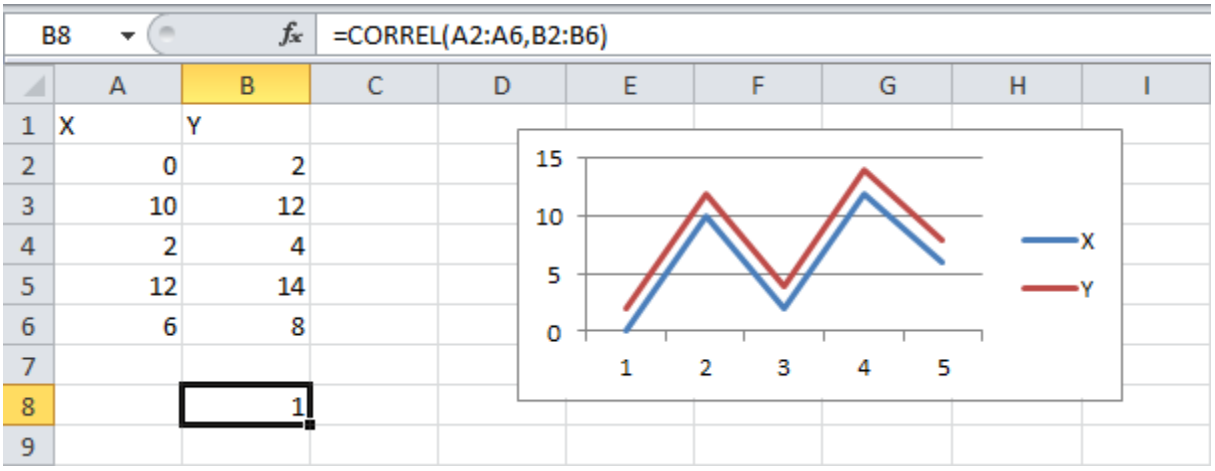
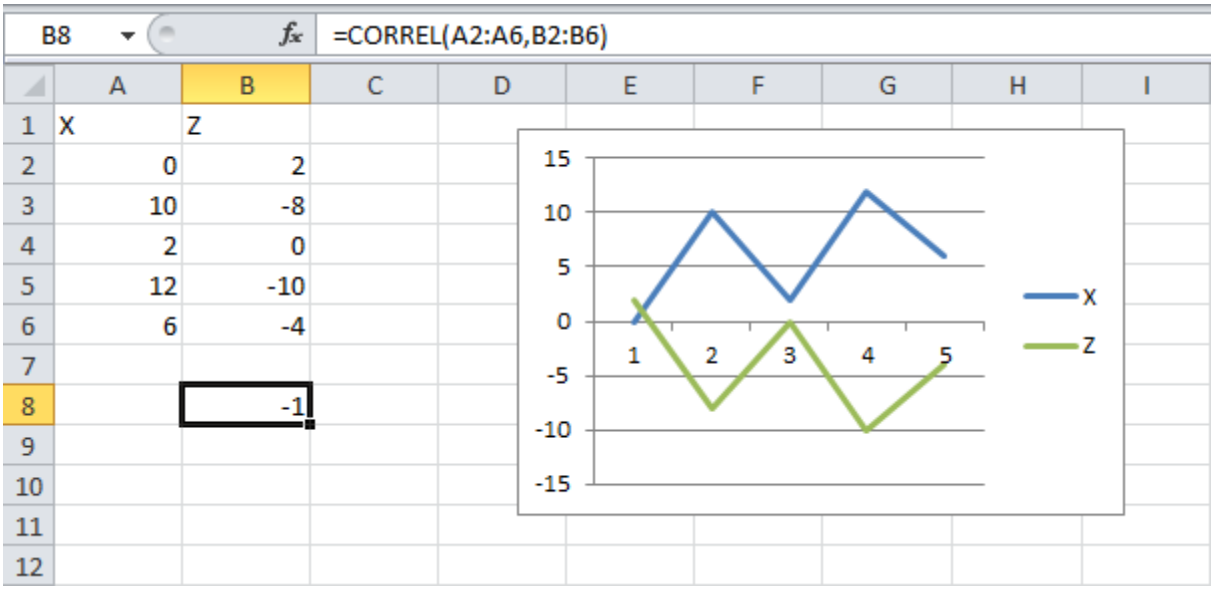
6- Correlation

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.



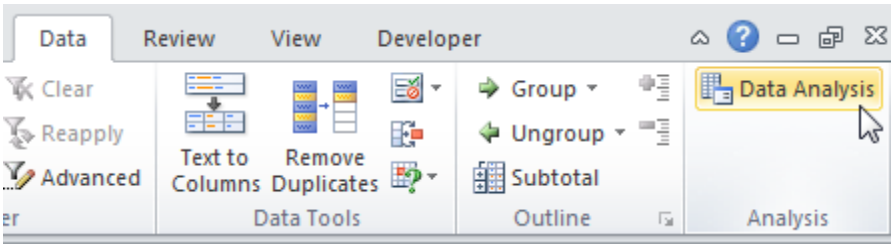
- A correlation coefficient of -1 indicates a perfect negative correlation. As variable X increases, variable Z decreases. As variable X decreases, variable Z increases.



- A correlation coefficient near 0 indicates no correlation.

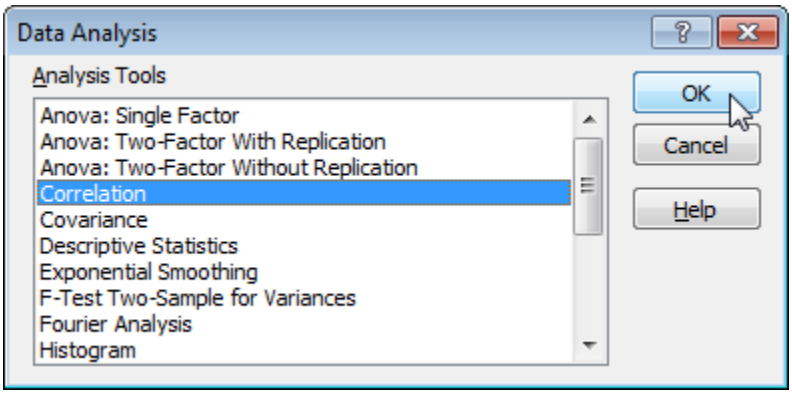
To use the Analysis Toolpak add-in in Excel to quickly generate correlation coefficients between multiple variables, execute the following steps.

1. On the Data tab, click Data Analysis.

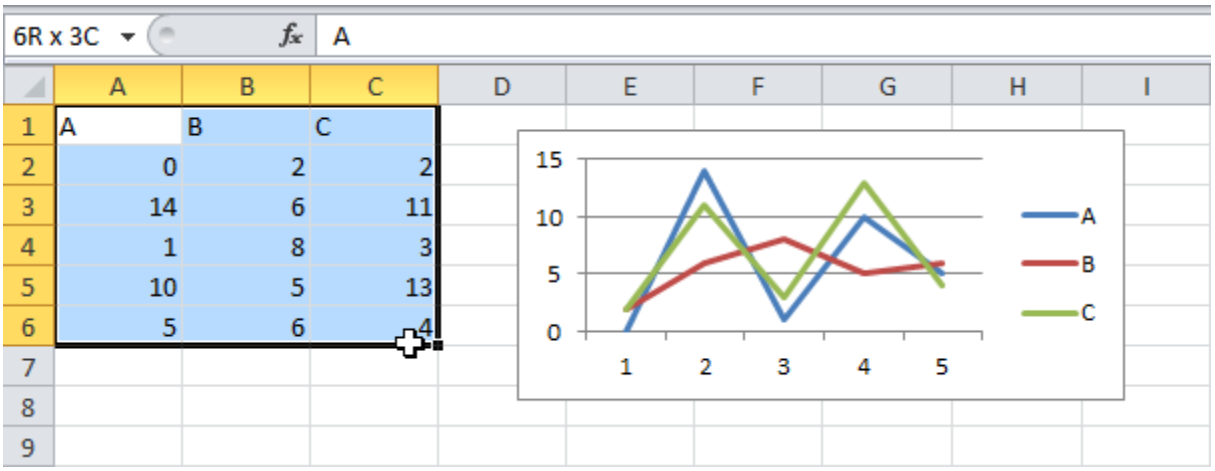


Note: can't find the Data Analysis button? Click here to load the Analysis ToolPak add-in.

2. Select Correlation and click OK.



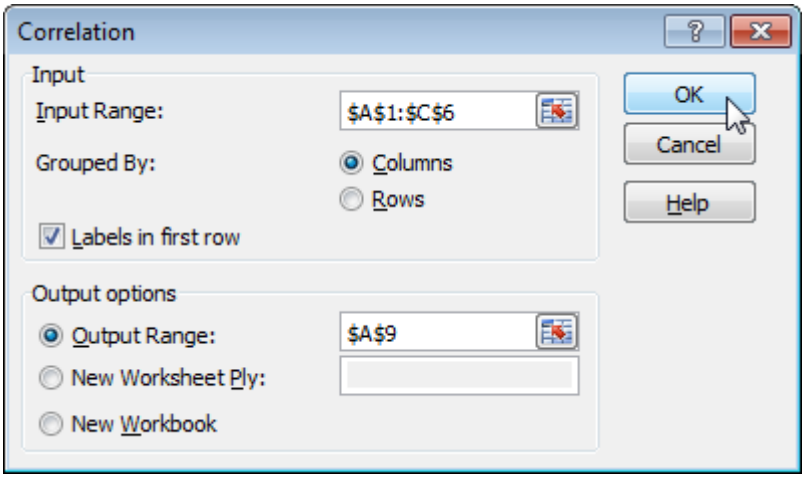
3. For example, select the range A1:C6 as the Input Range.



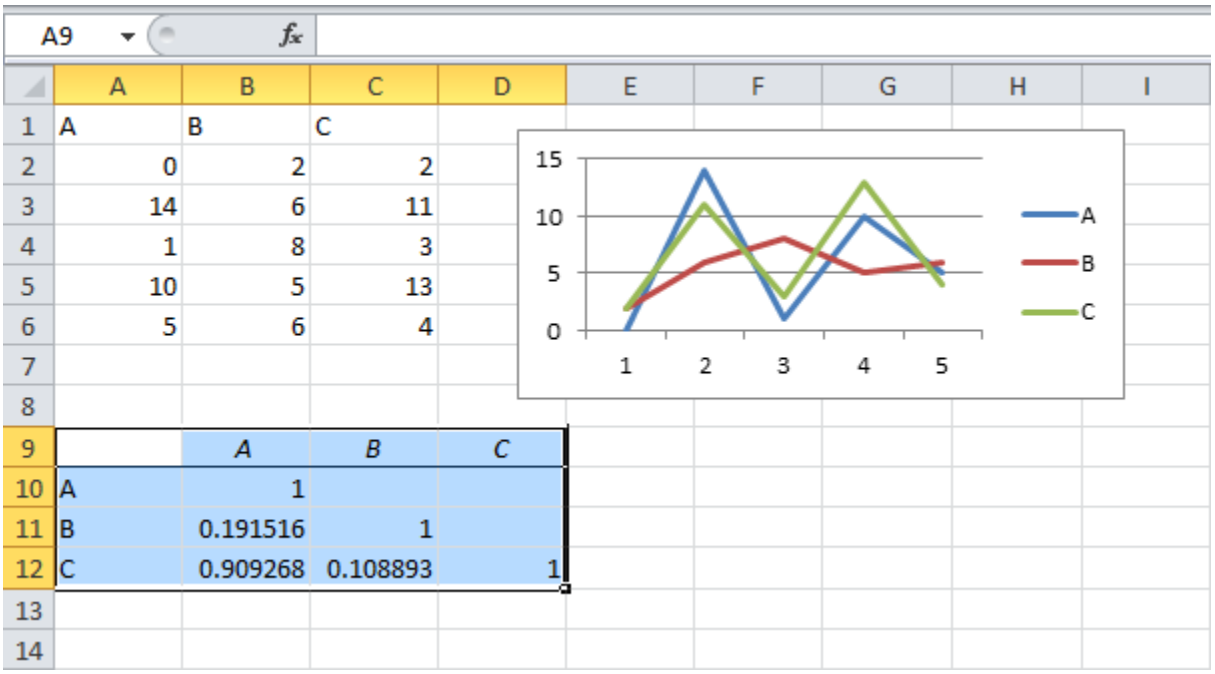
4. Check Labels in first row.

5. Select cell A9 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.

7-Regression

This example teaches you how to perform a regression analysis in Excel and how to interpret the Summary Output.

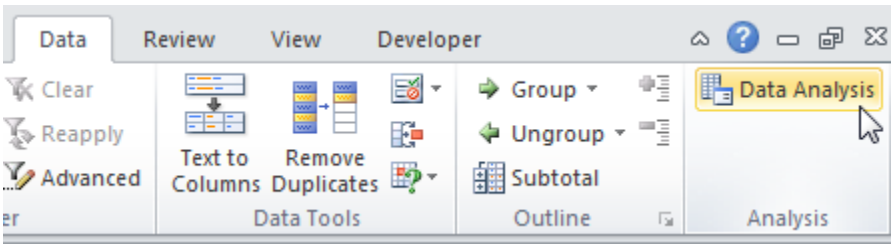
In this study a random sample of service call records for a computer repair operation were examined and the length of each call (in minutes) and the number of components repaired or replaced were recorded. The data is given below.

Minutes	Units
23	1
29	2
49	3
64	4
74	4
87	5
96	6
97	6
109	7
119	8
149	9
145	9
154	10
166	10

Below you can find our data. The big question is:

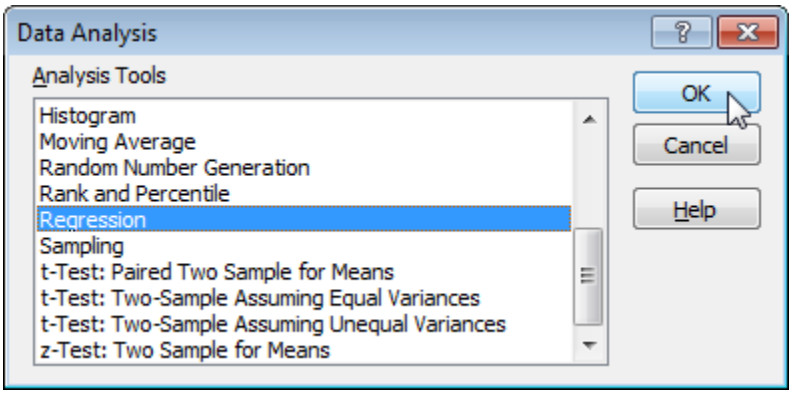
is there a relation between Minutes (Output) and Units (Input). In other words: can we predict Minutes if we know the unites?

1. On the Data tab, 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.



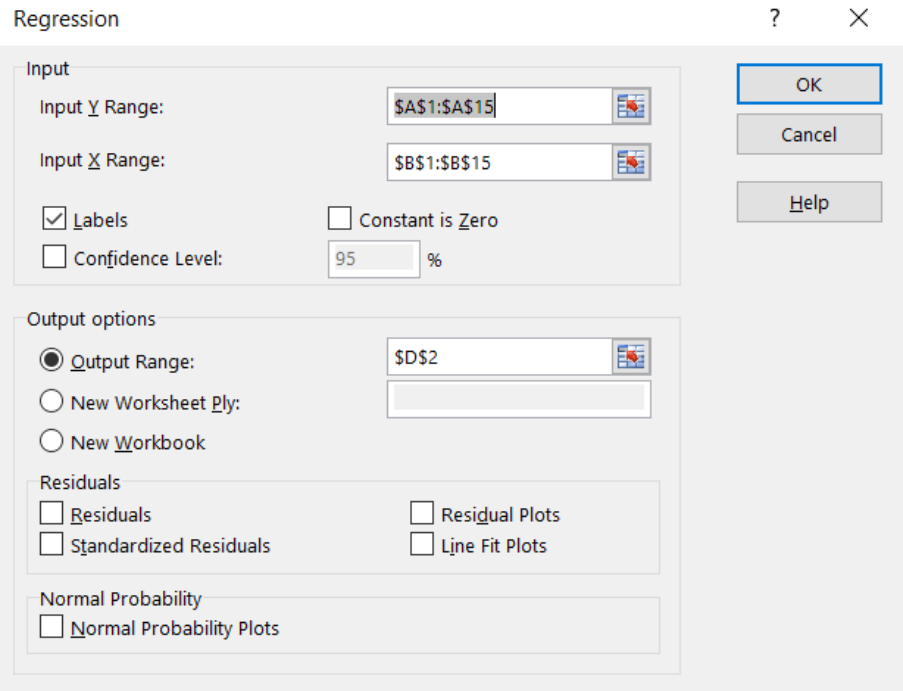
3. Select the Y Range (A2:A15). This is the predictor variable (also called dependent variable).

4. Select the X Range(B2:B15). These are the explanatory variables (also called independent variables). These columns must be adjacent to each other.

5. Check Labels.

6. Select an Output Range.

8. Click OK.



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

Minutes	Units
23	1
29	2
49	3
64	4
74	4
87	5
96	6
97	6
109	7
119	8
149	9
145	9
154	10
166	10

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.994
R Square	0.987
Adjusted R	0.986
Standard E	5.392
Observatic	14.000

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	1.000	27419.509	27419.509	943.201	0.000
Residual	12.000	348.848	29.071		
Total	13.000	27768.357			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	4.162	3.355	1.240	0.239	-3.148	11.472	-3.148	11.472
Units	15.509	0.505	30.712	0.000	14.409	16.609	14.409	16.609

Results

The regression line is: y = Minutes=4.162+15.509*Units. In other words, for increasing the units by one, the Time Minutes increases by 15.509, while there is 4.162 minutes does not depend on the unites.