DATA ANALYTICS TOOLKITS

Use the Analysis ToolPak to perform complex data analysis

If you need to develop complex statistical or engineering analyses, you can save steps and time by using the Analysis ToolPak. You provide the data and parameters for each analysis,and the tool uses the appropriate statistical or engineering macro functions to calculate and display the results in an output table. Some tools generate charts in addition to output tables.

The data analysis functions can be used on only one worksheet at a time. When you performdata analysis on grouped worksheets, results will appear on the first worksheet and empty formatted tables will appear on the remaining worksheets. To perform data analysis on the remainder of the worksheets, recalculate the analysis tool for each worksheet.

The Analysis ToolPak includes the tools described in the following sections. To access thesetools, click **Data Analysis** in the **Analysis** group on the **Data** tab. If the **Data Analysis** command is not available, you need to load the Analysis ToolPak add-in program.

Load and activate the Analysis ToolPak

1. Click the **File** tab, click **Options**, and then click the **Add-Ins** category.
2. In the **Manage** box, select **Excel Add-ins** and then click **Go**.
   If you're using Excel for Mac, in the file menu go to **Tools** > **Excel Add-ins.**
3. In the **Add-Ins** box, check the **Analysis ToolPak** check box, and then click **OK**.
   - If **Analysis ToolPak** is not listed in the **Add-Ins available** box, click **Browse**
     to locate it.

- If you are prompted that the Analysis ToolPak is not currently installed on your computer, click **Yes** to install it.

Note: To include Visual Basic for Application (VBA) functions for the Analysis ToolPak, you can load the Analysis ToolPak - VBA Add-in the same way that you load the Analysis ToolPak. In the **Add-ins available** box, select the **Analysis ToolPak - VBA** check box.

Anova

The Anova analysis tools provide different types of variance analysis. The tool that you should use depends on the number of factors and the number of samples that you have from the populations that you want to test.

Anova: Single Factor

This tool performs a simple analysis of variance on data for two or more samples. The analysis provides a test of the hypothesis that each sample is drawn from the same underlying probability distribution against the alternative hypothesis that underlying probability distributions are not the same for all samples. If there are only two samples, you can use the worksheet function **T.TEST**. With more than two samples, there is no convenient generalization of **T.TEST**, and the Single Factor Anova model can be called upon instead.

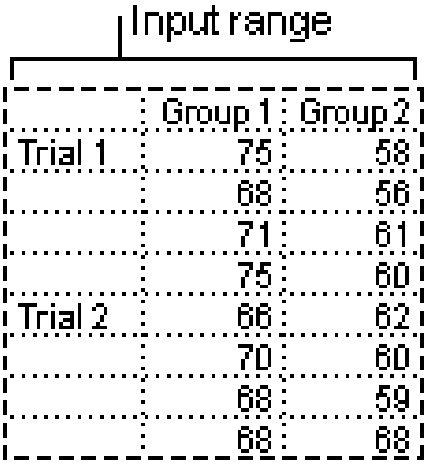
Anova: Two-Factor with Replication

This analysis tool is useful when data can be classified along two different dimensions. For example, in an experiment to measure the height of plants, the plants may be given different brands of fertilizer (for example, A, B, C) and might also be kept at different temperatures (for example, low, high). For each of the six possible pairs of {fertilizer, temperature}, we have an equal number of observations of plant height. Using this Anova tool, we can test:

- Whether the heights of plants for the different fertilizer brands are drawn from the same underlying population. Temperatures are ignored for this analysis.

- Whether the heights of plants for the different temperature levels are drawn from the same underlying population. Fertilizer brands are ignored for this analysis.

Whether having accounted for the effects of differences between fertilizer brands found in the first bulleted point and differences in temperatures found in the second bulleted point, the six samples representing all pairs of {fertilizer, temperature} values are drawn from the same population. The alternative hypothesis is that there are effects due to specific {fertilizer, temperature} pairs over and above the differences that are based on fertilizer alone or on temperature alone.



Anova: Two-Factor Without Replication

This analysis tool is useful when data is classified on two different dimensions as in the Two-Factor case With Replication. However, for this tool it is assumed that there is only a single observation for each pair (for example, each {fertilizer, temperature} pair in the preceding example).

Correlation

The **CORREL** and **PEARSON** worksheet functions both calculate the correlation coefficient between two measurement variables when measurements on each variable are observed for each of N subjects. (Any missing observation for any subject causes that subject to be ignored in the analysis.) The

Correlation analysis tool is particularly useful when there are more than two measurement variables for each of N subjects. It provides an output table, a correlation matrix, that shows the value of **CORREL** (or **PEARSON**) applied to each possible pair of measurement variables.

The correlation coefficient, like the covariance, is a measure of the extent to which two measurement variables "vary together." Unlike the covariance, the correlation coefficient is scaled so that its value is independent of the units in which the two measurement variables are expressed. (For example, if the two measurement variables are weight and height, the value of the correlation coefficient is unchanged if weight is converted from pounds to kilograms.) The value of any correlation coefficient must be between -1 and +1 inclusive.

You can use the correlation analysis tool to examine each pair of measurement variables to determine whether the two measurement variables tend to move together — that is, whether large values of one variable tend to be associated with large values of the other (positive correlation), whether small values of one variable tend to be associated with large values of the other (negative correlation), or whether values ofboth variables tend to be unrelated (correlation near 0 (zero)).

Covariance

The Correlation and Covariance tools can both be used in the same setting, when you have N different measurement variables observed on a set of individuals. The Correlation and Covariance tools each give an output table, a matrix, that shows the correlation coefficient or covariance, respectively, between each pair of measurement variables. The difference is that correlation coefficients are scaled to lie between -1 and +1 inclusive. Corresponding covariances are not scaled. Both the correlation coefficient and the covariance are measures of the extent to which two variables "vary together."

The Covariance tool computes the value of the worksheet function **COVARIANCE.P** for each pair of measurement variables. (Direct use of COVARIANCE.P rather than the Covariance tool is a reasonable alternative when there are only two measurement variables, that is, N=2.) The entry on the diagonal

Covariance tool's output table in row i, column i is the covariance of the i-th measurement variable with itself. This is just the population variance for that variable, as calculated by the worksheet function **VAR.P**.

You can use the Covariance tool to examine each pair of measurement variables to determine whether thetwo measurement variables tend to move together — that is, whether large values of one variable tend to be associated with large values of the other (positive covariance), whether small values of one variable tend to be associated with large values of the other (negative covariance), or whether values of both variables tend to be unrelated (covariance near 0 (zero)).

Descriptive Statistics

The Descriptive Statistics analysis tool generates a report of univariate statistics for data in the input range, providing information about the central tendency and variability of your data.

Exponential Smoothing

The Exponential Smoothing analysis tool predicts a value that is based on the forecast for the prior period, adjusted for the error in that prior forecast. The tool uses the smoothing constant a, the magnitude of whichdetermines how strongly the forecasts respond to errors in the prior forecast.

Note: Values of 0.2 to 0.3 are reasonable smoothing constants. These values indicate that the current forecast should be adjusted 20 percent to 30 percent for error in the prior forecast. Larger constants yield a faster response but can produce erratic projections. Smaller constants can result in long lags for forecast values.

F-Test Two-Sample for Variances

The F-Test Two-Sample for Variances analysis tool performs a two-sample F-test to compare two population variances.

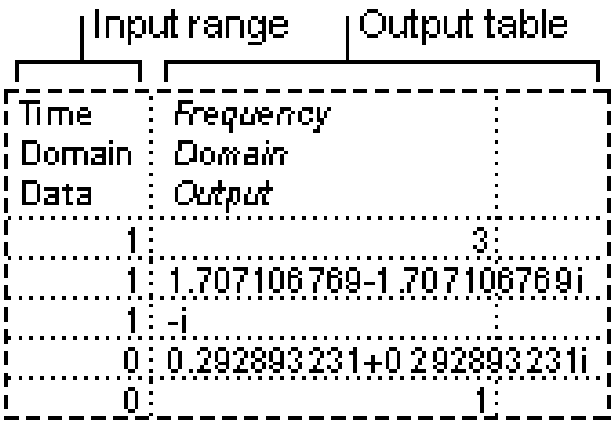
For example, you can use the F-Test tool on samples of times in a swim meet for each of two teams. The tool provides the result of a test of the null hypothesis that these two samples come from distributions with equal variances, against the alternative that the variances are not equal in the underlying distributions.

The tool calculates the value f of an F-statistic (or F-ratio). A value of f close to 1 provides evidence that the underlying population variances are equal. In the output table, if f < 1 "P(F <= f) one-tail" gives the probability of observing a value of the F-statistic less than f when population variances are equal, and "F Critical one-tail" gives the critical value less than 1 for the chosen significance level, Alpha. If f > 1, "P(F

<= f) one-tail" gives the probability of observing a value of the F-statistic greater than f when population variances are equal, and "F Critical one-tail" gives the critical value greater than 1 for Alpha.

Fourier Analysis

The Fourier Analysis tool solves problems in linear systems and analyzes periodic data by using the Fast Fourier Transform (FFT) method to transform data. This tool also supports inverse transformations, in which the inverse of



transformed data returns the original data.

Histogram

The Histogram analysis tool calculates individual and cumulative frequencies for a cell range of data and data bins. This tool generates data for the number of occurrences of a value in a data set.

For example, in a class of 20 students, you can determine the distribution of scores in letter-grade categories. A histogram table presents the letter-grade boundaries and the number of scores between the lowest bound and the current bound. The single most-frequent score is the mode of the data.

Tip: In Excel 2016, you can now create a histogram or Pareto chart.

Moving Average

The Moving Average analysis tool projects values in the forecast period, based on the average value of the variable over a specific number of preceding periods. A moving average provides trend information that a simple average of all historical data would mask. Use this tool to forecast sales, inventory, or other trends. Each forecast value is based on the following formula.

$$F_{(t+1)} = \frac{1}{N}\sum_{j=1}^{N} A_{t-j+1}$$

where:

- N is the number of prior periods to include in the moving average
- A j is the actual value at time j
- F j is the forecasted value at time j

Random Number Generation

The Random Number Generation analysis tool fills a range with independent random numbers that are drawn from one of several distributions. You can characterise the subjects in a population with a probability distribution. For example, you can use a normal distribution to characterise the population of individuals' heights, or you can use a Bernoulli distribution of two possible outcomes to characterise the population of coin-flip results.

Rank and Percentile

The Rank and Percentile analysis tool produces a table that contains the ordinal and percentage rank of each value in a data set. You can analyze the relative standing of values in a data set. This tool uses the worksheet functions **RANK.EQ** and **PERCENTRANK.INC**. If you want to account for tied values, use the **RANK.EQ** function, which treats tied values as having the same rank, or use the **RANK.AVG** function, which returns the average rank for the tied values.

Regression

The Regression analysis tool performs linear regression analysis by using the "least squares" method to fit a line through a set of observations. You can analyse how a single dependent variable is affected by the values of one or more independent variables. For example, you can analyse how an athlete's performance is affected by such factors as age, height, and weight. You can apportion shares in the performance measure to each of these three factors, based on a set of performance data, and then use the results to predict the performance of a new, untested athlete.

The Regression tool uses the worksheet function **LINEST**.

Sampling

The Sampling analysis tool creates a sample from a population by treating the input range as a population. When the population is too large to process or chart, you can use a representative sample. You can also create a sample that contains only the values from a particular part of a cycle if you believe that the input data is periodic. For example, if the input range contains quarterly sales figures, sampling with a periodic rate of four places the values from the same quarter in the output range.

---

t-Test

The Two-Sample t-Test analysis tools test for equality of the population means that underlie each sample. The three tools employ different assumptions: that the population variances are equal, that the population variances are not equal, and that the two samples represent before-treatment and after-treatment observations on the same subjects.

For all three tools below, a t-Statistic value, t, is computed and shown as "t Stat" in the output tables. Depending on the data, this value, t, can be negative or nonnegative. Under the assumption of equal underlying population means, if t < 0, "P(T <= t) one-tail" gives the probability that a value of the t- Statistic would be observed that is more negative than t. If t >=0, "P(T <= t) one-tail" gives the probability that a value of the t-Statistic would be observed that is more positive than t. "t Critical one-tail" gives the cutoff value, so that the probability of observing a value of the t-Statistic greater than or equal to "t Critical one-tail" is Alpha.

"P(T <= t) two-tail" gives the probability that a value of the t-Statistic would be observed that is larger in absolute value than t. "P Critical two-tail" gives the cutoff value, so that the probability of an observed t- Statistic larger in absolute value than "P Critical two-tail" is Alpha.

t-Test: Paired Two Sample For Means

You can use a paired test when there is a natural pairing of observations in the samples, such as when a sample group is tested twice — before and after an experiment. This analysis tool and its formula perform a paired two-sample Student's t-Test to determine whether observations that are taken before a treatment and observations taken after a treatment are likely to have come from distributions with equal population means. This t-Test form does not assume that the variances of both populations are equal.

Note: Among the results that are generated by this tool is pooled variance, an accumulated measure of the spread of data about the mean, which is derived from the following formula.

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$$

t-Test: Two-Sample Assuming Equal Variances

This analysis tool performs a two-sample student's t-Test. This t-Test form assumes that the two data sets came from distributions with the same variances. It is referred to as a homoscedastic t-Test. You can use this t-Test to determine whether the two samples are likely to have come from distributions with equal population means.

t-Test: Two-Sample Assuming Unequal Variances

This analysis tool performs a two-sample student's t-Test. This t-Test form assumes that the two data sets came from distributions with unequal variances. It is referred to as a heteroscedastic t-Test. As with the preceding Equal Variances case, you can use this t-Test to determine whether the two samples are likely to have come from distributions with equal population means. Use this test when there are distinct subjects in the two samples. Use the Paired test, described in the follow example, when there is a single set of subjects and the two samples represent measurements for each subject before and after a

treatment.

The following formula is used to determine the statistic value t.

$$t' = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

The following formula is used to calculate the degrees of freedom, df. Because the result of the calculation is usually not an integer, the value of df is rounded to the nearest integer to obtain a critical value from the t table. The Excel worksheet function **T.TEST** uses the calculated df value without rounding, because it is possible to compute a value for **T.TEST** with a noninteger df. Because of these different approaches to determining the degrees of freedom, the results of **T.TEST** and this t-Test tool will differ in the Unequal Variances case.

$$df = \frac{\left(\frac{S_1^2}{m} + \frac{S_2^2}{n}\right)^2}{\frac{\left(S_1^2/m\right)^2}{m-1} + \frac{\left(S_2^2/n\right)^2}{n-1}}$$

---

z-Test

The z-Test: Two Sample for Means analysis tool performs a two sample z-Test for means with known variances. This tool is used to test the null hypothesis that there is no difference between two population means against either one-sided or two-sided alternative hypotheses. If variances are not known, the worksheet function **Z.TEST** should be used instead.

When you use the z-Test tool, be careful to understand the output. "P(Z <= z) one-tail" is really P(Z >= ABS(z)), the probability of a z-value further from 0 in the same direction as the observed z value when there is no difference between the population means. "P(Z <= z) two-tail" is really P(Z >= ABS(z) or Z <=

-ABS(z)), the probability of a z-value further from 0 in either direction than the observed z-value when there is no difference between the population means. The two-tailed result is just the one-tailed result multiplied by 2. The z-Test tool can also be used for the case where the null hypothesis is that there is a specific nonzero value for the difference between the two population means. For example, you can use this test to determine differences between the performances of two car models.

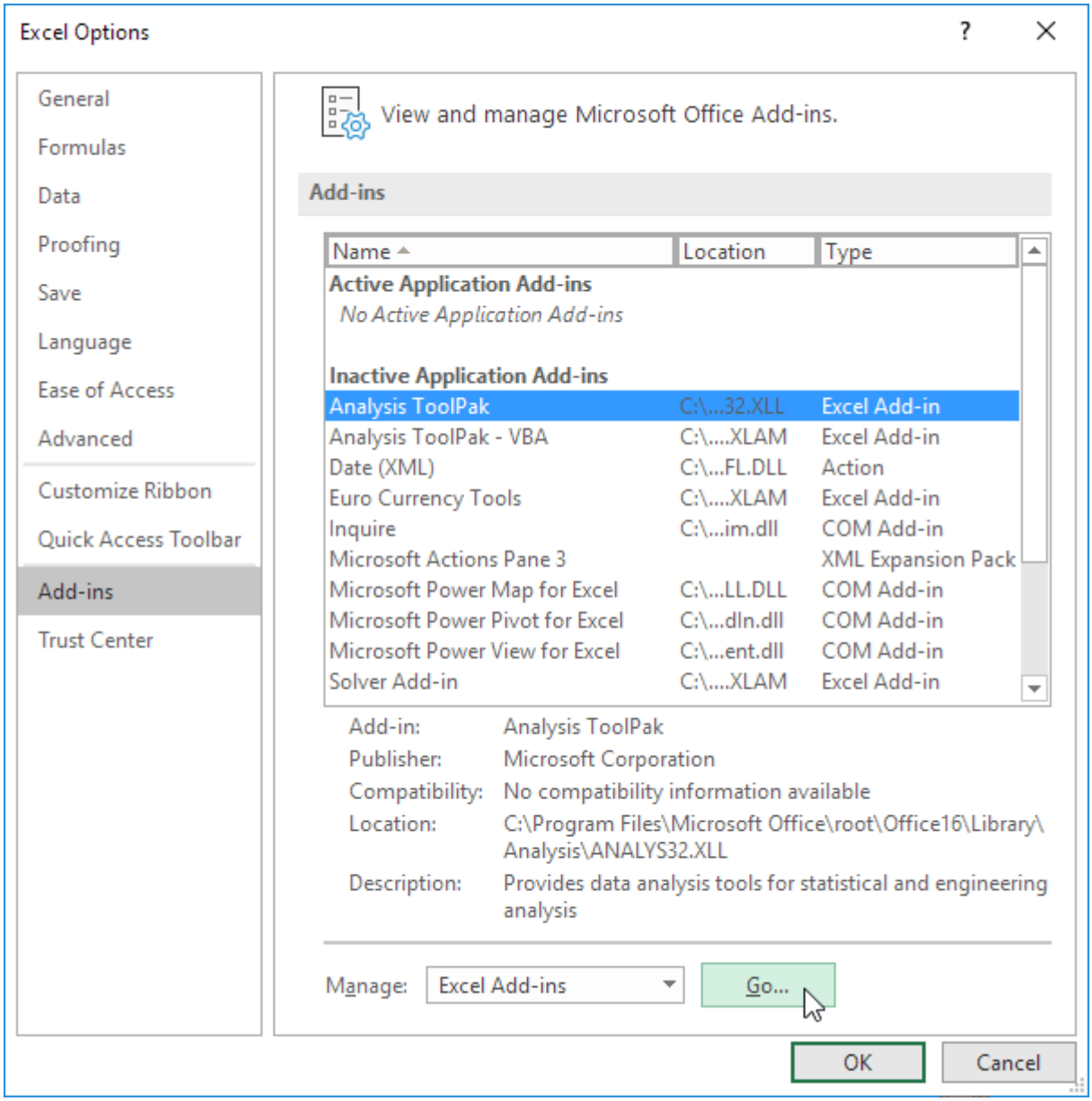
Analysis ToolPak

The Analysis ToolPak is an Excel add-in program that provides data analysis tools for financial, statistical and engineering data analysis.

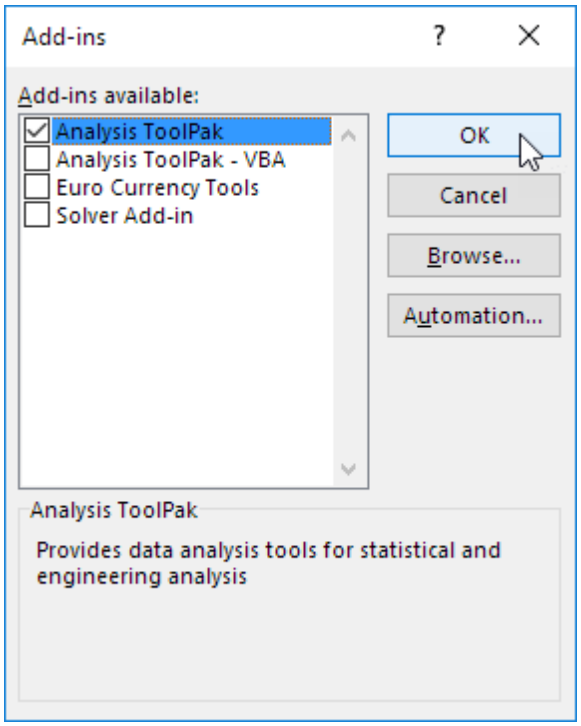
To load the Analysis ToolPak add-in, execute the following steps.

1. On the File tab, click Options.

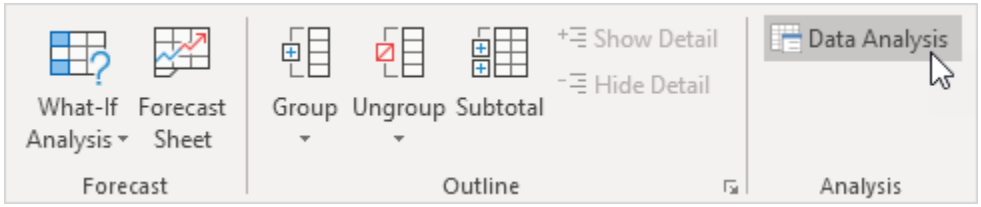
2. Under Add-ins, select Analysis ToolPak and click on the Go button.



3. Check Analysis ToolPak and click on OK.

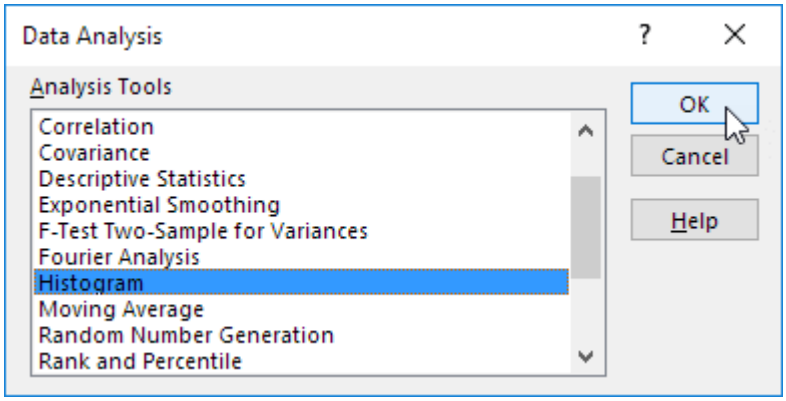


4. On the Data tab, in the Analysis group, you can now click on Data Analysis.



The following dialog box below appears.

5. For example, select Histogram and click OK to create a Histogram in Excel.



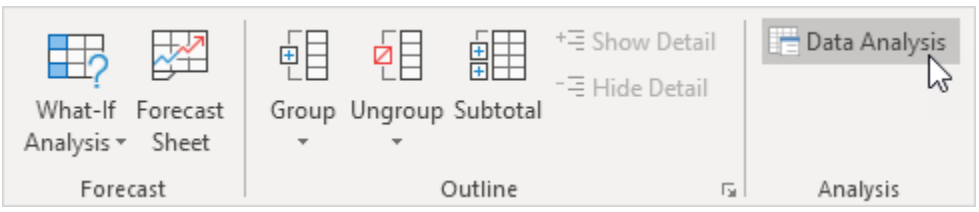
Histogram

This example teaches you how to make a histogram in Excel.

1. First, enter the bin numbers (upper levels) in the range C4:C8.

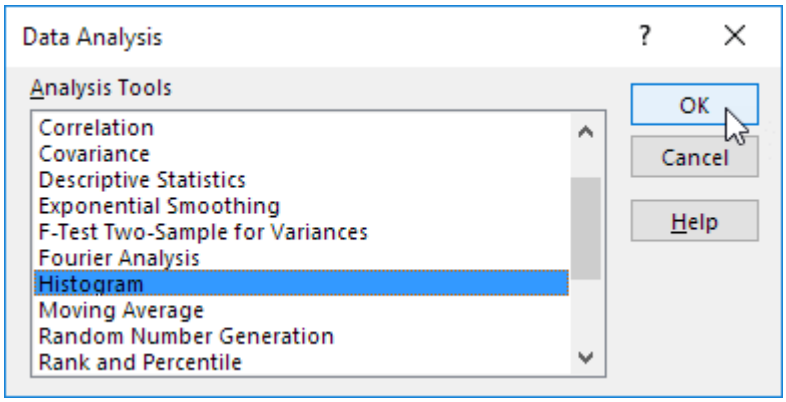
	A	B	C	D
1	Number of students			
2	22			
3	29			
4	40		20	
5	30		25	
6	48		30	
7	24		35	
8	21		40	
9	19			
10	24			
11	22			
12	25			
13	52			
14	35			
15	40			
16	31			
17	37			
18	21			
19	23			
20				

2. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

3. Select Histogram and click OK.

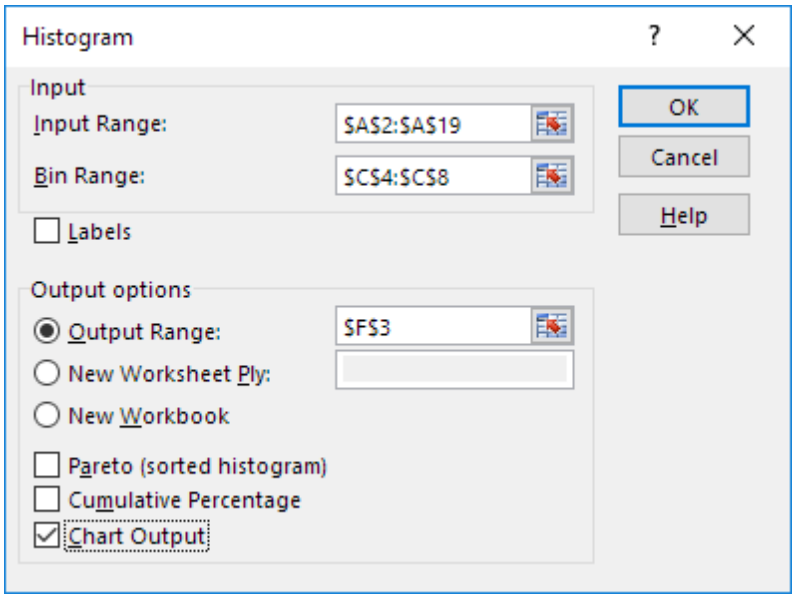


4. Select the range A2:A19.

5. Click in the Bin Range box and select the range C4:C8.

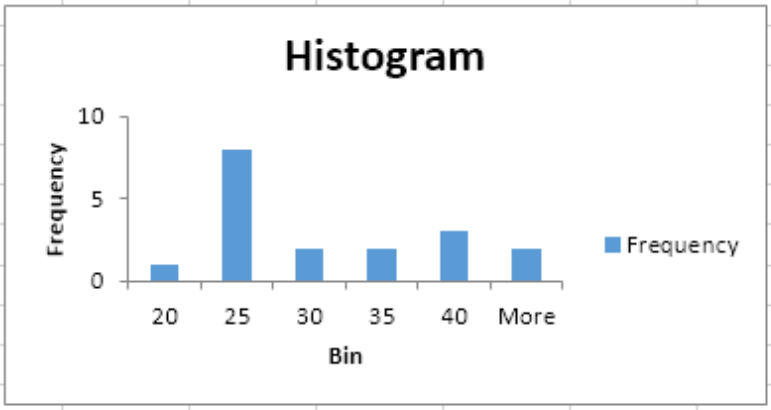
6. Click the Output Range option button, click in the Output Range box and select cell F3.

7. Check Chart Output.



8. Click OK.

Bin	Frequency
20	1
25	8
30	2
35	2
40	3
More	2



9. Click the legend on the right side and press Delete.

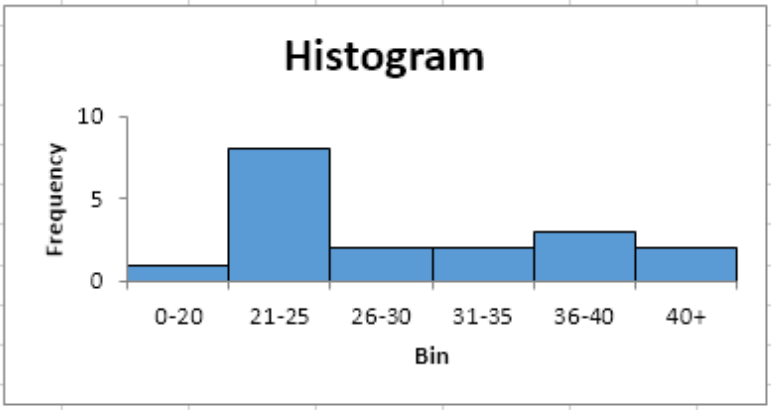
10. Properly label your bins.

11. To remove the space between the bars, right click a bar, click Format Data Series and change the Gap Width to 0%.

12. To add borders, right click a bar, click Format Data Series, click the Fill & Line icon, click Border and select a color.

Result:

Bin	Frequency
0-20	1
21-25	8
26-30	2
31-35	2
36-40	3
40+	2

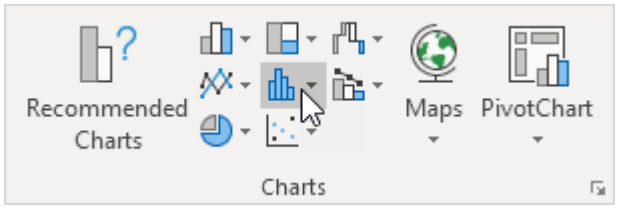


If you have Excel 2016 or later, simply use the Histogram chart type.

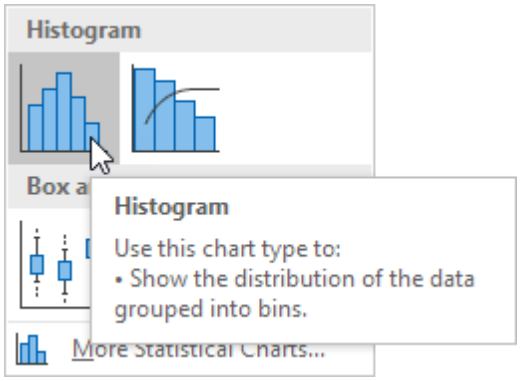
13. Select the range A1:A19.

	A	B
1	Number of students	
2	22	
3	29	
4	40	
5	30	
6	48	
7	24	
8	21	
9	19	
10	24	
11	22	
12	25	
13	52	
14	35	
15	40	
16	31	
17	37	
18	21	
19	23	
20		

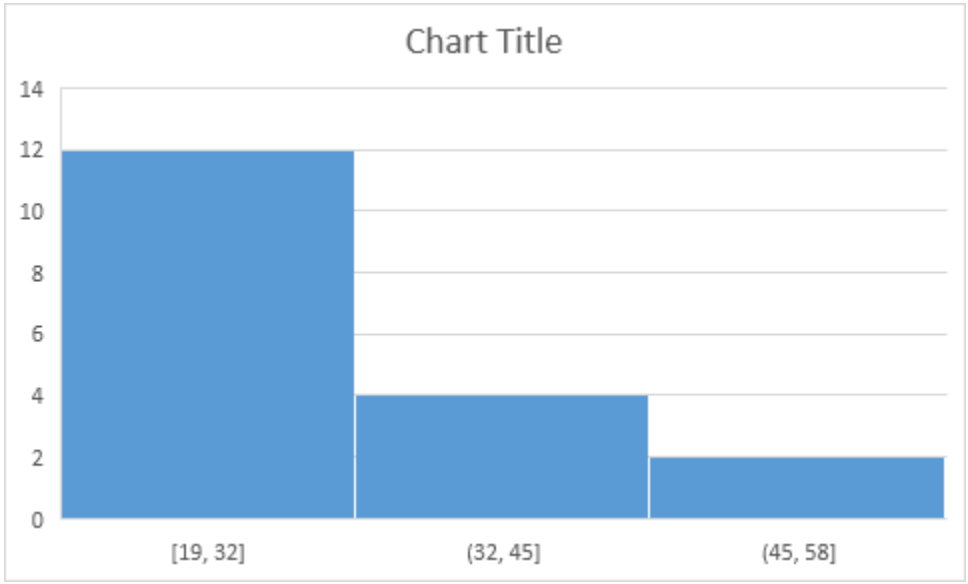
14. On the Insert tab, in the Charts group, click the Histogram symbol.



15. Click Histogram.

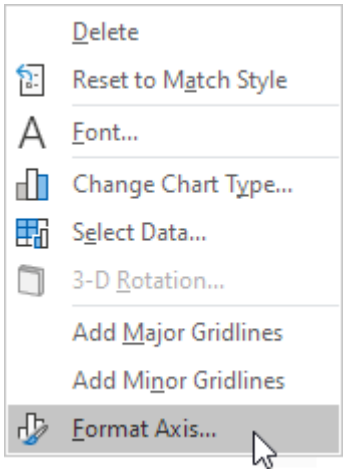


Result. A histogram with 3 bins.



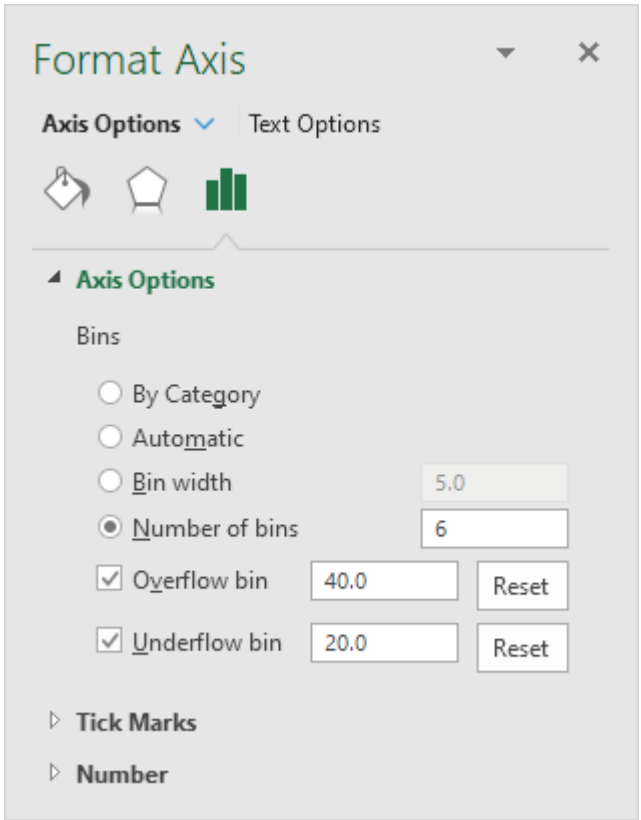
Note: Excel uses Scott's normal reference rule for calculating the number of bins and the bin width.

16. Right click the horizontal axis, and then click Format Axis.

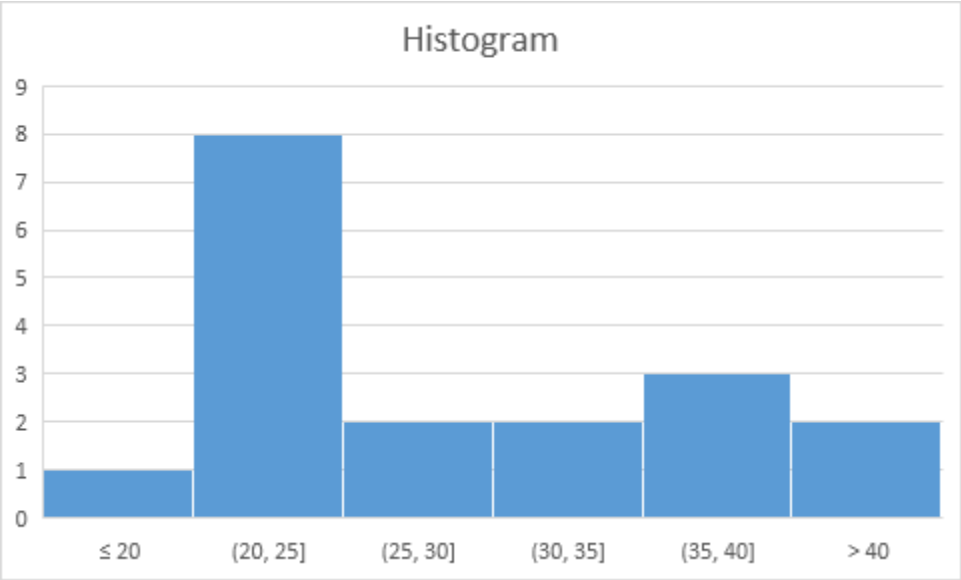


The Format Axis pane appears.

17. Define the histogram bins. We'll use the same bin numbers as before (see first picture on this page). Bin width: 5. Number of bins: 6. Overflow bin: 40. Underflow bin: 20.

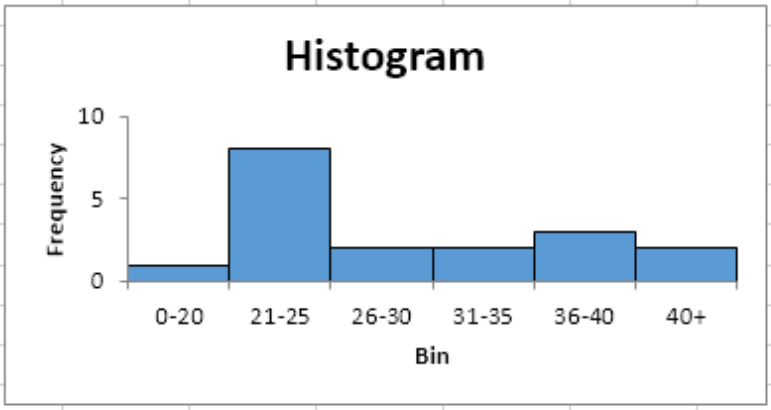


Result:



Recall, we made the following histogram using the Analysis ToolPak

Bin	Frequency
0-20	1
21-25	8
26-30	2
31-35	2
36-40	3
40+	2



Conclusion: the bin labels look different, but the histograms are the same. ≤20 is the same as 0-20, (20, 25] is the same as 21-25, etc.

Tip: you can also use pivot tables to easily create a frequency distribution in Excel.

Create a Macro

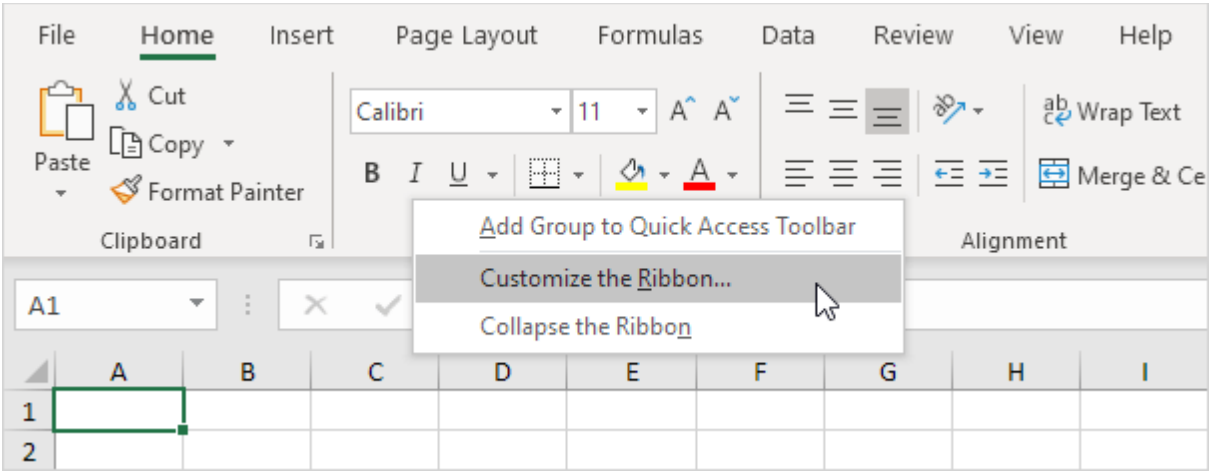
Developer Tab | Command Button | Assign a Macro | Visual Basic Editor

With Excel VBA you can automate tasks in Excel by writing so called macros. In this chapter, learn how to create a simple macro which will be executed after clicking on a command button. First, turn on the Developer tab.

Developer Tab

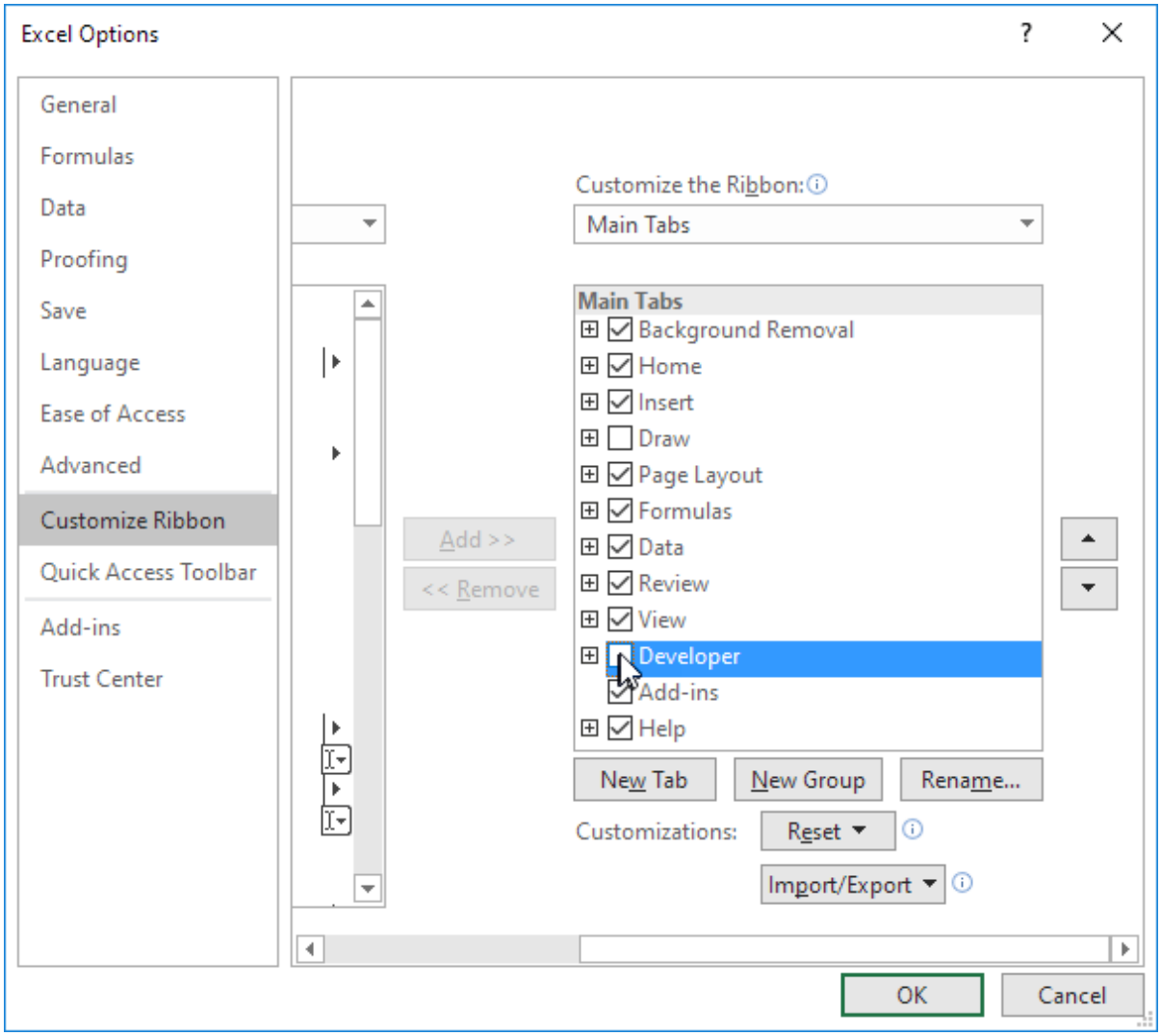
To turn on the Developer tab, execute the following steps.

1. Right click anywhere on the ribbon, and then click Customize the Ribbon.



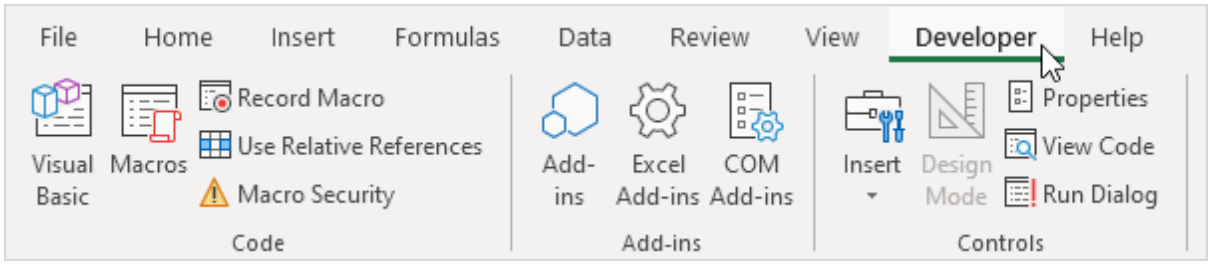
2. Under Customize the Ribbon, on the right side of the dialog box, select Main tabs (if necessary).

3. Check the Developer check box.



4. Click OK.

5. You can find the Developer tab next to the View tab.

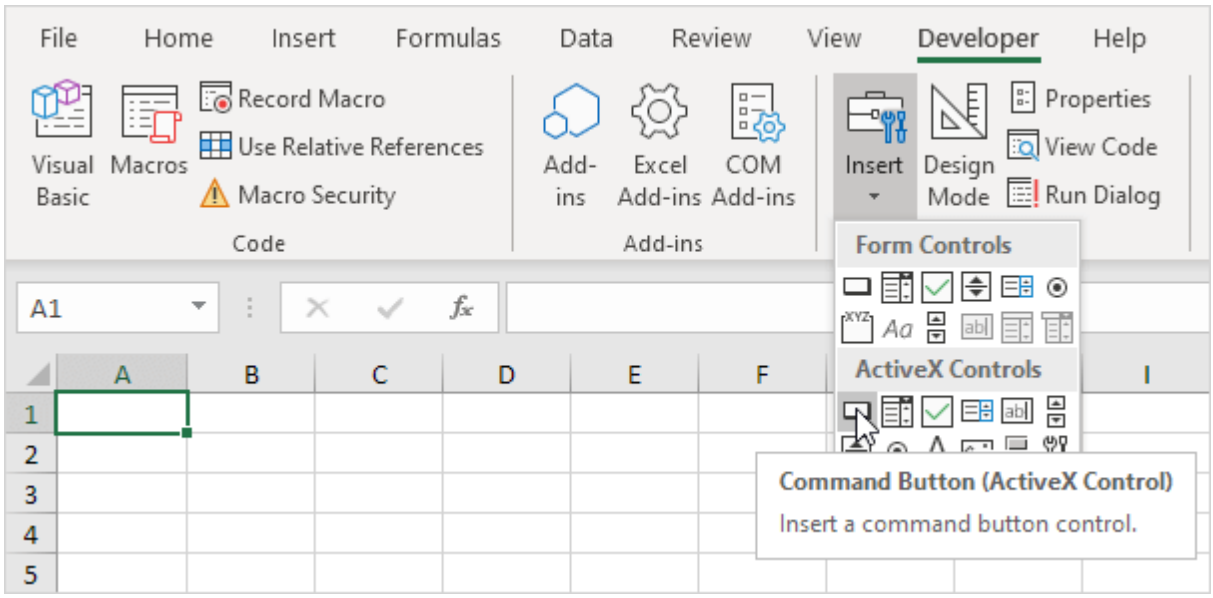


Command Button

To place a command button on your worksheet, execute the following steps.

1. On the Developer tab, click Insert.

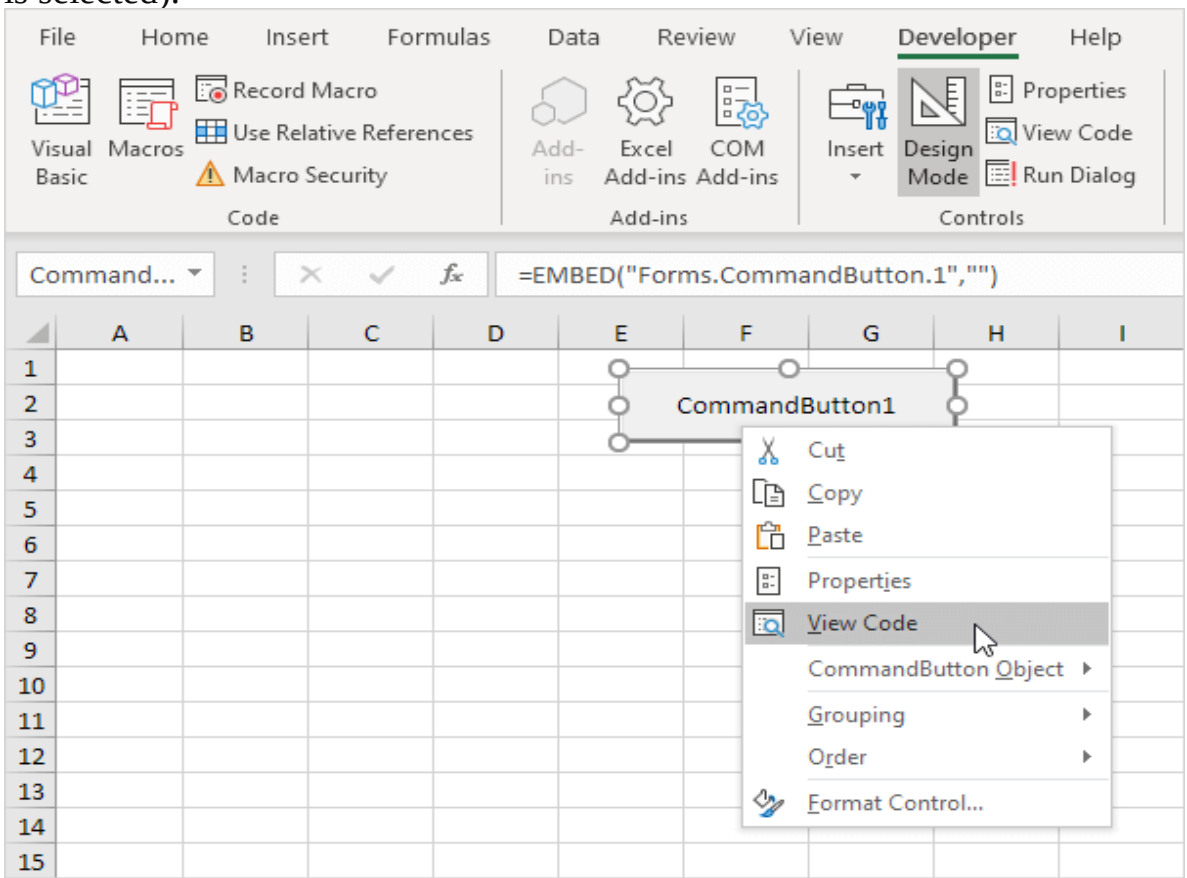
2. In the ActiveX Controls group, click Command Button.



3. Drag a command button on your worksheet.

Assign a Macro

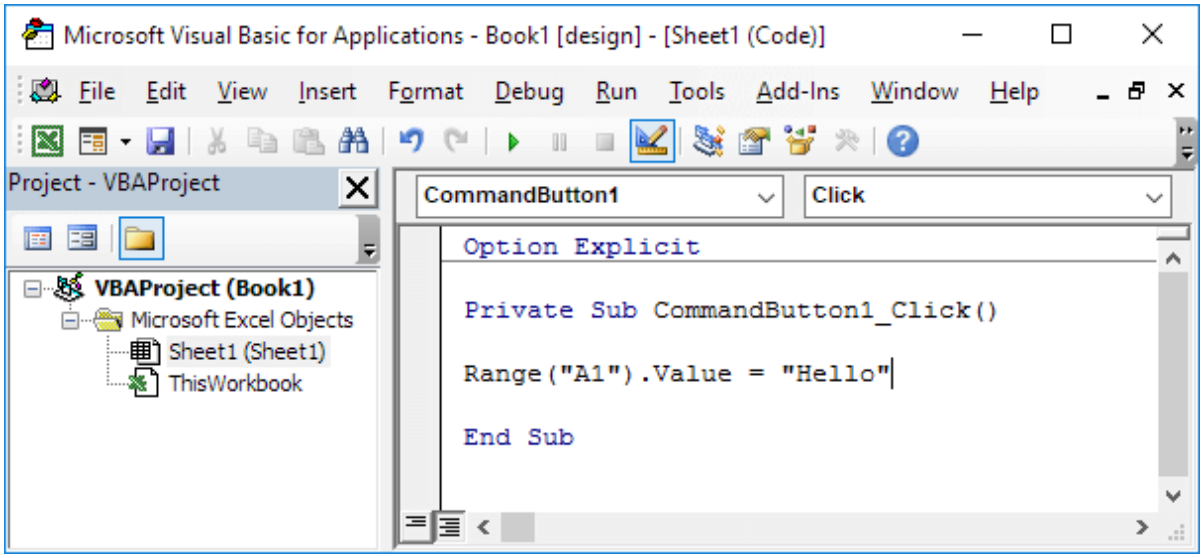
To assign a macro (one or more code lines) to the command button, execute the following steps.1. Rightclick CommandButton1 (make sure Design Mode is selected).



The Visual Basic Editor appears.

3. Place your cursor between Private Sub CommandButton1_Click() and End Sub.

4. Add the code line shown below.

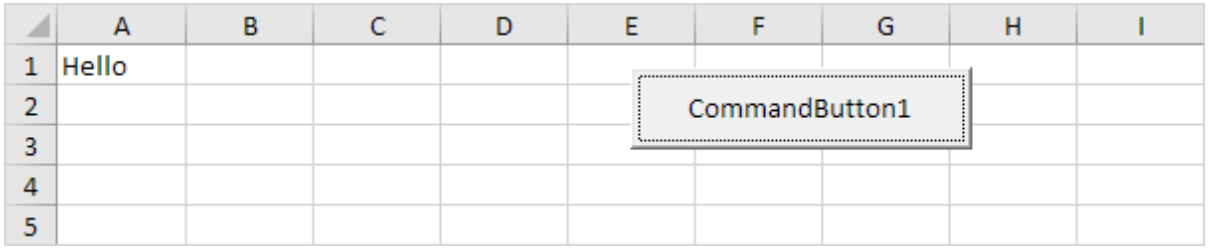


Note: the window on the left with the names Sheet1 (Sheet1) and ThisWorkbook is called the ProjectExplorer. If the Project Explorer is not visible, click View, Project Explorer. If the Code window for Sheet1 is not visible, click Sheet1 (Sheet1). You can ignore the Option Explicit statement for now.

5. Close the Visual Basic Editor.

6. Click the command button on the sheet (make sure Design Mode is deselected).

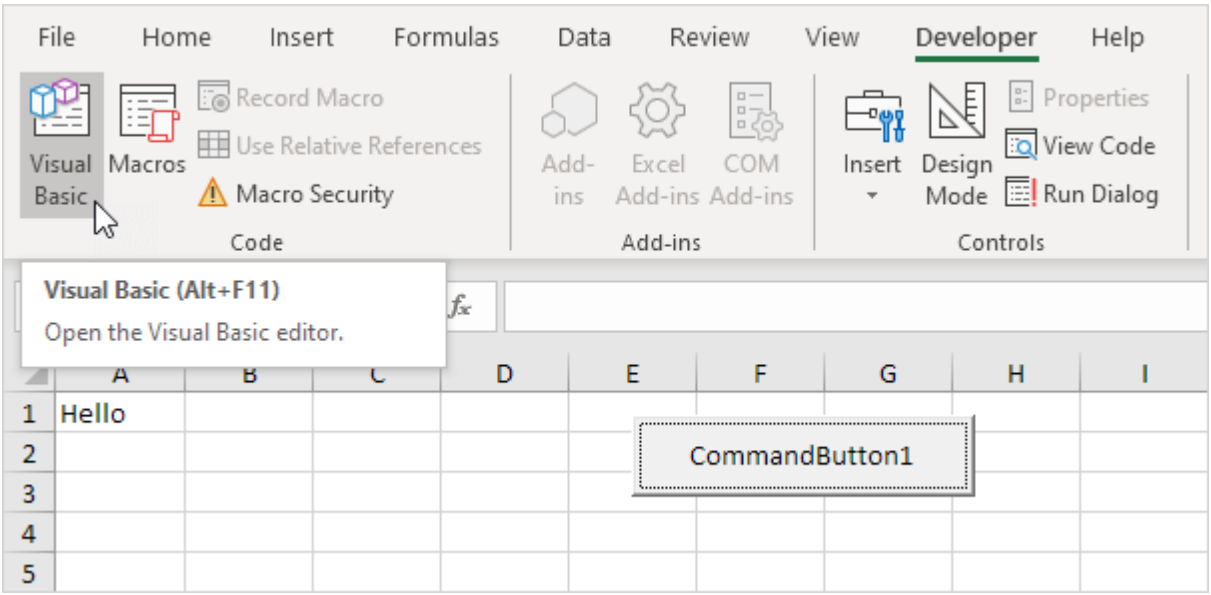
Result:



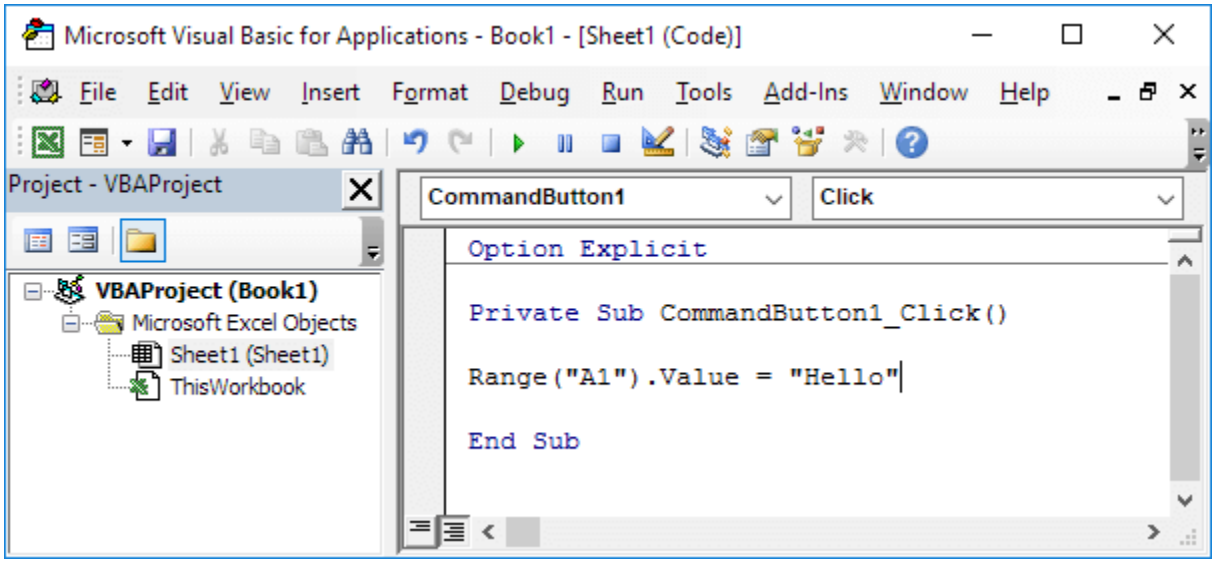
Congratulations. You've just created a macro in Excel!

Visual Basic Editor

To open the Visual Basic Editor, on the Developer tab, click Visual Basic.



The Visual Basic Editor appears.



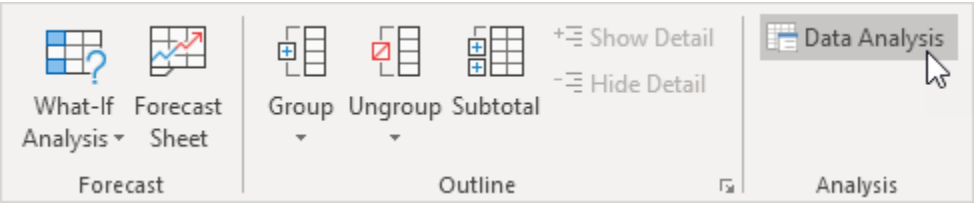
Descriptive Statistics

You can use the Analysis Toolpak add-in to generate descriptive statistics. For example, you may have thescores of 14 participants for a test.

	A	B
1	Scores	
2	82	
3	93	
4	91	
5	69	
6	96	
7	61	
8	88	
9	58	
10	59	
11	100	
12	93	
13	71	
14	78	
15	98	
16		

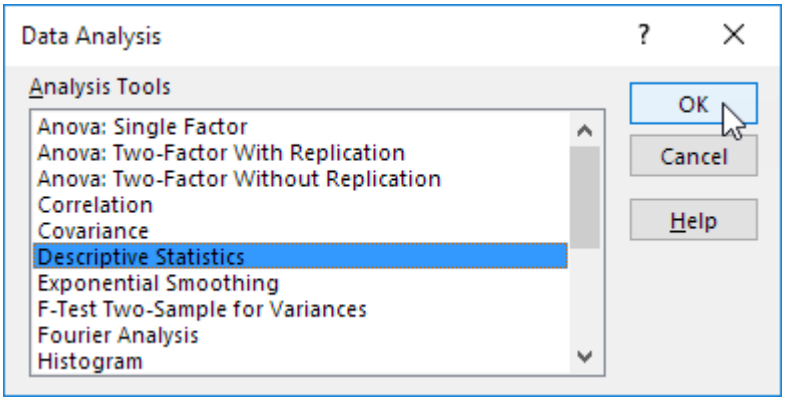
To generate descriptive statistics for these scores, execute the following steps.

1. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

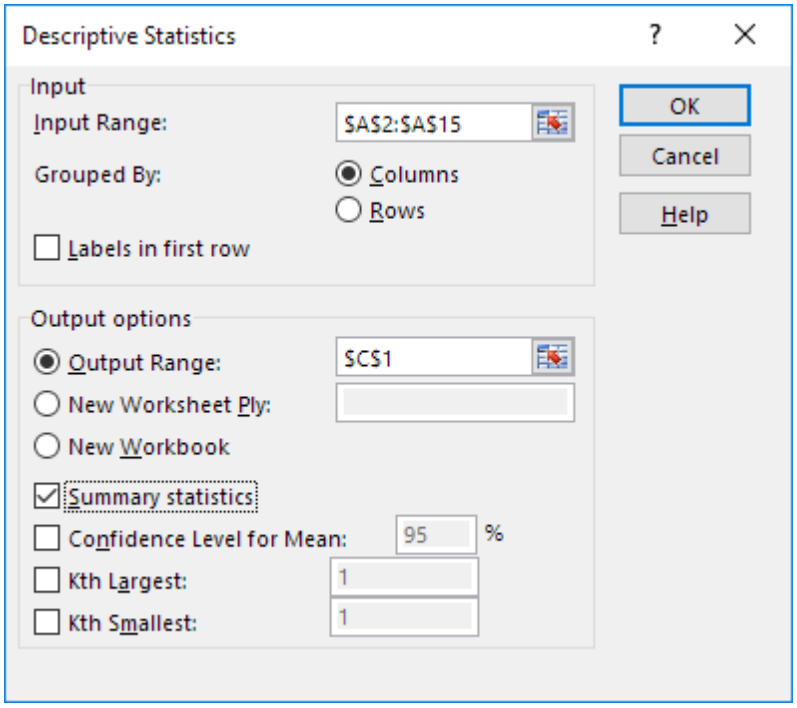
2. Select Descriptive Statistics and click OK.



3. Select the range A2:A15 as the Input Range.

4. Select cell C1 as the Output Range.

5. Make sure Summary statistics is checked.



6. Click OK.

	A	B	C	D	E
1	Scores		*Column1*		
2	82				
3	93		Mean	81.21428571	
4	91		Standard Error	4.045318243	
5	69		Median	85	
6	96		Mode	93	
7	61		Standard Deviation	15.13619489	
8	88		Sample Variance	229.1043956	
9	58		Kurtosis	-1.426053506	
10	59		Skewness	-0.402108004	
11	100		Range	42	
12	93		Minimum	58	
13	71		Maximum	100	
14	78		Sum	1137	
15	98		Count	14	
16					

Anova

This example teaches you how to perform a single factor ANOVA (analysis of variance) in Excel. A single factor or one-way ANOVA is used to test the null hypothesis that the means of several populations are all equal.Below you can find the salaries of people who have a degree in economics, medicine or history.

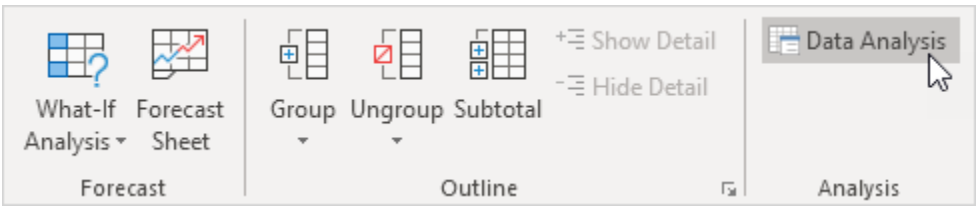
H0: μ1 = μ2 = μ3

H1: at least one of the means is different.

	A	B	C	D
1	economics	medicine	history	
2	42	69	35	
3	53	54	40	
4	49	58	53	
5	53	64	42	
6	43	64	50	
7	44	55	39	
8	45	56	55	
9	52		39	
10	54		40	
11				

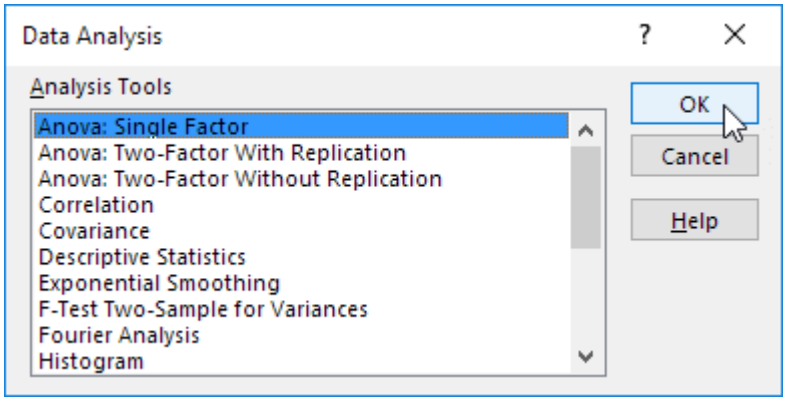
To perform a single factor ANOVA, execute the following steps.

1. On the Data tab, in the Analysis group, click Data Analysis.



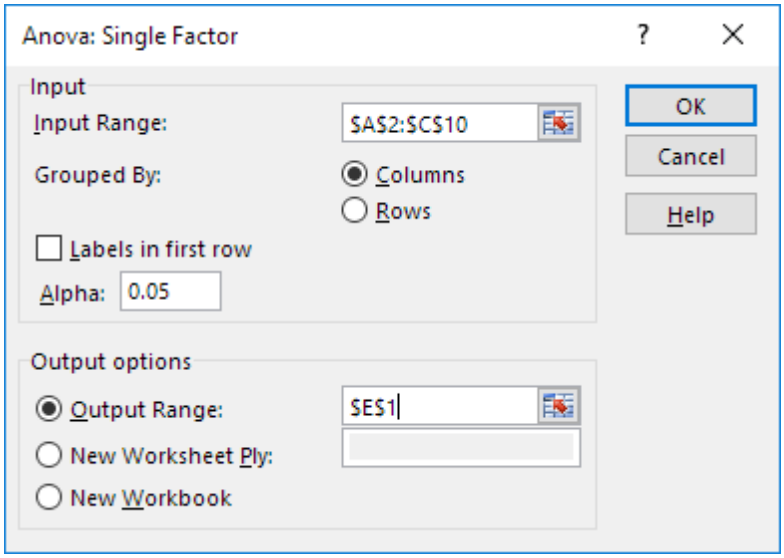
Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

2. Select Anova: Single Factor and click OK.



3. Click in the Input Range box and select the range A2:C10.

4. Click in the Output Range box and select cell E1.



5. Click

OK.

Result:

E	F	G	H	I	J	K
Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1	9	435	48.33333	23.5		
Column 2	7	420	60	32.33333		
Column 3	9	393	43.66667	50.5		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1085.84	2	542.92	15.19623	7.16E-05	3.443357
Within Groups	786	22	35.72727			
Total	1871.84	24				

Conclusion: if F > F crit, we reject the null hypothesis. This is the case, 15.196 > 3.443. Therefore, we reject the null hypothesis. The means of the three populations are not all equal. At least one of the means is different. However, the ANOVA does not tell you where the difference lies. You need a t-Test to test each pair of means.

F-Test

This example teaches you how to perform an F-Test in Excel. The F-Test is used to test the null hypothesis that the variances of two populations are equal.

Below you can find the study hours of 6 female students and 5 male students.

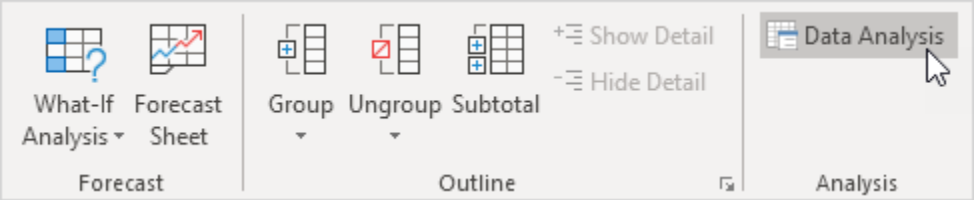
$H0: \sigma_1^2 = \sigma_2^2$

$H1: \sigma_1^2 \neq \sigma_2^2$

	A	B	C
1	Female	Male	
2	26	23	
3	25	30	
4	43	18	
5	34	25	
6	18	28	
7	52		
8			

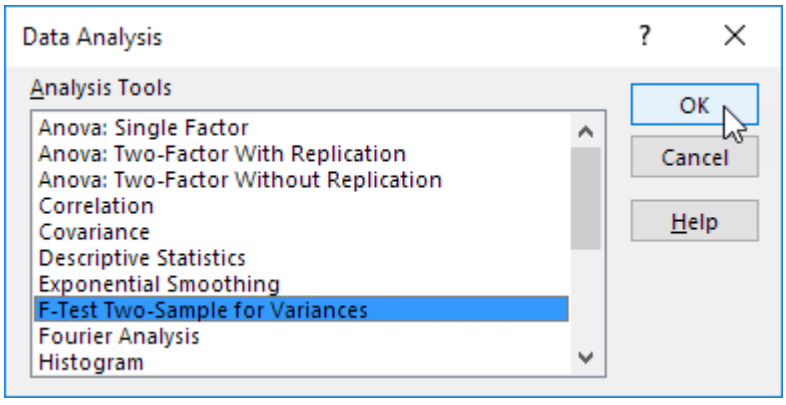
To perform an F-Test, execute the following steps.

1. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

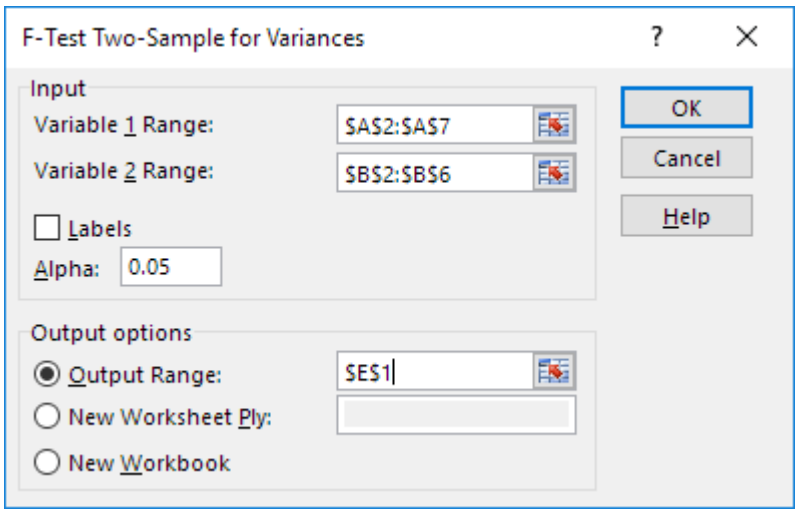
2. Select F-Test Two-Sample for Variances and click OK.



3. Click in the Variable 1 Range box and select the range A2:A7.

4. Click in the Variable 2 Range box and select the range B2:B6.

5. Click in the Output Range box and select cell E1.



6. Click

OK.

Result:

E	F	G
F-Test Two-Sample for Variances		
	Variable 1	*Variable 2*
Mean	33	24.8
Variance	160	21.7
Observations	6	5
df	5	4
F	7.373271889	
P(F<=f) one-tail	0.037888376	
F Critical one-tail	6.256056502	

Important: be sure that the variance of Variable 1 is higher than the variance of Variable 2. This is the case, 160 > 21.7. If not, swap your data. As a result, Excel calculates the correct F value, which is the ratio of Variance 1 to Variance 2 (F = 160 / 21.7 = 7.373).

Conclusion: if F > F Critical one-tail, we reject the null hypothesis. This is the case, 7.373 > 6.256. Therefore, we reject the null hypothesis. The variances of the two populations are unequal.

Go to Next Chapter: Create a Macro

t-Test

This example teaches you how to perform a t-Test in Excel. The t-Test is used to test the null hypothesis that the means of two populations are equal.

Below you can find the study hours of 6 female students and 5 male students.

H0: μ1 -

μ2 = 0 H1:

μ1 - μ2 ≠

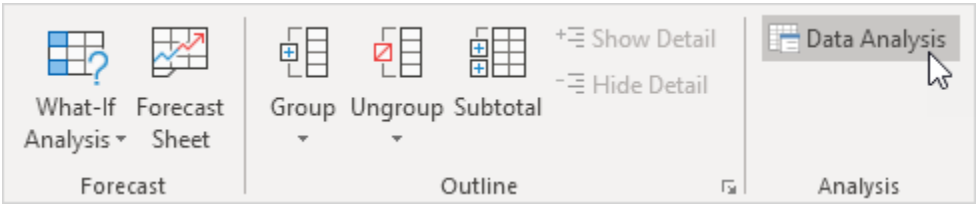
0

	A	B	C
1	Female	Male	
2	26	23	
3	25	30	
4	43	18	
5	34	25	
6	18	28	
7	52		
8			

To perform a t-Test, execute the following steps.

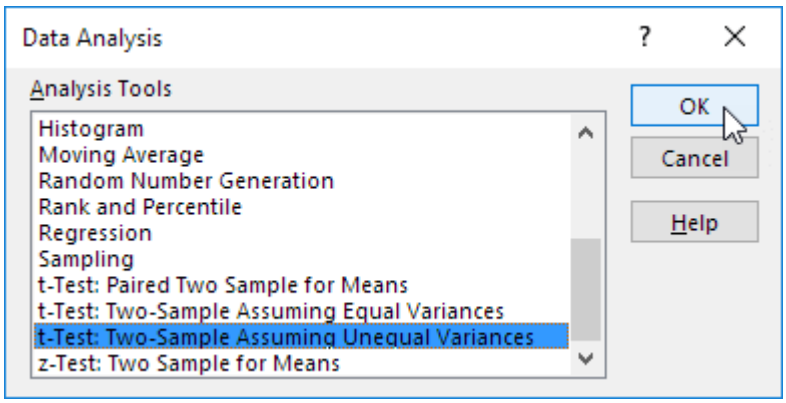
1. First, perform an F-Test to determine if the variances of the two populations are equal. This is not the case.

2. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

3. Select t-Test: Two-Sample Assuming Unequal Variances and click OK.

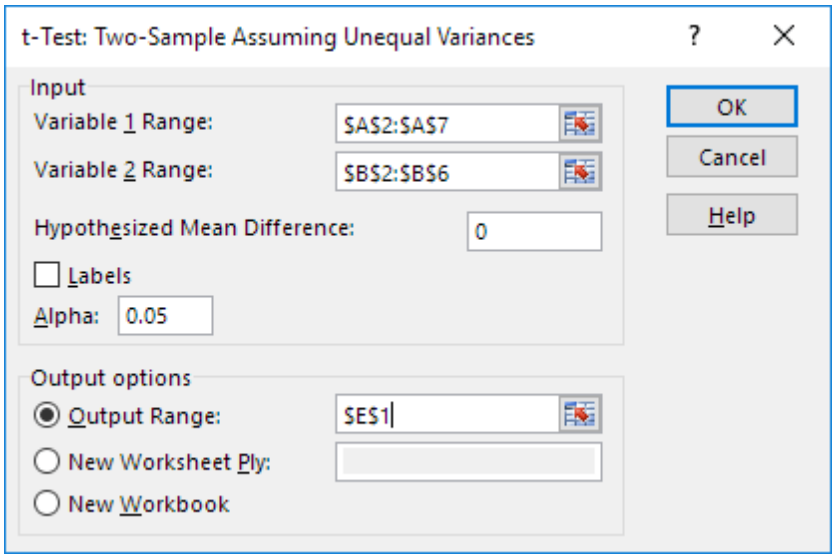


4. Click in the Variable 1 Range box and select the range A2:A7.

5. Click in the Variable 2 Range box and select the range B2:B6.

6. Click in the Hypothesized Mean Difference box and type 0 (H0: μ1 - μ2 = 0).

7. Click in the Output Range box and select cell E1.



8. Click

OK.

Result:

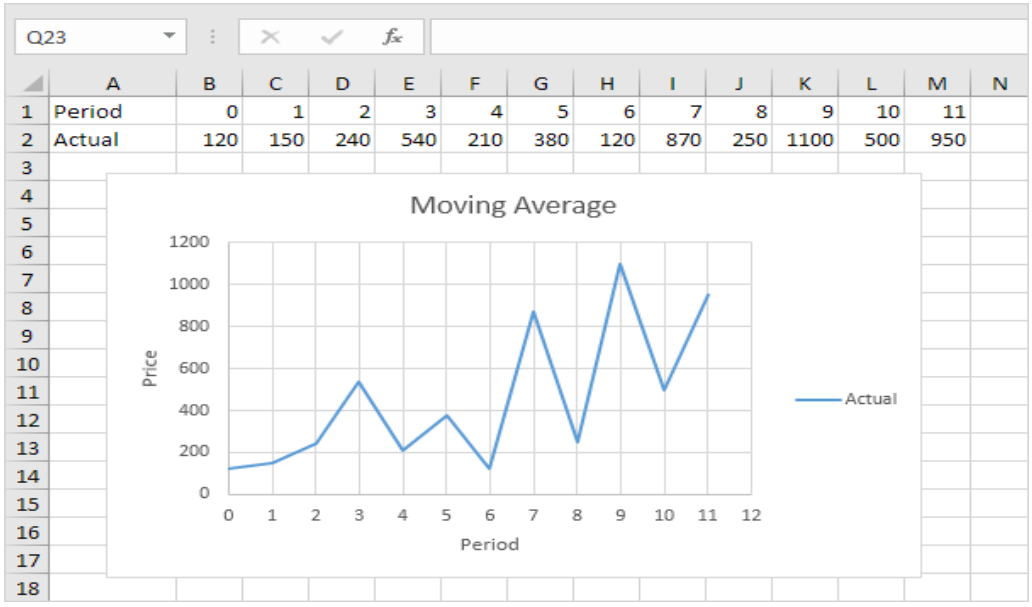
E	F	G
t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	*Variable 2*
Mean	33	24.8
Variance	160	21.7
Observations	6	5
Hypothesized Mean Difference	0	
df	7	
t Stat	1.47260514	
P(T<=t) one-tail	0.092170202	
t Critical one-tail	1.894578605	
P(T<=t) two-tail	0.184340405	
t Critical two-tail	2.364624252	

Conclusion: We do a two-tail test (inequality). lf t Stat < -t Critical two-tail or t Stat > t Critical two-tail, we reject the null hypothesis. This is not the case, -2.365 < 1.473 < 2.365. Therefore, we do not reject thenull hypothesis. The observed difference between the sample means (33 - 24.8) is not convincing enoughto say that the average number of study hours between female and male students differ significantly.

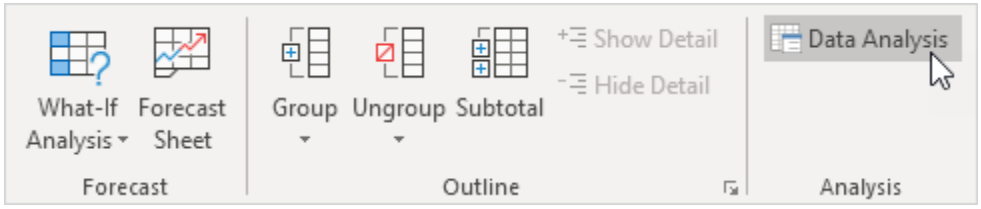
Moving Average

This example teaches you how to calculate the moving average of a time series in Excel. A movingaverage is used to smooth out irregularities (peaks and valleys) to easily recognize trends.

1. First, let's take a look at our time series.

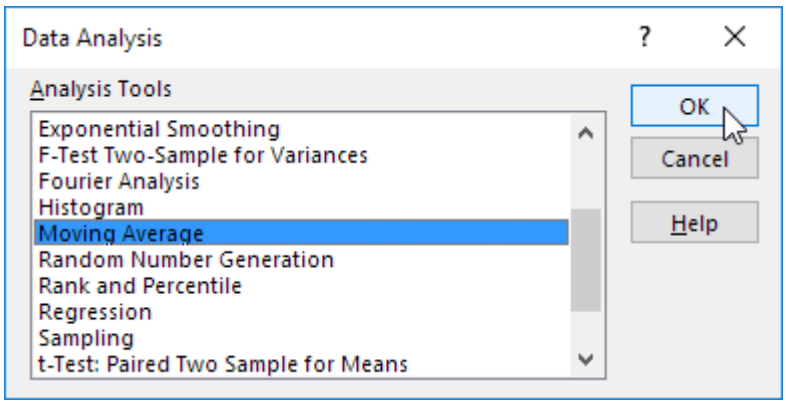


2. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

3. Select Moving Average and click OK.

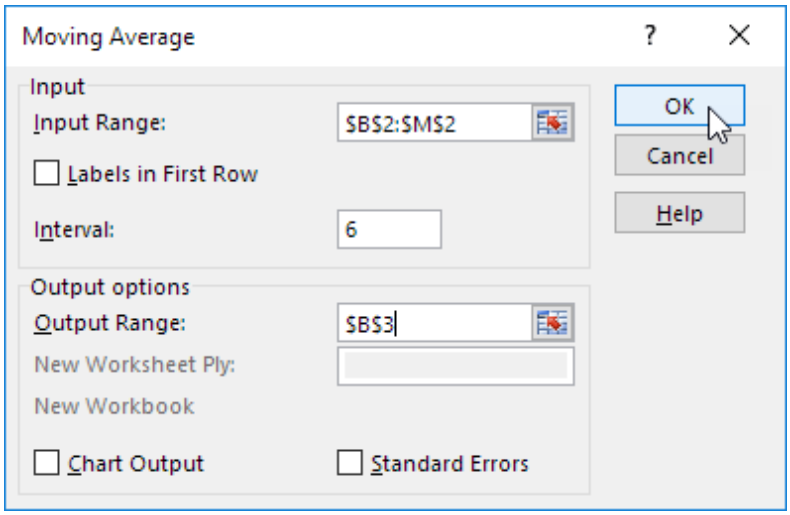


4. Click in the Input Range box and select the range B2:M2.

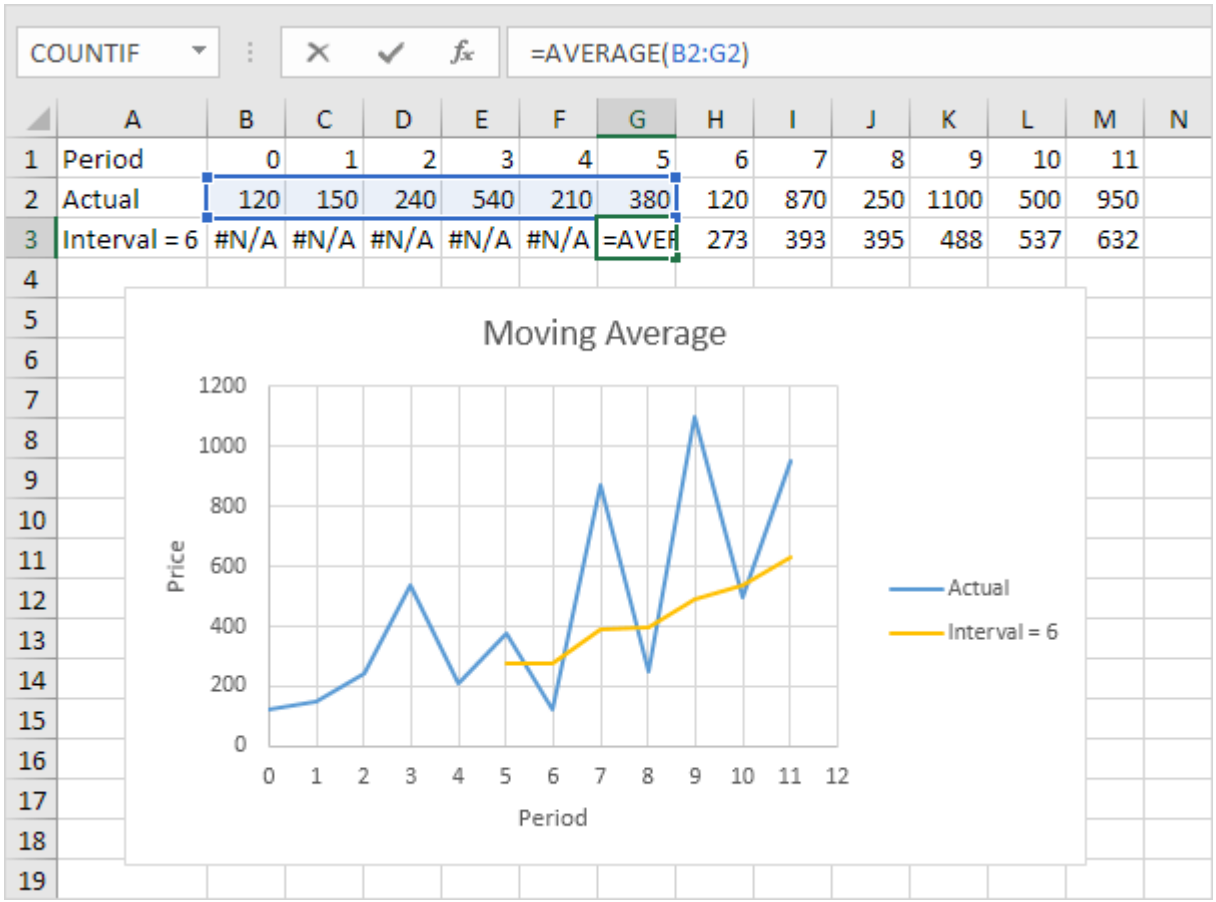
5. Click in the Interval box and type 6.

6. Click in the Output Range box and select cell B3.

7. Click OK.



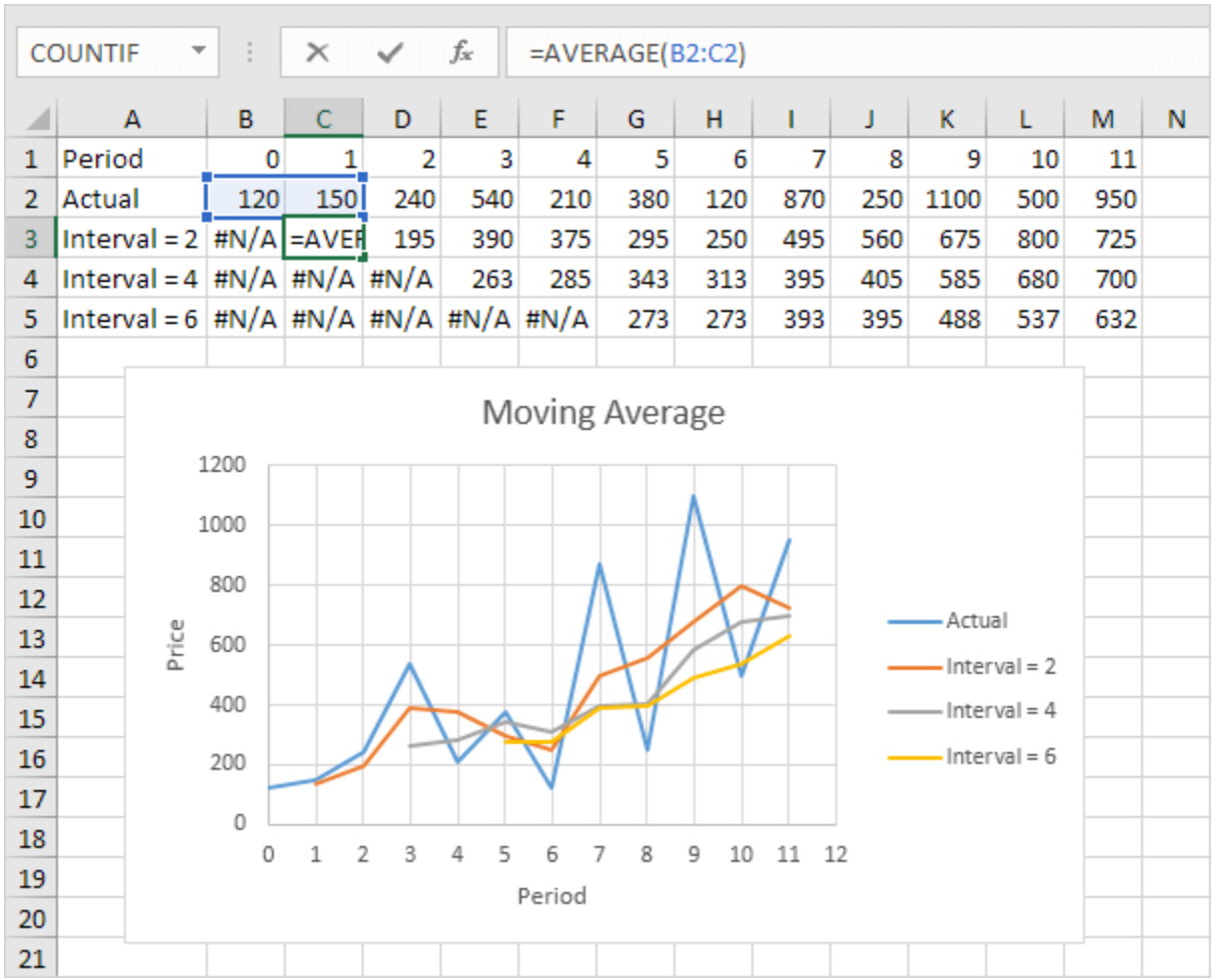
8. Plot a graph of these values.



Explanation: because we set the interval to 6, the moving average is the average

of the previous 5 data points and the current data point. As a result, peaks and valleys are smoothed out. The graph shows an increasing trend. Excel cannot calculate the moving average for the first 5 data points because there are not enough previous data points.

9. Repeat steps 2 to 8 for interval = 2 and interval = 4.



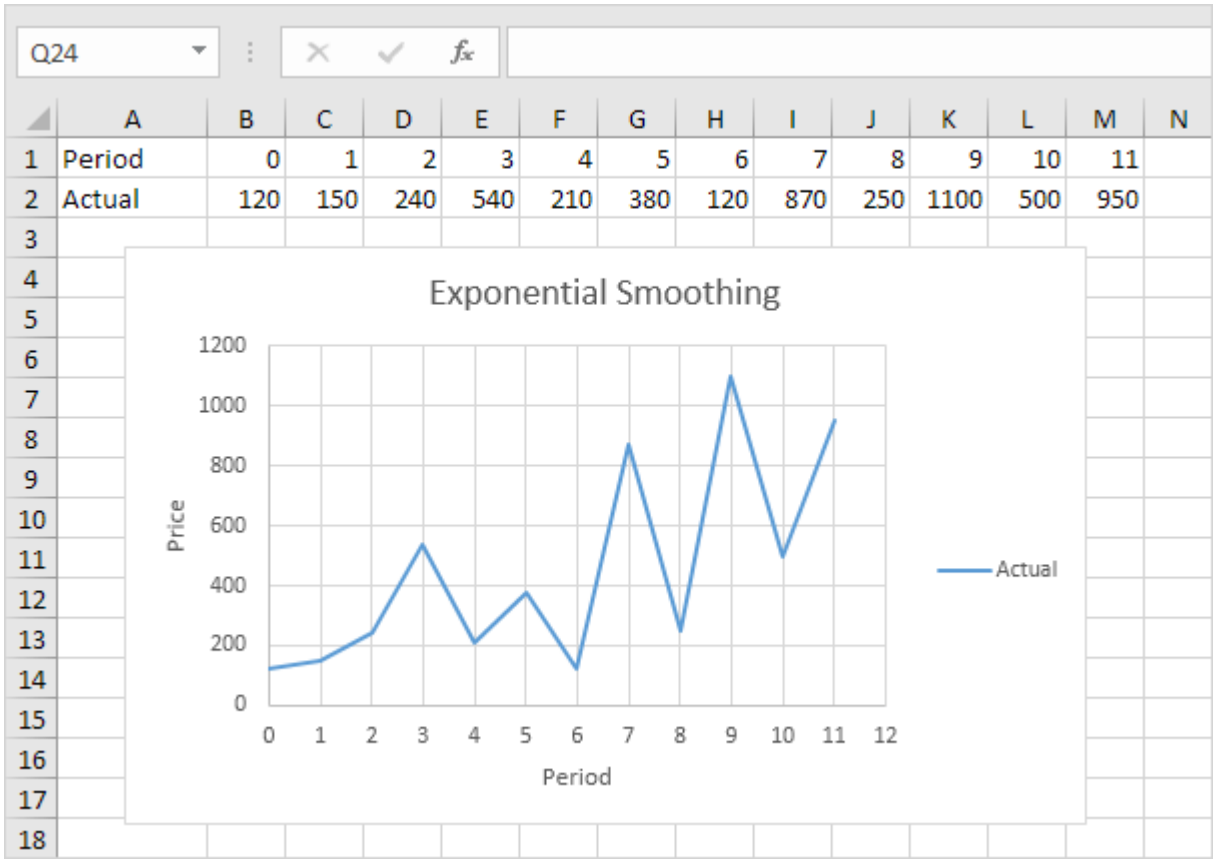
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Period	0	1	2	3	4	5	6	7	8	9	10	11
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950
3	Interval = 2	#N/A	=AVER	195	390	375	295	250	495	560	675	800	725
4	Interval = 4	#N/A	#N/A	#N/A	263	285	343	313	395	405	585	680	700
5	Interval = 6	#N/A	#N/A	#N/A	#N/A	#N/A	273	273	393	395	488	537	632

Conclusion: The larger the interval, the more the peaks and valleys are smoothed out. The smaller the interval, the closer the moving averages are to the actual data points.

Exponential Smoothing

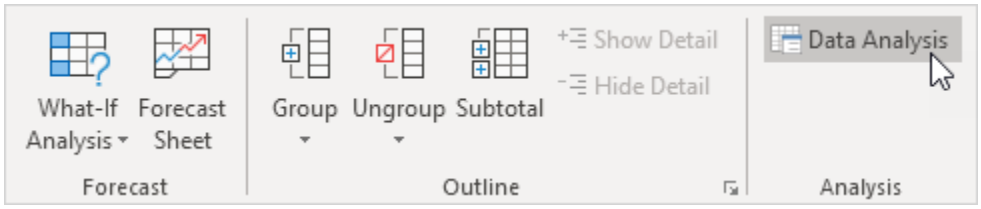
This example teaches you how to apply exponential smoothing to a time series in Excel. Exponential smoothing is used to smooth out irregularities (peaks and valleys) to easily recognize trends.

1. First, let's take a look at our time series.



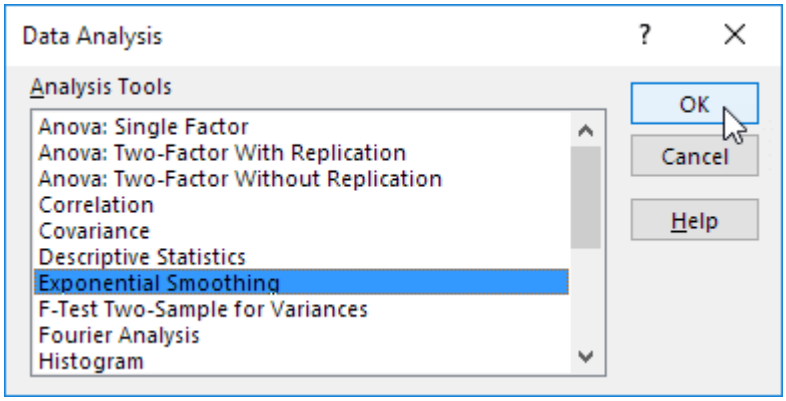
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Period	0	1	2	3	4	5	6	7	8	9	10	11
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950

2. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

3. Select Exponential Smoothing and click OK.

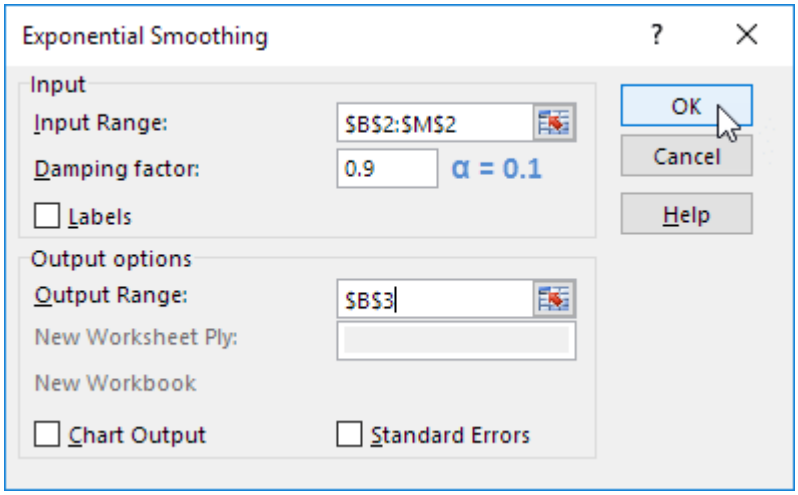


4. Click in the Input Range box and select the range B2:M2.

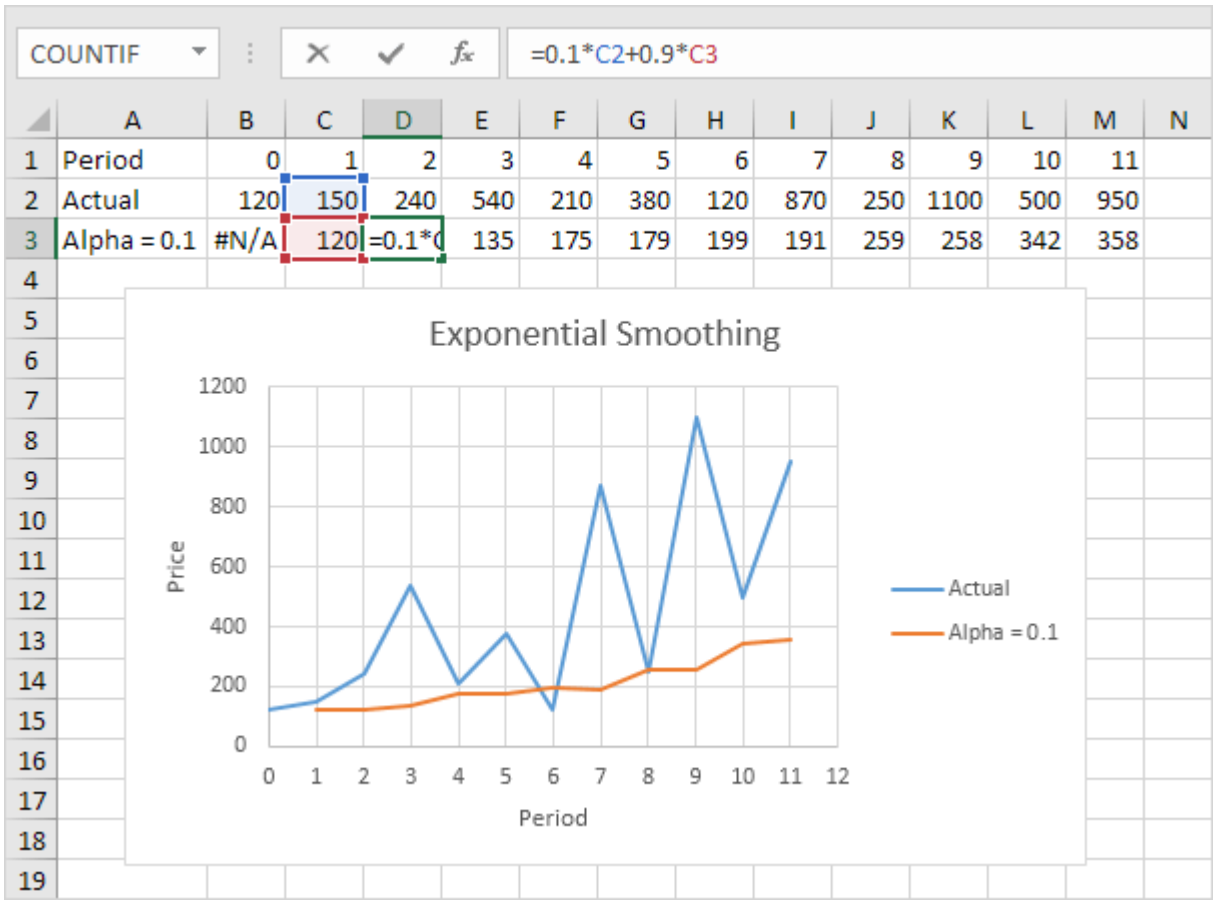
5. Click in the Damping factor box and type 0.9. Literature often talks about the smoothing constant α(alpha). The value (1- α) is called the damping factor.

6. Click in the Output Range box and select cell B3.

7. Click OK.



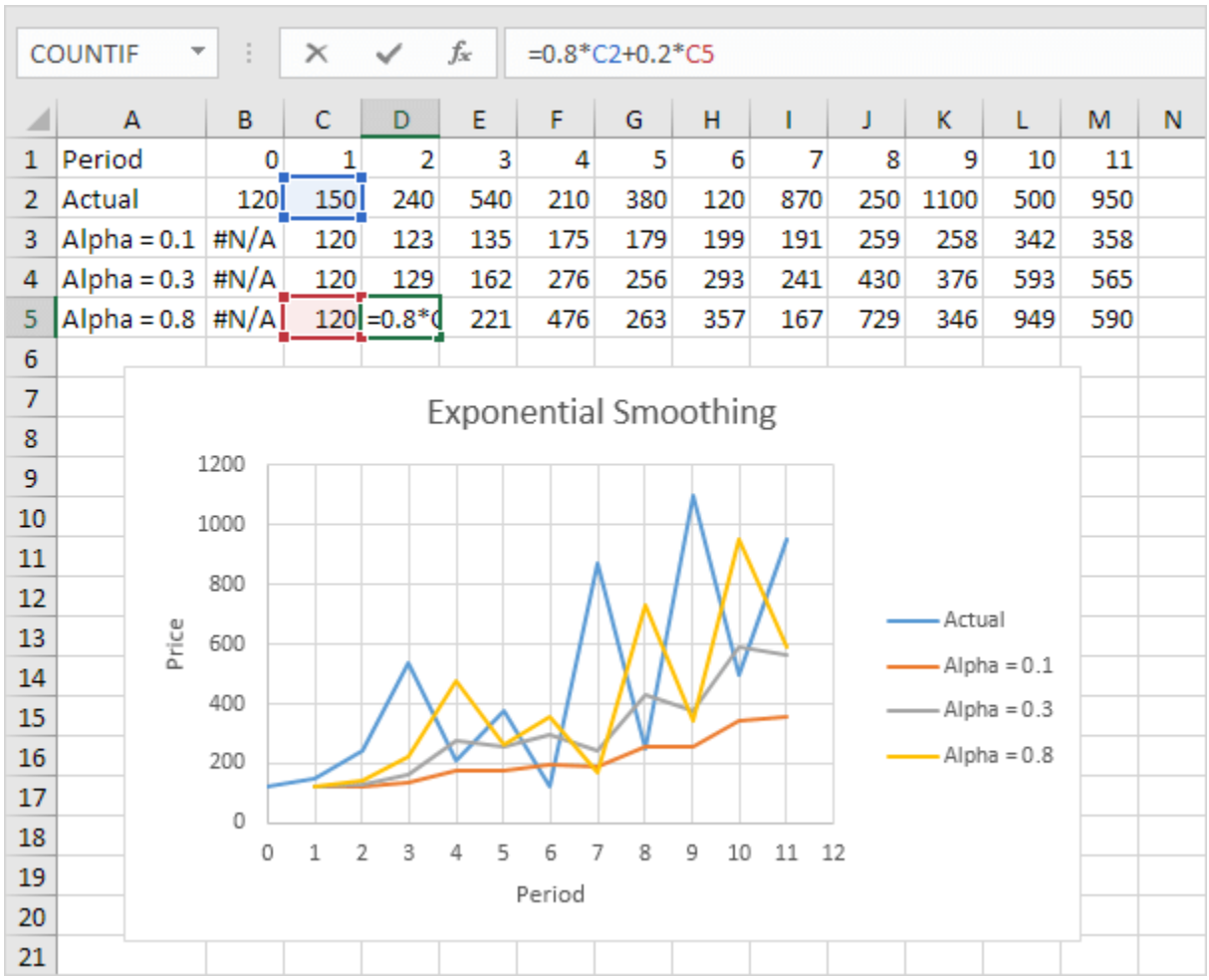
8. Plot a graph of these values.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Period	0	1	2	3	4	5	6	7	8	9	10	11
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950
3	Alpha = 0.1	#N/A	120	=0.1*C	135	175	179	199	191	259	258	342	358

Explanation: because we set alpha to 0.1, the previous data point is given a relatively small weight while the previous smoothed value is given a large weight (i.e. 0.9). As a result, peaks and valleys are smoothed out. The graph shows an increasing trend. Excel cannot calculate the smoothed value for the first data pointbecause there is no previous data point. The smoothed value for the second data point equals the previous data point.

9. Repeat steps 2 to 8 for alpha = 0.3 and alpha = 0.8.



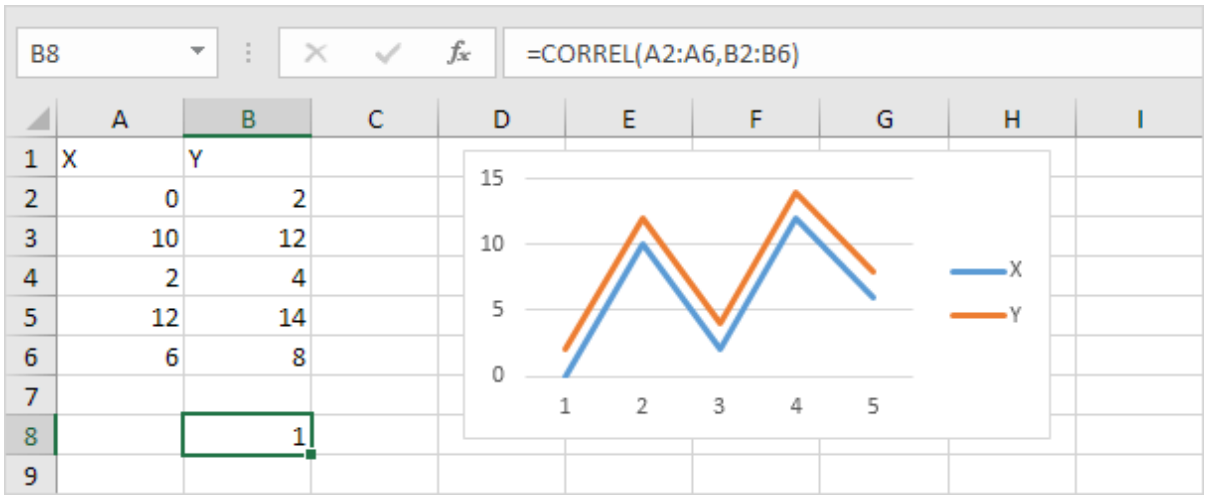
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Period	0	1	2	3	4	5	6	7	8	9	10	11
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950
3	Alpha = 0.1	#N/A	120	123	135	175	179	199	191	259	258	342	358
4	Alpha = 0.3	#N/A	120	129	162	276	256	293	241	430	376	593	565
5	Alpha = 0.8	#N/A	120	=0.8*C	221	476	263	357	167	729	346	949	590

Conclusion: The smaller alpha (larger the damping factor), the more the peaks and valleys are smoothed out. The larger alpha (smaller the damping factor), the closer the smoothed values are to the actual data points.

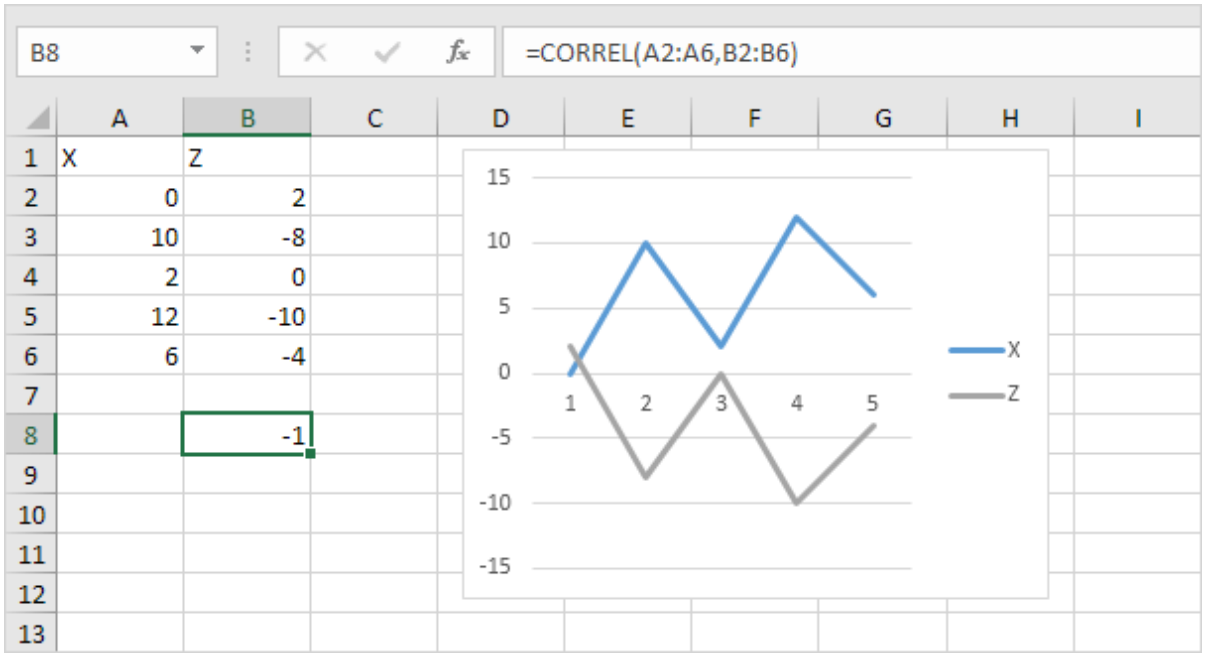
Correlation

The correlation coefficient (a value between -1 and +1) tells you how strongly two variables are related to each other. We can use the CORREL function or the Analysis Toolpak add-in in Excel to find the correlation coefficient between two variables.

- A correlation coefficient of +1 indicates a perfect positive correlation. As variable X increases, variable Y increases. As variable X decreases, variable Y decreases.



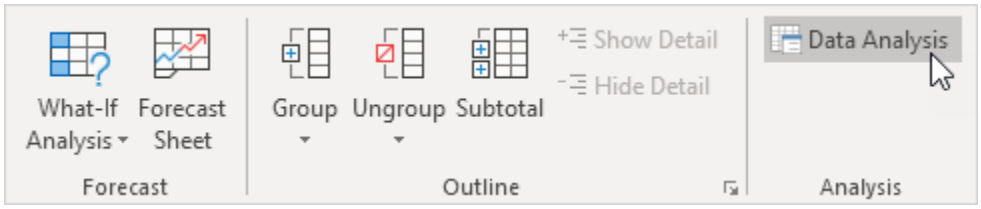
- A correlation coefficient of -1 indicates a perfect negative correlation. As variable X increases, variable Z decreases. As variable X decreases, variable Z increases.



- A correlation coefficient near 0 indicates no correlation.

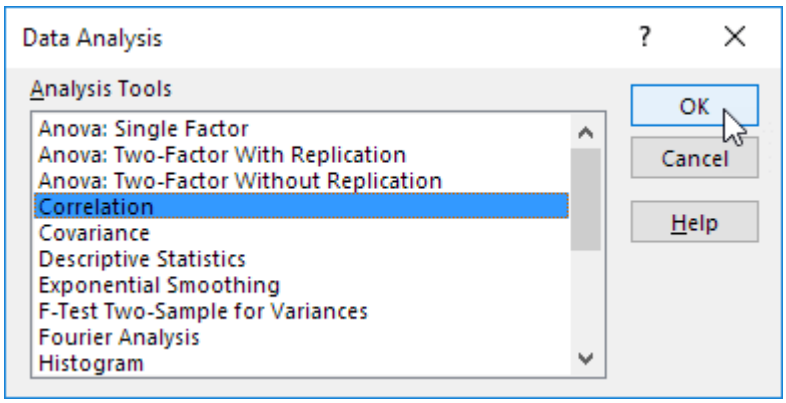
To use the Analysis Toolpak add-in in Excel to quickly generate correlation coefficients between multiple variables, execute the following steps.

1. On the Data tab, in the Analysis group, click Data Analysis.

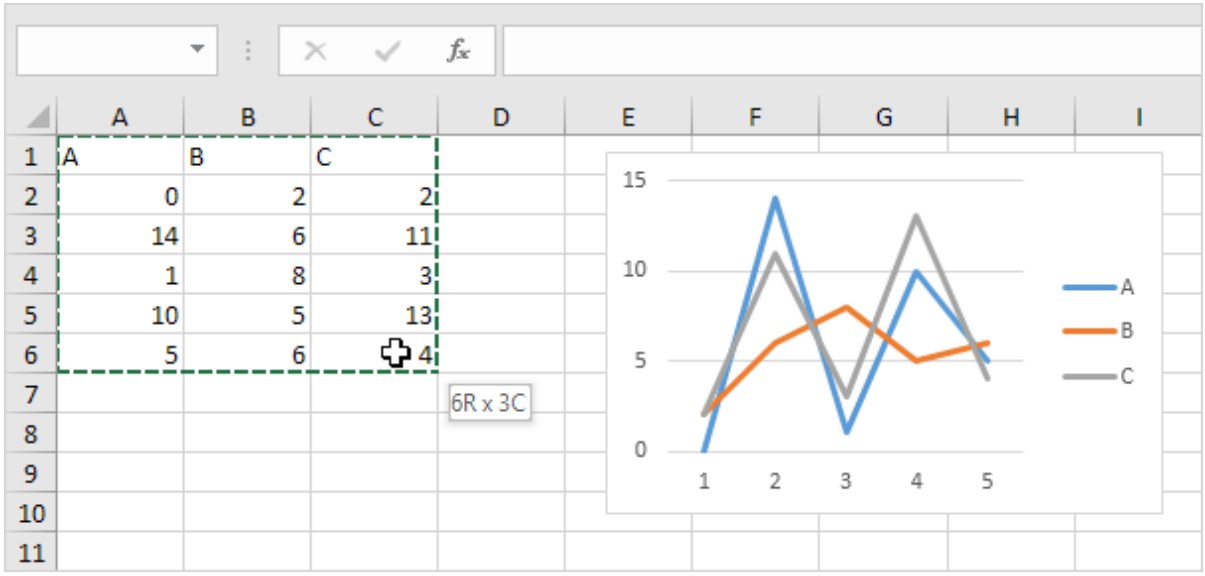


Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

2. Select Correlation and click OK.



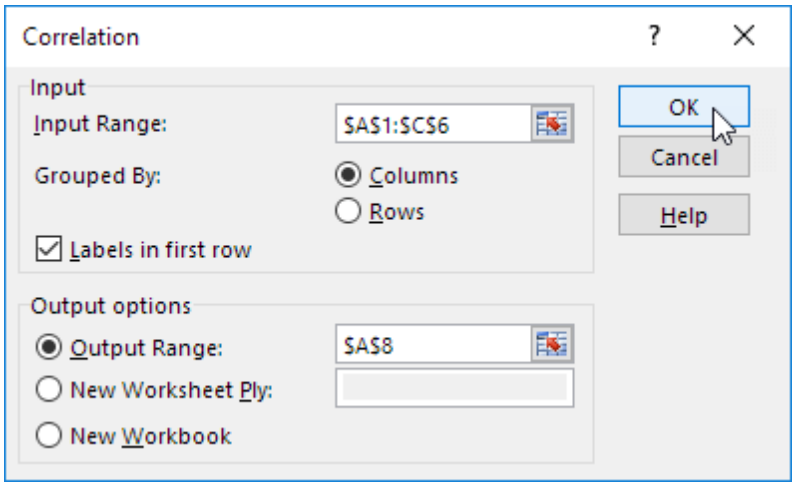
3. For example, select the range A1:C6 as the Input Range.



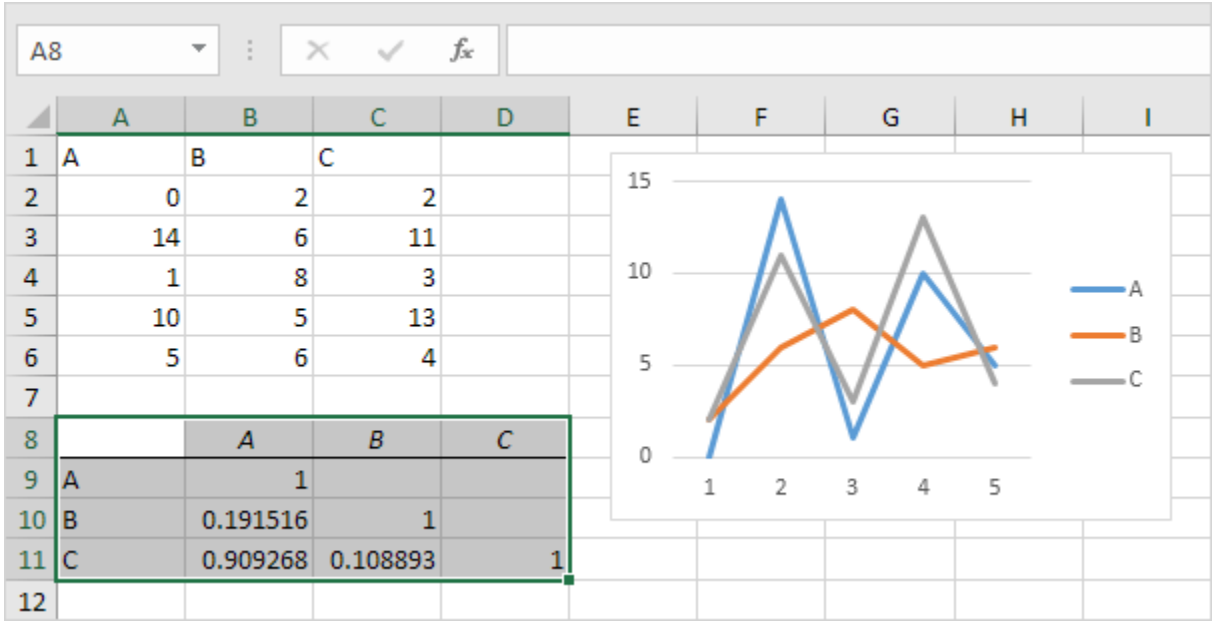
4. Check Labels in first row.

5. Select cell A8 as the Output Range.

6. Click OK.



Result.



Conclusion: variables A and C are positively correlated (0.91). Variables A and B are not correlated (0.19). Variables B and C are also not correlated (0.11) . You can verify these conclusions by looking at the graph.

Regression

R Square | Significance F and P-Values | Coefficients | Residuals

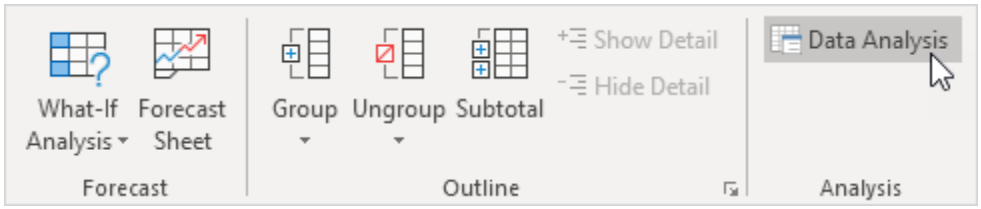
This example teaches you how to run a linear regression analysis in Excel and

how to interpret the Summary Output.

Below you can find our data. The big question is: is there a relation between Quantity Sold (Output) and Price and Advertising (Input). In other words: can we predict Quantity Sold if we know Price and Advertising?

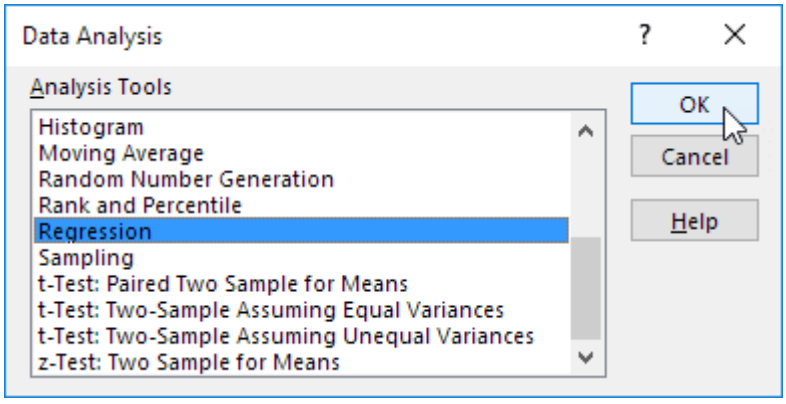
	A	B	C	D
1	Quantity Sold	Price	Advertising	
2	8500	$2	$2,800	
3	4700	$5	$200	
4	5800	$3	$400	
5	7400	$2	$500	
6	6200	$5	$3,200	
7	7300	$3	$1,800	
8	5600	$4	$900	
9				

1. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

2. Select Regression and click OK.



3. Select the Y Range (A1:A8). This is the predictor variable (also called dependent variable).

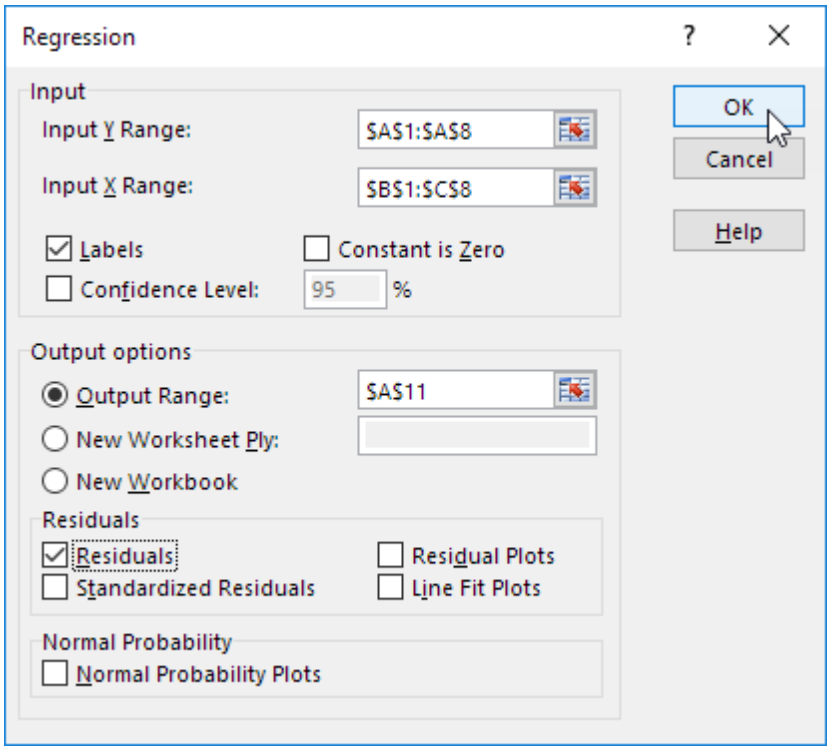
4. Select the X Range(B1:C8). These are the explanatory variables (also called independent variables). These columns must be adjacent to each other.

5. Check Labels.

6. Click in the Output Range box and select cell A11.

7. Check Residuals.

8. Click OK.



Excel produces the following Summary Output (rounded to 3 decimal places).

R Square

R Square equals 0.962, which is a very good fit. 96% of the variation in

Quantity Sold is explained by the independent variables Price and Advertising.

The closer to 1, the better the regression line (read on) fits the data.

11	SUMMARY OUTPUT	
12		
13	*Regression Statistics*	
14	Multiple R	0.981
15	R Square	0.962
16	Adjusted R Square	0.943
17	Standard Error	310.524
18	Observations	7
19		

Significance F and P-values

To check if your results are reliable (statistically significant), look at Significance F (0.001). If this value is less than 0.05, you're OK. If Significance F is greater than 0.05, it's probably better to stop using this set of independent variables. Delete a variable with a high P-value (greater than 0.05) and rerun the regression until Significance F drops below 0.05.

Most or all P-values should be below below 0.05. In our example this is the case. (0.000, 0.001 and 0.005).

20	ANOVA						
21		*df*	*SS*	*MS*	*F*	*Significance F*	
22	Regression	2	9694299.568	4847149.784	50.269	0.001	
23	Residual	4	385700.432	96425.108			
24	Total	6	10080000.000				
25							
26		*Coefficients*	*Std Error*	*t Stat*	*P-values*	*Lower 95%*	*Upper 95%*
27	Intercept	8536.214	386.912	22.062	0.000	7461.975	9610.453
28	Price	-835.722	99.653	-8.386	0.001	-1112.404	-559.041
29	Advertising	0.592	0.104	5.676	0.005	0.303	0.882
30							

Coefficients

The regression line is: y = Quantity Sold = 8536.214 -835.722 * Price + 0.592 * Advertising. In other words, for each unit increase in price, Quantity Sold decreases with 835.722 units. For each unit increase in Advertising, Quantity Sold increases with 0.592 units. This is valuable information.

You can also use these coefficients to do a forecast. For example, if price equals $4 and Advertising equals

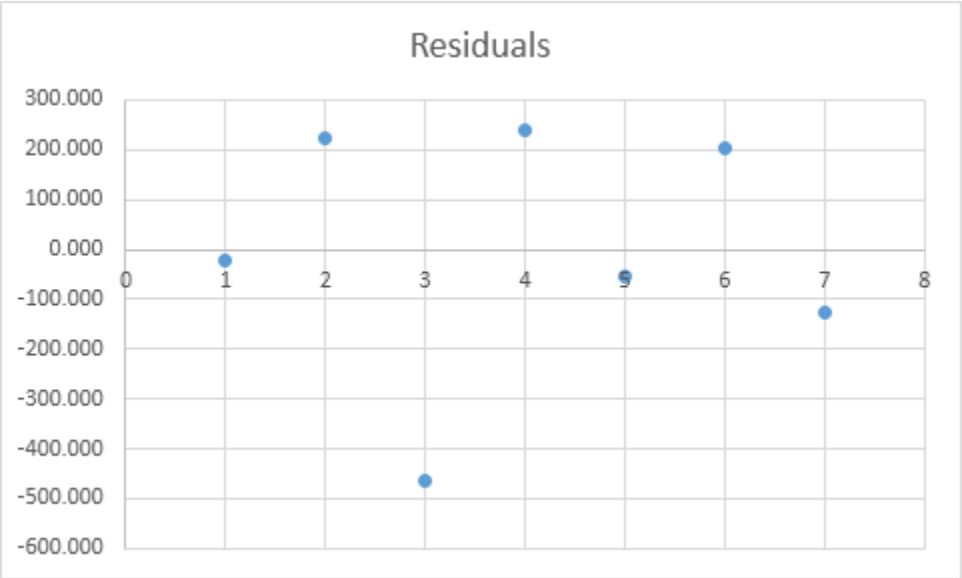
$3000, you might be able to achieve a Quantity Sold of 8536.214 -835.722 * 4 + 0.592 * 3000 = 6970.

Residuals

The residuals show you how far away the actual data points are fom the predicted data points (using the equation). For example, the first data point equals 8500. Using the equation, the predicted data point equals 8536.214 -835.722 * 2 + 0.592 * 2800 = 8523.009, giving a residual of 8500 - 8523.009 = -23.009.

33	RESIDUAL OUTPUT		
34			
35	*Observation*	*Predicted Quantity Sold*	*Residuals*
36	1	8523.009	-23.009
37	2	4476.048	223.952
38	3	6265.938	-465.938
39	4	7160.883	239.117
40	5	6252.733	-52.733
41	6	7095.058	204.942
42	7	5726.330	-126.330
43			

You can also create a scatter plot of these residuals.



Create a Macro

Developer Tab | Command Button | Assign a Macro | Visual Basic Editor

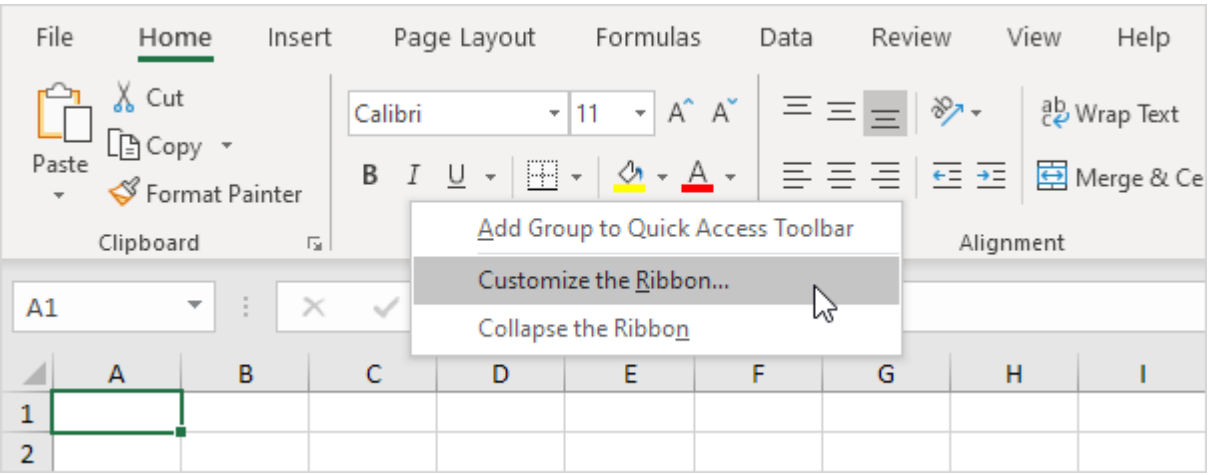
With Excel VBA you can automate tasks in Excel by writing so called macros.

In this chapter, learn howto create a simple macro which will be executed after clicking on a command button. First, turn on the Developer tab.

Developer Tab

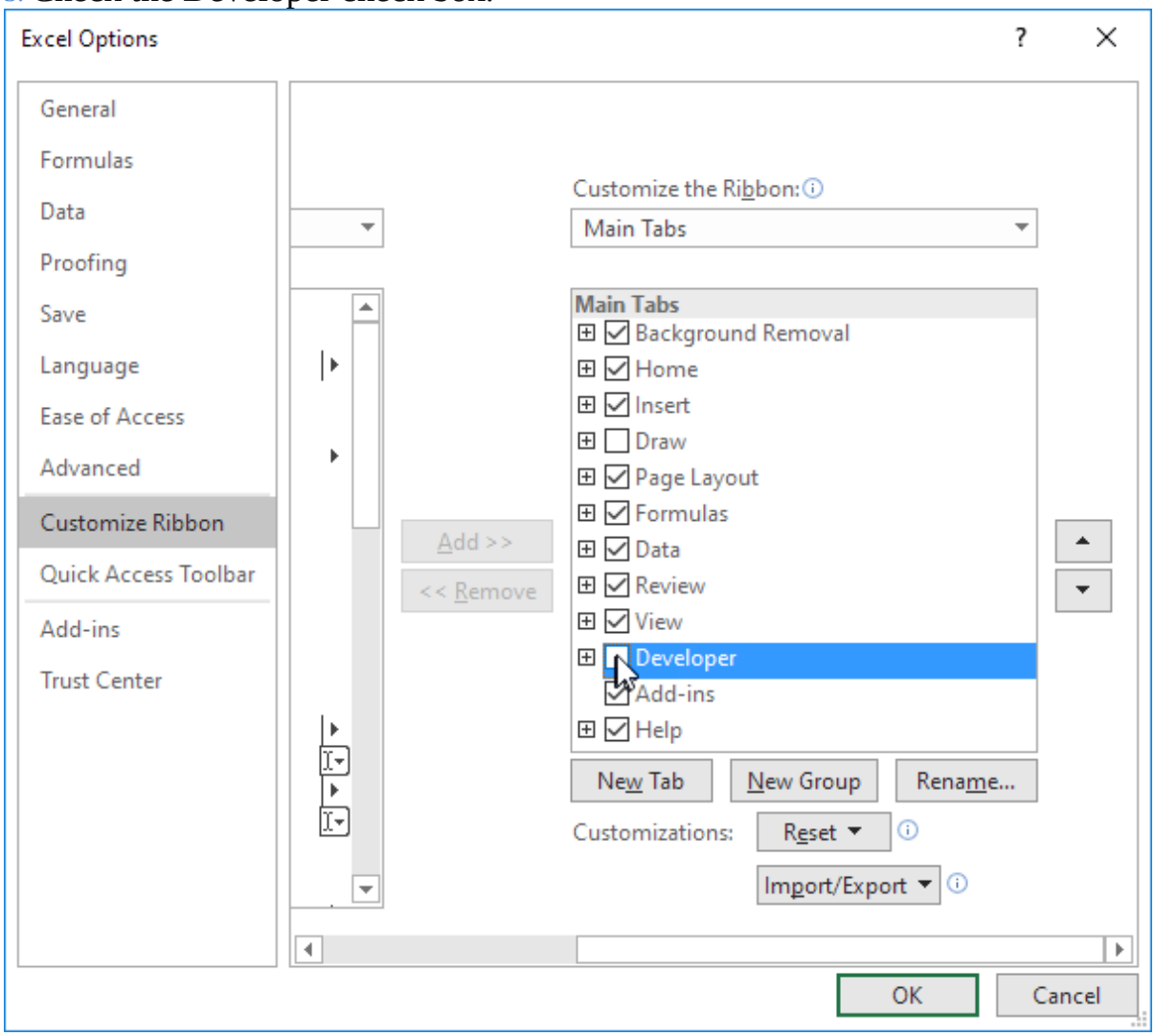
To turn on the Developer tab, execute the following steps.

1. Right click anywhere on the ribbon, and then click Customize the Ribbon.



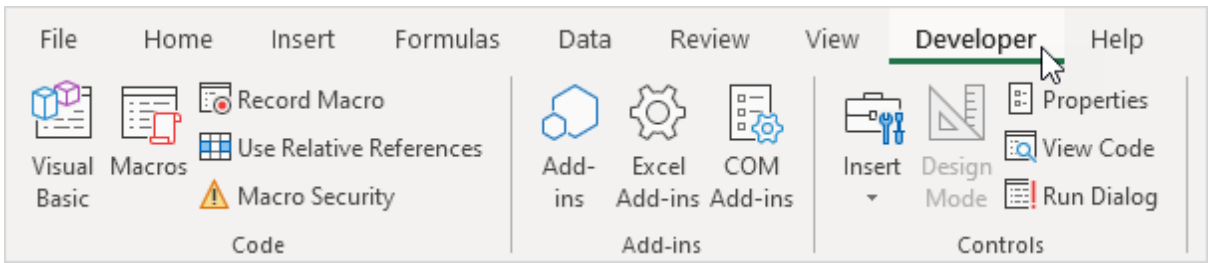
2. Under Customize the Ribbon, on the right side of the dialog box, select Main tabs (if necessary).

3. Check the Developer check box.



4. Click OK.

5. You can find the Developer tab next to the View tab.

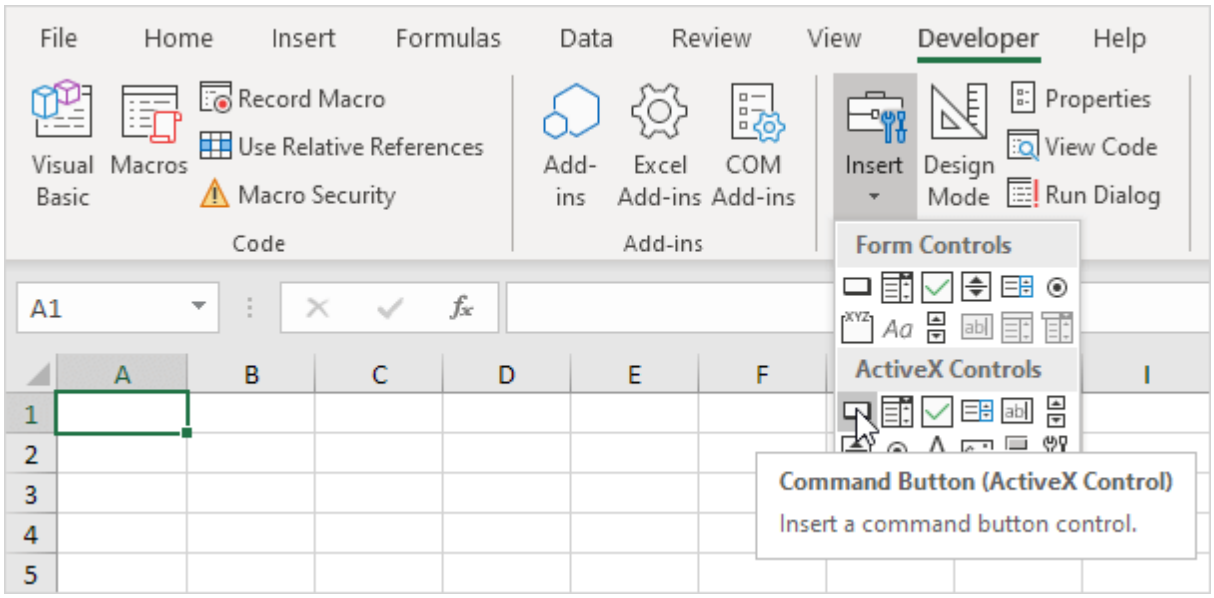


Command Button

To place a command button on your worksheet, execute the following steps.

1. On the Developer tab, click Insert.

2. In the ActiveX Controls group, click Command Button.



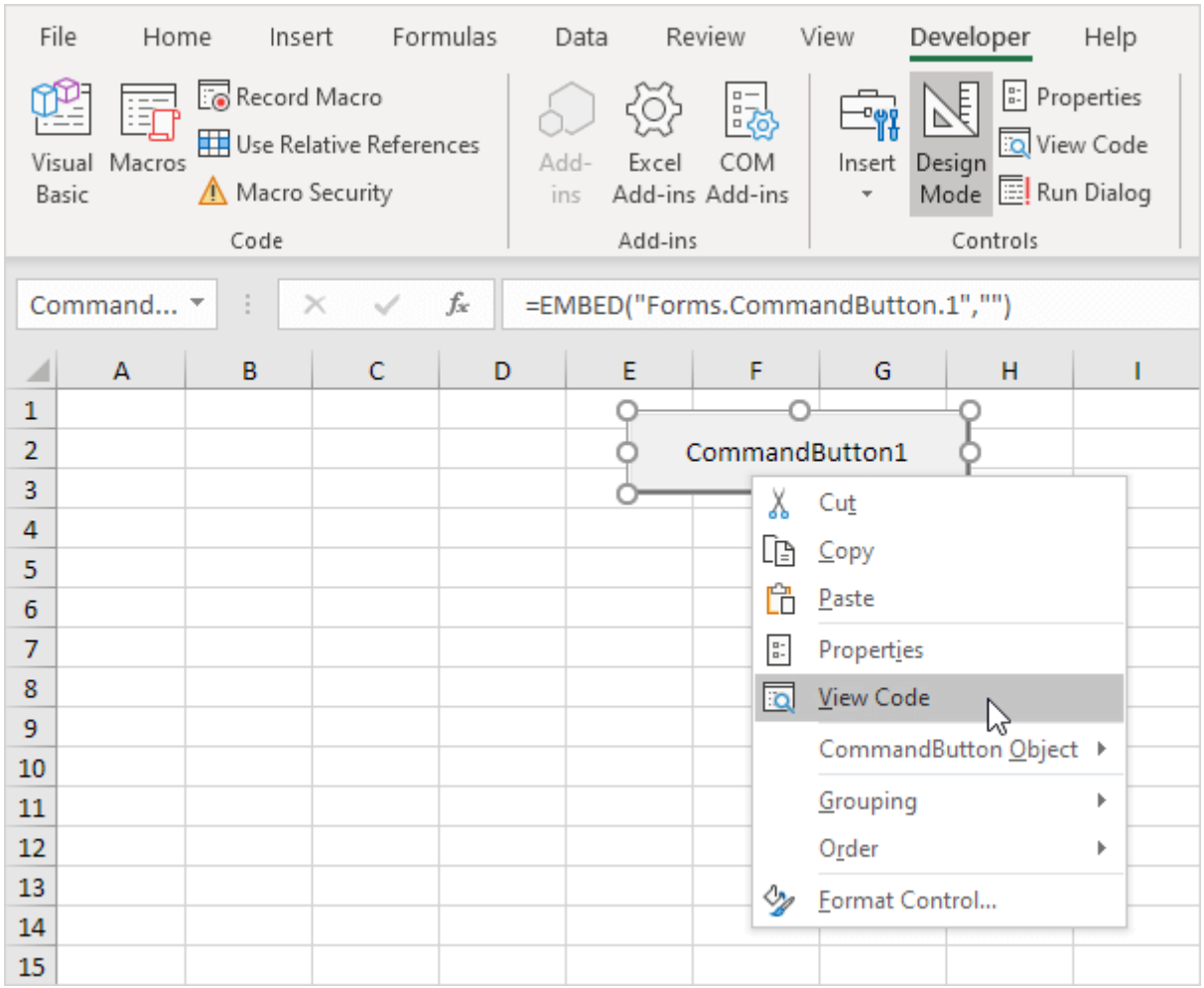
3. Drag a command button on your worksheet.

Assign a Macro

To assign a macro (one or more code lines) to the command button, execute the following steps.

1. Right click CommandButton1 (make sure Design Mode is selected).

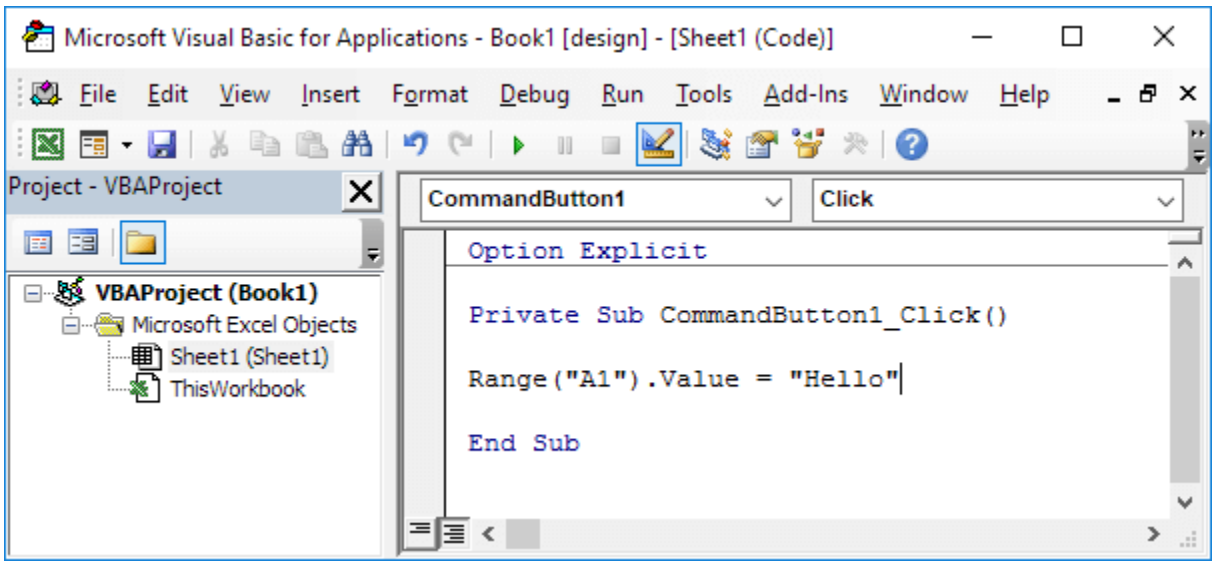
2. Click View Code.



The Visual Basic Editor appears.

3. Place your cursor between Private Sub CommandButton1_Click() and End Sub.

4. Add the code line shown below.



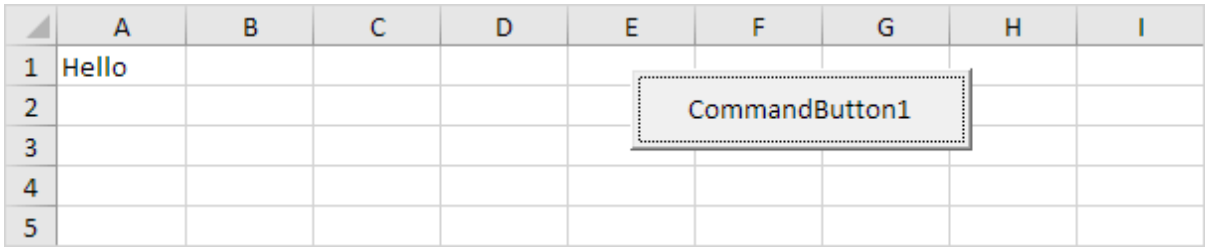
Note: the window on the left with the names Sheet1 (Sheet1) and

ThisWorkbook is called the ProjectExplorer. If the Project Explorer is not visible, click View, Project Explorer. If the Code window for Sheet1 is not visible, click Sheet1 (Sheet1). You can ignore the Option Explicit statement for now.

5. Close the Visual Basic Editor.

6. Click the command button on the sheet (make sure Design Mode is deselected).

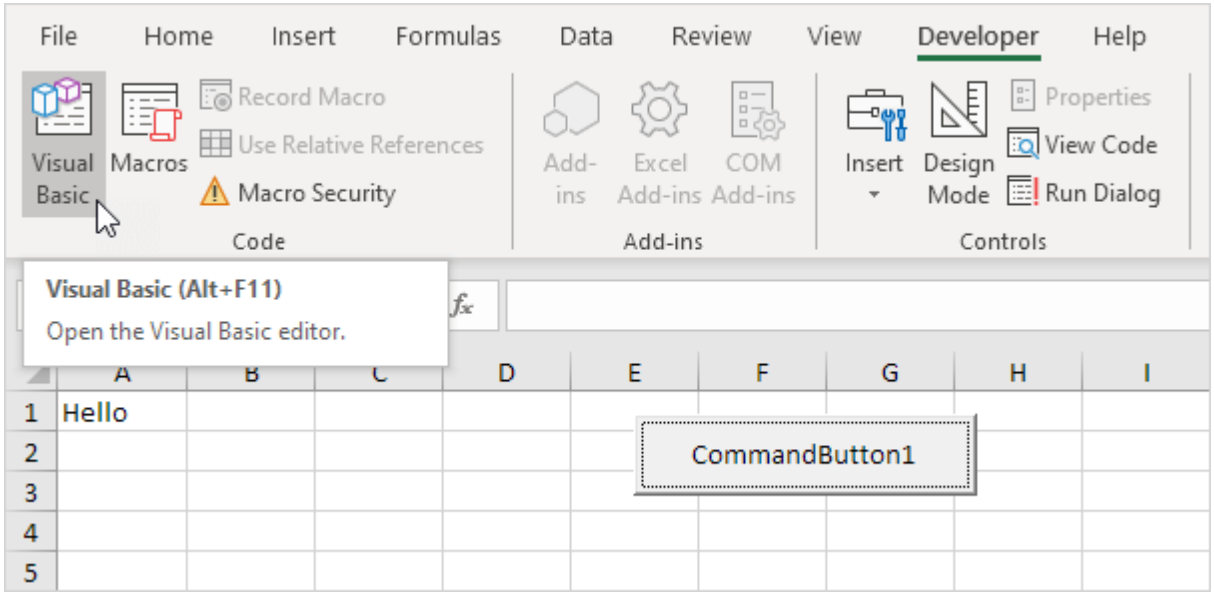
Result:



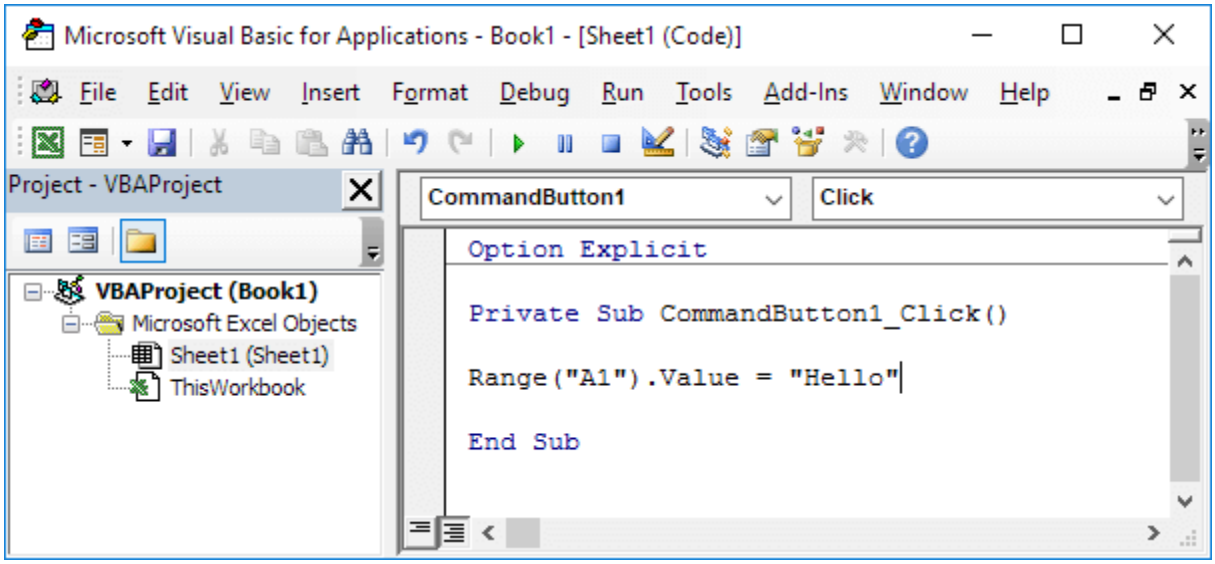
Congratulations. You've just created a macro in Excel!

Visual Basic Editor

To open the Visual Basic Editor, on the Developer tab, click Visual Basic.



The Visual Basic Editor appears.

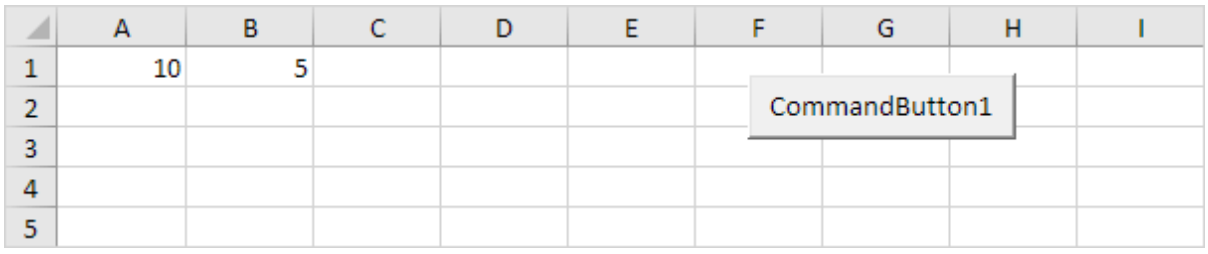


Swap Values

This example teaches you how to swap two values in Excel VBA. You will often need this structure in more complicated programs as we will see later.

Situation:

Two values on your worksheet.



Place a command button on your worksheet and add the following code lines:

1. First, we declare a variable called temp of type Double.

```
Dim temp As Double
```

2. We initialize the variable temp with the value of cell A1.

```
temp = Range("A1").Value
```

3. Now we can safely write the value of cell B1 to cell A1 (we have stored the

```
Range("A1").Value = Range("B1").Value
```

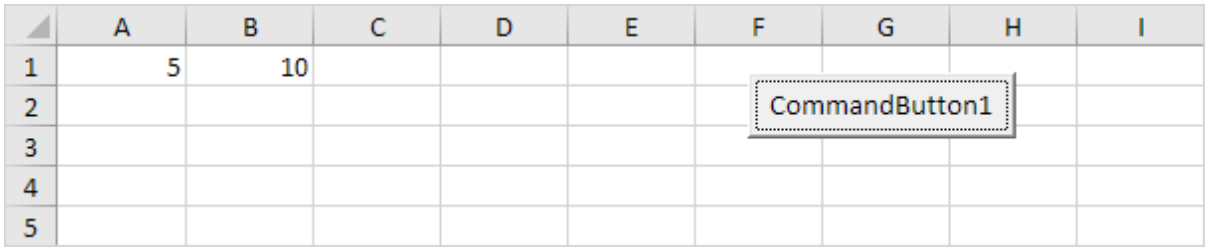
value of cell A1 to temp sowe will not lose it).

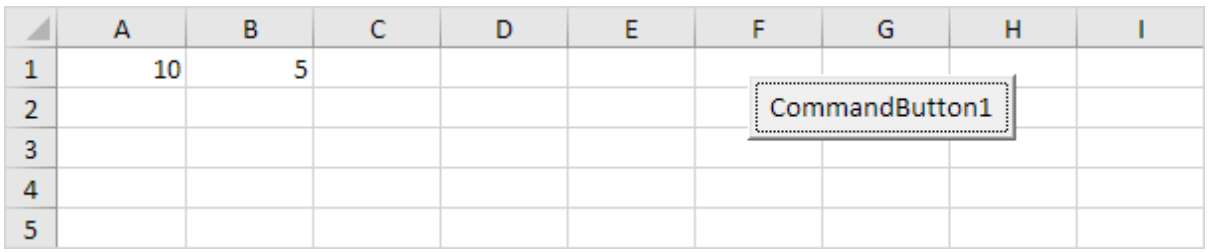
4. Finally, we write the value of cell A1 (written to temp) to cell B1.

```
Range("B1").Value = temp
```

5. Click the command button two times.

Result:



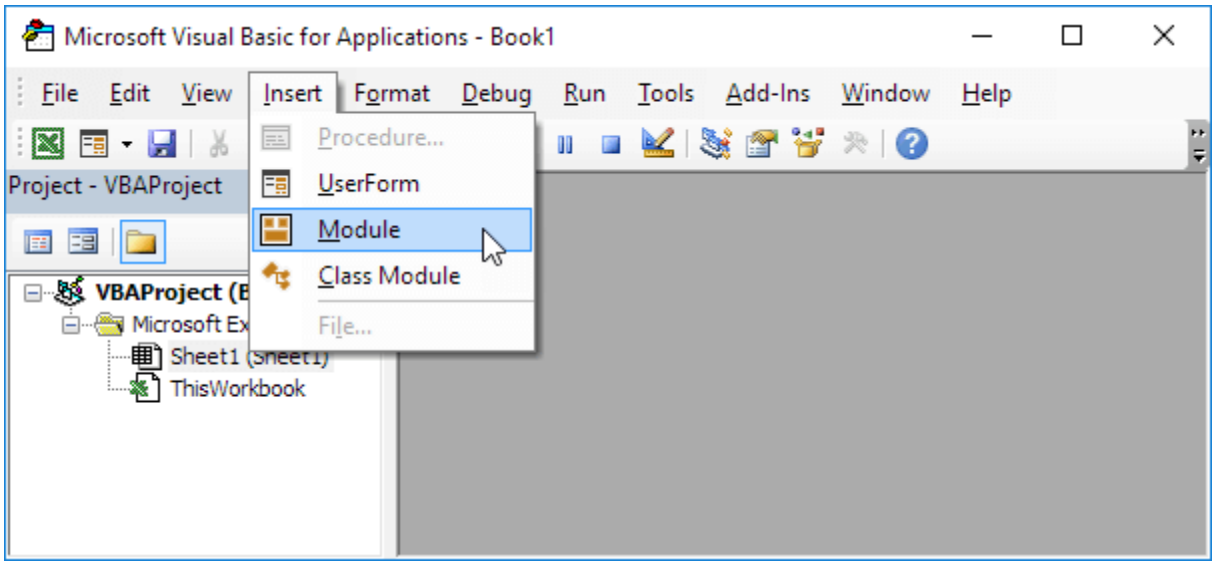


Run Code from a Module

As a beginner to Excel VBA, you might find it difficult to decide where to put your VBA code. The Create a Macro chapter illustrates how to run code by clicking on a command button. This example teachesyou how to run code from a module.

1. Open the Visual Basic Editor.

2. Click Insert, Module.



3. Create a procedure (macro) called Cyan.

```
Sub Cyan()

End Sub
```

Note: a procedure is either a sub or a function. Learn more about functions and subs here, if you like.

4. The sub changes the background color of your worksheet to cyan. To

```
Cells.Interior.ColorIndex = 28
```

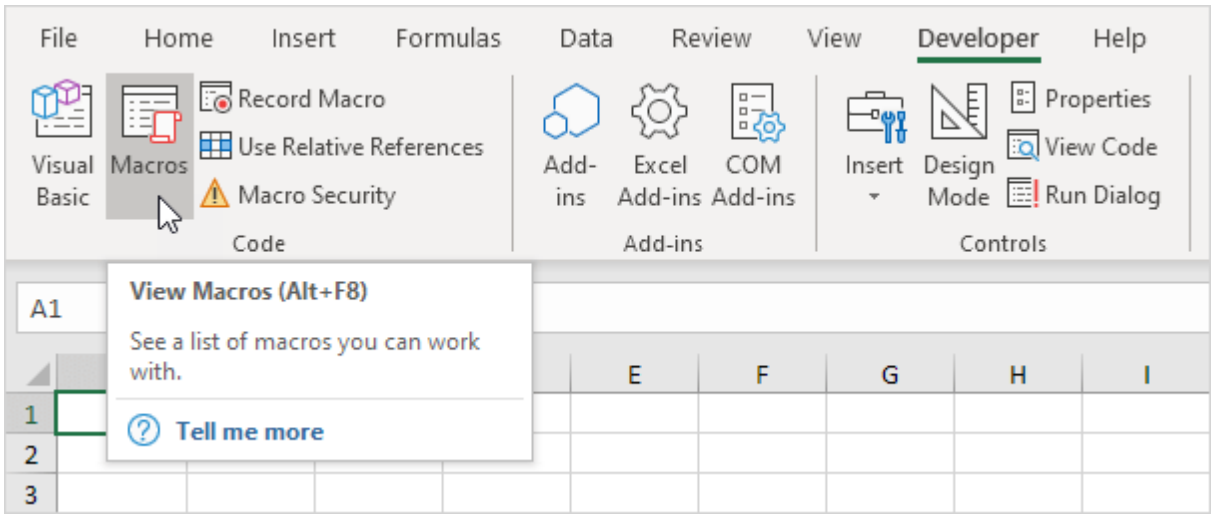
achieve this, add the following code line.

Note: instead of ColorIndex number 28 (cyan), you can use any

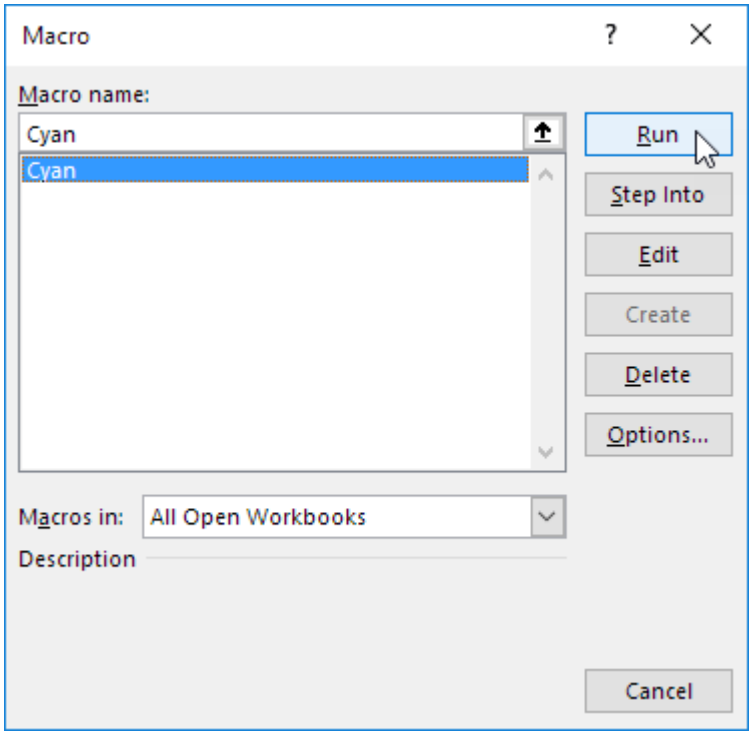
ColorIndex number. To run the procedure, execute the following

steps.

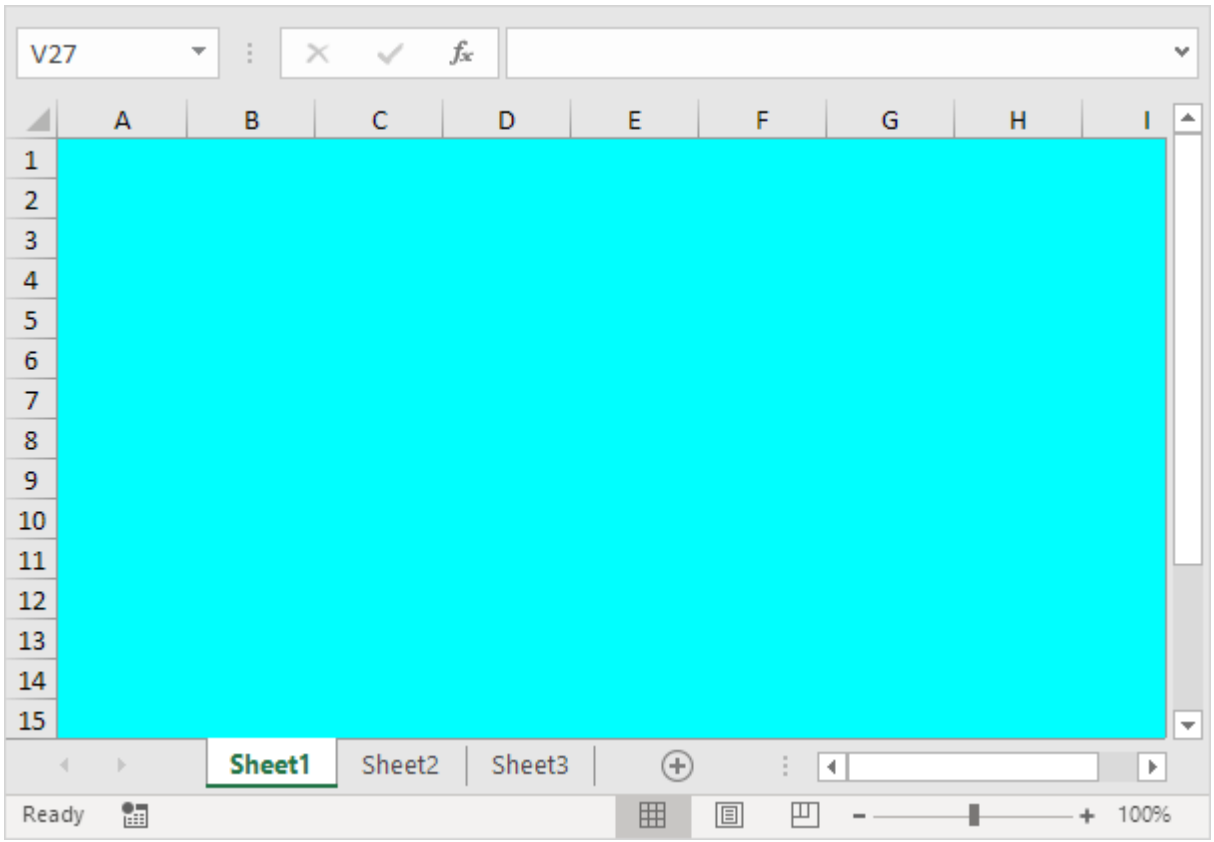
5. Click Macros.



6. Select Cyan and click Run.



Result:



Note: code placed into a module is available to the whole workbook. That means you can select Sheet2 or Sheet3 and change the background color of these sheets as well. The Add a Macro to the Toolbar program illustrates how to make a macro available to all your workbooks (Excel files). Remember, code placed on a sheet (assigned to a command button) is only available for that particular sheet.

Macro Recorder

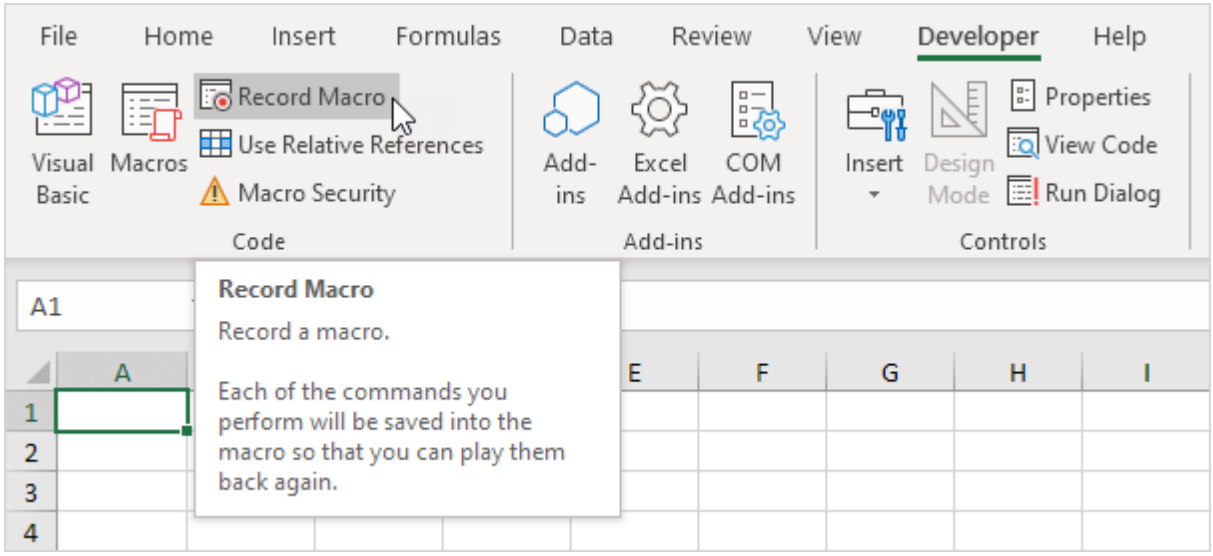
Record a Macro | Run a Recorded Macro | See the Macro

The Macro Recorder, a very useful tool included in Excel VBA, records every task you perform with Excel. All you have to do is record a specific task once. Next, you can execute the task over and over withthe click of a button. The Macro Recorder is also a great help when you don't know how to program a specific task in Excel VBA. Simply open the Visual Basic Editor after recording the task to see how it canbe programmed.

Unfortunately, there are a lot of things you cannot do with the Macro Recorder. For example, you cannot loop through a range of data with the Macro Recorder. Moreover, the Macro Recorder uses a lot more code than is required, which can slow your process down.

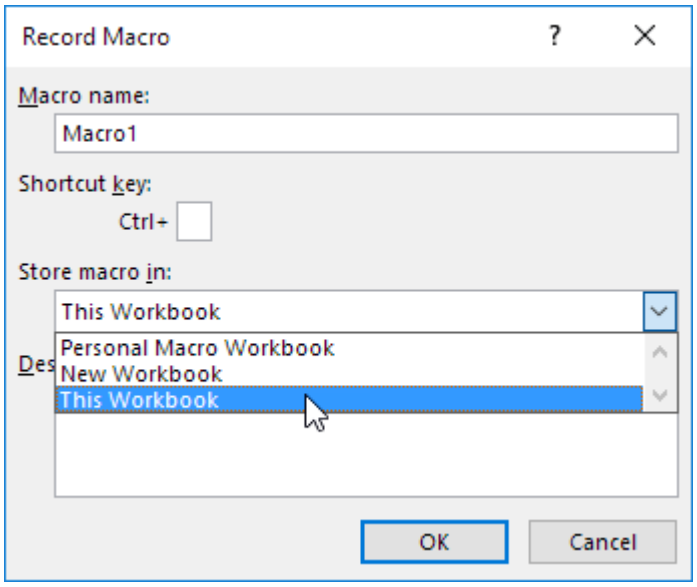
Record a Macro

1. On the Developer tab, click Record Macro.



2. Enter a name.

3. Select This Workbook from the drop-down list. As a result, the macro will only be available in the current workbook.

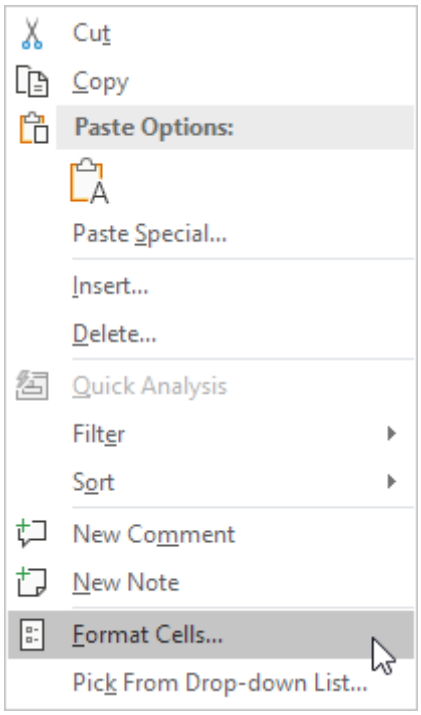


4.

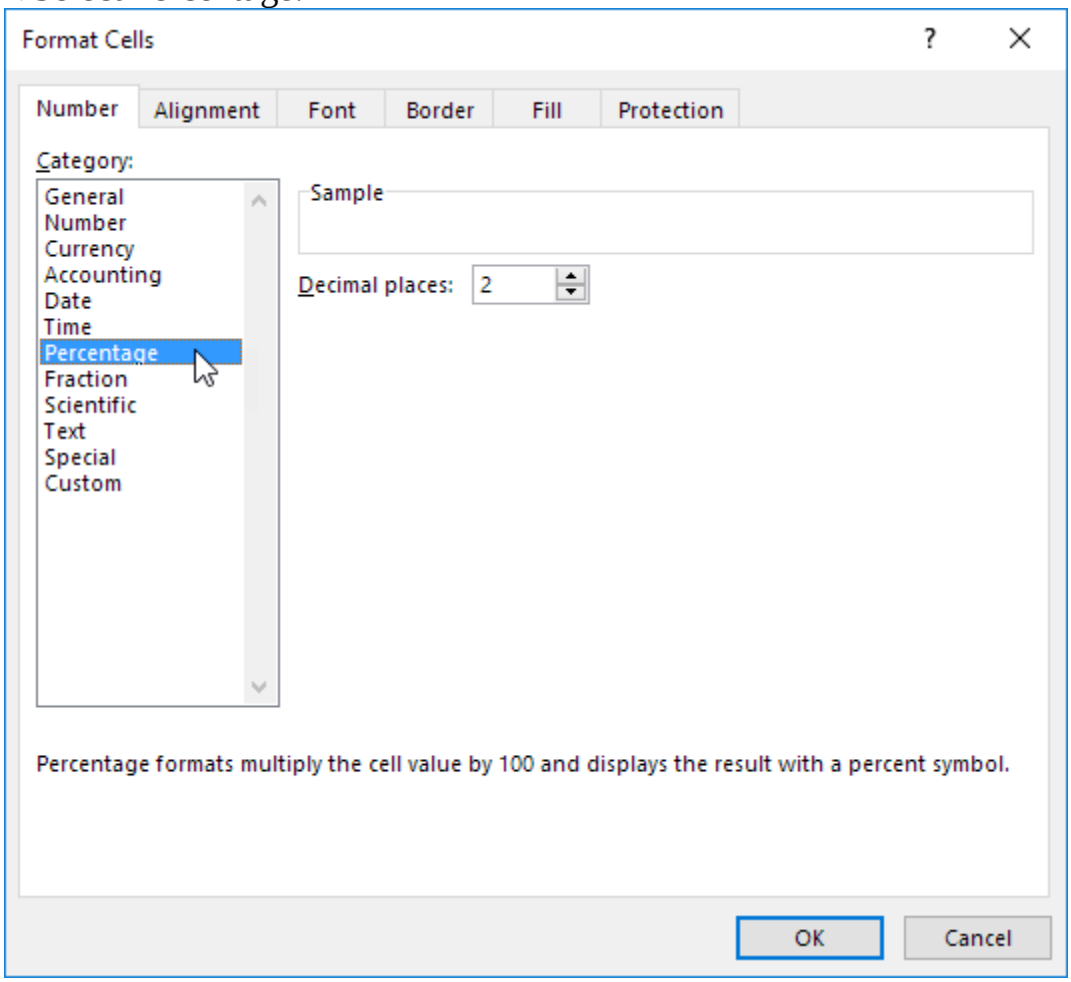
Note: if you store your macro in Personal Macro Workbook, the macro will be available to all your workbooks (Excel files). This is possible because Excel stores your macro in a hidden workbook that opens automatically when Excel starts. If you store your macro in New Workbook, the macro will only be available in an automatically new opened workbook.

5. Click OK.

6. Right mouse click on the active cell (selected cell). Be sure not to select any other cell! Next, click Format Cells.

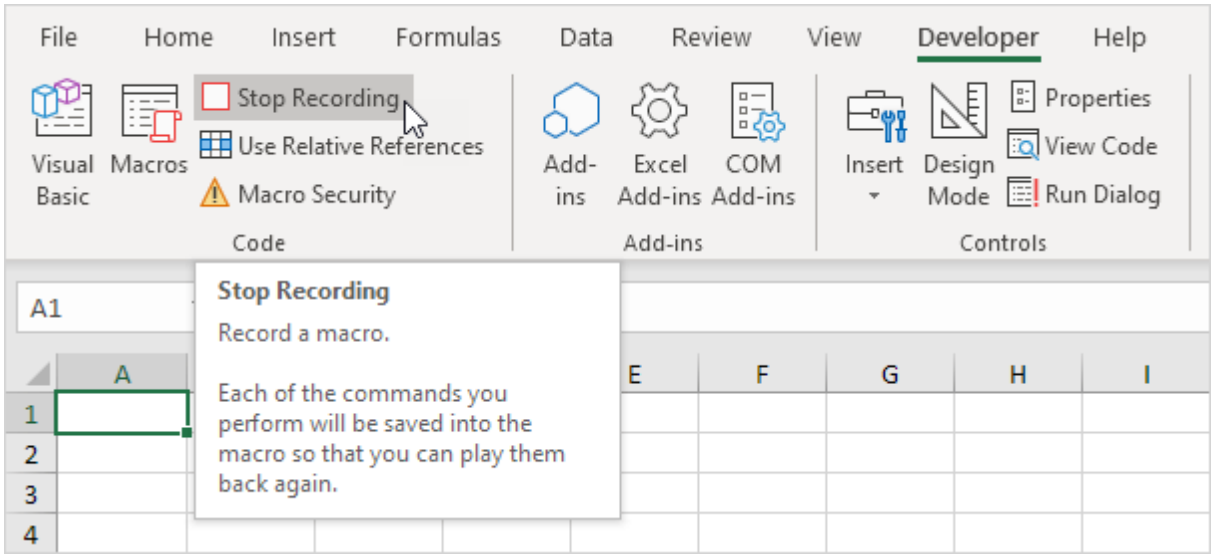


7. Select Percentage.



8. Click OK.

9. Finally, click Stop Recording.



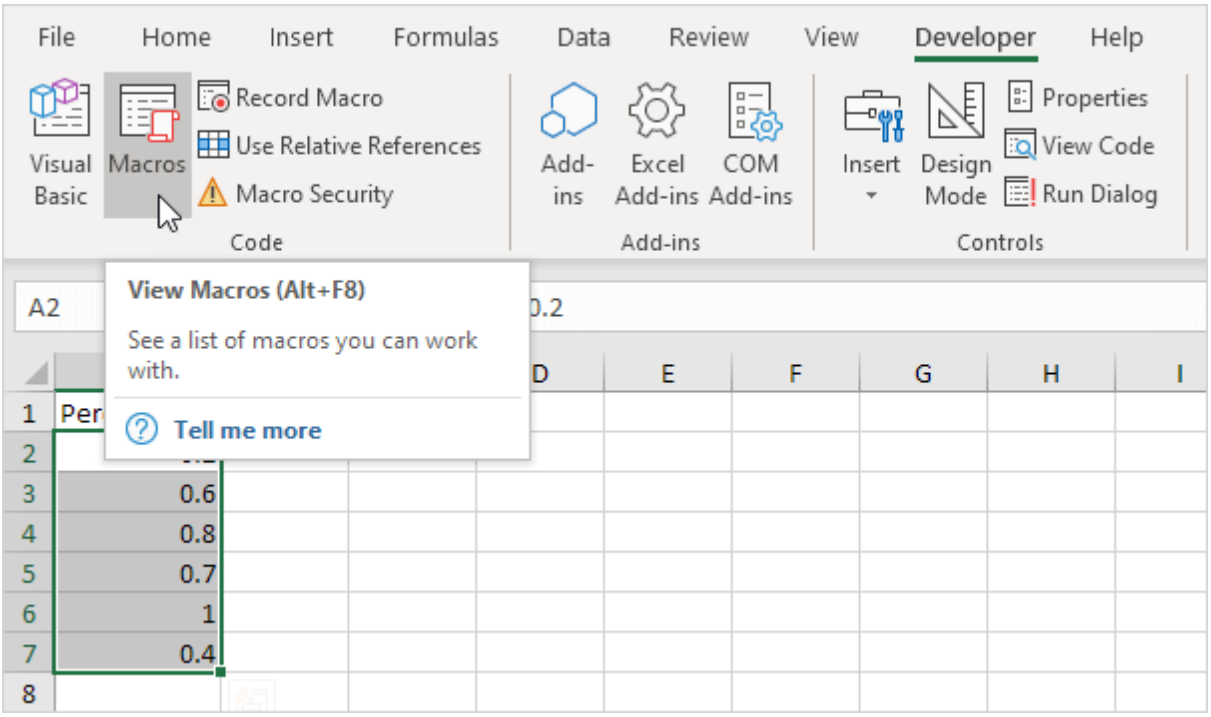
Now we'll test the macro to see if it can change the number format to Percentage.

1. Enter some numbers between 0 and 1.

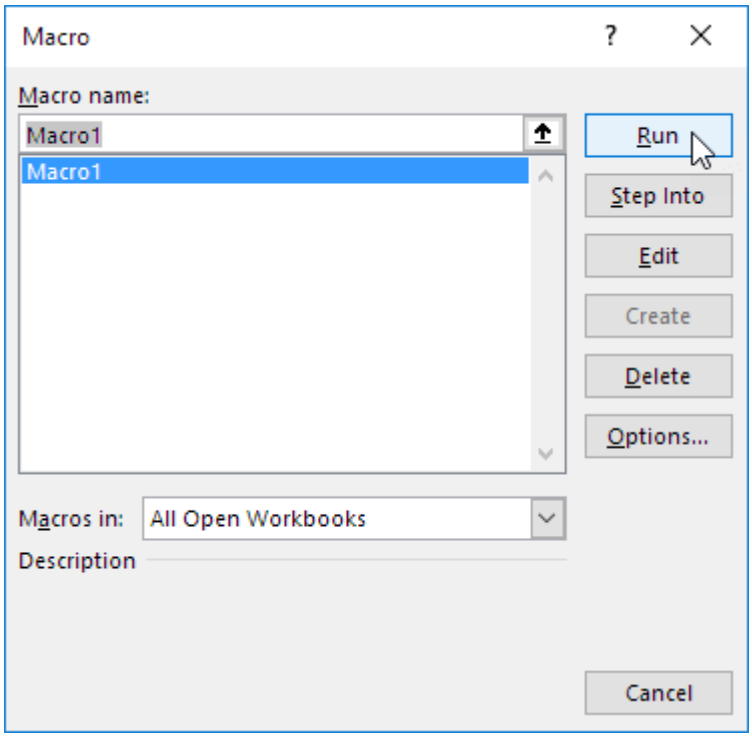
2. Select the numbers.

	A	B
1	Percentages	
2	0.2	
3	0.6	
4	0.8	
5	0.7	
6	1	
7	0.4	
8		

3. On the Developer tab, click Macros.



4. Click Run.

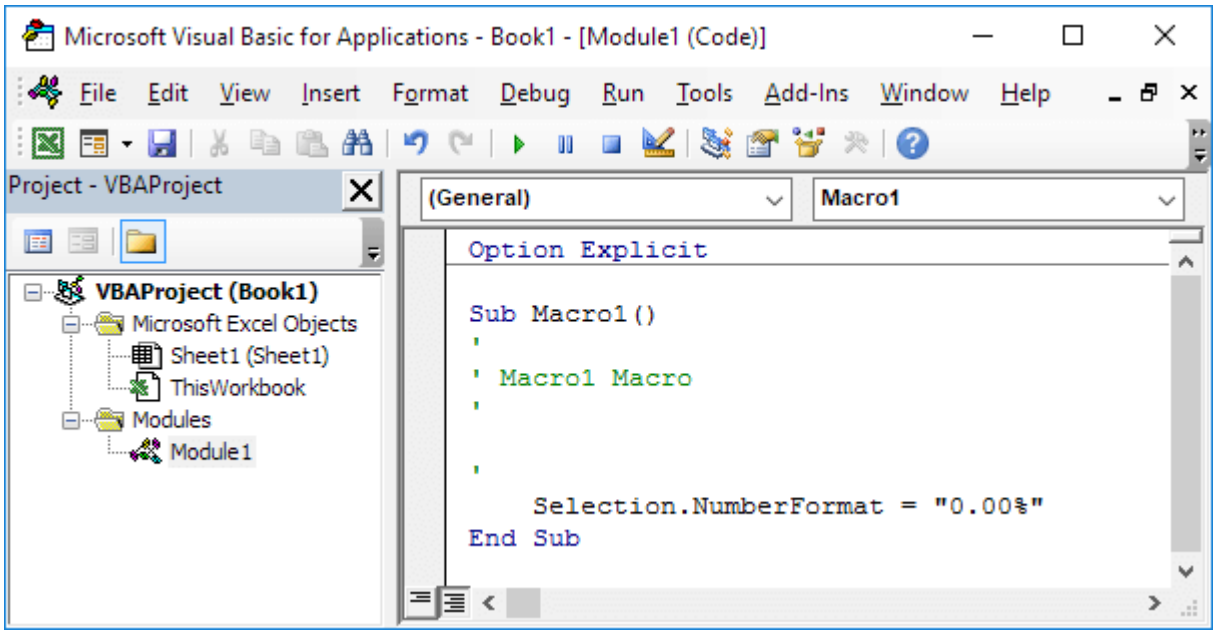


Result:

	A	B
1	Percentages	
2	20.00%	
3	60.00%	
4	80.00%	
5	70.00%	
6	100.00%	
7	40.00%	
8		

See the Macro

To take a look at the macro, open the Visual Basic Editor.



Note: the macro has been placed into a module called Module1. Code placed into a module is available tothe whole workbook. That means you change the number format of cells on other sheets as well.

Remember, code placed on a sheet (assigned to a command button) is only available for that particularsheet. You can ignore the Option Explicit statement for now.

Use Relative References

By default, Excel records macros in absolute mode. However, sometimes it is

useful to record macros inrelative mode. This program teaches you how to do this. If you don't know how to record a macro, we highly recommend you to read this example first.

Recording in Absolute Mode

To record a macro in absolute mode, execute the following steps.

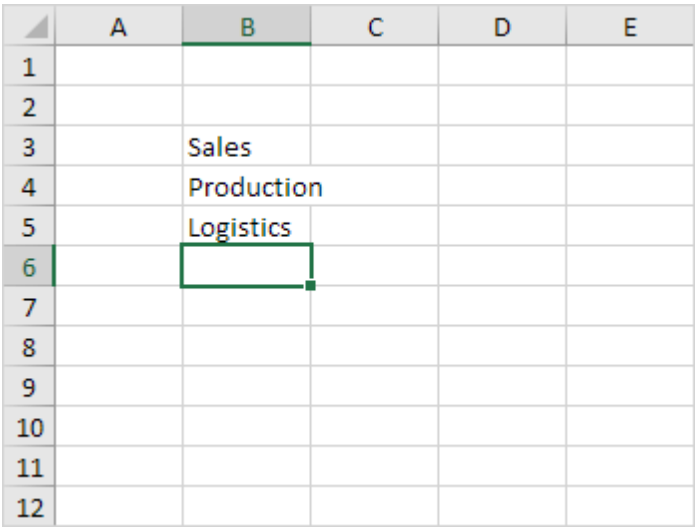
1. First, click Record Macro.

2. Next, select cell B3. Type Sales and press enter.

3. Type Production and press enter.

4. Type Logistics and press enter.

Result:



5. Click Stop Recording.

6. Empty Range("B3:B5").

7. Select any cell on the sheet and run the recorded macro.

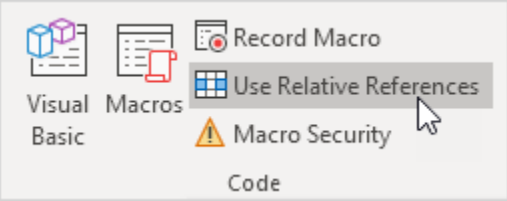
Result:

	A	B	C	D	E
1					
2					
3		Sales			
4		Production			
5		Logistics			
6					
7					
8					
9					
10					
11					
12					

A macro recorded in absolute mode always produces the same result.

Recording in Relative Mode Wouldn't it be nice to place these words anywhere on the sheet automatically? Not just Range("B3:B5"). This would make the macro much more flexible. Solution: record the macro in relative mode.

1. Select "Use Relative References".



2. First, select any single cell (for example, cell B8).

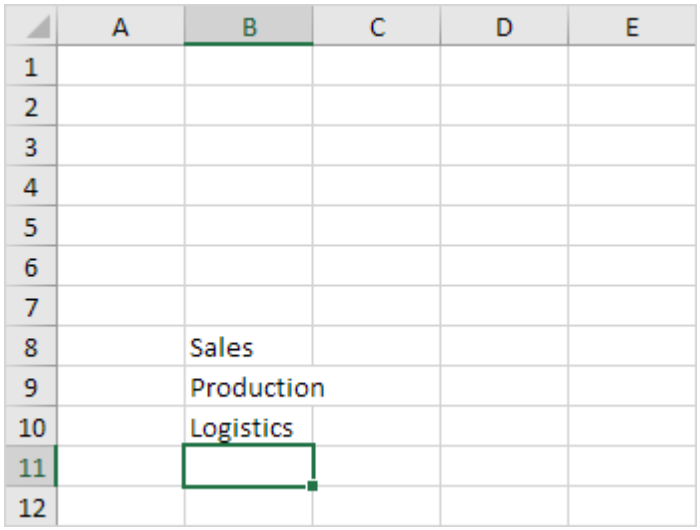
3. Next, click Record Macro.

4. Type Sales and press enter.

5. Type Production and press enter.

6. Type Logistics and press enter.

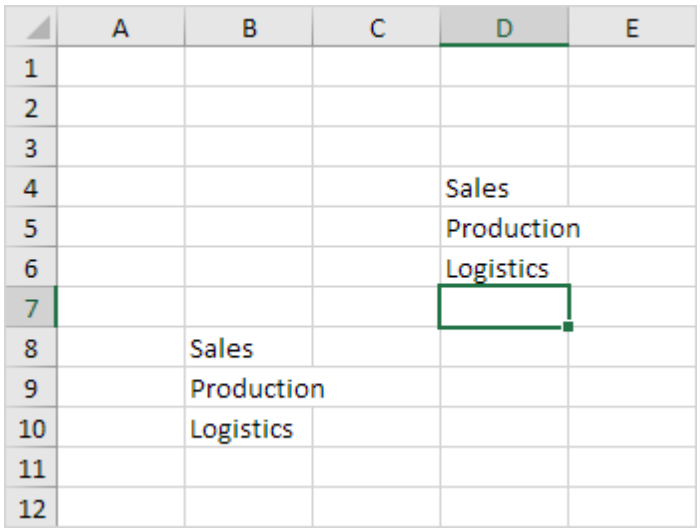
Result:



7. Click Stop Recording.

8. Select any other cell (for example, cell D4) and run the recorded macro.

Result:



Excel places the words relative to the initial selected cell. That's why it's called recording in relative mode.

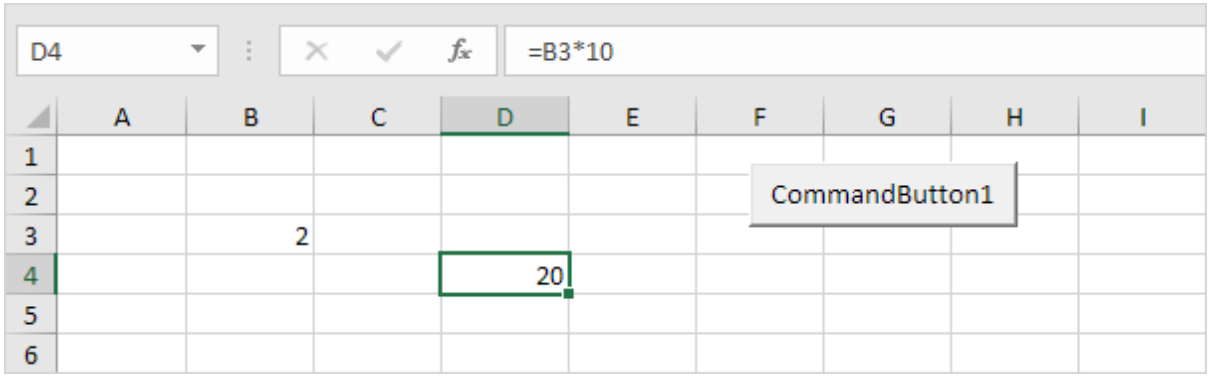
FormulaR1C1

This example illustrates the difference between A1, R1C1 and R[1]C[1] style in Excel VBA.

1. Place a command button on your worksheet and add the following code line (A1 style):

```
Range("D4").Formula = "=B3*10"
```

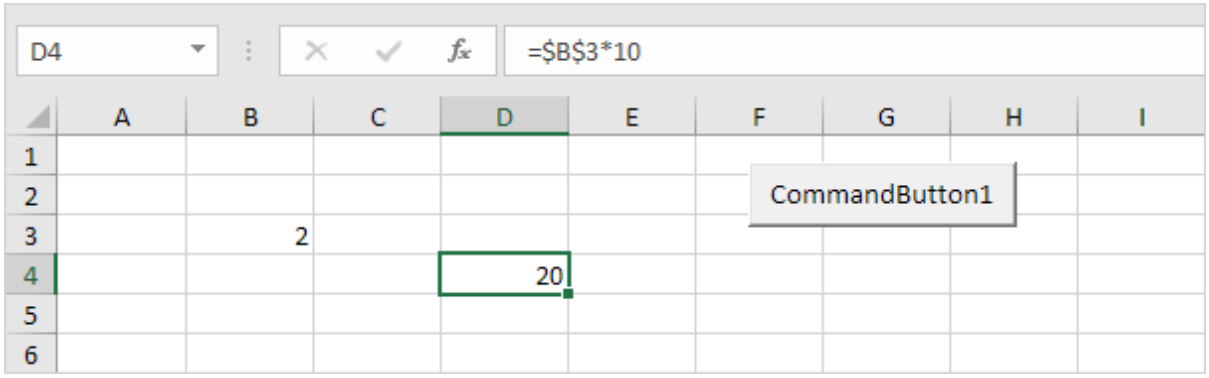
Result:



2. Add the following code line (R1C1 style):

```
Range("D4").FormulaR1C1 = "=R3C2*10"
```

Result:

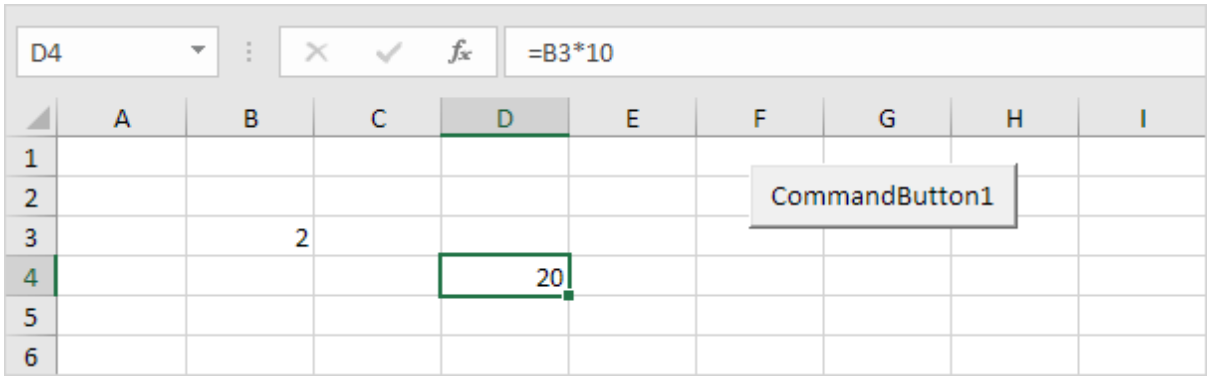


Explanation: cell D4 references cell B3 (row 3, column 2). This is an absolute reference ($ symbol in front of the row number and column letter).

3. Add the following code line (R[1]C[1] style):

```
Range("D4").FormulaR1C1 = "=R[-1]C[-2]*10"
```

Result:



Explanation: cell D4 references cell B3 (one row above and 2 columns to the left). This is a relative reference. This code line gives the exact same result as the code line used at step 1.

4. Why learning about this? Because the Macro Recorder uses the FormulaR1C1 property (R[1]C[1] style). The Macro Recorder creates the following code lines if you enter the formula =B3*10 into cell D4.

```
Sub Macro1()
'
' Macro1 Macro
'

'
    Range("D4").Select
    ActiveCell.FormulaR1C1 = "=R[-1]C[-2]*10"
    Range("D5").Select
End Sub
```

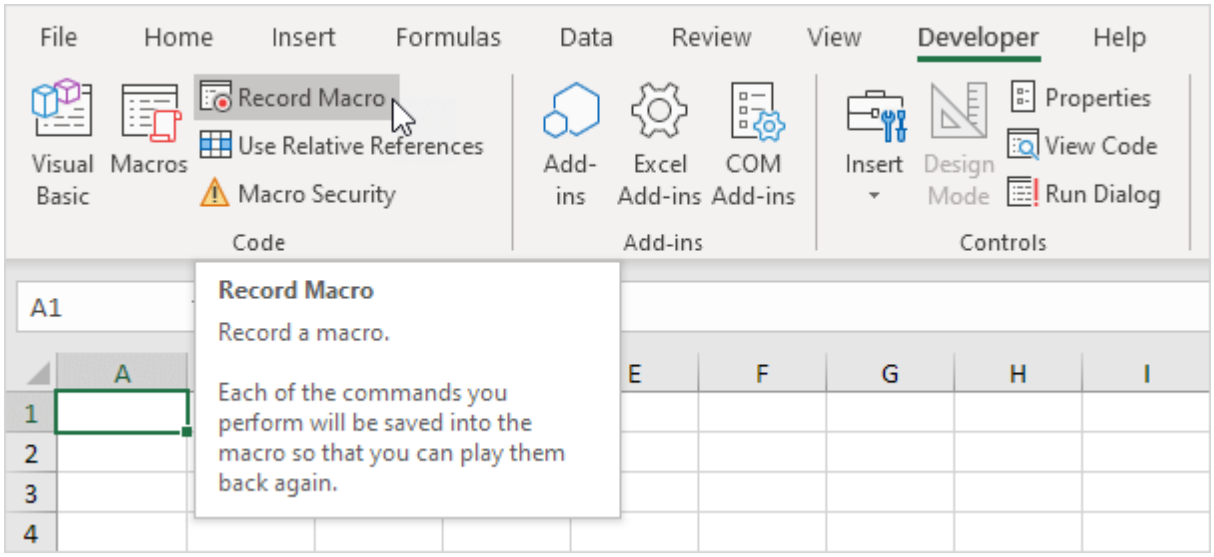
Explanation: you can see that this is the exact same code line used at step 3.

Add a Macro to the Toolbar

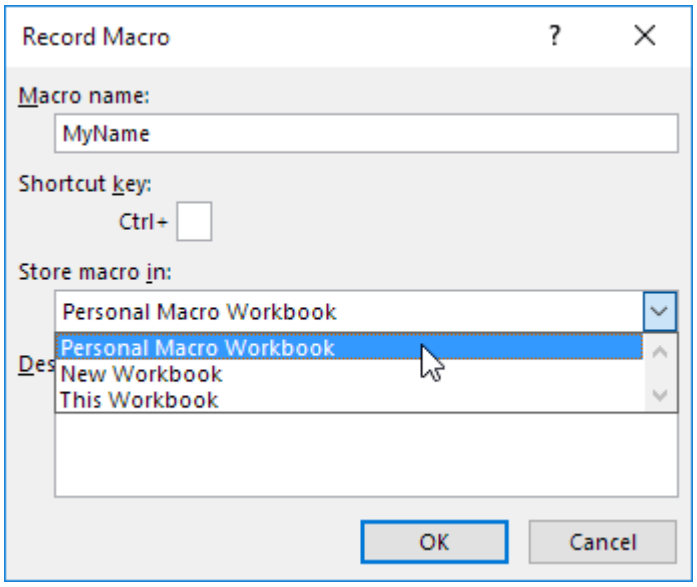
If you use an Excel macro frequently, you can add it to the Quick Access

Toolbar. This way you canquickly access your macro. First, we record an empty macro.

1. On the Developer tab, click Record Macro.

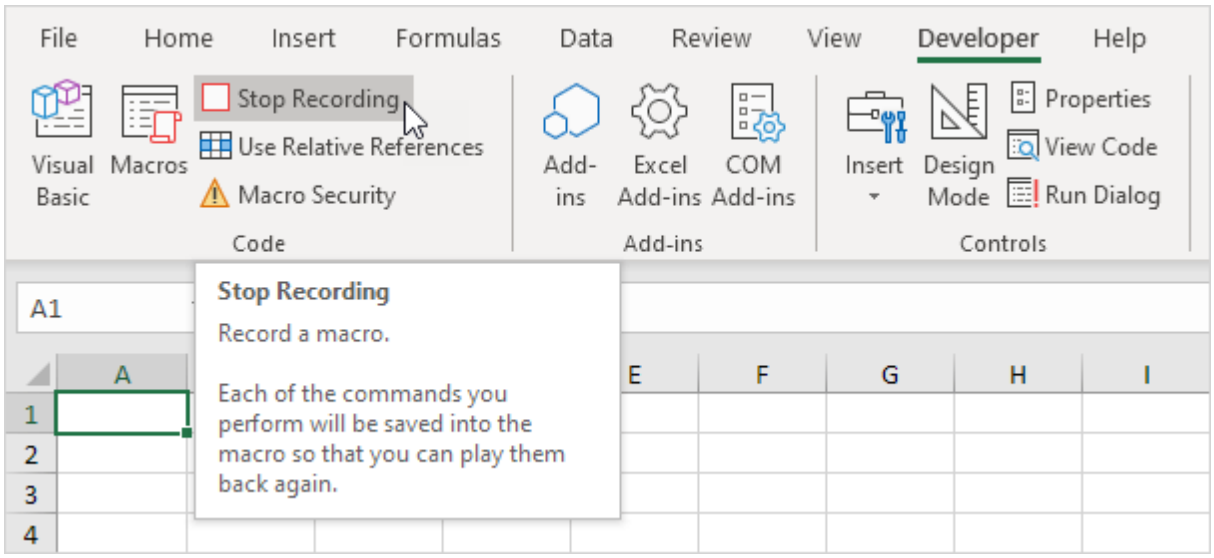


2. Name the macro MyName. Choose to store the macro in Personal Macro Workbook. This way the macro will be available to all your workbooks (Excel files). This is possible because Excel stores yourmacro in a hidden workbook that opens automatically when Excel starts.



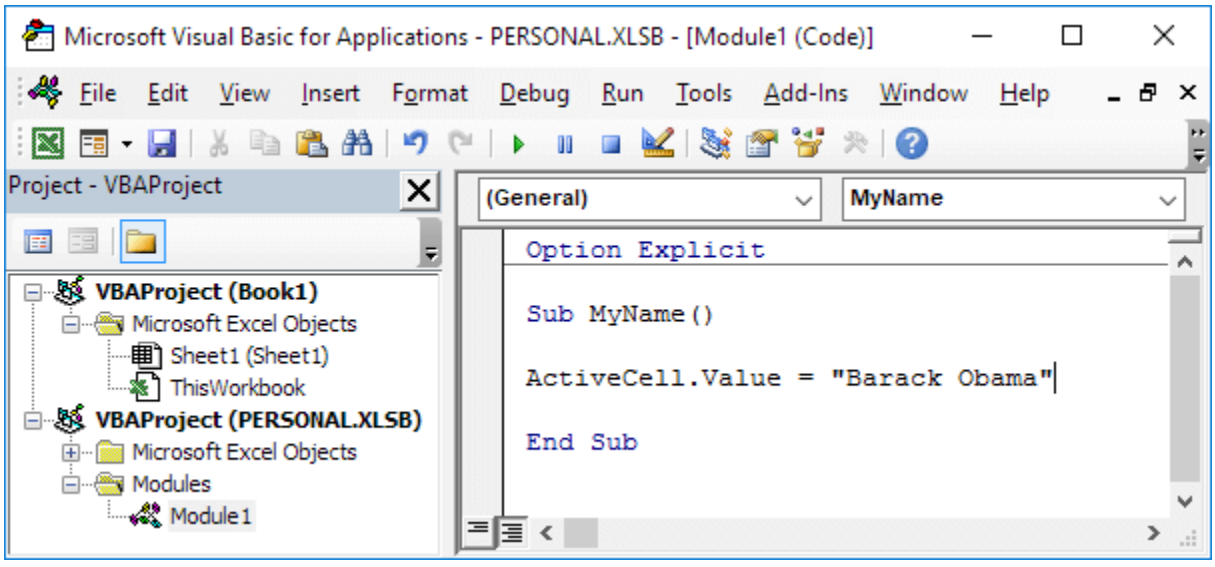
3. Click OK.

4. Click Stop Recording.



5. Open the Visual Basic Editor.

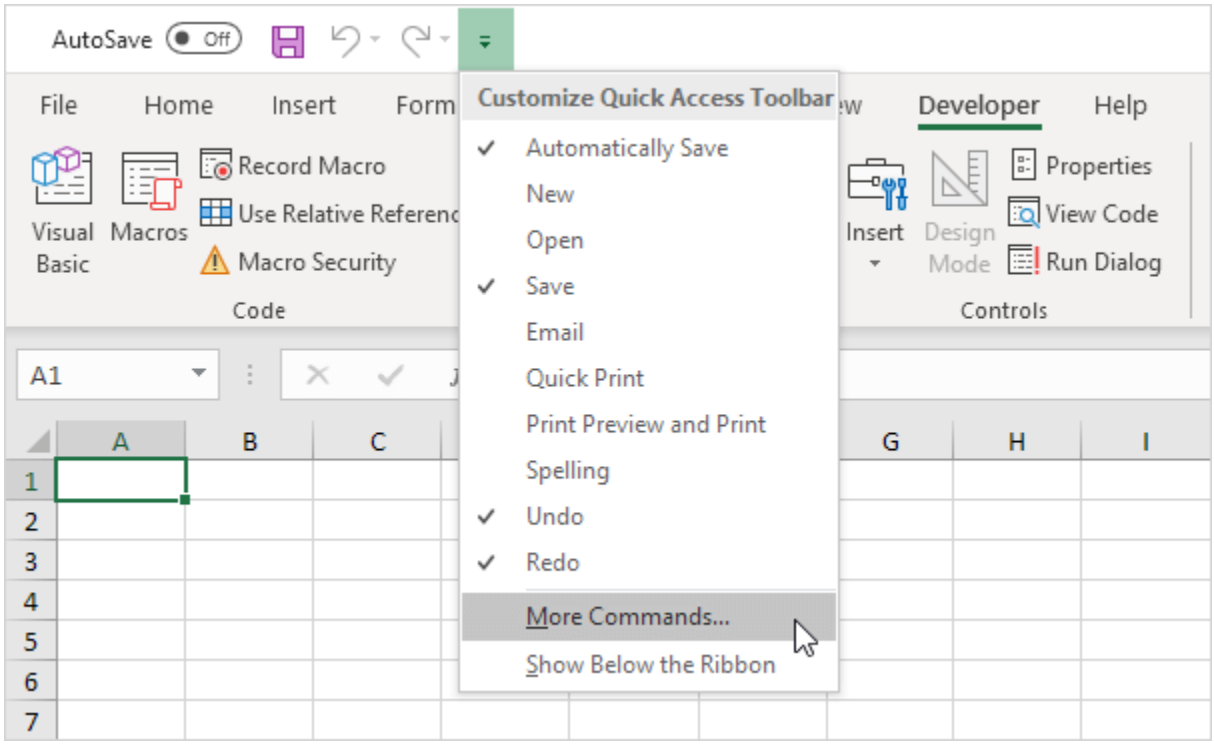
6. Create the macro:



This macro places your name in the Active Cell.

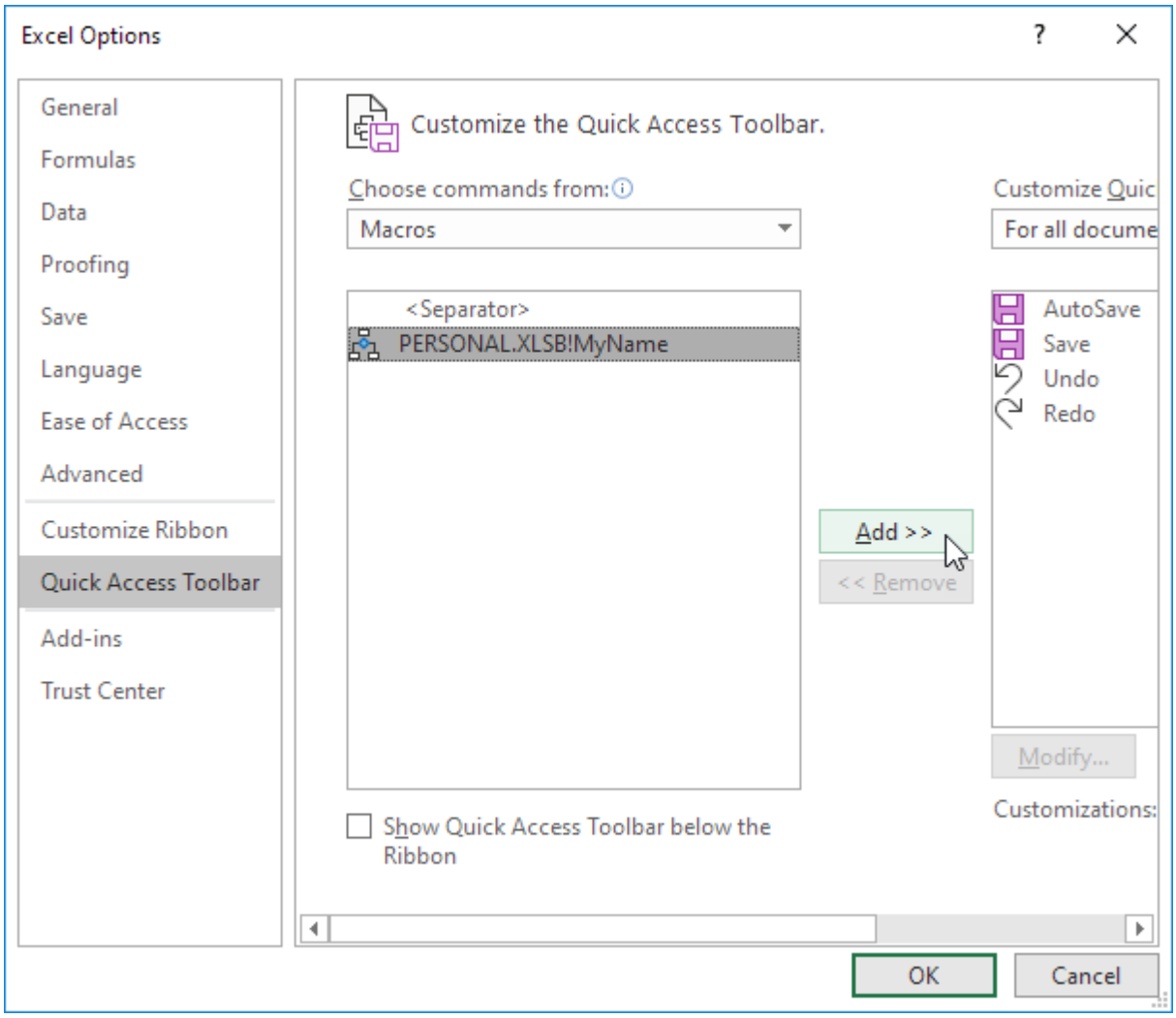
7. Close the Visual Basic Editor.

8. Now we can add this macro to the Quick Access Toolbar. Click the down arrow and click More Commands.

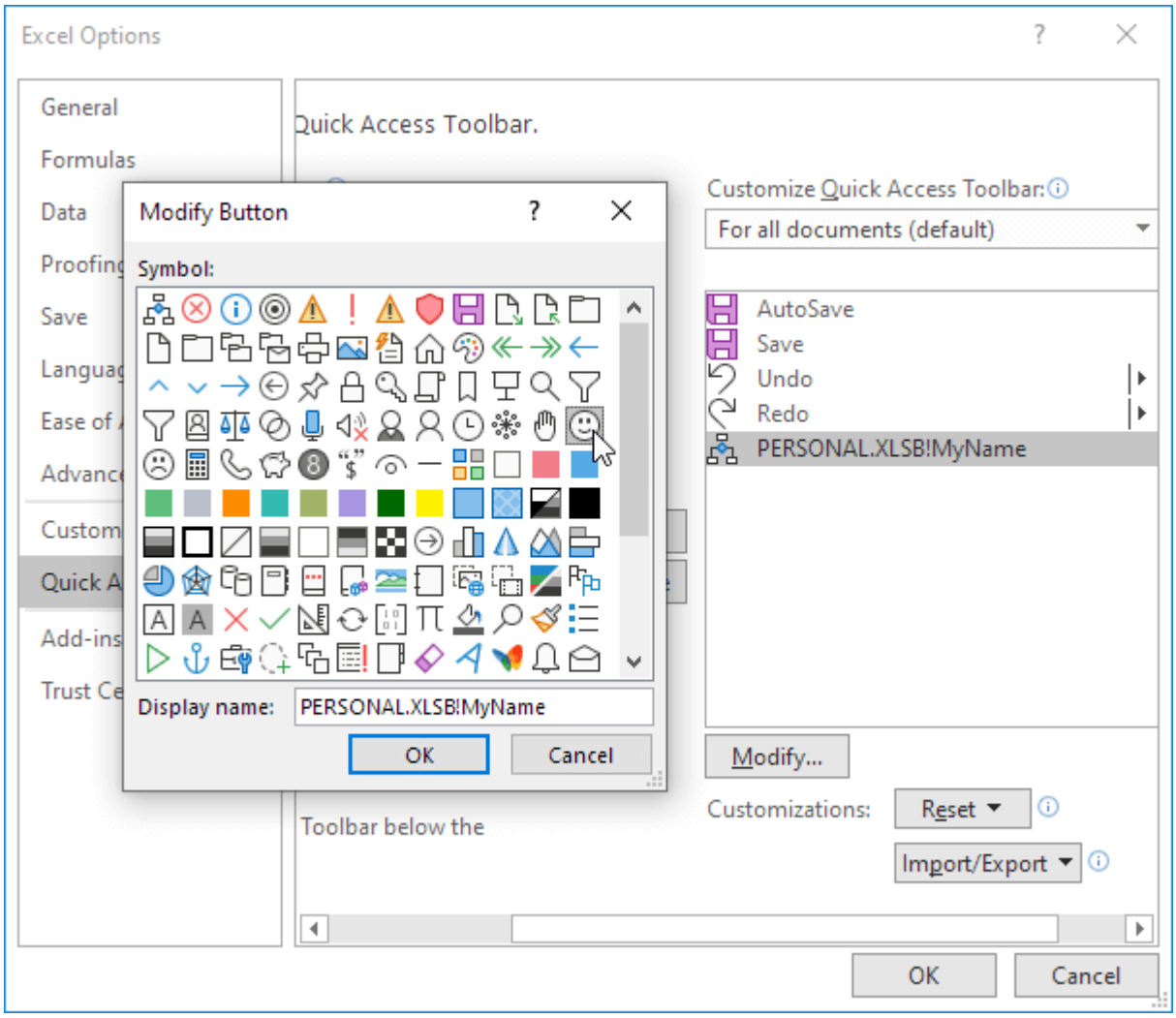


9. Under Choose commands, select Macros.

10. Select the macro and click Add.



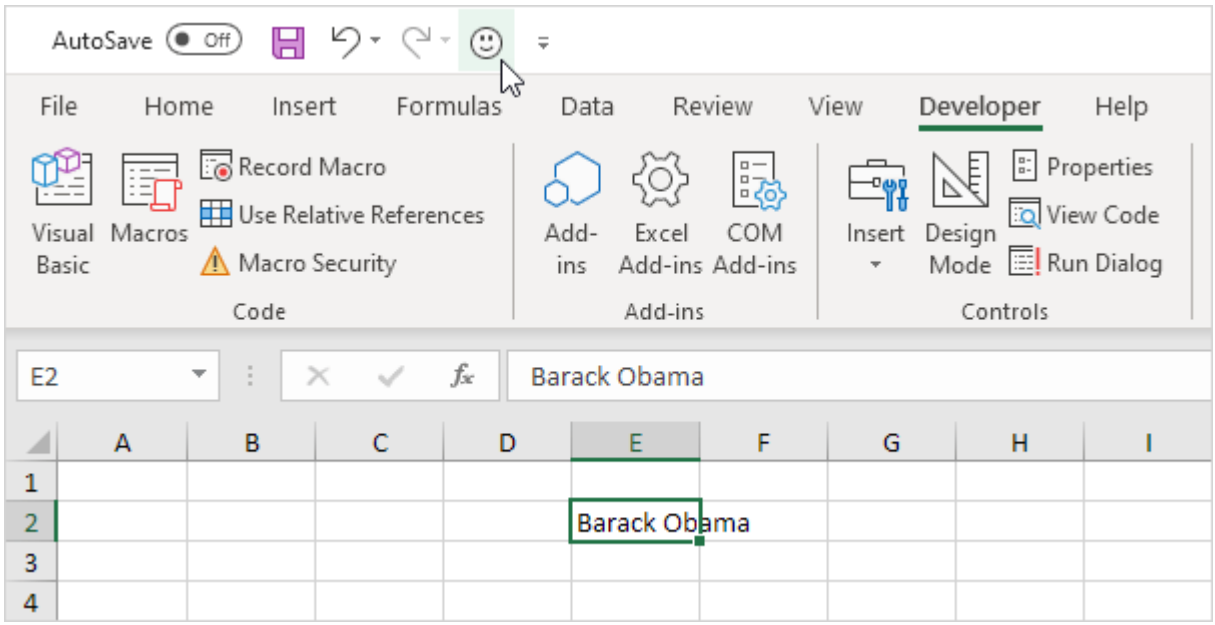
11. You can modify the button that will be added to the Quick Access Toolbar by clicking on Modify. Forexample, choose a smiley.



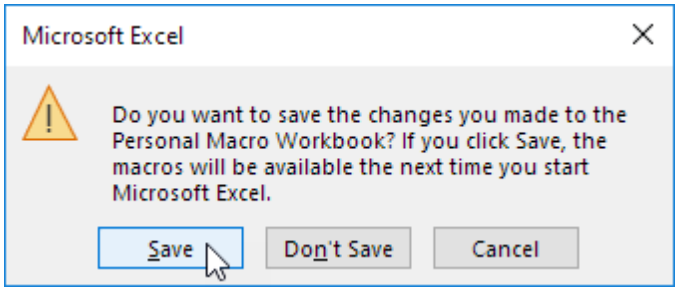
12. Click OK twice.

13. You can now execute the macro. For example, select cell E2 and click on the smiley button added tothe Quick Access Toolbar.

Result:



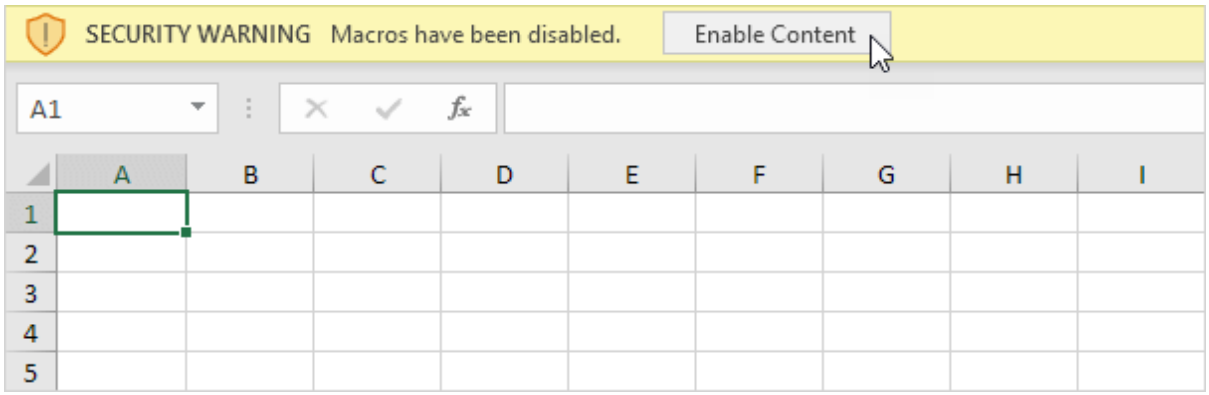
14. When you close Excel, Excel asks you to save the changes you made to the Personal Macro Workbook. Click Save to store this macro in a hidden workbook that opens automatically when Excel starts. This way the macro will be available to all your workbooks (Excel files).



Enable Macros

Enable macros in Excel when the message bar appears. Change your macro security settings in the Trust Center. To create macros, turn on the Developer tab.

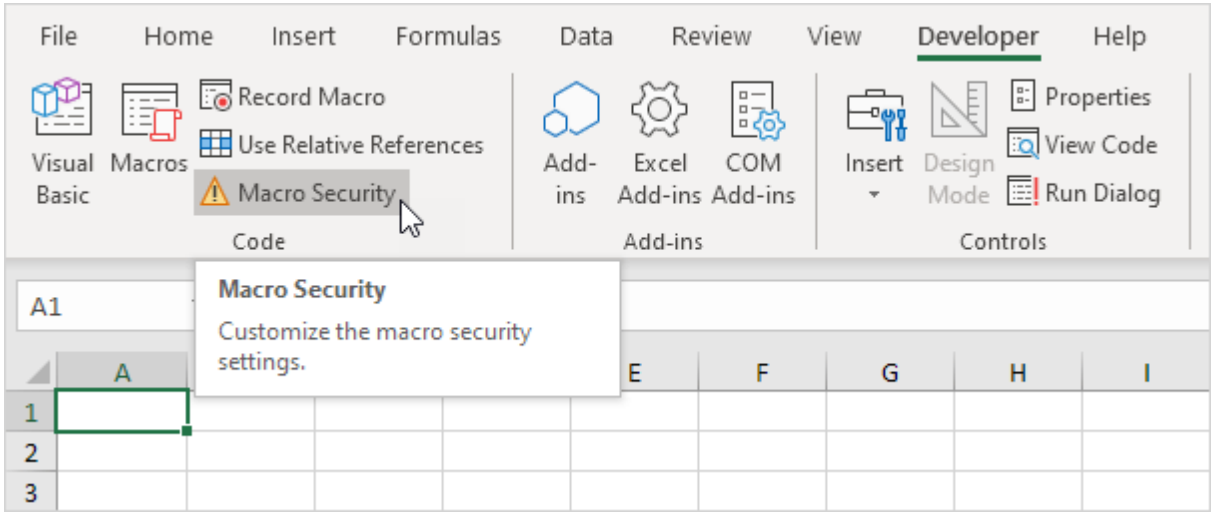
1. When the message bar appears, click Enable Content to enable macros.



Note: by clicking Enable Content, the Excel file becomes a trusted document.

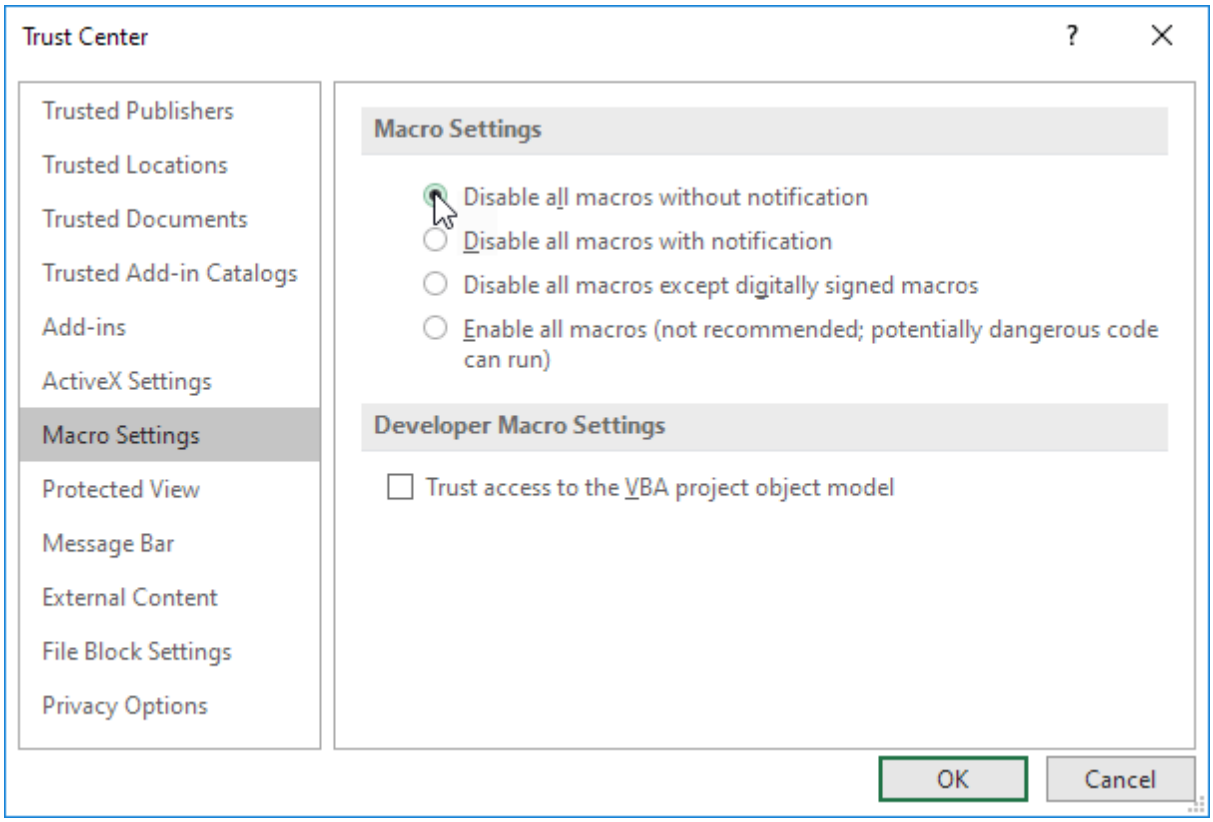
As a result, you won't seethe Security Warning again when you open this specific Excel file in the future.

2. To change your macro security settings, on the Developer tab, click Macro Security.

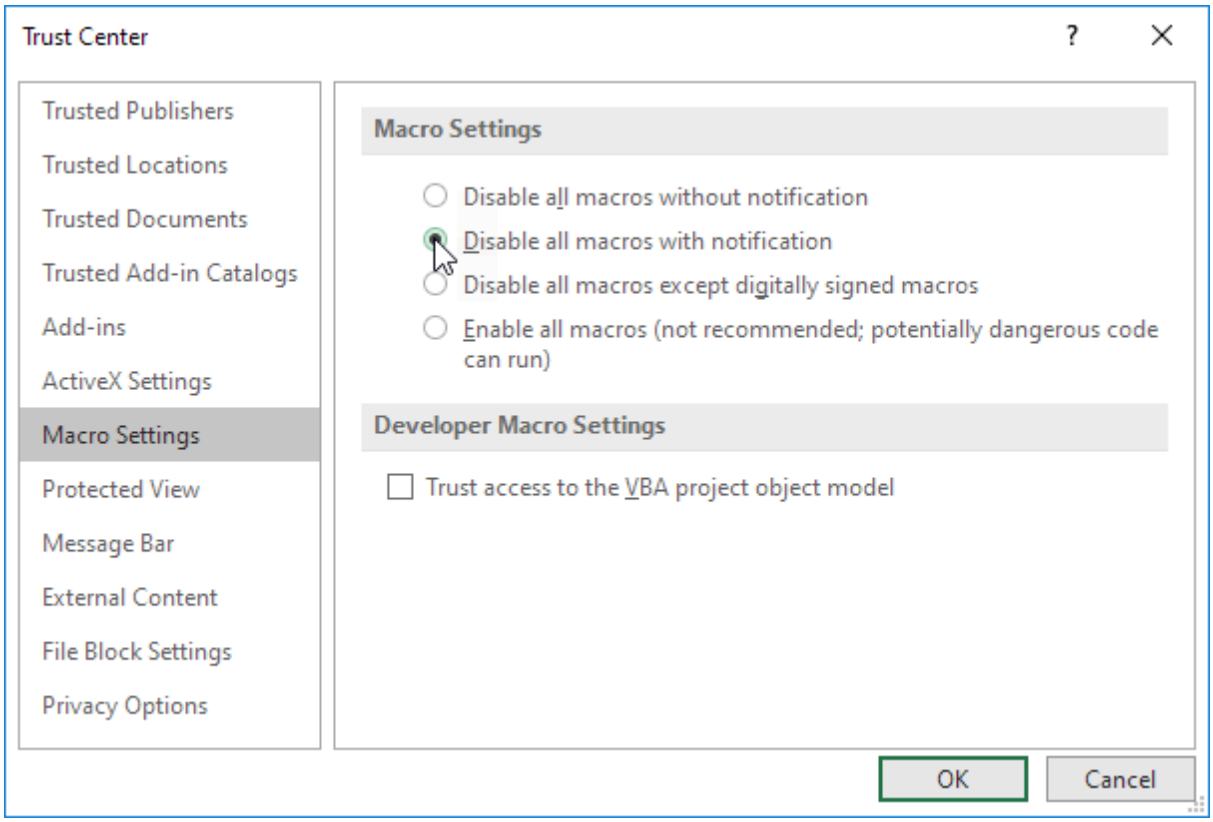


The Trust Center opens.

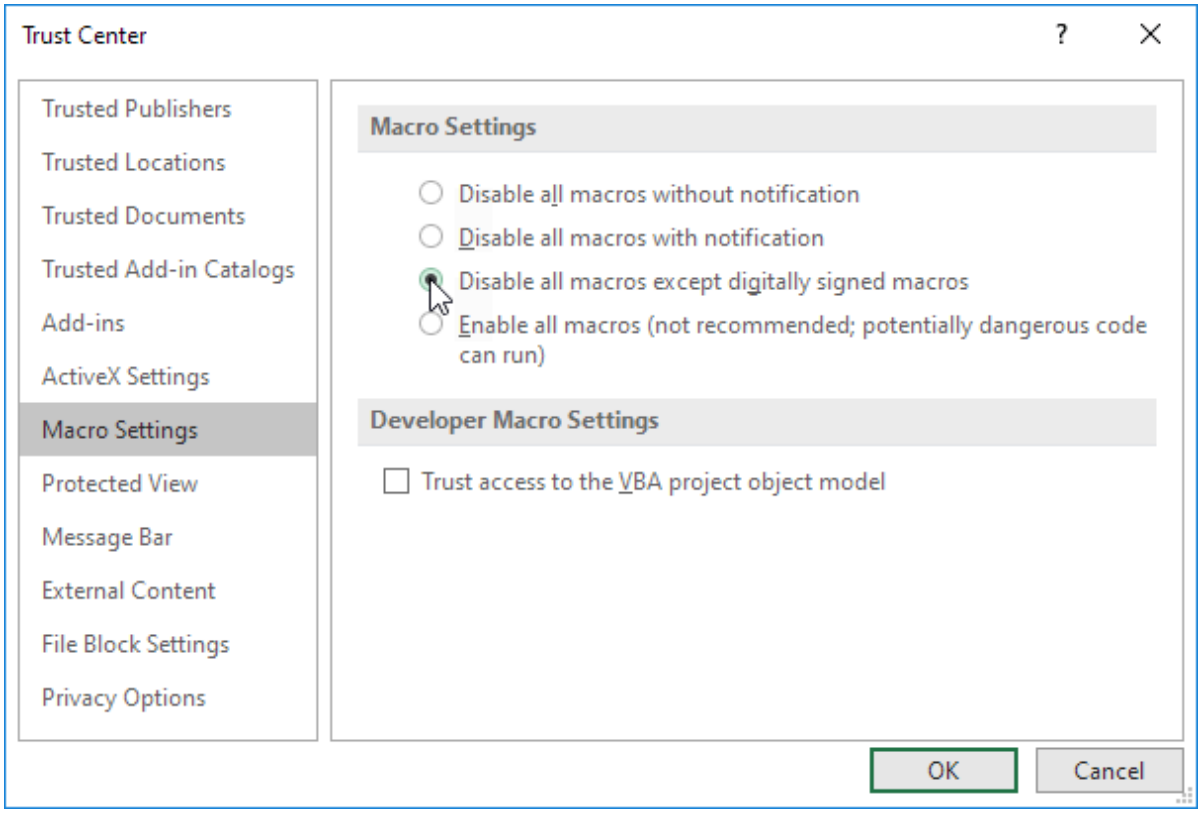
1. The first option disables all macros.



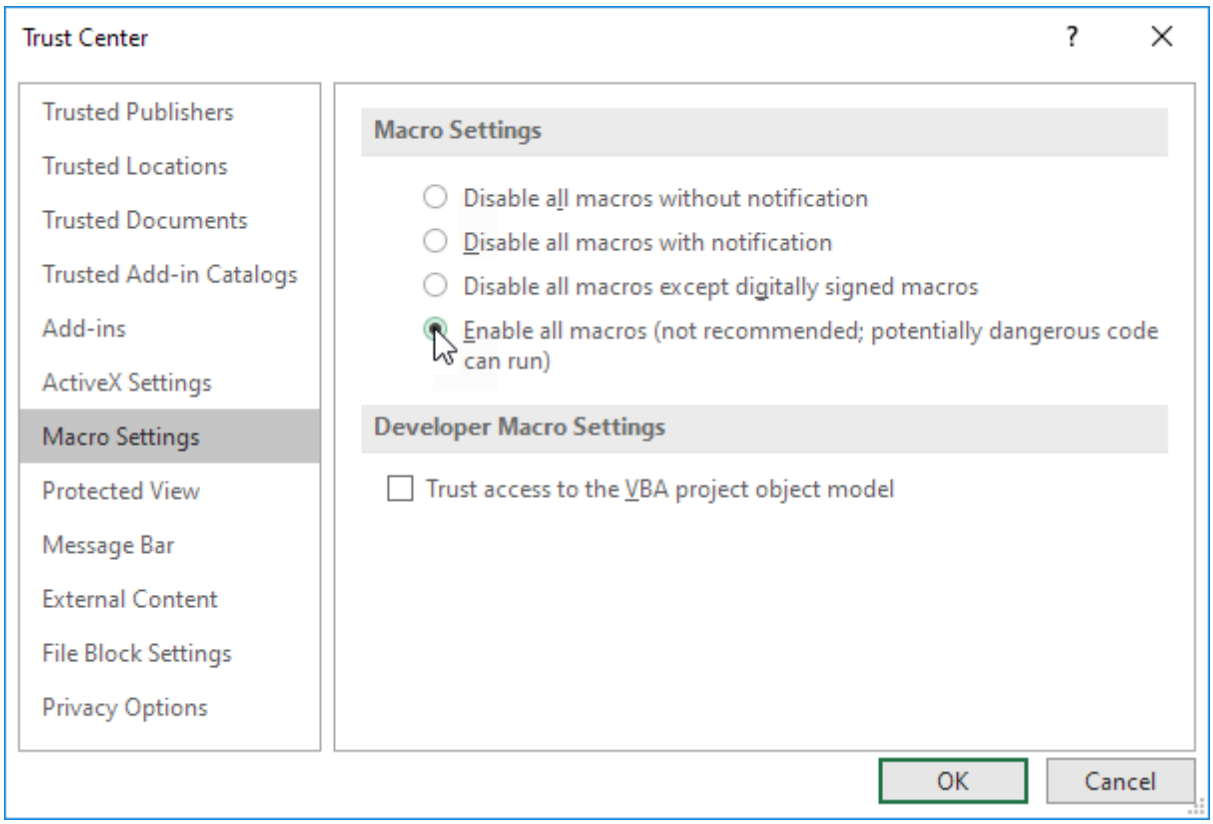
2. The second option always asks you to enable a macro. Use this security level if you are downloading a lot of Excel files from the internet. Don't click Enable Content (see first screenshot on this page) if you don't trust the owner of the Excel file.



3. The third option only allows macros with a digital signature to run, and asks you to enable others.



4. The fourth option enables all macros. Use this security level if you are a beginner and only typing yourown macros at the moment. With this security level you don't have to enable macros all the time.



5. If you're new to Excel VBA, let's create a simple macro.

Protect Macro

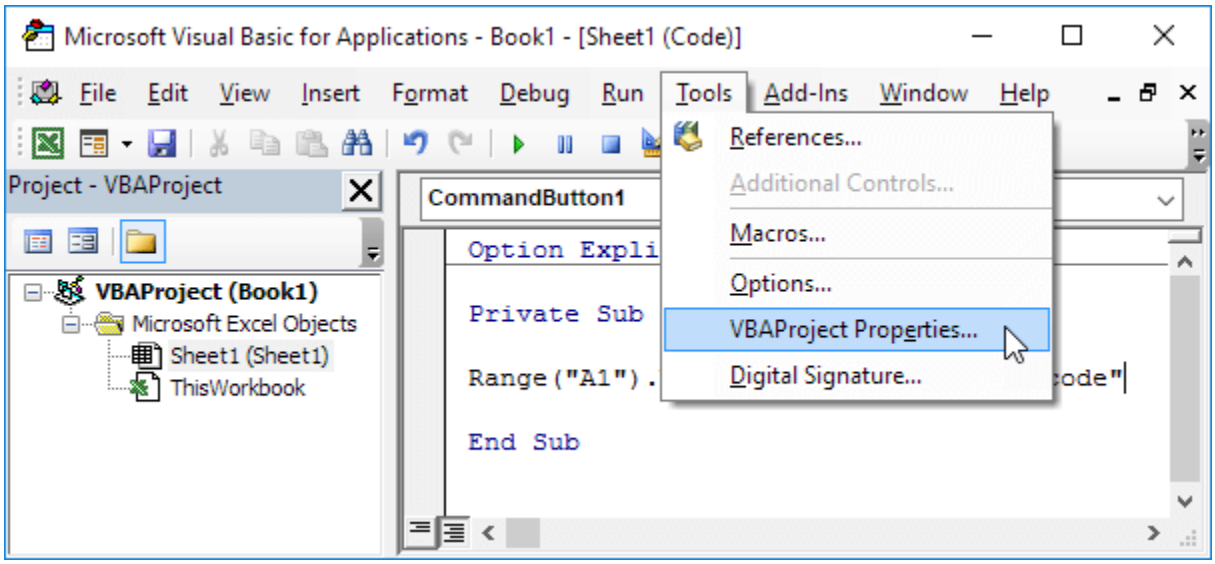
Just like you can password protect workbooks and worksheets, you can password protect a macro in Excelfrom being viewed (and executed).

Place a command button on your worksheet and add the following code lines:

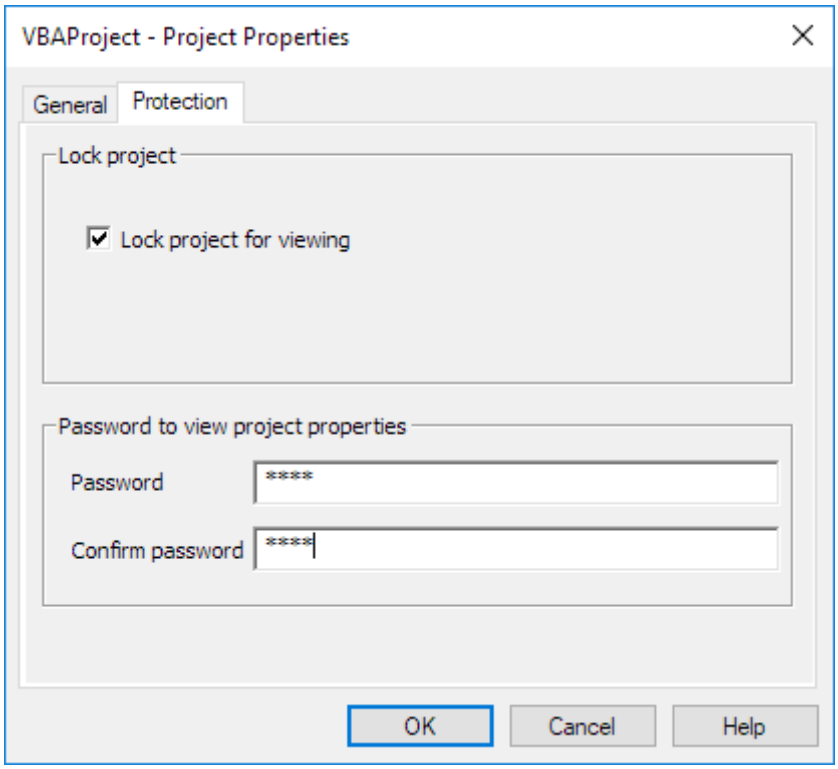
1. First, create a simple macro that you want to protect.

```
Range("A1").Value = "This is secret code"
```

2. Next, click Tools, VBAProject Properties.



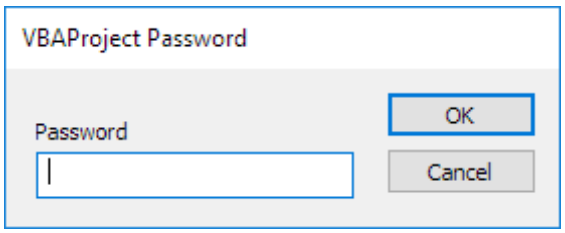
3. On the Protection tab, check "Lock project for viewing" and enter a password twice.



4. Click OK.

5. Save, close and reopen the Excel file. Try to view the code.

The following dialog box will appear:

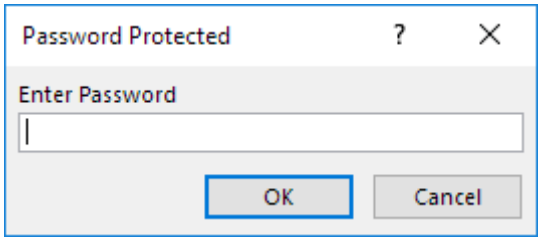


You can still execute the code by clicking on the command button but you cannot view or edit the code anymore (unless you know the password). The password for the downloadable Excel file is "easy".

6. If you want to password protect the macro from being executed, add the following code lines:

```
Dim password As Variant

password = Application.InputBox("Enter Password", "Password Protected")



Select Case password

    Case Is = False

      'do  nothing

   Case Is = "easy"

      Range("A1").Value = "This is secret code"

   Case Else

      MsgBox "Incorrect Password"

End Select
```

Result when you click the command button on the sheet:



Explanation: The macro uses the InputBox method of the Application object. If the users clicks Cancel, this method returns False and nothing happens (InputBox disappears). Only when the user knows the password ("easy" again), the secret code will be executed. If the entered password is incorrect, a MsgBox is displayed. Note that the user cannot take a look at the password in the Visual Basic Editor because the project is protected from being viewed.

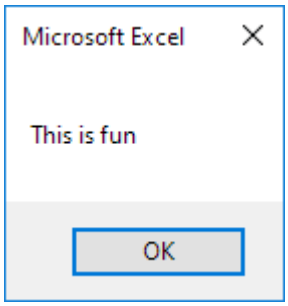
MsgBox

The MsgBox is a dialog box in Excel VBA you can use to inform the users of your program. Place a command button on your worksheet and add the following code lines:

1. A simple message.

```
MsgBox "This is fun"
```

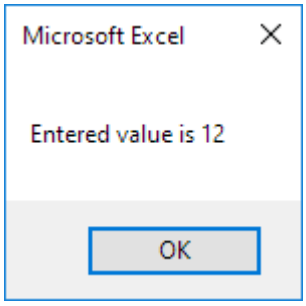
Result when you click the command button on the sheet:



2. A little more advanced message. First, enter a number into cell A1.

```
MsgBox "Entered value is " & Range("A1").Value
```

Result when you click the command button on the sheet:

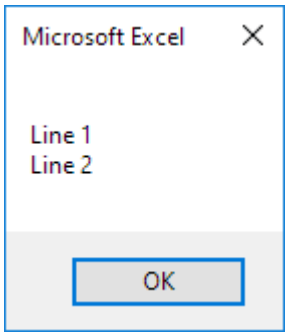


Note: we used the & operator to concatenate (join) two strings. Although Range("A1").value is not a string, it works here.

3. To start a new line in a message, use vbNewLine.

```
MsgBox "Line 1" & vbNewLine & "Line 2"
```

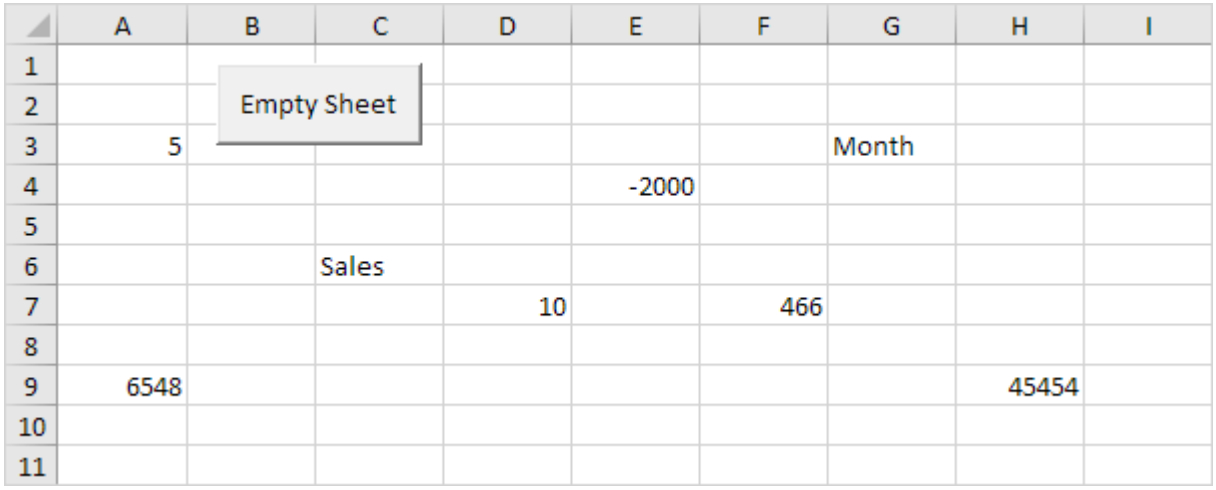
Result when you click the command button on the sheet:



MsgBox Function

The MsgBox function in Excel VBA can return a result while a simple

MsgBox cannot. Situation:



Place a command button on your worksheet and add the following code lines:

1. First, we declare a variable called answer of type Integer.

```
Dim answer As Integer
```

2. We use the MsgBox function to initialize the variable answer with the input from the user.

The MsgBox function, when using parentheses, has three arguments. The first part is used for the message in the message box. Use the second part to specify which buttons and icons you want to appear in the message box. The third part is displayed in the title bar of the message box.

```
answer = MsgBox("Are you sure you want to empty the sheet?", vbYesNo + vbQuestion, "Empty Sheet")
```

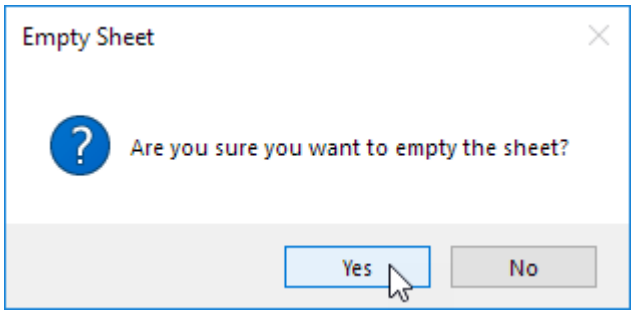
Note: Place your cursor on vbYesNo in the Visual Basic Editor and click F1 to see which other buttons and icons you can use. Instead of the constants vbYesNo and vbQuestion, you can also use the corresponding values 4 and 32.

3. If the user clicks the Yes button, Excel VBA empties the sheet. If the user clicks the No button, nothing happens. Add the following code lines to achieve this.

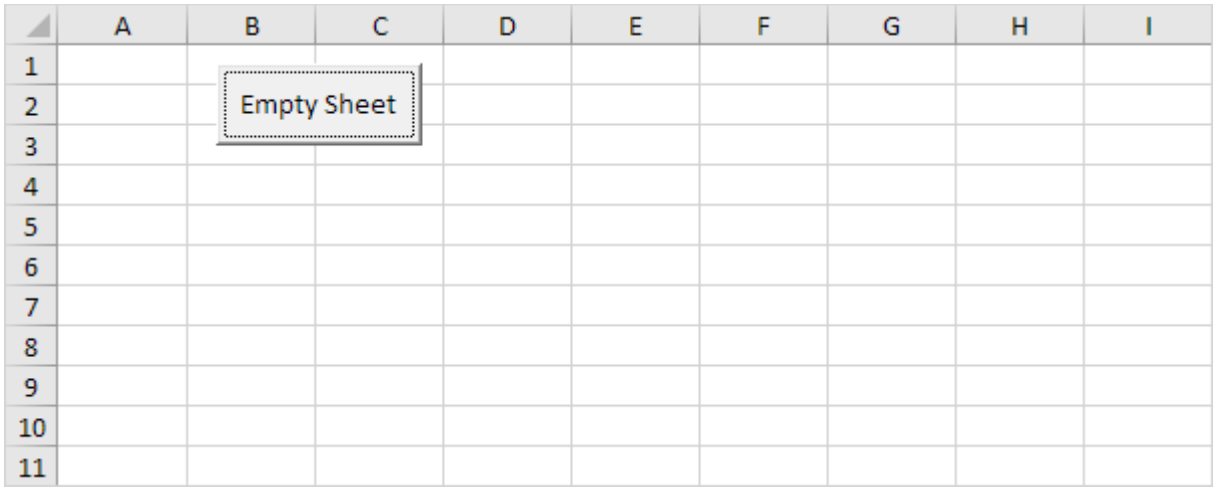
```
If answer = vbYes Then
    Cells.ClearContents
Else
    'do nothing
End If
```

4. Click the command button on the sheet.

5. Click Yes.



Result:



InputBox Function

You can use the InputBox function in Excel VBA to prompt the user to enter a value. Place a command button on your worksheet and add the following code lines:

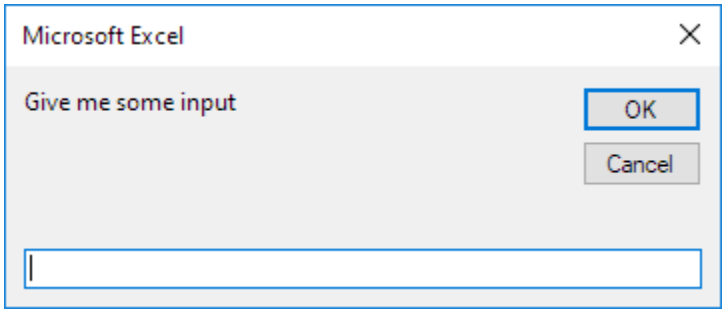
1. First, declare the variable myValue of type Variant.

```
Dim myValue As Variant
```

Note: we use a variable of type Variant here because a Variant variable can hold any type of value. This way the user can enter text, numbers, etc.

2. Add the following code line to show the input box.

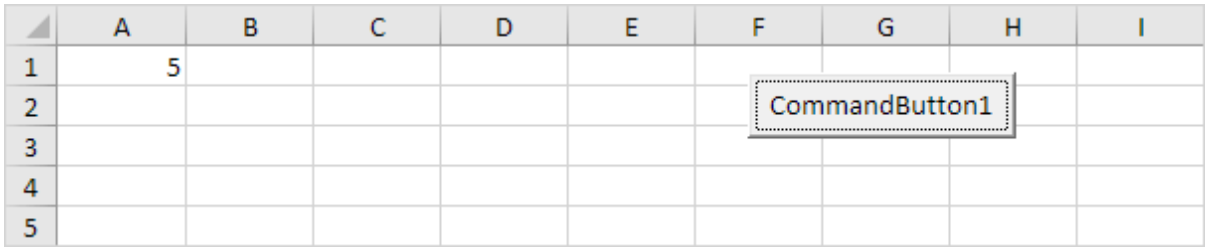
```
myValue = InputBox("Give me some input")
```



3. Write the value of myValue to cell A1.

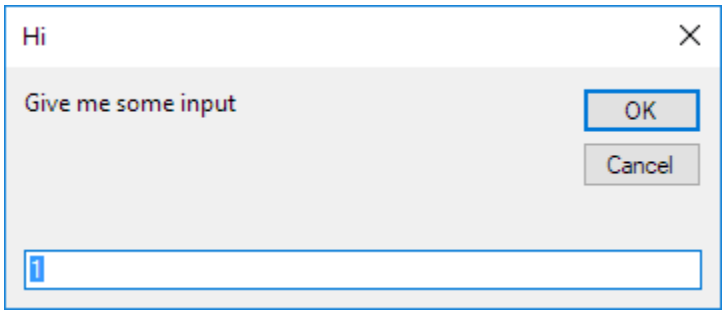
```
Range("A1").Value = myValue
```

Result when the user enters the value 5 and clicks the OK button.



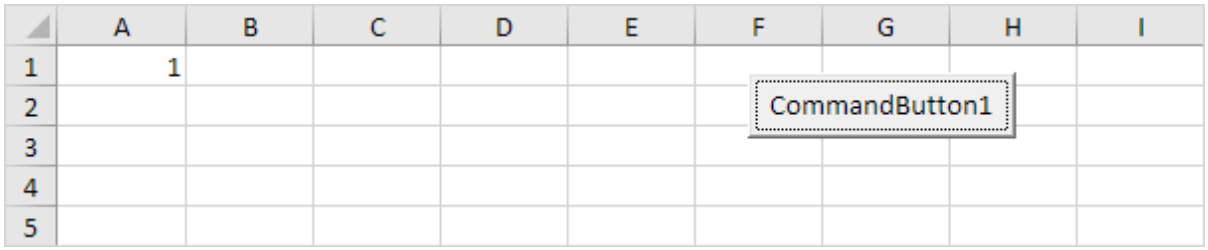
4. The InputBox function has more optional arguments. The following code line shows an input box with a title displayed in the title bar and has a default value.

```
myValue = InputBox("Give me some input", "Hi", 1)
```



The default value will be used if no other input is provided.

Result when the user only clicks the OK button.



Note: Place your cursor on InputBox in the Visual Basic Editor and click F1 for help on the other optional arguments.