

Calculating Statistics using Excel

This document was put together following the third edition of *Quantitative Research Methods for Communication: A Hands-On Approach* by Wrench, Thomas-Maddox, Richmond, and McCroskey. In the 2nd and 3rd editions, we discussed calculating statistics using Microsoft Excel. Starting in the 4th edition, we've moved more information to the online website (<https://oup-arc.com/wrench>) including numerous walk-through videos for running the statistics discussed in our book. We wanted to create this special addendum to include the step-by-step walk throughs of using Excel on its own to calculate statistics (as was seen in the 2nd and 3rd editions). Starting in the 4th edition, we've recommended using Statistician for Excel (<http://statsaddin.com/>) because it makes calculating statistics in Excel easier. However, as this document demonstrates, it's still very possible to calculate a wide range of statistics using Excel on its own.

EXCEL AND CRONBACH'S ALPHA

Unfortunately, there is not an easy way to calculate Cronbach's alpha in Excel. As such, you must use a variety of formulas for the calculations of alpha. Thankfully, a psychometrician at the University of Leicester in the United Kingdom named Professor Glenn Fulcher has created an easy-to-use Excel template to conduct an alpha reliability (<http://languagetesting.info/statistics/excel.html/>). When you download the alpha reliability file, it will open with all of the calculations in place, but there will be no actual data within the file yet. The Excel file is built for up to 200 scale items and 3,005 individual cases (Figure Excel.1). Note that there are two tabs to this Excel file. The one currently selected in Figure Excel.1 is where the data will be pasted. The arrow in Figure Excel.1 is pointing to the tab where you can see the alpha reliability after you have pasted the data.

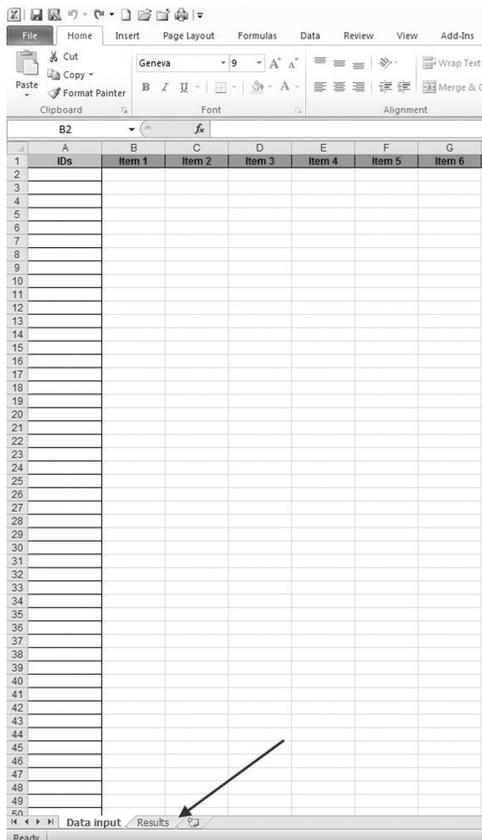


Figure Excel.1
Fulcher's Excel File

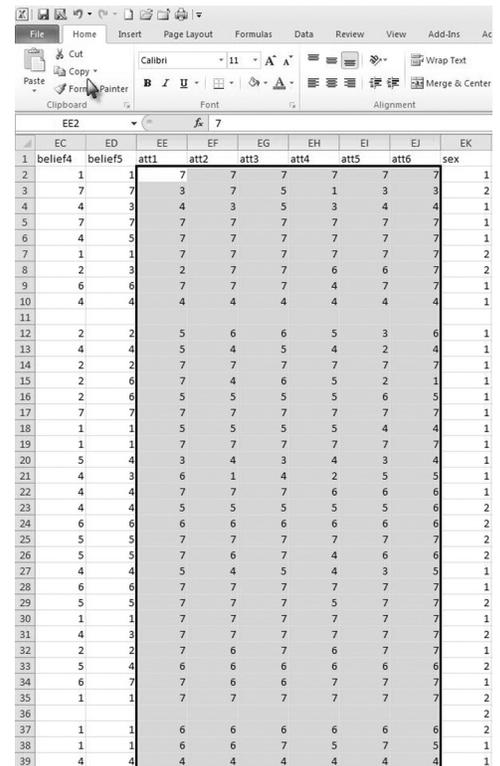


Figure Excel.2
Select Data in Recoded Dataset

To paste the data, find the data you want to include in the alpha reliability test in the Recoded Dataset (Figure Excel.2). Note that you do not select the variable names at the top of each column. Then go to copy function and click “copy” (note the pointer in Figure Excel.2). Once you have copied the data, you can then go to the Alpha Reliability Excel File and paste that

copied data into the file. Select the square located at B2 and paste the data (Figure Excel.3). Once you have pasted the data, just click on the “Results” worksheet at the bottom of the screen and you can view the alpha reliability calculated (Figure Excel.4). Please note that it is using a slightly different formula for calculating alpha reliability than the one used by SPSS, which has resulted in a slightly elevated alpha reliability. To learn more about the specific calculations utilized in this alpha reliability test, check out Fulcher’s discussion on conducting alpha reliability in Fulcher and Davidson (2007) or Fulcher (2010).

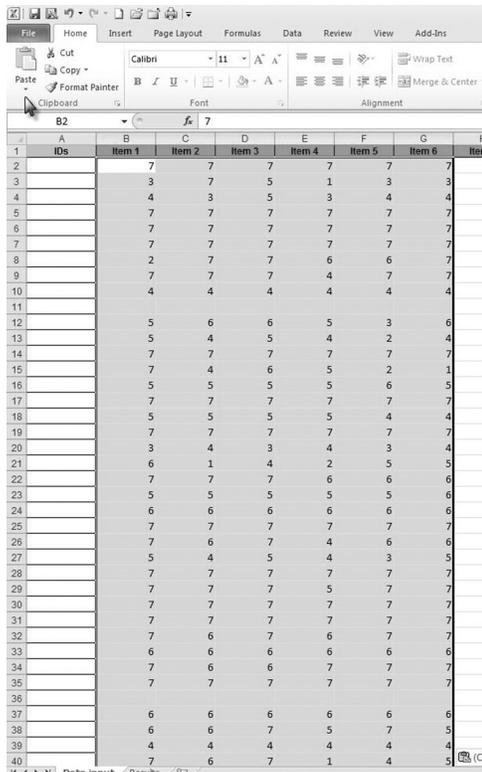


Figure Excel.3
Pasted Scale Items into Excel

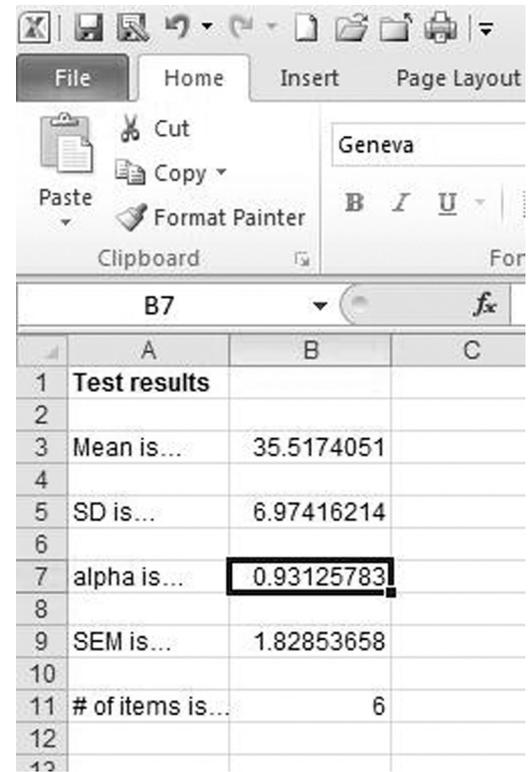


Figure 8.4
Excel Alpha Reliability Results

EXCEL AND FREQUENCY DISTRIBUTIONS

At this point, we will examine how to calculate a frequency chart using Excel as well. Given that Microsoft Excel is readily available to most students, we will include information on using this program to calculate statistics. Although Excel may be a convenient program for analysis, keep in mind that the primary purpose for the program is not statistical analysis. Thus, the process for computing even basic statistics often involves more steps than in other software programs, and the options for post hoc analyses are often more limited than in programs such as SPSS.

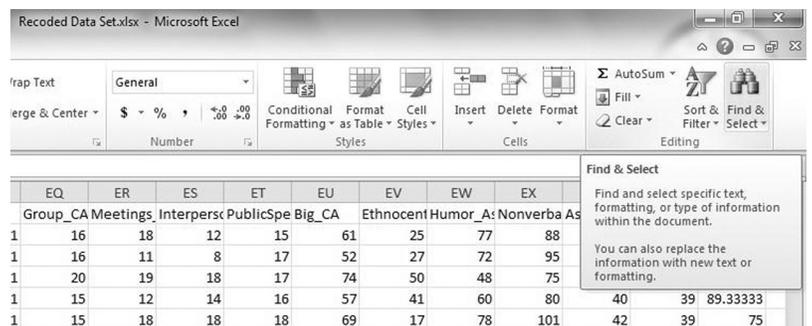


Figure Excel.5
Find and Select

Before you begin using Excel, consider the following caveats. First, missing data must be entered as a blank cell in Excel. The program is unable to read any other format for missing data. If you are importing a database that automatically fills an empty cell with data, utilize the “Find and Select” feature to convert all missing data (see Figure Excel.5). The imported “Recorded Dataset” file may include several cells with “#NULL!” to represent missing

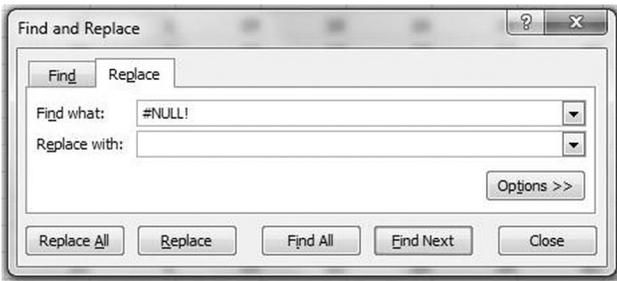


Figure Excel.6
Replacing #Null!

values. To convert the data, click on the “Replace” tab at the top of the screen, type the value that appears in the cells for missing data in the “Find What” box, and leave the “Replace with” box blank. Click on “Replace All” to convert all missing data to empty cells (Figure Excel.6).

Second, the steps for calculating statistics may differ depending on the version of Microsoft Excel that you are using (e.g., Excel 2007 vs. Excel 2010). In addition, the data analysis process for statistical computation varies depending on

whether you are using a Mac versus a PC for data analysis. Finally, Excel does not come preinstalled with the data analysis tools that are required to complete statistical analyses. However, this is easily remedied with a free download. To determine whether your computer has the data analysis tool kit installed, click on “Help” and type “Data Analysis.”

If you are using a PC, the Data Analysis ToolPak is available from the Microsoft website (<http://www.office.microsoft.com/>). The Mac platform will require you to download a program called “StatPlus” (<http://www.analystsoft.com/>). In this text, we will describe the procedures for calculating statistics using the PC version of Excel 2010.

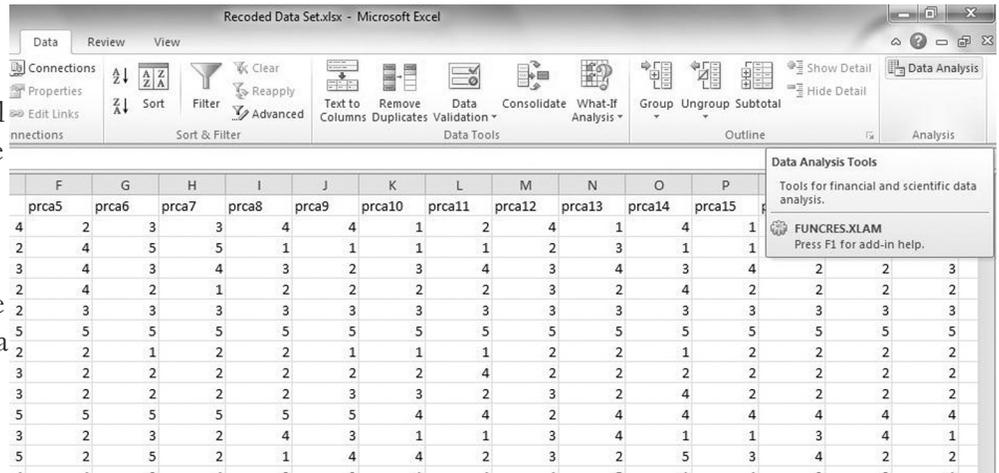


Figure Excel.7
Excel Data Menu

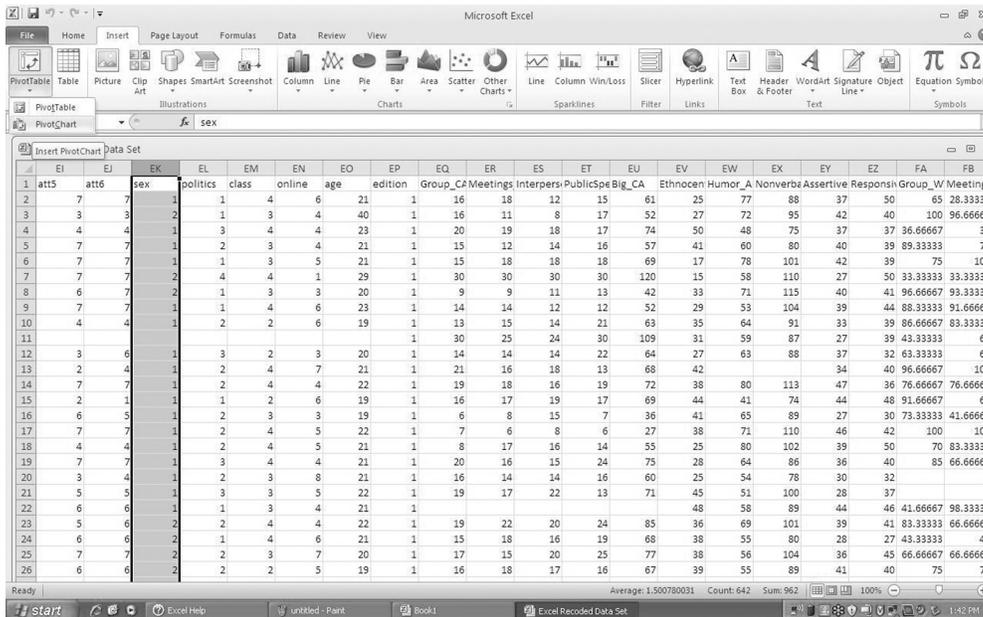


Figure Excel.8
Pivot Table

view your data and output, it is recommended that you select and work with one column or variable at a time. To compute the frequency counts for each value, Excel provides the PivotTable and PivotChart options. Begin by highlighting the “Sex” column in the spreadsheet and clicking on the “Insert” tab at the top of the screen. Click

Once you have downloaded the data analysis package to Excel, you will see a tab on the top of the spreadsheet screen titled “Data.” Clicking will provide access to an area labeled “Analysis.” This area will enable you to calculate a variety of basic statistics in Excel (Figure Excel.7).

To calculate frequencies in Excel, open the data file titled “Recorded Dataset” from the textbook site. Remember—if the dataset contains any missing data, you will need to utilize the “Find and Select” function to convert any data that are used to represent missing values into blank cells. To be able to easily

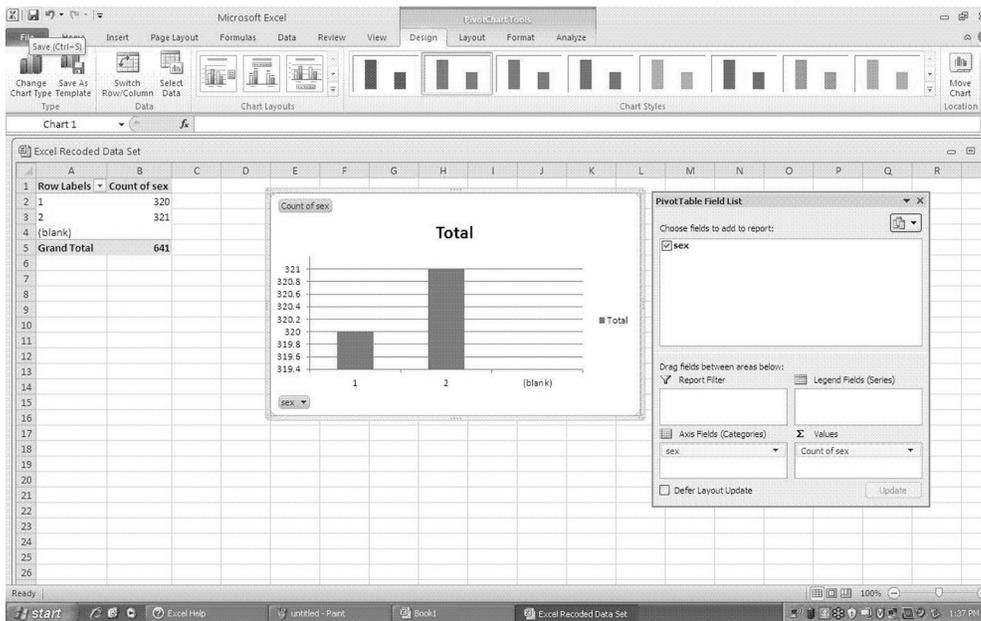


Figure Excel.9
Frequency Distribution in Excel

on “PivotTable” at the left side of the screen, and the range of values that is highlighted should appear in the box. Check the box for “New Worksheet” and select “OK” (Figure Excel.8). A new sheet will open, and you will be asked to indicate the fields that you wish to add to your report. Drag and drop “Age” from the list into both the “Row Labels” and the “Values” boxes. A frequency table with counts for each age reported will appear on the screen. If you would like to add a histogram, select the “PivotChart” option and follow the same process by dragging “Sex” into the “Axis Field” and

“Values” boxes. A frequency table and histogram will appear on the new sheet (Figure Excel.9).

Excel and Descriptive Statistics

To calculate the descriptive statistics for the CA scores for participant, open the “Recoded Dataset” file. Select the “Big CA” field and highlight all values in the column. Click on the “Data” tab at the top of the sheet, select “Data Analysis” from the upper right side of the screen, and choose “Descriptive Statistics” from the menu of options. (Figure Excel.10) The range of highlighted values should appear in the “Input Range” box. Select “New Worksheet Ply” to create a new sheet for the output. Select the “Summary Statistics” option and indicate the “Confidence Level for the Mean” in the box provided (Figure Excel.11). Click “OK” and the output will appear on a new sheet (Figure Excel.12).

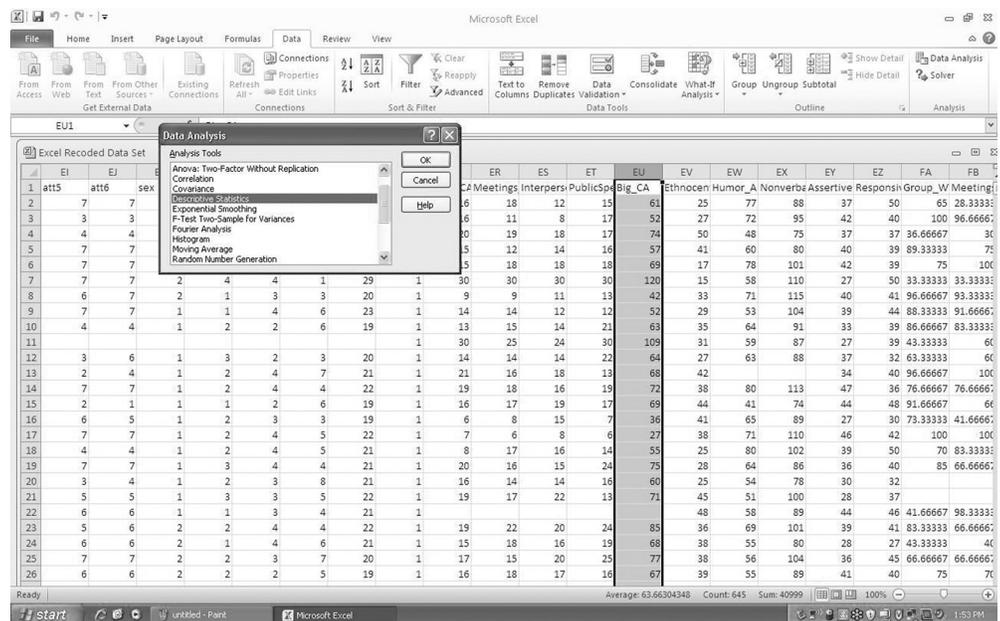


Figure Excel.10
Excel’s Descriptive Statistics Dialogue Box

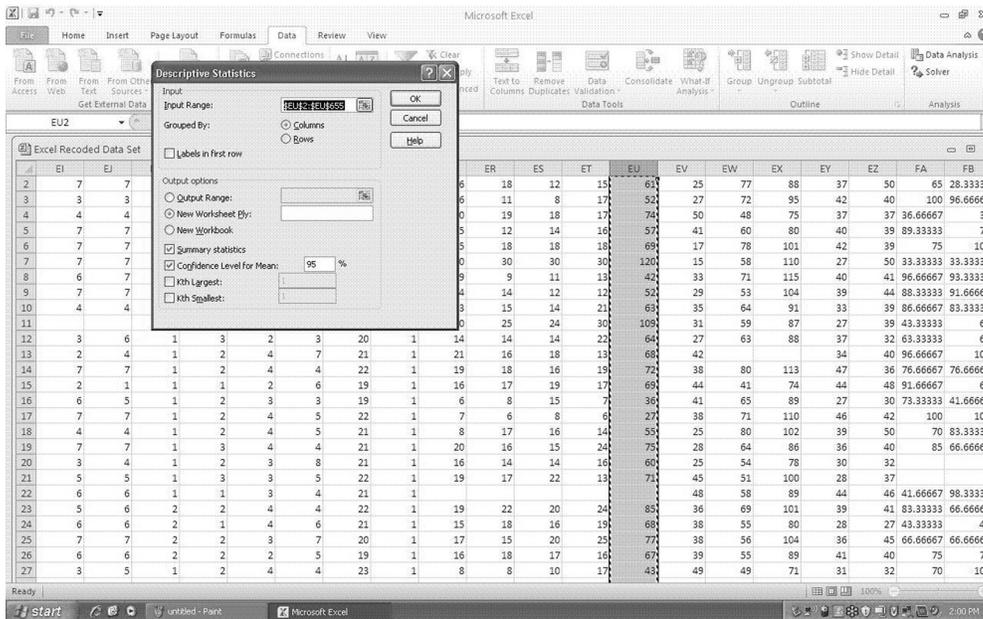


Figure Excel.11
Means in Excel

	A	B	C
1	Column1		
2			
3	Mean	63.66304348	
4	Standard Error	0.67005403	
5	Median	64	
6	Mode	67	
7	Standard Deviation	17.00406504	
8	Sample Variance	289.1382277	
9	Kurtosis	0.334931041	
10	Skewness	0.203134612	
11	Range	96	
12	Minimum	24	
13	Maximum	120	
14	Sum	40999	
15	Count	644	
16	Confidence Level(95.0%)	1.315758431	
17			
18			

Figure Excel.12
Descriptive Statistics Calculated by Excel

EXCEL AND CHI-SQUARES

To calculate the chi-square using Excel, select the Excel file titled “Chi-square” from the textbook’s website. Click the “Insert” file tab at the top of the screen and choose the “Pivot-Table” function (Figure Excel.13).

The data for both the “sex” and the “class” variables should be highlighted. The range for the data will appear in the “Table/Range” box. Select “New Worksheet” from the menu to create a new page for the output (Figure Excel.14). Drop and drag “sex” to the section titled “Row Labels” and drag “class” to the box titled “Column Labels.”

The sum of scores for “sex” should automatically appear in the “Values” section (Figure Excel.15). Click on the “X” to close out the dialogue box. The output table includes the total number of responses for each

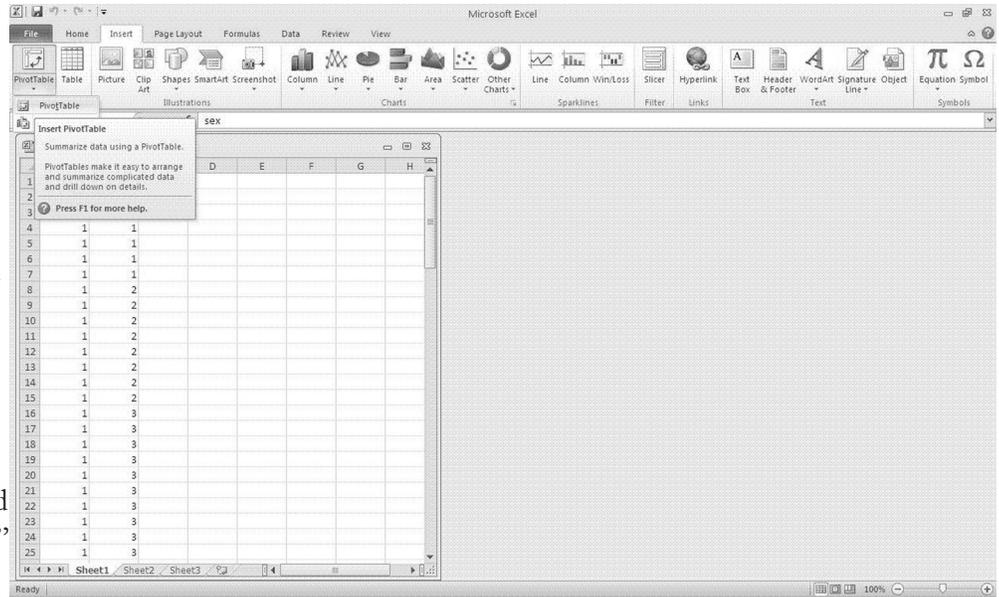


Figure Excel.13
Pivot Table

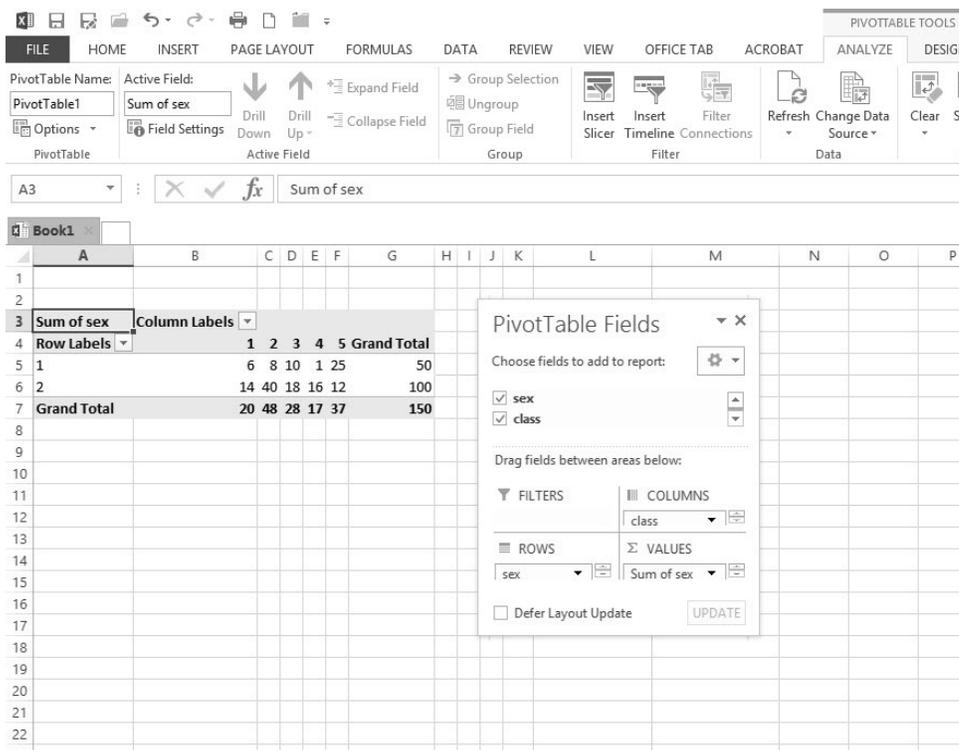


Figure Excel.14
New Pivot Table

category we identified (e.g., males in the freshman class). Be careful, Excel often does a great job of bringing in the counts for the first row but not for the second row. As such, you may find yourself with a percentage value and not a count value. Look at Figure Excel.16. You will note that the first number in the second row is 14 when it should have been

7. To change this, right click on the first number and scroll down to “Summarize Values By.” The first one, which will be checked, is Sum. You want to uncheck Sum and check Count (Figure Excel.17). This will correct that row to counts instead of summed totals.

Computing the chi-square statistic involves a couple of steps. First, copy all of the information from the PivotTable. Click on an

empty cell below the table, right click, choose “Paste Special,” and choose the “Paste Values” option. This will duplicate the table directly below the original PivotTable (Figure Excel.18). Doing so will enable you to keep the data separate while you compute the expected values versus the observed values. In the duplicate table of values that you created, highlight and delete the values for each of the cells. Be careful not to delete any of the totals from

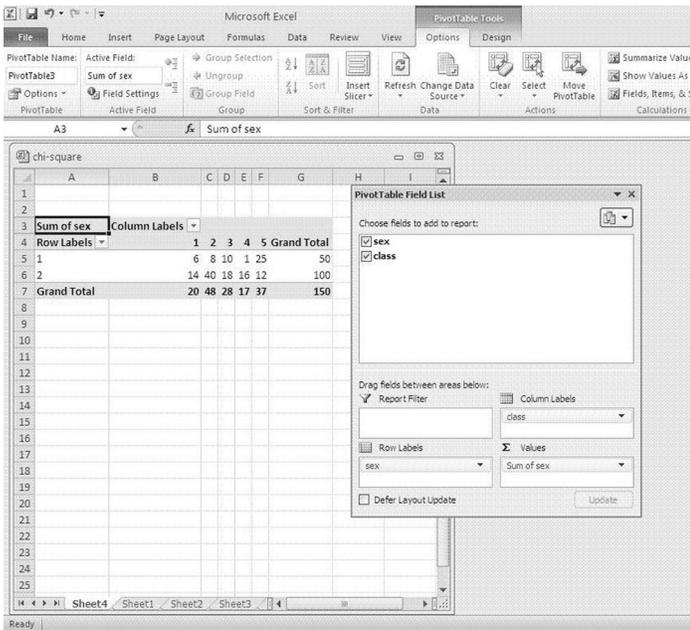


Figure Excel.15
Pivot Table Field List Dialogue Box

the table.

To compute the expected values into the table, click on the first empty cell and type the formula “=column total value * row total value/grand total value” (Figure Excel.19). Repeat these steps for each empty cell in the table until you have computed the expected values for each cell (Figure Excel.20). Finally, select an empty cell below the two tables and click. In the cell, type the Excel formula for a chi-square analysis, which is =CHITEST (actual values, expected values). To enter the range of values, simply highlight the block of expected values from the top table and click Enter (Figure Excel.21). Complete the same steps for the expected values in the table below. Click “Enter,” and the probability for the chi-square value will appear in the designated cell (Figure Excel.2).

Please note that the actual chi-square statistic is not actually calculated by Microsoft Excel. Instead, Excel calculates your *p* value, so you would still need to calculate the chi-square statistic by hand. Now that we have examined the chi-square test itself, let’s see how a

researcher would write up this finding using APA style in a results section of an article.

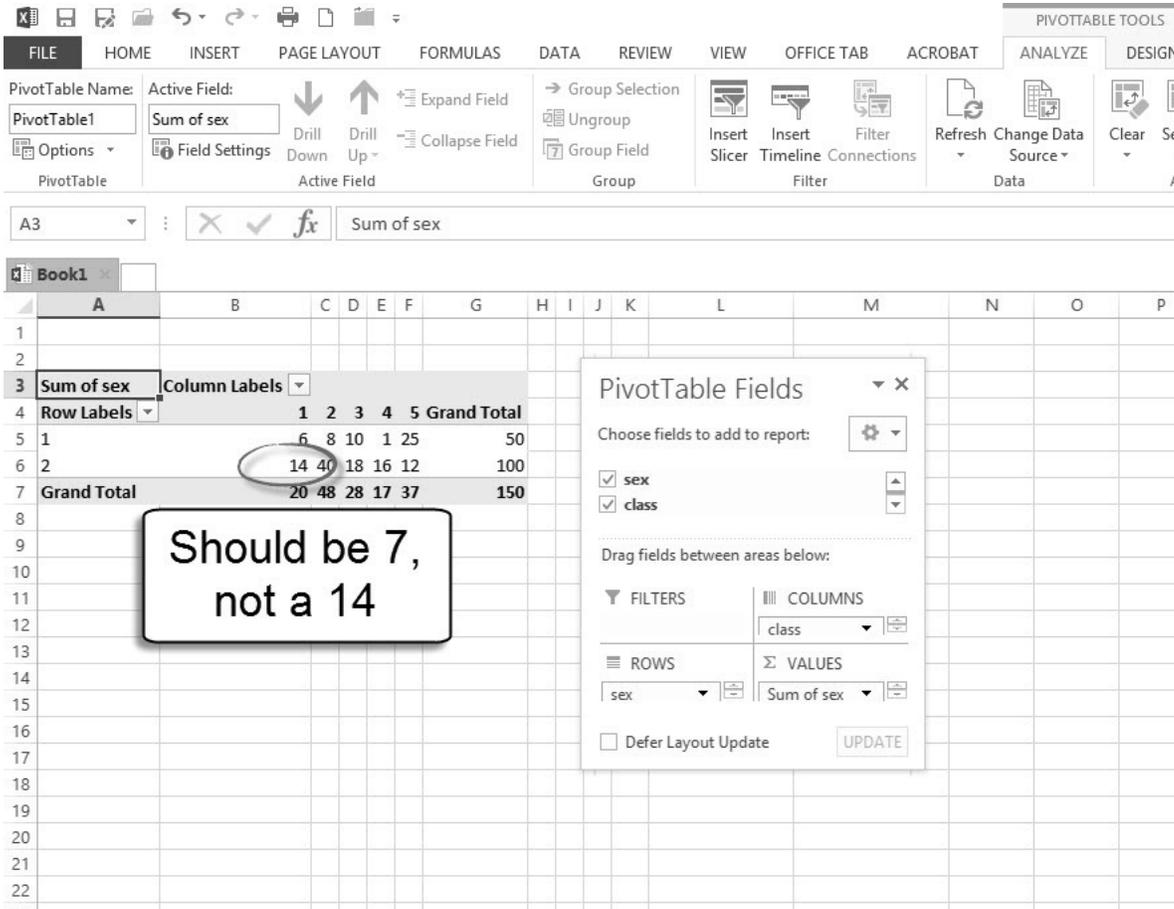


Figure Excel.16
Summed Total Problem

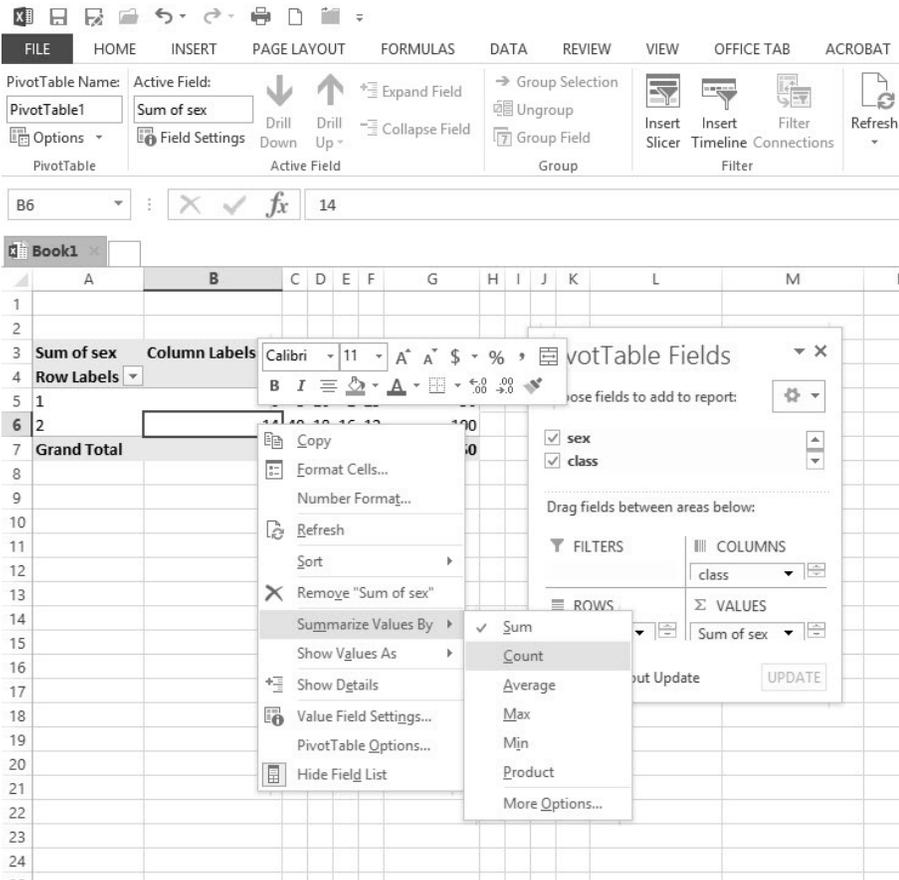


Figure Excel.17
Summarize Values By

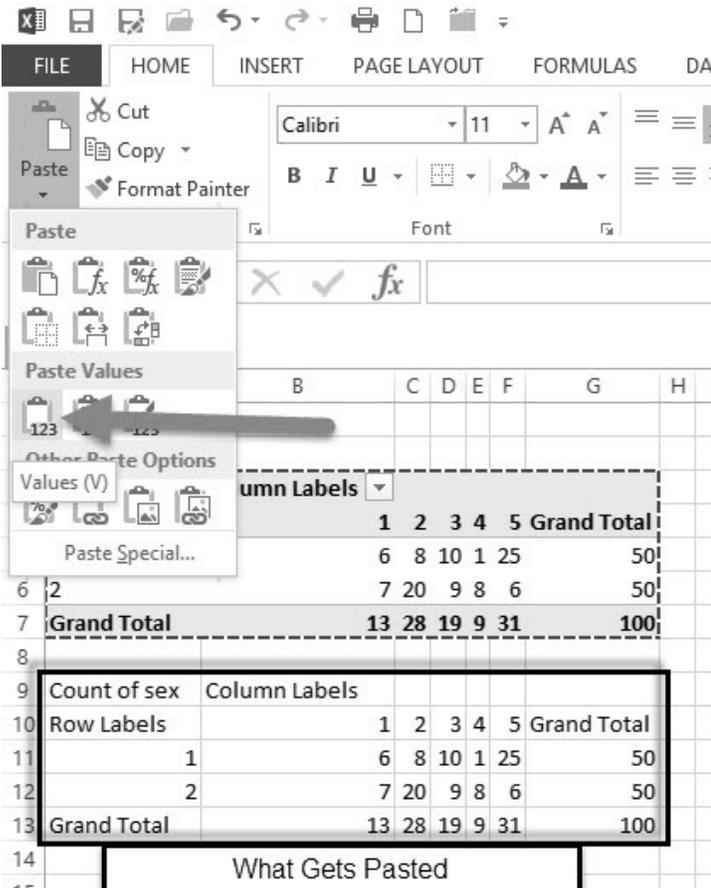


Figure Excel.18
Paste as Values

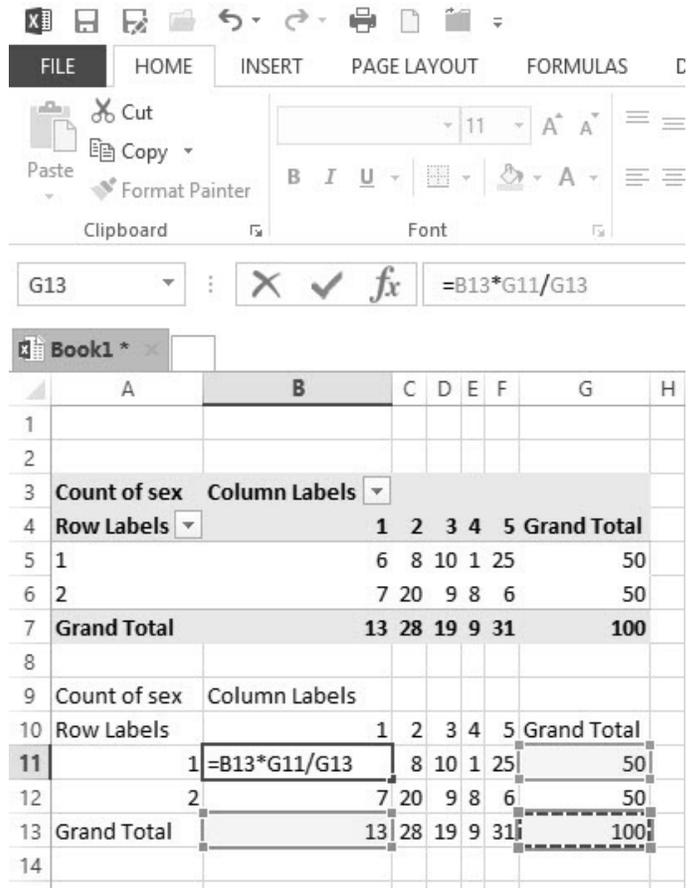


Figure Excel.19
Computing Expected Values

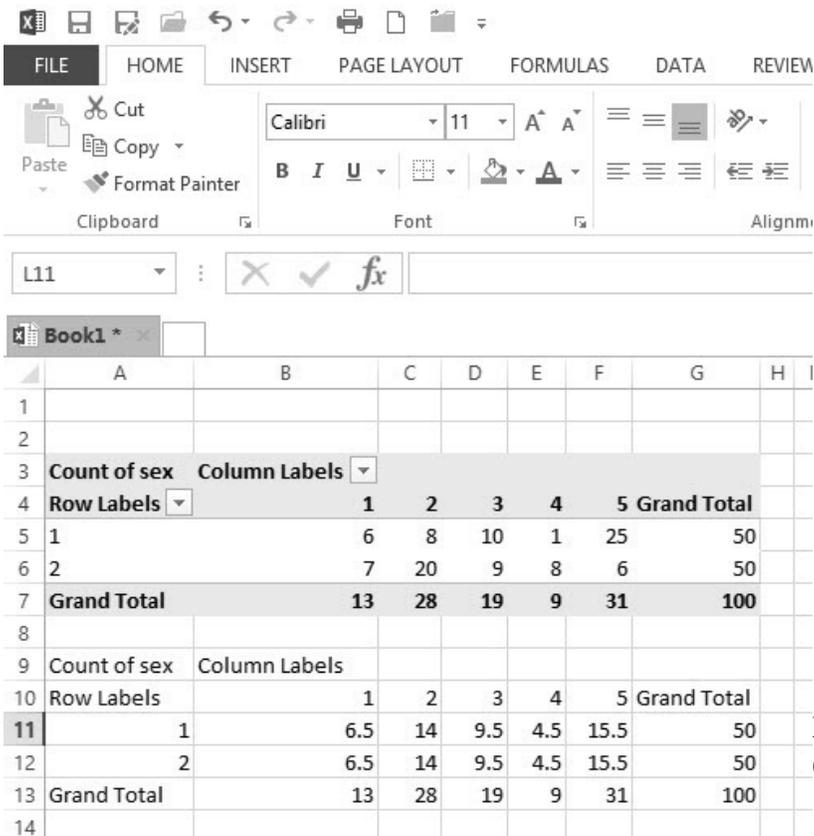


Figure Excel.20
Computed Expected Values

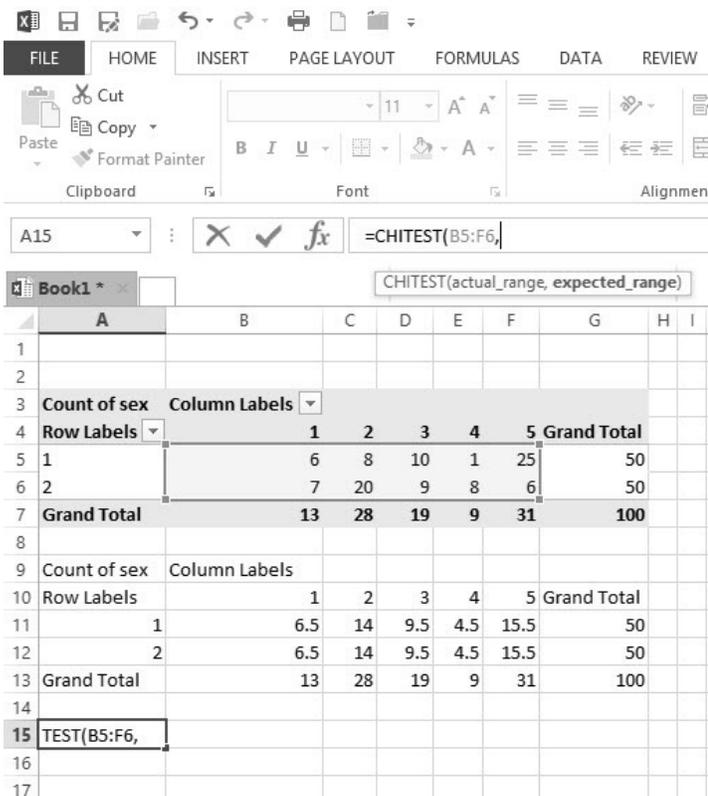


Figure Excel.21
Entering Actual Frequencies

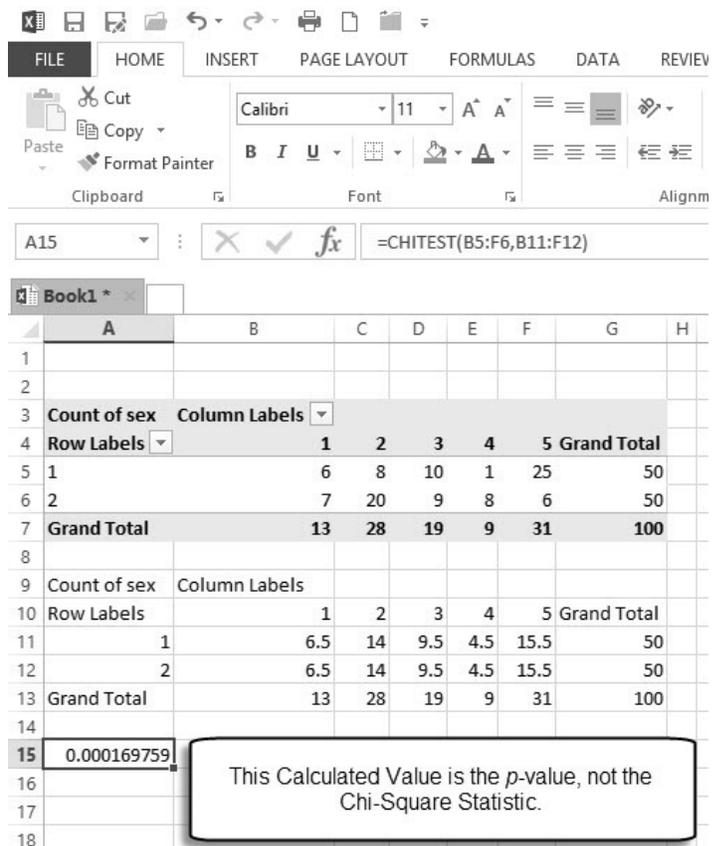


Figure Excel.22
Chi-square Test Results

EXCEL AND t TESTS

To compare the means for independent samples in Excel, first open the “ t -Test” file from the textbook’s website. Once the file is opened, you will see data for the two dress groups. As we explained earlier, Excel uses the numeric representations of 1 and 2 for each of these groups (Class A and Class B). Click on the “Data” tab at the top of the spreadsheet, select “Data Analysis” from the right side of the screen, and choose the “ t -Test: Two Sample Assuming Equal Variances” option from the menu (Figure Excel.23). To fill in the data ranges under the Input category, click the box labeled “Variable 1 Range” and highlight the values in cells B2 through B6 because this range corresponds with those in the first group. Next, click in the box titled “Variable 2 Range” and highlight the values in cells B7 through B11 because these scores correspond with those in the second group. Click the “Labels” box

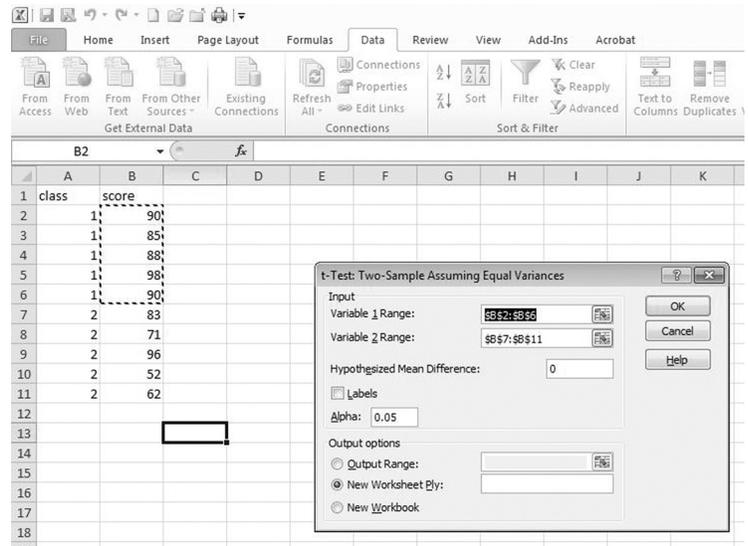


Figure Excel.23

t Test: Two Sample Assuming Equal Variances

	Variable 1	Variable 2
Mean	1.5	81.5
Variance	0.277778	227.1667
Observations	10	10
Pooled Variance	113.7222	
Hypothesized Mean Difference	0	
df	18	
t Stat	-16.7746	
P(T<=t) one-tail	9.79E-13	
t Critical one-tail	1.734064	
P(T<=t) two-tail	1.96E-12	
t Critical two-tail	2.100922	

Figure Excel.24

Computed t Test in Excel

and the

“New Worksheet Ply” box. A default of 0.05 is set for alpha, but this can be adjusted if desired. Click “OK” and the data output will appear in a new screen (Figure Excel.24). In addition, the z -Test option is provided in the menu of options.

Please note that Excel does not provide you with the Levene test for equality of variances, so there is no way to ascertain whether this assumption of the t test is violated. You can repeat the test and run it where you assume that equal variances has been violated, but without Levene’s test for equality of variances, it is hard to know whether the assumption has been violated.

APA WRITE-UP (EXCEL)

A paired two sample for means t test was conducted to determine whether an instructor’s dress in a classroom (fashionable [$M = 90.20$, $SD = 4.82$] vs. out-of-date dress [$M = 72.80$, $SD = 7.73$]) affected the scores students received on their midterm examinations. Overall, no significant difference was noted between the two groups, $t(8) = 2.31$, $p = 0.06$.

EXCEL AND ONE-WAY ANOVAS

To compute the one-way ANOVA in Excel, begin by locating the “One-Way ANOVA” spreadsheet on the textbook’s website. After opening the file, you will see two columns of data. The “color” column represents the three room colors (1 = white, 2 = yellow, and 3 = blue). The column labeled “time” represents the number of minutes required to complete a 10-item quiz. First, you must separate out the time it took to take the test into three distinct columns. Excel needs the three groups separated into distinct columns (Figure Excel.25). Next, click on the “Data” file tab at the top of the spreadsheet, select “Data Analysis” from the right side of the screen, and choose “Anova: Single Factor” from the drop-down menu (Figure Excel.26). Highlight all the data included in cells D1

through F5. The data range will appear in the box labeled “Input Range.” Be sure the “Columns,” “Labels in First Row,” and “New Worksheet Ply” options are all selected (Figure 18.15). Excel sets the Alpha default at 0.05, but this can be adjusted if you choose to do so. Click “OK” to proceed to the output worksheet (Figure Excel.27).

	A	B	C	D	E	F	G
1	color	time		White	Yellow	Blue	
2		1	5		5	5	8
3		1	6		6	6	10
4		1	5		5	10	10
5		1	4		4	7	8
6		2	5				
7		2	6				
8		2	10				
9		2	7				
10		3	8				
11		3	10				
12		3	10				
13		3	8				

Figure Excel.25
Select ANOVA: Single Factor

Anova: Single Factor

Input Range: \$D\$1:\$F\$5

Grouped By: Columns Rows

Labels in First Row

Alpha: 0.05

Output options:
 Output Range:
 New Worksheet Ply:
 New Workbook

Figure Excel.26
“ANOVA: Single Factor” Dialogue Box

Groups	Count	Sum	Average	Variance
White	4	20	5	0.666667
Yellow	4	28	7	4.666667
Blue	4	36	9	1.333333

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	32	2	16	7.2	0.013571	4.256495
Within Groups	20	9	2.222222			
Total	52	11				

Figure Excel.27
ANOVA Summary Table

EXCEL AND PEARSON PRODUCT-MOMENT CORRELATIONS

To calculate a correlation using Excel, open the file titled “correlation” in the Excel data file on the textbook’s website. In this file, data for two variables will be listed—“CA” and “hrchange.” “CA” represents the CA scores for each respondent, and “hrchange” reflects the change in heart rate when the respondent is asked to give an impromptu speech. The process for calculating the correlation is similar to other Excel statistical analyses. Click on the “Data” tab at the top of the spreadsheet, select “Data Analysis,” and select “Correlation” from the menu. Click on “OK” and you will be prompted to enter the values in the “Input Range” box. Click on the box and highlight all of the data in cells A1 through B21—this includes the variable labels included in the first row of the spreadsheet. Check the options for “Labels in First Row” and “New Worksheet Ply” before clicking “OK” (Figure Excel.28). The data output will appear in a new sheet (Figure Excel.28). With the Excel file, you will have to look at the critical value table in (SEE BOOK) to determine whether the test is statistically significant.

Again, we have bolded the significant correlation on the results to make it stand out more easily. Excel correlation tables can become somewhat confusing because they are

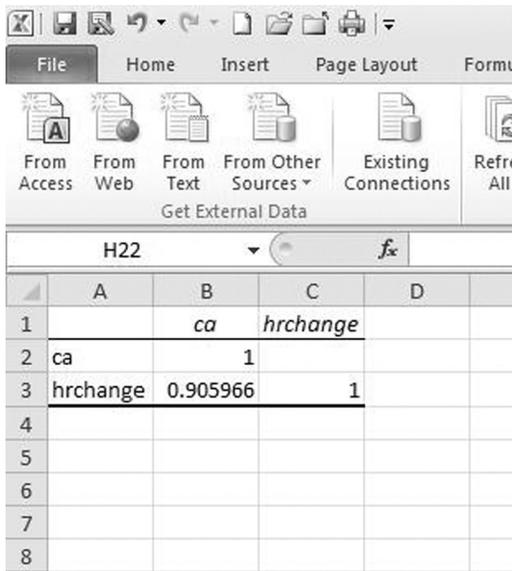


Figure Excel.29
Excel Correlation Output

EXCEL AND THE SIMPLE LINEAR REGRESSION

To compute a regression in Excel, begin by opening the “Regression” data file on the textbook’s website. In this spreadsheet, you will see two variables, “CA” and “hrchange”. These are the same variables used in our calculation of correlations from the last chapter. Recall that “CA” represents an individual’s scores on the PRCA-24, and “hrchange” reflects the change in an individual’s heart rate while giving an impromptu speech. To begin, select the “Data” tab at the top of the spreadsheet, click on “Data Analysis” at the right side of the screen, and select “Regression” from the drop-down menu (Figure Excel.30). Next, click on the “Input Y Range” box and highlight the data in cells A1 through A21. Select “Input X Range” and highlight the data in B1 through B21. Excel sets a default confidence interval of 95 percent for the regression statistic. Click on “New Worksheet Ply” and “Labels” before selecting “OK” (Figure Excel.31). The output data will appear in a new sheet (Figure Excel.32).

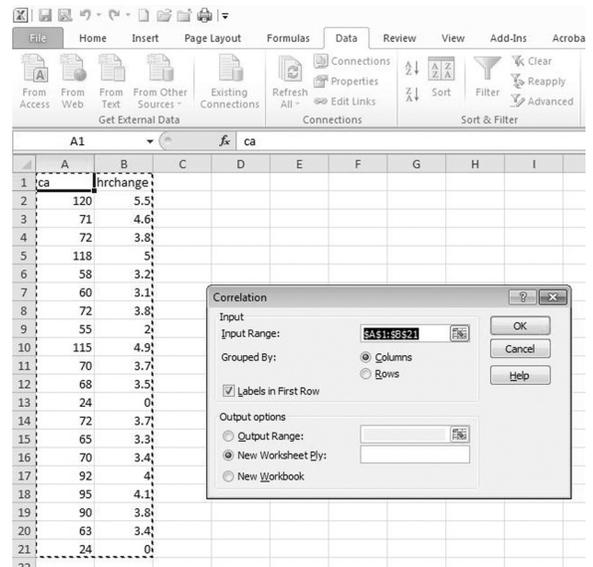


Figure Excel.28
Excel Correlation Dialogue Box

not placed in a nice chart format like in SPSS. Although you may not notice this problem here, in a few minutes when we look at a larger example you will see it more clearly. Again, to make these tables easier to understand, we recommend finding the diagonal lines of 1.0 and marking a line through them with a highlighter to make it easier to view.

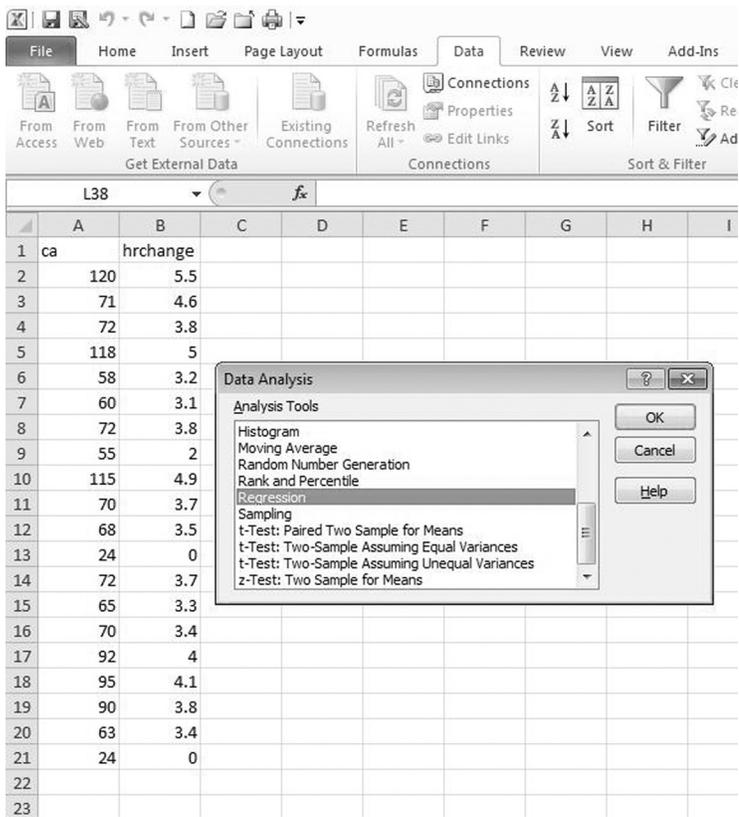


Figure Excel.30
Select Regression from the Data Analysis Dialogue Box

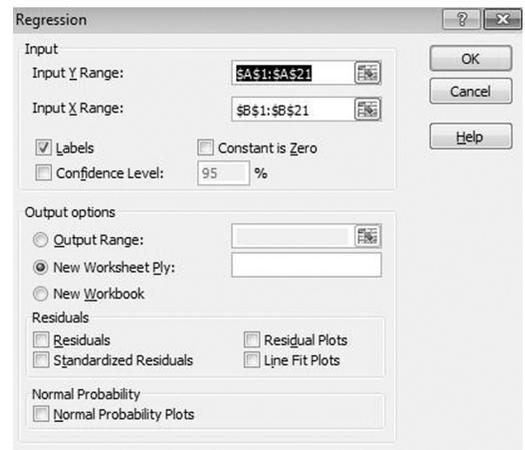


Figure Excel.31
Regression Dialogue Box

Regression Statistics	
Multiple R	0.905966025
R Square	0.820774439
Adjusted R Square	0.810817464
Standard Error	11.31410048
Observations	20

ANOVA					
	df	SS	MS	F	Significance F
Regression	1	10552.04034	10552.04034	82.43210307	3.86484E-08
Residual	18	2304.159655	128.0088697		
Total	19	12856.2			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	15.89349176	6.851128069	2.31983574	0.032302983	1.499805796	30.28717772	1.499805796	30.28717772
hrchange	16.80421751	1.850845231	9.07921269	3.86484E-08	12.91573597	20.69269905	12.91573597	20.69269905

Figure Excel.32
Excel Regression Output

