, we reject the null hypothesis of equal variances (the significance of *F* is .017 < .05 alpha). The *t* obtained is 7.324 (equal variances not assumed) with a probability of .000 (smaller than .05 or .01). We can reject the null hypothesis of no difference and conclude that men work significantly more hours per week than women. The difference of 7.36 hours is significant at the .000 level.

What if we wanted to do a one-tailed test instead? SPSS does not directly list the probability for a one-tailed test, but it is easy to calculate. If we had specified a directional research hypothesis—such as that men work more hours than women—we would simply take the probability reported by SPSS and divide it into half for a one-tailed test. Because the probability is so large in this case, our conclusion will be the same whether we do a one- or a two-tailed test.

The last portion of output on each line is the 95% confidence interval for the mean difference in hours worked between the two groups. (Confidence intervals were reviewed in Chapter 7.) It is helpful information when testing mean differences because the actual mean difference will vary from sample to sample. The 95% confidence interval gives us a range over which the sample mean differences are likely to vary.

SPSS PROBLEMS [GSS18SSDS-A]

S1. Use the GSS file to investigate whether or not Americans use the Internet at least 7 hours per week (estimating an hour per day). Perform the One Sample T Test procedure (as presented in SPSS Demonstration 1) to do this test with the variable WWWHR. Do the test at the .01 significance level. What did you find? Do Americans use the Internet 7 hours per week, more or less?

S2. The GSS includes a measure of highest educational degree completed (DEGREE). Test whether there is a significant difference between those with less than high school (coded 0) and those with a bachelor's degree (coded 3) in the number of hours on the Internet per week (WWWHR). Assume α is .05 for a two-tailed test. Summarize your findings.

S3. Use the variable PRES16 as your independent or grouping variable (1 = *Clinton* and
2 = *Trump*). Investigate whether there is a significant difference between these two groups in terms of their age (AGE), education (EDUC), and number of children (CHILDS). Assume that α is .05 for a two-tailed test. Based on your analysis, write three Step 5–type statements summarizing your findings.

S4. For this analysis, use the variable BIBLE as your independent variable, comparing individuals who believe the Bible is the word of God (1) or a book of fables (3). Use the same dependent variables, AGE, EDUC, and CHILDS, to estimate *t* tests. Assume α is .05 for a two-tailed test. Prepare a statement to summarize your findings.

EXCEL DEMONSTRATIONS [GSS18SSDS-E]

Demonstration 1: A One-Sample T Test

In this demonstration, we will use Excel to conduct a One-Sample T Test of a respondent's ideal number of children (CHLDIDEL) to see if American adults feel the ideal number of children is greater than two children. We will be doing a one-tailed *t* test. Copy CHLDIDEL data from the protected Data View sheet and paste it into a new Excel sheet (see Figure 7.11).

Although Excel's *Data Analysis* function does not have an option to conduct a One-Sample T Test, we can easily work around this shortcoming with a few extra, but easy, steps. The process begins by creating a fake variable, otherwise known in statistics as a dummy variable, next to CHLDIDEL. We will label this fake variable "DUMMY"—which you can see in Figure 7.12. Next, we will enter a "0" in cells B2 and B3. You don't need to enter a "0" in any other cells of our dummy variable (see Figure 7.12).

Now we can use Excel's *Data Analysis* function. Navigate to Excel's Data tab and select *Data Analysis*. A window of Analysis Tools will appear. Select *t-Test: Two-Sample Assuming Unequal Variances* and then click *OK*.

Figure 7.11

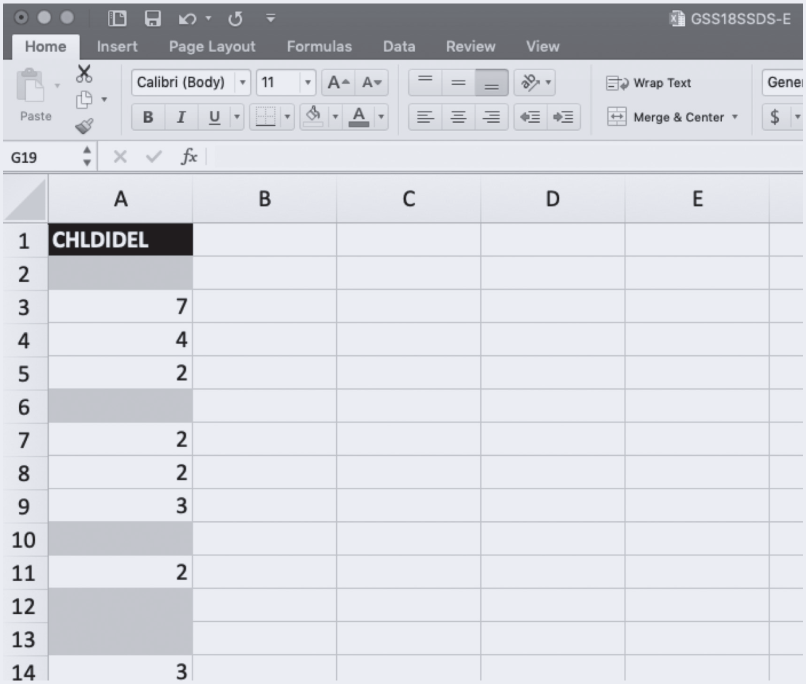
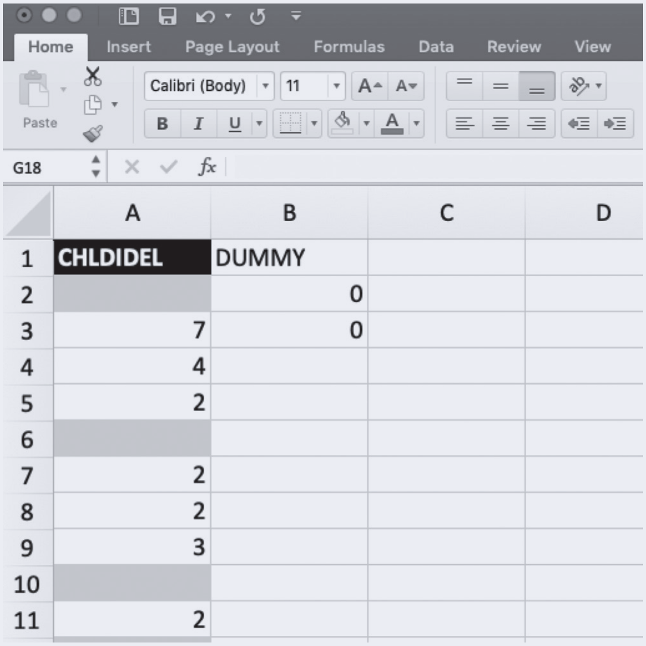


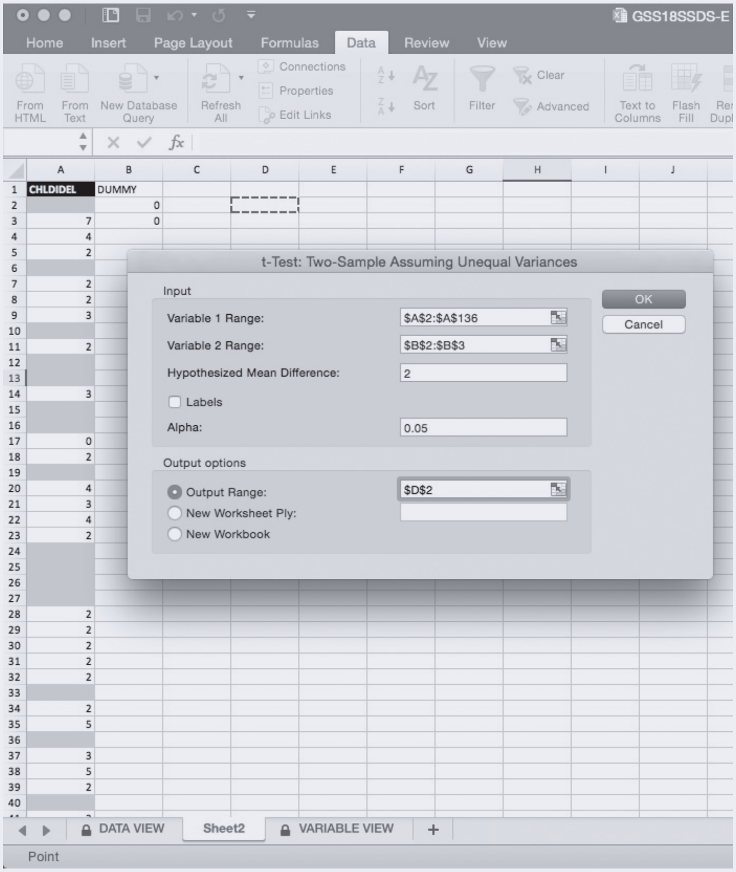
Figure 7.12



Click in the empty box next to "Variable 1 Range" and highlight the column of CHLDIDEL data from A2 to A136. Do not select A1 for it contains the variable name, CHLDIDEL. Click in the empty box next to "Variable 2 Range" and highlight the column of DUMMY data from B2 to B3. Do not select B1 for it contains the variable name, DUMMY. In the empty box next to "Hypothesized Mean Difference," enter "2." This will be our test value and allow us to test if our respondents report a different amount of ideal number of children. If there is not a significant difference in the ideal number of children our respondents report and our test value of "2," our *p* value will be greater than our alpha. In our example, Excel has automatically set our alpha as ".05." Click in the empty box next to "Output Range," and then select any cell in the current sheet you are working in. This will tell Excel where to place the *t*-test data analysis table it will generate. In our example, we've chosen for the output table to begin in cell D2. Click *OK* (see Figure 7.13).

We can delete the column labeled "Variable 2," for that is our DUMMY variable. You can do that by highlighting the column and clicking on *Edit* and then

Figure 7.13

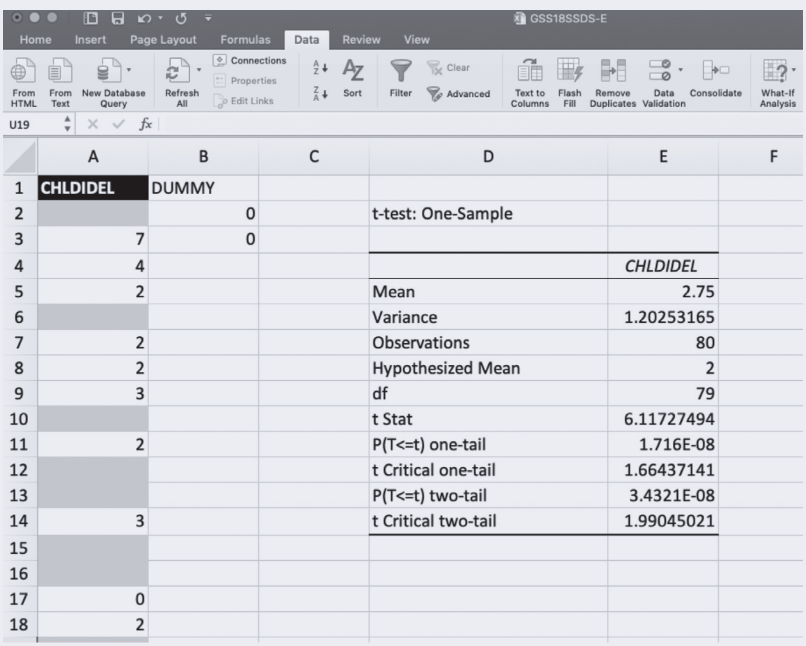


Clear in the mainExcel toolbar. For organizational purposes, we will widen column D. In cell E4, we will replace "Variable 1" with "CHLDIDEL." In cell D8, we will change "Hypothesized Mean Difference" to "Hypothesized Mean." And, finally, in the title for the table, replace "t-Test: Two-Sample Assuming Unequal Variances" with "t-Test: One-Sample." We've successfully tricked Excel into conducting a One-Sample T Test (see Figure 7.14).

A total of 80 people (listed in our table as "observations") answered the question about ideal number of children. The mean CHLDIDEL is 2.75, with a variance of 1.20 (rounded). To obtain the standard deviation, all you need to do is take the square root of the variance. The standard deviation of CHLDIDEL is 1.10 (rounded). Next to "Hypothesized Mean," we see our test value (2), followed by our degrees of freedom (79). Our *t* value is listed next to "t Stat" as 6.12 (rounded). Excel then gives us the *p* value for a one-tail *t* test followed by the critical *t* value for a one-tail *t* test. It then does the same for a two-tail *t* test. Note the *p* value for the one-tail *t* test (cell E11) and two-tail *t* test (cell E13) is noted in exponential notation. 1.72E-08 is the same as .0000000172. Furthermore, 3.43E-08 is the same as .0000000343. Both *p* values are essentially .000.

We would interpret our findings as follows: Our *t* statistic is 6.12 (rounded), with 79 degrees of freedom. At the .05 significance level, we would reject the null hypothesis and conclude that American adults feel the ideal number of children is greater than two because our *p* value (.000) for a one-tailed *t* test is less than our alpha (.05).

Figure 7.14



	A	B	C	D	E	F
1	CHLDIDEL	DUMMY				
2		0		t-test: One-Sample		
3	7	0				
4	4				*CHLDIDEL*	
5	2			Mean	2.75	
6				Variance	1.20253165	
7	2			Observations	80	
8	2			Hypothesized Mean	2	
9	3			df	79	
10				t Stat	6.11727494	
11	2			P(T<=t) one-tail	1.716E-08	
12				t Critical one-tail	1.66437141	
13				P(T<=t) two-tail	3.4321E-08	
14	3			t Critical two-tail	1.99045021	
15						
16						
17	0					
18	2					

Demonstration 2: A Test of Mean Differences

We will now conduct a one-tailed *t* test of mean differences of CHLDIDEL by respondent's sex (SEX) to see if females report a significantly lower ideal number of children than males. Copy both CHLDIDEL and SEX data from the protected Data View sheet and paste it into a new Excel sheet. Sort the SEX data by selecting all of the data in the CHLDIDEL and SEX columns (include the first row, which is the variable label for each column). Navigate to Excel's Data tab and select *Sort*. A Sort window will appear. Under "Column," we will instruct Excel to sort our data by SEX. Under "Sort On," we will instruct Excel to sort by "Values." And, under "Order," we will instruct Excel to sort from A to Z. Select *OK*. The data will be sorted by SEX, which will make it easier for us to conduct a test of mean differences.

Navigate to Excel's Data tab and select *Data Analysis*. A window of Analysis Tools will appear. Select *t-Test: Two-Sample Assuming Equal Variances* and then *OK*.

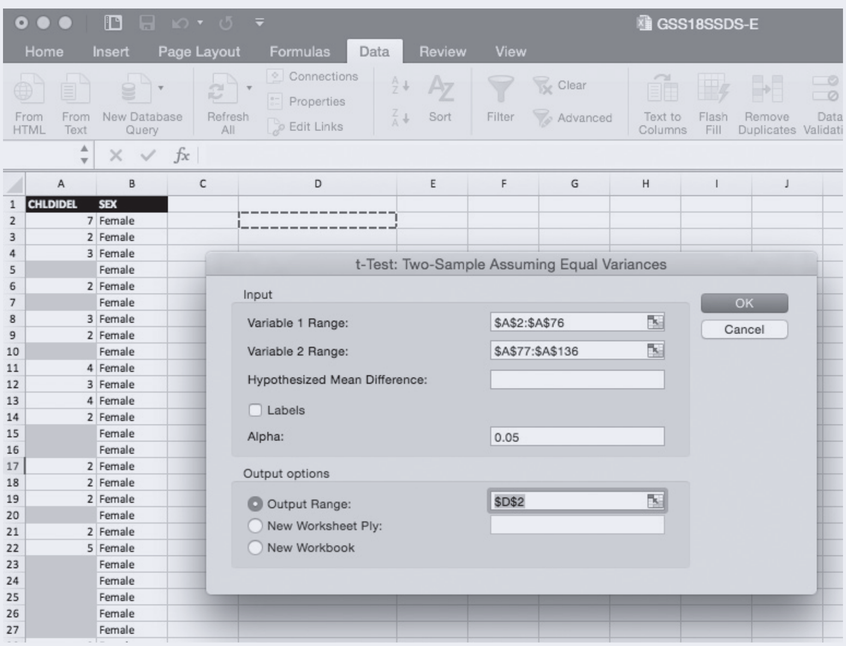
Click in the empty box next to "Variable 1 Range" and highlight the column of CHLDIDEL data from A2 to A76. This represents females. Do not select A1 for it contains the variable name, CHLDIDEL.

Click in the empty box next to "Variable 2 Range" and highlight the column of CHLDIDEL data from A77 to A136. This represents males.

In our example, Excel has automatically set our alpha as ".05." We will leave it as such.

Click in the empty box next to "Output Range," and then select any cell in the current sheet you are working in. This will tell Excel where to place the *t*-test data analysis table it will generate. In our example, we've chosen for the output table to begin in cell D2. Click *OK* (see Figure 7.15).

Figure 7.15



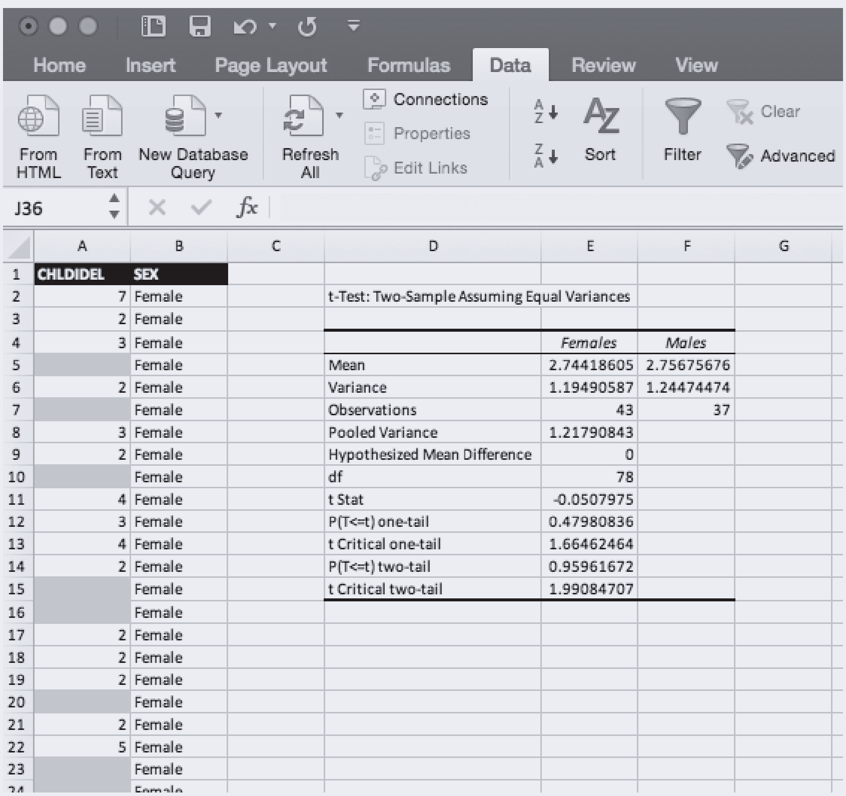
For organizational purposes, we will widen column D. In cell E4, we will replace "Variable 1" with "Females." In cell F4, we will replace "Variable 2" with "Males."

A total of 43 females and 37 males (listed in our table as "observations") answered the question about ideal number of children. The mean CHLDIDEL for females is 2.74, with a variance of 1.19. To obtain the standard deviation, all you need to do is take the square root of the variance. The standard deviation of CHLDIDEL for females is 1.09. The mean CHLDIDEL for males is 2.76, with a variance of 1.24. To obtain the standard deviation, all you need to do is take the square root of the variance. The standard deviation of CHLDIDEL for males is 1.11. The degrees of freedom is 78 (43 + 37 – 2).

Our *t* value is listed next to "t Stat" as –0.05. Excel then gives us the *p* value for a one-tail *t* test followed by the critical *t* value for a one-tail *t* test. It then does the same for a two-tail *t* test.

We would interpret our findings as follows: Our *t* statistic is –0.05 with 78 degrees of freedom. At the .05 significance level, we would fail to reject the null hypothesis that females have a lower ideal number of children than males because our *p* value (.48) for a one-tailed *t* test is greater than our alpha (.05).

Figure 7.16



EXCEL PROBLEMS [GSS18SSDS-E]

E1. Investigate whether or not Americans have 12 years of education.

a. Perform a one-sample *t* test with the variable EDUC (highest year of school completed). See Excel Demonstration 1 as an example.

b. What is the mean value of EDUC?

c. What is the obtained *t* value?

d. What is the *p* value for a two-tailed *t* test?

e. Are you able to reject the null hypothesis that Americans have 12 years of education? Set alpha at .05.

E2. Compare the mean years of education (EDUC) by sex. Do females complete more years of education than males?

a. Test the mean difference of EDUC (highest year of school completed) by SEX (respondent's sex). See Excel Demonstration 2 as an example.

b. What is the mean value of EDUC for females? For males?

c. What is the obtained *t* value?

d. What is the *p* value for a one-tailed *t* test?

e. Are you able to reject the null hypothesis that females and males complete a similar number of years of education? Set alpha at .05.