MICROSOFT EXCEL 2007-2010

FORECASTING AND DATA ANALYSIS

NOTE

Unless otherwise stated, screenshots in this book were taken using Excel 2007 with a blue colour scheme and running on Windows Vista. There may, therefore, be minor visual differences if you are using a different colour scheme, using Excel 2010, or if you are running on a different version of Windows.

Separate screenshots and instructions are given where there is a significant difference between the versions or the operating systems. In all other cases concepts, discussions, procedures and functionality are the same.

CONTENTS

LESSON 1 - FORECASTING

In this lesson, you will learn how to:

- Work with Linear Regression
  - Use the **Forecast** function
  - Use the **Trend** function
  - Use the **Slope** and **Intercept** Function
- Work with Exponential Regression
  - Use the **Growth** function
- Use Exponential smoothing
- Use Data Analysis Tool Exponential Smoothing
- Use naïve forecasting
- Use Moving Average to forecast
- Use weighted Moving Average to forecast
- Use the Data Analysis tool Moving Average
- Work with Seasonal Forecasting

CONCEPT AND TERMS

Discussion

Forecasting is required in many situations to be able to do effective and efficient planning. A future electricity demands forecast requires to planning building new power stations; scheduling staff in a call centre next week requires forecasts of call volume; just imaging how much forecasting the planners behind London 2012 Olympic Games must have done. Forecasts can be required years in advance, or only a few minutes beforehand. Whatever the circumstances or time horizons involved, accurate forecasts are essential part of planning.

Often in forecasting, a key step is knowing when something can be forecast accurately, and when forecasts are no better than tossing a coin. Good forecasts capture the genuine patterns and relationships which exist in the historical data, but do not replicate past events that will not occur again.

In all environments where numbers are collected and people make use of these numbers the ability to forecast or extrapolate data may be required. It doesn't really matter if you are talking about sales figures, expenses, man-hours, growth, market shares etc. to create a budget you need to forecast. To be able to forecast in Excel you must have historical data. In this workbook time series data is used for the forecasts. The periods can be quarters, months, minutes, hours, and years. The predictability of an event or a quantity depends on several factors including:

- How well you understand the factors that contribute to it;
- How much data are available;
- Whether the forecasts can affect the thing you are trying to forecast.

It is not possible to forecast everything. It is easy to forecast the time of the sunrise tomorrow, but it is impossible to forecast tomorrow's lottery numbers and difficult to forecast, who is going to win this year's European Song Contest

Can you trust a forecast created using Excel?

In Excel you have different methods to calculate a forecast, but Excel can only base the calculations on known data. You can use What-if analysis tools to manipulate your forecasts based on knowledge you have about the future (you will get a new product line, you have to close down 5 stores, you will spend money on marketing).

In forecasting you are going to look at the trends that the data have and use these trends to help forecast future values or values outside the measured data. The trends can also be used to infill data where data is missing in the collected data. You will also forecast seasonal data and also look at data which are not suitable for forecasting.

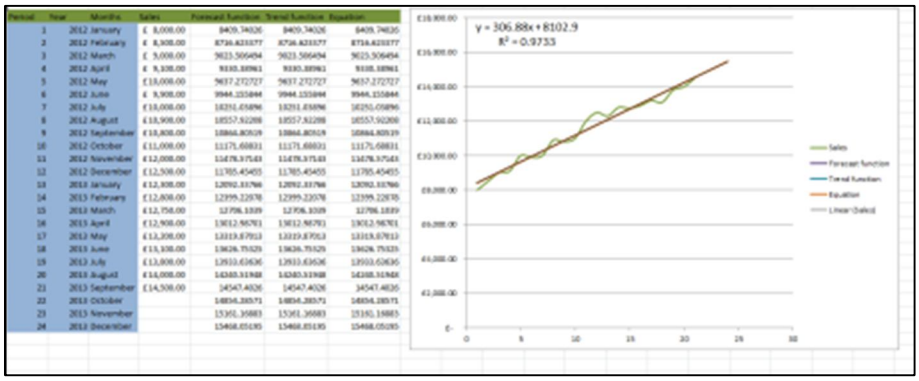
LINEAR REGRESSION

Discussion

Linear Regression analysis is a statistical technique for estimating the relationships among variables. In other words how do the sales figures change over time?

If the goal is prediction, or forecasting, linear regression can be used to fit a predictive model to an observed data set of y and x values or known actual data (y) over time (x) (time series data). After developing such a model, if an additional value of x (a new period) is then given without its accompanying value of y, the fitted model can be used to make a prediction of the value of y.

The sales figures (y) are known for a number of periods (x) it makes it possible forecasting sales (y) for future periods (x).



In Excel the linear regression can be calculated using the **Forecast** function, The **Trend** function, the **Fill-Handle**, by calculating the equation: **Y = mX + c**, and by adding a **Trendline** to a chart.

The Data Analysis Tool **Regression** is an analysis tool to return important information working with linear regression such as the **Slope,** the **Y-Interceptor, R-square**, and other statistical useful information. The different terms will be explained later in this workbook.

You can also calculate the **Slope** and the Y **Interceptor** using the functions **Slope** and **Intercept**.

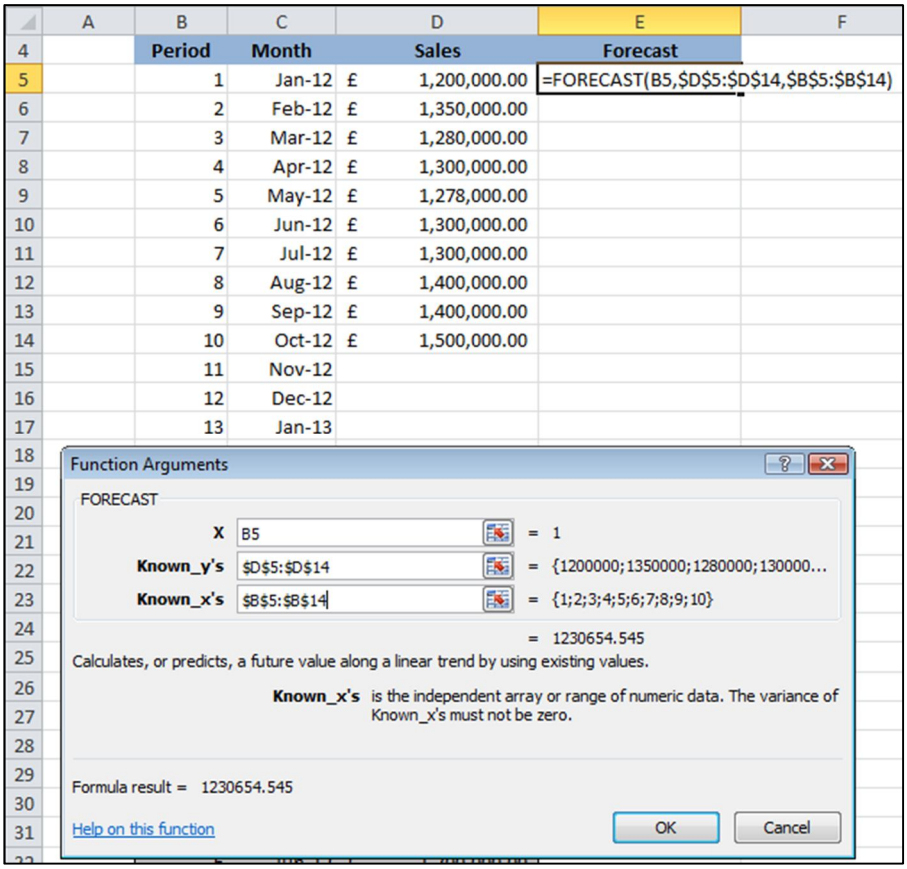
THE FORECAST FUNCTION

Discussion

Microsoft Excel provides a variety of functions you can use to calculate forecasts. One of them is the **Forecast** function. The **Forecast** function can calculate trend, a linear

forecast, or linear regression for as many future periods needed based on known actual data from previous periods.

The **Forecast** function consists of three required arguments, in the following order: **X**, **Known_y's**, and **Known_x's**. **X** is the period for which you want to calculate the forecast. **Known_y's** is the array with the known values. **Known_x's** is the array with the known periods (must be a numeric range and not dates, months or years).



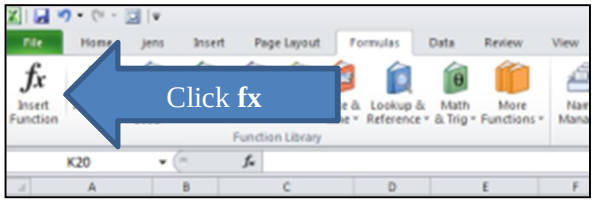
Procedures

1. To use a **Forecast** function, first create a data range containing known factual data and responding periods.

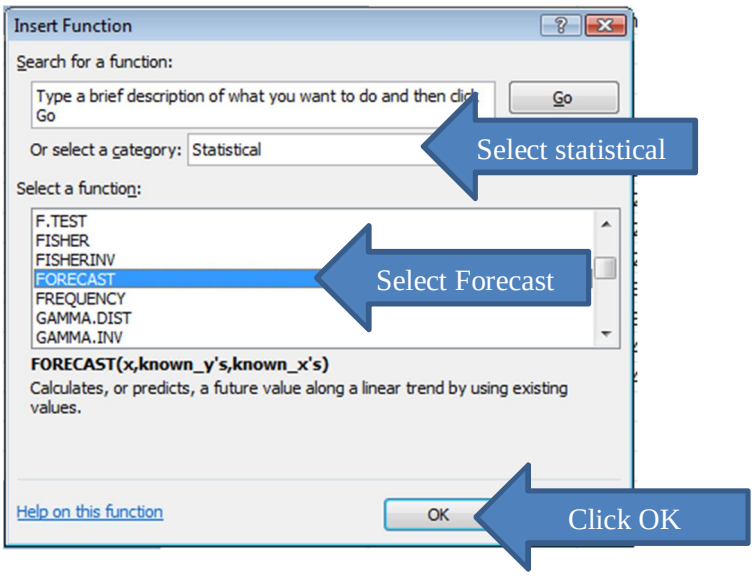
2. Click in the cell where you want to place the function.

3. Click on the **Formulas** tab.

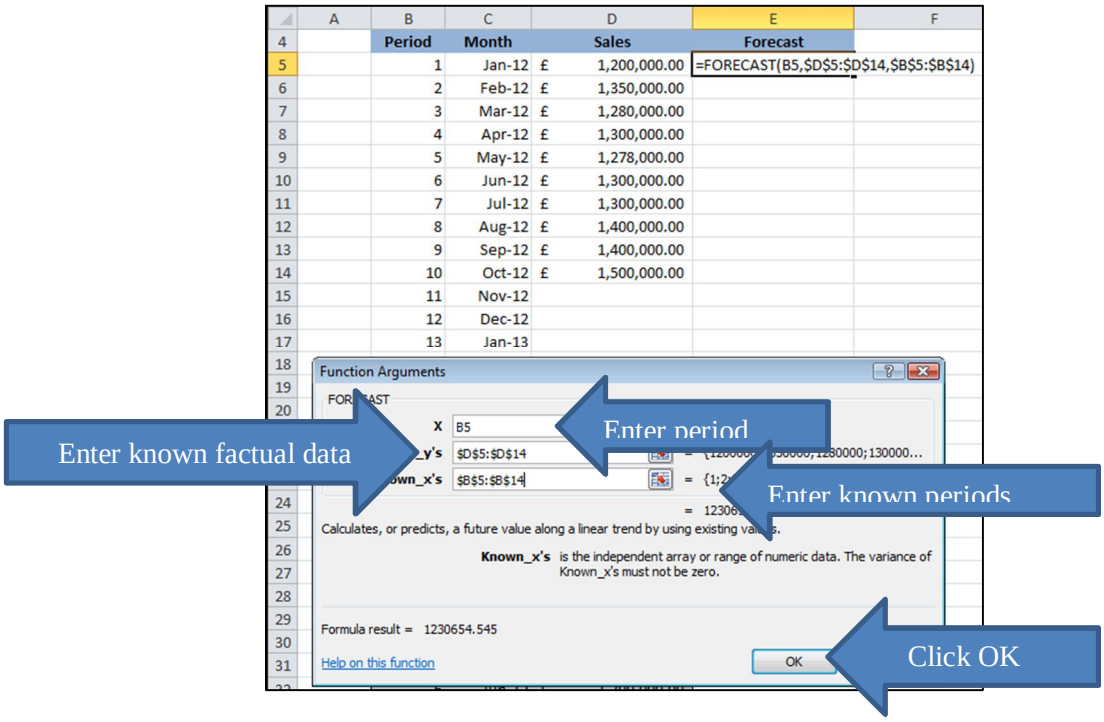
4. In the **Function Library** group, click on the **Insert Function** button.



5. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.
6. Click on **Forecast**.
7. Click on the **OK** button.



8. In the **Function Arguments** dialog box, click in the **X** box.
9. Enter the cell reference for the cell with the period information.
10. In the **Known_y's** box, select the cells containing the known values.
11. Make the cell references absolute (Press F4 or add the $ signs).
12. In the **Known_x's** box, select the cells containing the known periods.
13. Make the cell references absolute (Press F4 or add the $ signs).
14. Click **OK**.



15. Copy down the **Forecast** function to get the forecast for known periods and unknown periods.

Period	Month		Sales		Forecast
1	Jan-12	£	1,200,000.00	£	1,230,654.55
2	Feb-12	£	1,350,000.00	£	1,252,909.09
3	Mar-12	£	1,280,000.00	£	1,275,163.64
4	Apr-12	£	1,300,000.00	£	1,297,418.18
5	May-12	£	1,278,000.00	£	1,319,672.73
6	Jun-12	£	1,300,000.00	£	1,341,927.27
7	Jul-12	£	1,300,000.00	£	1,364,181.82
8	Aug-12	£	1,400,000.00	£	1,386,436.36
9	Sep-12	£	1,400,000.00	£	1,408,690.91
10	Oct-12	£	1,500,000.00	£	1,430,945.45
11	Nov-12			£	1,453,200.00
12	Dec-12			£	1,475,454.55

The forecast function is built in the **Fill handle**. Select the known periods' factual data and drag down the **Fill handle** and the **Fill** handle will forecast the unknown periods.

Period	Month	Sales	Forecast
1	Jan-12	£ 1,200,000.00	£ 1,243,381.82
2	Feb-12	£ 1,500,000.00	£ 1,255,030.30
3	Mar-12	£ 1,180,000.00	£ 1,266,678.79
4	Apr-12	£ 1,300,000.00	£ 1,278,327.27
5	May-12	£ 1,278,000.00	£ 1,289,975.76
6	Jun-12	£ 1,100,000.00	£ 1,301,624.24
7	Jul-12	£ 1,300,000.00	£ 1,313,272.73
8	Aug-12	£ 1,200,000.00	£ 1,324,921.21
9	Sep-12	£ 1,400,000.00	£ 1,336,569.70
10	Oct-12	£ 1,500,000.00	£
11	Nov-12		£
12	Dec-12		£ 1,371,515.15

Drag fill handle

Period	Month	Sales	Forecast
1	Jan-12	£ 1,200,000.00	£ 1,243,381.82
2	Feb-12	£ 1,500,000.00	£ 1,255,030.30
3	Mar-12	£ 1,180,000.00	£ 1,266,678.79
4	Apr-12	£ 1,300,000.00	£ 1,278,327.27
5	May-12	£ 1,278,000.00	£ 1,289,975.76
6	Jun-12	£ 1,100,000.00	£ 1,301,624.24
7	Jul-12	£ 1,300,000.00	£ 1,313,272.73
8	Aug-12	£ 1,200,000.00	£ 1,324,921.21
9	Sep-12	£ 1,400,000.00	£ 1,336,569.70
10	Oct-12	£ 1,500,000.00	£ 1,348,218.18
11	Nov-12	£ 1,359,866.67	£ 1,359,866.67
12	Dec-12	£ 1,371,515.15	£ 1,371,515.15

If you compare the result using the fill handle in the Sales column and using the **Forecast** function in the Forecast column you can see that both methods will return the same result for November and December. Exactly as the **Forecast** function the fill handle can predict as many future periods needed based on known actual data.

THE TREND FUNCTION

Discussion

The **Trend** can calculate trend, a linear forecast, or linear regression for as many future periods needed based on data from previous periods.

The **Trend** function is an array function and the steps to use it are different from "normal" functions in Excel. The whole range where you want the result to be displayed must be selected and after the needed arguments have been entered the function dialog box the keys **Ctrl Shift Enter** must be pressed instead of pressing **Enter**. Then the function will return the result for the whole range in the selected range.

The **Trend** function consists of one required argument and three optional arguments, in the following order: **Known_y's**, **Known_x's**, **New_x's**, and **Const**. **Known_y's** is the array with the known values. **Known_x's** is the array with the known periods (must be a numeric range and not dates, months or years). **New_x's** is the array with future periods if the **Trend** function is used for forecasting. **Const** is a logical value specifying whether to force the constant **c** (the Y interceptor) to equal 0. In other words by forcing the Y interceptor to equal 0 the trendline will have a start from 0,0 (X=0 and y=0).

Procedures

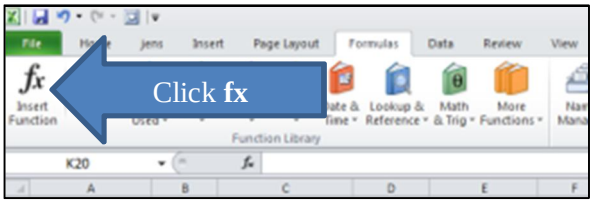
1. To use a **Trend** function, first create a data range containing known factual data and responding periods.

	A	B	C	D
1	Forecasting			
2				
3				
4	Period	Month	Sales	Trend
5	1	Jan-12	£ 1,200,000.00	
6	2	Feb-12	£ 1,350,000.00	
7	3	Mar-12	£ 1,280,000.00	
8	4	Apr-12	£ 1,300,000.00	
9	5	May-12	£ 1,278,000.00	
10	6	Jun-12	£ 1,300,000.00	
11	7	Jul-12	£ 1,300,000.00	
12	8	Aug-12	£ 1,400,000.00	
13	9	Sep-12	£ 1,400,000.00	
14	10	Oct-12	£ 1,500,000.00	
15	11	Nov-12		
16	12	Dec-12		

2. Select the range where you want the result of the function.

3. Click on the **Formulas** tab.

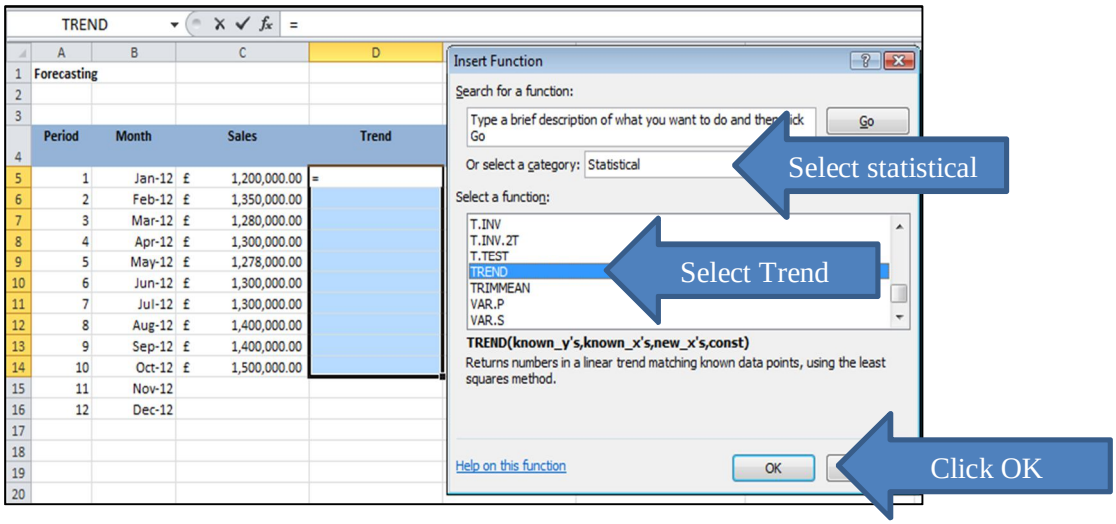
4. In the **Function Library** group, click on the **Insert Function** button.



5. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.

6. Click on **Trend**.

7. Click on the **OK** button.



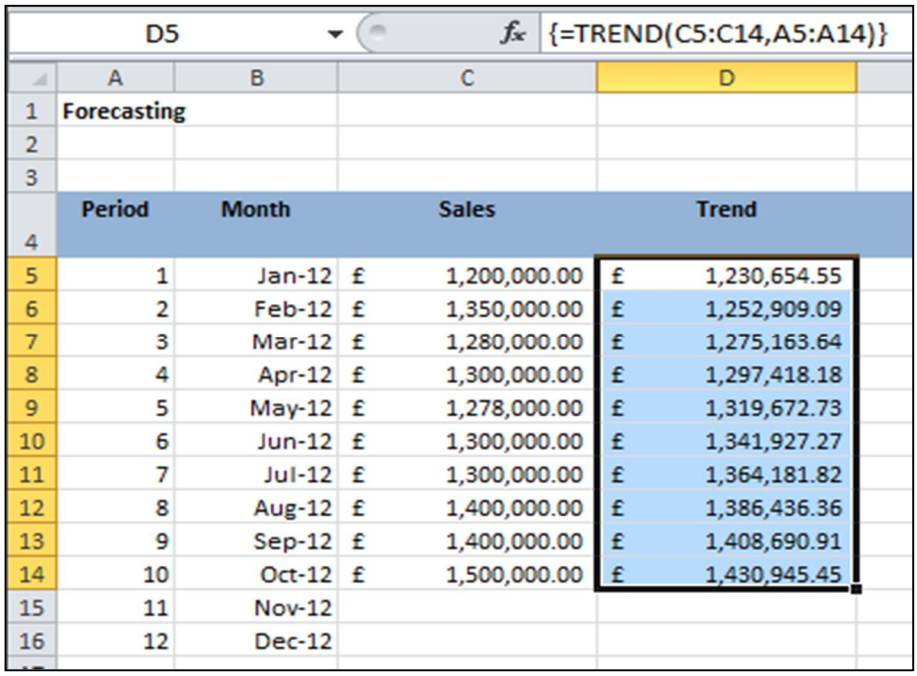
8. In the **Function Arguments** dialog box, click in the **Known_y's** box.

9. Select the cells containing the known values.

10. In the **Known_x's** box, select the cells containing the known periods.

11. Press **Ctrl Shift Enter.**

The array functions will add the result to the whole selected array and will look different in the formula bar. Excel will display the formula enclosed in curly brackets { }.



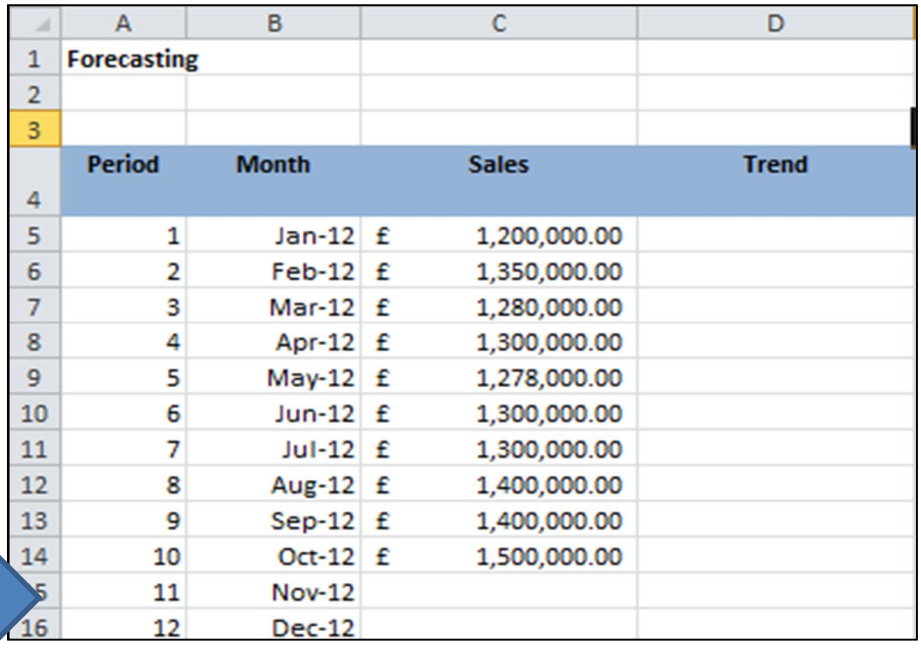
D5 {=TREND(C5:C14,A5:A14)}

	A	B	C	D
1	Forecasting			
2				
3				
4	Period	Month	Sales	Trend
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45
15	11	Nov-12		
16	12	Dec-12		

If the Trend function is used for forecasting:

Procedures

1. To use a **Trend** function for forecasting, first create a data range containing known factual data and responding periods and add the periods for which you want to forecast.



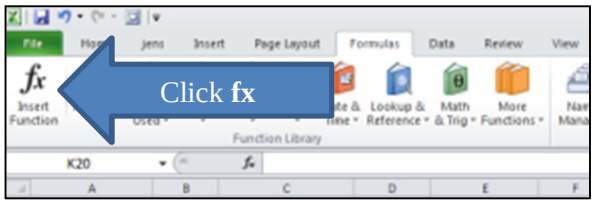
	A	B	C	D
1	Forecasting			
2				
3				
4	Period	Month	Sales	Trend
5	1	Jan-12	£ 1,200,000.00	
6	2	Feb-12	£ 1,350,000.00	
7	3	Mar-12	£ 1,280,000.00	
8	4	Apr-12	£ 1,300,000.00	
9	5	May-12	£ 1,278,000.00	
10	6	Jun-12	£ 1,300,000.00	
11	7	Jul-12	£ 1,300,000.00	
12	8	Aug-12	£ 1,400,000.00	
13	9	Sep-12	£ 1,400,000.00	
14	10	Oct-12	£ 1,500,000.00	
15	11	Nov-12		
16	12	Dec-12		

Period 11 and 12

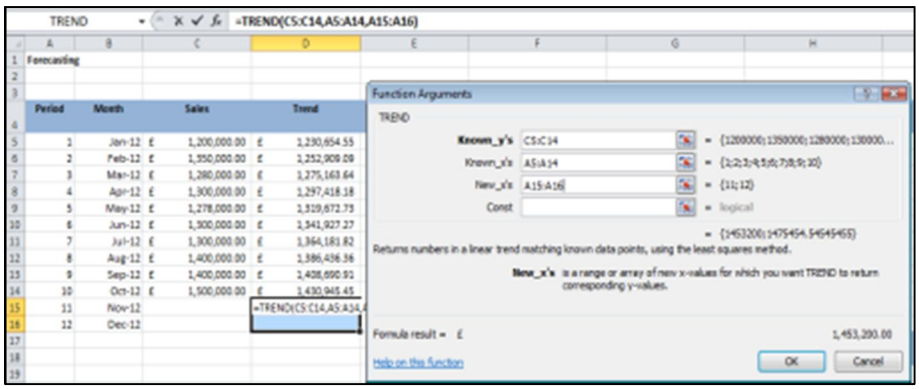
2. Select the range where you want the result of the function.

3. Click on the **Formulas** tab.

4. In the **Function Library** group, click on the **Insert Function** button.



5. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.
6. Click on **Trend**.
7. Click on the **OK** button.
8. In the **Function Arguments** dialog box, click in the **Known_y's** box.
9. Select the cells containing the known values.
10. In the **Known_x's** box, select the cells containing the known periods.
11. In the **New_x's box**, select the cells containing the unknown periods.
12. Press **Ctrl Shift Enter**.



> *The **Trend** function will return exactly the same result as the forecast function and as the **Forecast** function able to forecast as many future periods needed.*

THE SLOPE AND INTERCEPT FUNCTION

Discussion

To make it easy to calculate the trend or forecast equation **Y = mX + c** you need the slope and y interceptor.

First the equation:

Y = mX + c

Where:

Y is the unknown measure (the value you want to forecast)

X is the X value (the period in time series data) for which you want to know the value of Y

Where:

m is the gradient of the line or the slope

Where:

c is the Y interceptor of the line (or Y value when there is no X value or X =0)

The slope equation looks like this:

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

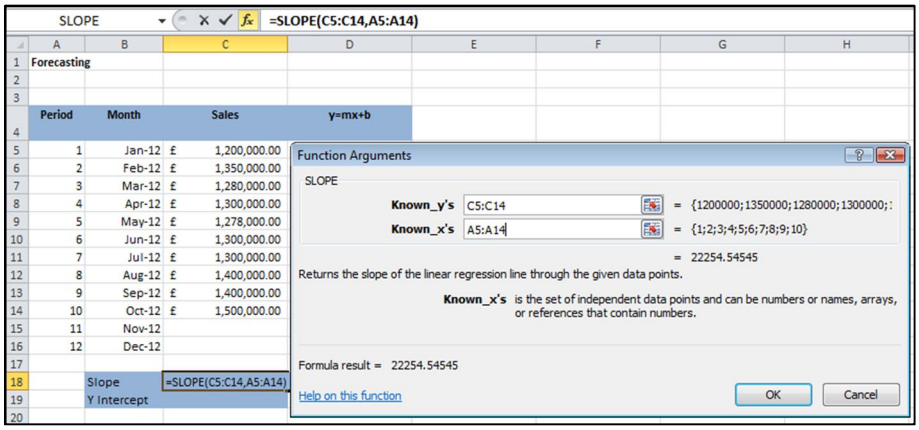
The **Slope** function consists of two required arguments, in the following order: **Known_y's**, and **Known_x's**. **Known_y's** is the array with the known values. **Known_x's** is the array with the known responding periods (must be a numeric range and not dates, months or years).

Procedures

1. Select the cell where the result is to be displayed.
2. Click on the **Formulas** tab.
3. In the **Function Library** group, click on the **Insert Function** button.
4. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.
5. Click on **Slope**.
6. Click on the OK button.
7. In the **Known_y's** box, select the cells containing the known values.

8. In the **Known_x's** box, select the cells containing the responding known periods.

9. Click on the **OK** button.



The **Intercept** function consists of two required arguments, in the following order: **Known_y's**, and **Known_x's**. **Known_y's** is the array with the known values. **Known_x's** is the array with the known responding periods (must be a numeric range and not dates, months or years).

Procedures

1. Select the cell where the result is to be displayed.

2. Click on the **Formulas** tab.

3. In the **Function Library** group, click on the **Insert Function** button.

4. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.

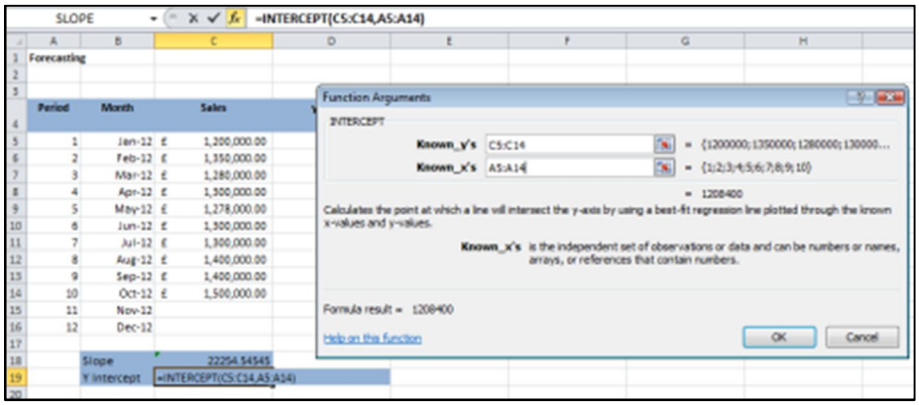
5. Click on **Intercept**.

6. Click on the OK button.

7. In the **Known_y's** box, select the cells containing the known values.

8. In the **Known_x's** box, select the cells containing the responding known periods.

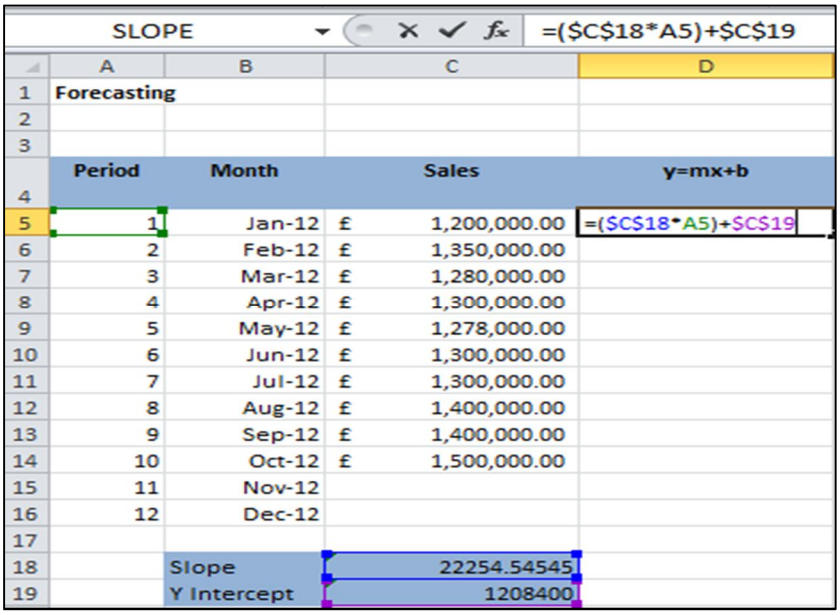
9. Click on the **OK** button.



The **Slope** and **Y interceptor** can now be used to calculate the trend and forecast using the trend / forecast equation:

Procedures

1. Select the cell where the result is to be displayed.
2. Type =(click in the cell with the **slope**. Press F4 to lock the cell reference.
3. Type * click in the cell with the period number (x). Type)
4. Type + click in the cell with **y interceptor**. Press F4 to lock the cell reference.
5. Press **enter**.
6. Copy down the equation.



The **Forecast** function, the **Trend** function and the trend / forecast equation will return exactly the same result. It does not really matter which method used. The equation can be used to calculate the forecast for as many future periods needed.

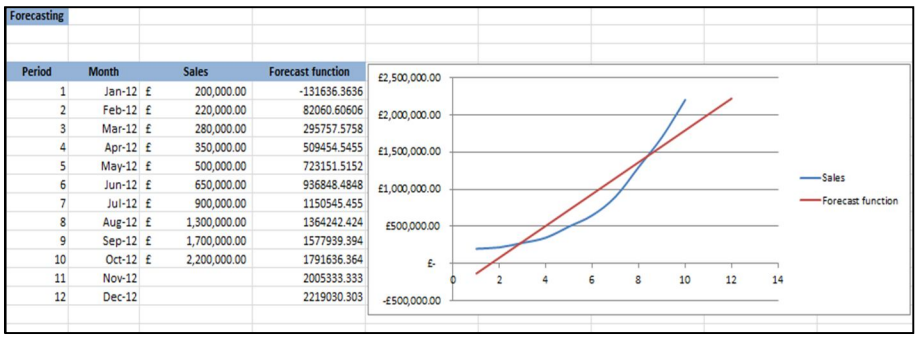
	A	B	C	D	E	F
1	Forecasting					
2						
3						
4	Period	Month	Sales	y=mx+b	Forecast function	Trend function
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	£ 1,230,654.55	£ 1,230,654.55
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	£ 1,252,909.09	£ 1,252,909.09
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	£ 1,275,163.64	£ 1,275,163.64
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	£ 1,297,418.18	£ 1,297,418.18
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	£ 1,319,672.73	£ 1,319,672.73
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	£ 1,341,927.27	£ 1,341,927.27
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	£ 1,364,181.82	£ 1,364,181.82
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	£ 1,386,436.36	£ 1,386,436.36
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	£ 1,408,690.91	£ 1,408,690.91
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	£ 1,430,945.45	£ 1,430,945.45
15	11	Nov-12		£ 1,453,200.00	£ 1,453,200.00	£ 1,453,200.00
16	12	Dec-12		£ 1,475,454.55	£ 1,475,454.55	£ 1,475,454.55
17						
18		Slope	22254.54545			
19		Y Intercept	1208400			

EXPONENTIAL REGRESSION

Discussion

Sometimes the growth in a model is not linear, but it is exponential. If the growth is exponential Excel has forecasting tools to replace the **Forecast** and **Trend** function.

In the example below the sales has an exponential growth rate. The **Forecast** function forecasts period 11 to 2005333.333. This is not a realistic forecast, because the known sales for period 10 are already £ 2,200,000.00.



In this section you will see how you can forecast an exponential growth.

THE GROWTH FUNCTION

Discussion

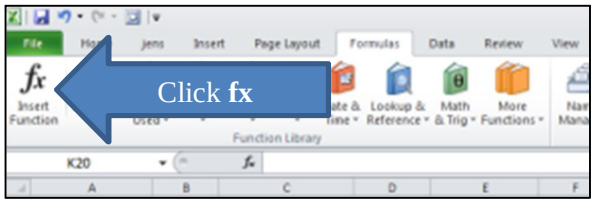
The **Growth** function can calculate exponential growth and an exponential growth forecast. The **Growth** function is an array function and the steps to use it are different from "normal" functions in Excel. The whole range where you want the result to be displayed must be selected and after the needed arguments have been entered the **Growth** function dialog box the keys **Ctrl Shift Enter** must be pressed instead of pressing **Enter**. Then the function will return the result for the whole range in the selected range.

The **Growth** function consists of one required argument and three optional arguments, in the following order: **Known_y's**, **Known_x's**, **New_x's**, and **Const**. **Known_y's** is the array with the known values. **Known_x's** is the array with the known periods (must be a numeric range and not dates, months or years). **New_x's** is the array with future periods if the **Growth** function is used for forecasting. **Const** is a logical value specifying whether to force the constant **c** (the Y interceptor) to equal 0. In other words by forcing the Y interceptor to equal 0 the trendline will have a start from 0,0 (X=0 and y=0).Procedures

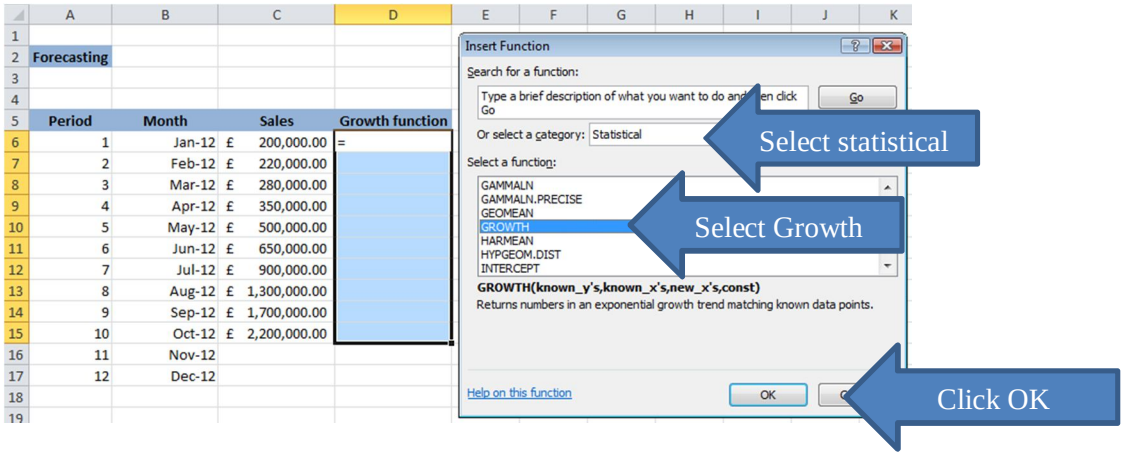
1. To use a **Growth** function, first create a data range with factual known data and responding periods.

5	Period	Month	Sales
6	1	Jan-12	£ 200,000.00
7	2	Feb-12	£ 220,000.00
8	3	Mar-12	£ 280,000.00
9	4	Apr-12	£ 350,000.00
10	5	May-12	£ 500,000.00
11	6	Jun-12	£ 650,000.00
12	7	Jul-12	£ 900,000.00
13	8	Aug-12	£ 1,300,000.00
14	9	Sep-12	£ 1,700,000.00
15	10	Oct-12	£ 2,200,000.00
16	11	Nov-12	
17	12	Dec-12	

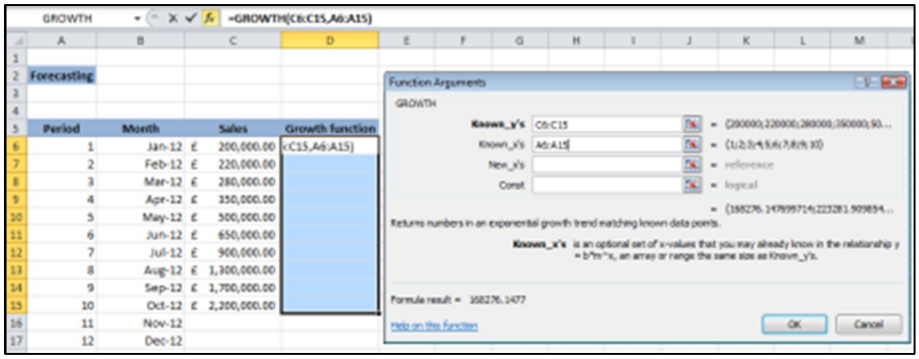
2. Select the array where you want the result of the function.
3. Click on the **Formulas** tab.
4. In the **Function Library** group, click on the **Insert Function** button.



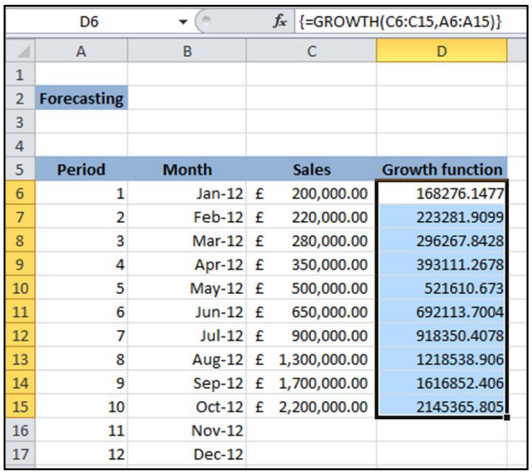
5. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.
6. Click on **Growth**.
7. Click on the **OK** button.



8. In the **Function Arguments** dialog box, click in the **Known_y's** box.
9. Select the cells containing the known values.
10. In the **Known_x's** box, select the cells containing the known periods.
11. Press **Ctrl Shift Enter.**



The array functions will add the result to the whole selected array and will look different in the formula bar. Excel will display the formula enclosed in curly brackets { }.

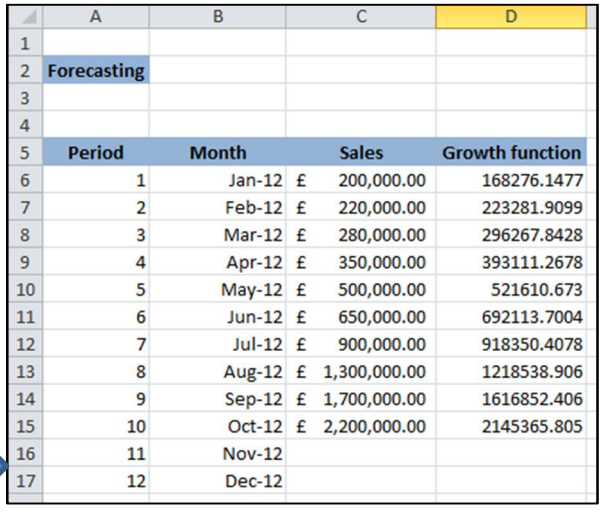


	A	B	C	D
1				
2	Forecasting			
3				
4				
5	Period	Month	Sales	Growth function
6	1	Jan-12	£ 200,000.00	168276.1477
7	2	Feb-12	£ 220,000.00	223281.9099
8	3	Mar-12	£ 280,000.00	296267.8428
9	4	Apr-12	£ 350,000.00	393111.2678
10	5	May-12	£ 500,000.00	521610.673
11	6	Jun-12	£ 650,000.00	692113.7004
12	7	Jul-12	£ 900,000.00	918350.4078
13	8	Aug-12	£ 1,300,000.00	1218538.906
14	9	Sep-12	£ 1,700,000.00	1616852.406
15	10	Oct-12	£ 2,200,000.00	2145365.805
16	11	Nov-12		
17	12	Dec-12		

The Growth function procedures if it is used for forecasting:

Procedures

1. To use a **Growth** function for forecasting first create a data range with factual known data and responding periods. Add the periods for the future you want to forecast.

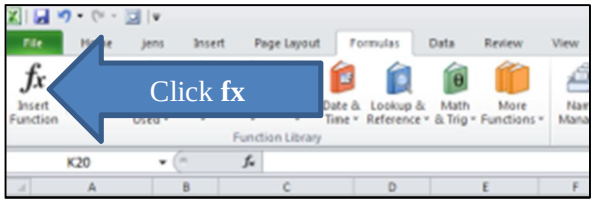


	A	B	C	D
1				
2	Forecasting			
3				
4				
5	Period	Month	Sales	Growth function
6	1	Jan-12	£ 200,000.00	168276.1477
7	2	Feb-12	£ 220,000.00	223281.9099
8	3	Mar-12	£ 280,000.00	296267.8428
9	4	Apr-12	£ 350,000.00	393111.2678
10	5	May-12	£ 500,000.00	521610.673
11	6	Jun-12	£ 650,000.00	692113.7004
12	7	Jul-12	£ 900,000.00	918350.4078
13	8	Aug-12	£ 1,300,000.00	1218538.906
14	9	Sep-12	£ 1,700,000.00	1616852.406
15	10	Oct-12	£ 2,200,000.00	2145365.805
16	11	Nov-12		
17	12	Dec-12		

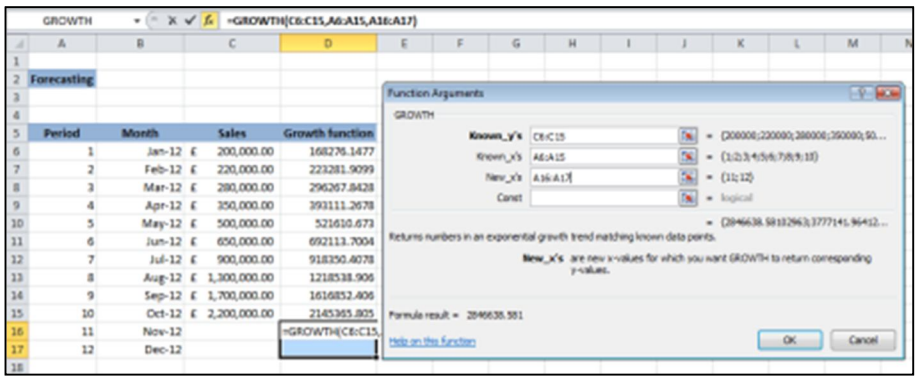
2. Select the array where you want the result of the function.

3. Click on the **Formulas** tab.

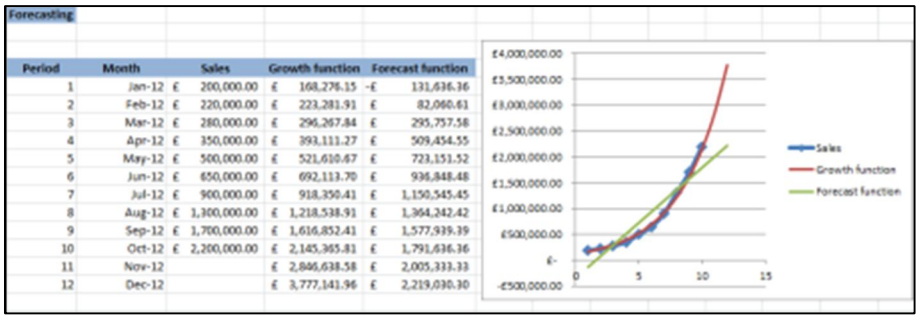
4. In the **Function Library** group, click on the **Insert Function** button.



5. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.
6. Click on **Growth**.
7. Click on the **OK** button.
8. In the **Function Arguments** dialog box, click in the **Known_y's** box.
9. Select the cells containing the known values.
10. In the **Known_x's** box, select the cells containing the known periods.
11. In the **New_x's box**, select the cells containing the unknown periods.
12. Press **Ctrl Shift Enter**.



If you compare the **Growth** function with the **Forecast** function in a chart it is obvious that in this example you get a much more accurate forecast using the **Growth** function.



Period	Month	Sales	Growth function	Forecast function
1	Jan-12	£ 200,000.00	£ 168,276.15	-£ 131,636.36
2	Feb-12	£ 220,000.00	£ 223,281.91	£ 82,060.61
3	Mar-12	£ 280,000.00	£ 296,267.84	£ 295,757.58
4	Apr-12	£ 350,000.00	£ 393,111.27	£ 509,454.55
5	May-12	£ 500,000.00	£ 521,610.67	£ 723,151.52
6	Jun-12	£ 650,000.00	£ 692,113.70	£ 936,848.48
7	Jul-12	£ 900,000.00	£ 918,350.41	£ 1,150,545.45
8	Aug-12	£ 1,300,000.00	£ 1,218,538.91	£ 1,364,242.42
9	Sep-12	£ 1,700,000.00	£ 1,616,852.41	£ 1,577,939.39
10	Oct-12	£ 2,200,000.00	£ 2,145,365.81	£ 1,791,636.36
11	Nov-12		£ 2,846,638.58	£ 2,005,333.33
12	Dec-12		£ 3,777,141.96	£ 2,219,030.30

EXPONENTIAL SMOOTHING

Discussion

Exponential smoothing is another method to forecast actual time series data. The method uses a percentage of the previous period's actual data and a percentage of the previous period's forecast. The model can easily be adjusted to increase the accuracy when more actual data is known. When the model is created you need to decide a percentage (also called Alpha). May be you decide to set Alpha to 30%. Then the exponential smoothing forecast will use 30% of the previous period actual data and 70% of the previous forecast to forecast the next period. Because exponential smoothing is calculated from the previous period's forecast you must guess the first forecast or use the actual known value as forecast value. Exponential smoothing can only predict one future period because you need the previous period's actual known data.

Later you will see how you can use error measurement calculations and the **Solver** tool to increase the accuracy when you are working with exponential smoothing.

The math:

F=forecast

t=period

D=demand or actual known data

$$F_{t+1}=\text{alpha} * D_t + (1-\text{alpha}) * F_t$$

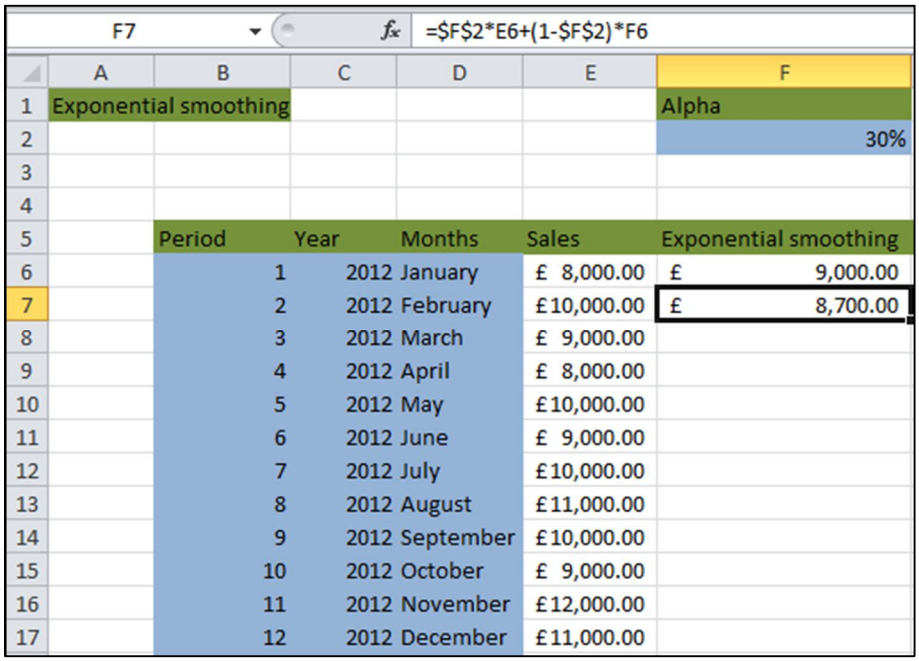
Example:

Alpha = 30%

Previous forecast = 9000

Previous actual known data = 8000

$$F_{t+1}=30\% * 8000 + (1-30\%) * 9000 = 8700$$



F7 =$F$2*E6+(1-$F$2)*F6

	A	B	C	D	E	F
1	Exponential smoothing					Alpha
2						30%
3						
4						
5		Period	Year	Months	Sales	Exponential smoothing
6		1	2012	January	£ 8,000.00	£ 9,000.00
7		2	2012	February	£10,000.00	£ 8,700.00
8		3	2012	March	£ 9,000.00	
9		4	2012	April	£ 8,000.00	
10		5	2012	May	£10,000.00	
11		6	2012	June	£ 9,000.00	
12		7	2012	July	£10,000.00	
13		8	2012	August	£11,000.00	
14		9	2012	September	£10,000.00	
15		10	2012	October	£ 9,000.00	
16		11	2012	November	£12,000.00	
17		12	2012	December	£11,000.00	

Procedures

1. Guess forecast for the first period or if you have data from the previous period use them as forecast for the previous period.

2. Select the cell where the result is to be displayed.

3. Type = select the cell with alpha. Type * select the cell with previous known actual data.

4. Type + (1- select the cell with alpha value)* select the cell with last forecast.

5. Press **enter**.

F7 =$F$2*E6+(1-$F$2)*F6

	A	B	C	D	E	F
1	Exponential smoothing					Alpha
2						30%
3						
4						
5		Period	Year	Months	Sales	Exponential smoothing
6		1	2012	January	£ 8,000.00	£ 9,000.00
7		2	2012	February	£10,000.00	£ 8,700.00
8		3	2012	March	£ 9,000.00	
9		4	2012	April	£ 8,000.00	
10		5	2012	May	£10,000.00	
11		6	2012	June	£ 9,000.00	
12		7	2012	July	£10,000.00	
13		8	2012	August	£11,000.00	
14		9	2012	September	£10,000.00	
15		10	2012	October	£ 9,000.00	
16		11	2012	November	£12,000.00	
17		12	2012	December	£11,000.00	

THE DATA ANALYSIS TOOL EXPONENTIAL SMOOTHING

Instead of doing the calculations yourself you can use the Data analysis tool **Exponential Smoothing**. You can find it by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Exponential Smoothing** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Exponential Smoothing** dialog box, you can get the result in the output range.

You can specify the following items in the **Exponential Smoothing** parameters dialog box:

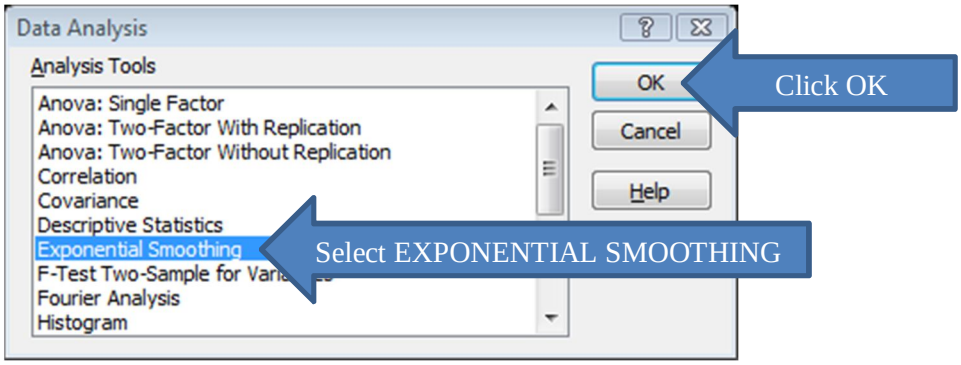
Parameter	Description
Input Range	The array where you have the data you want to examine
Damping factor	The damping factor is (1 – alpha) so to get an alpha value of 30% (30% of the last period actual data and 70% of the last forecast) 0.7 or 70% must be entered

Labels	If labels are selected in the input range the box must be ticked
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook
Chart Output	Tick the box to get a chart output
Standard Errors	Tick the box to get calculated standard errors

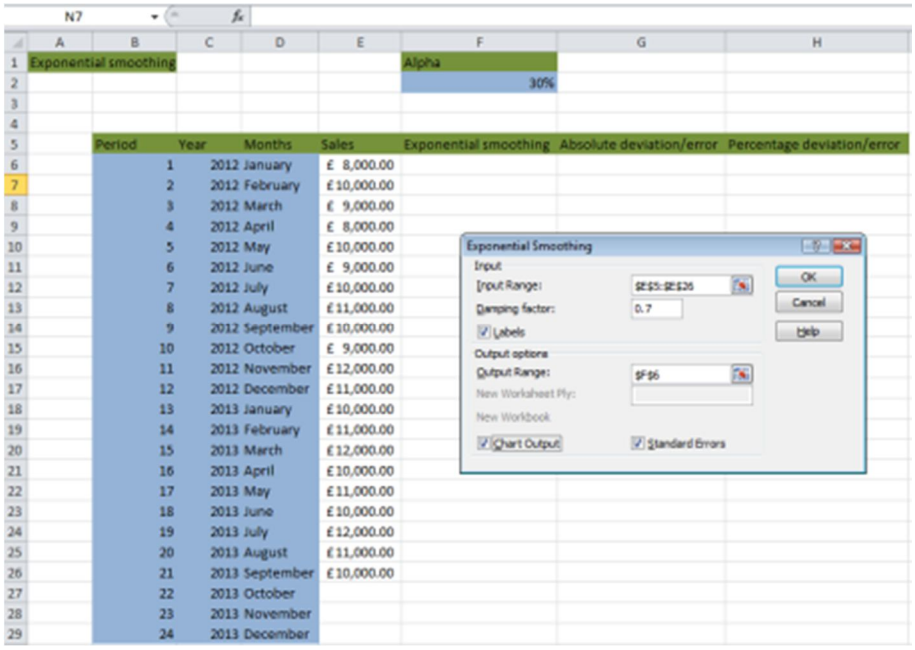
When you have finished entering the information click OK and you will get the output.

Procedures

1. To use the **Exponential Smoothing** tool, first click the Data tab.
2. In the Analysis group, click the Data Analysis button.
3. Select **Exponential Smoothing** from the list and click **OK**.



4. Select the Input Range and select the desired options.



5. Click OK

The output

Let us have a look at the output. You got the exponential smoothing forecast not for the first period as mentioned before you must have data from the previous period to work with exponential smoothing. Then you got standard error calculated based on the average from the previous 3 periods. That is why the standard error results starts from period 5. The error calculations are very important to make sure that the forecast is as accurate as possible and to make sure that you are using the right forecasting model. Later in this workbook you will see examples calculating forecast errors. The chart option was selected so you also got a chart in the output showing the factual known data and the exponential smoothing trend.

NAÏVE FORECASTING

Discussion

The Naïve Forecast definition:

Estimating technique in which the last period's actuals are used as this period's forecast, without adjusting them or attempting to establish causal factors. You are just using the last period's figures. The naïve forecast is normally not very accurate but can be useful to understand and develop other forecasting models. Because the last period is used it is only possible to forecast one future period.

Procedures

1. Select the cell where the result is to be displayed.
2. Type = and select the actual value from the previous period
3. Press enter.

Period	Month		Sales		Naïve Forecast
1	Jan-12	£	1,200,000.00		
2	Feb-12	£	1,350,000.00	£	1,200,000.00
3	Mar-12	£	1,280,000.00	£	1,350,000.00
4	Apr-12	£	1,300,000.00	£	1,280,000.00
5	May-12	£	1,278,000.00	£	1,300,000.00
6	Jun-12	£	1,300,000.00	£	1,278,000.00
7	Jul-12	£	1,300,000.00	£	1,300,000.00
8	Aug-12	£	1,400,000.00	£	1,300,000.00
9	Sep-12	£	1,400,000.00	£	1,400,000.00
10	Oct-12	£	1,500,000.00	£	1,400,000.00
11	Nov-12			£	1,500,000.00
12	Dec-12				

MOVING AVERAGE

Discussion

The Moving Average forecast is based on the average of known values from a number of periods. Let us first have a look at the math. If I want to forecast April using three months moving average the formula will look like this:

$$\text{"April"} = \frac{\text{"January"}+\text{"February"}+\text{"March"}}{3}$$

More scientific way at looking at the equation

$$F_{T+1} = \frac{A_t + A_{t-1} + A_{t-2}}{3}$$

Where **F** is the forecast **T** the period **A** the actual value divided with three because you are using the previous three periods.

The equation will look like this if you want a two period moving average:

$$F_{T+1} = \frac{A_t + A_{t-1}}{2}$$

The equation will look like this if you want a twelve period moving average:

$$F_{T+1} = \frac{A_t + A_{t-1} + A_{t-2} + A_{t-3} + A_{t-4} + A_{t-5} + A_{t-6} + A_{t-7} + A_{t-8} + A_{t-9} + A_{t-10} + A_{t-11} + A_{t-12}}{12}$$

The two months Moving Average forecast is based on 2 months known values. It can be the two previous month's values or it can be based on two months the year before whatever gives the most accurate forecast.

If you forecast one period using the average of the previous two periods you will only be able to forecast one future period. If you use the average of two periods one period back, you can forecast two future periods using this method etc…

In the example below (Moving Average (two periods) Previous two periods) the forecast for March is the average of January and February. You can forecast November but not December with this method. You need to know the November sales first. In the example (Moving Average (two periods) one period back) the forecast for April is also the average of January and February. Here you can forecast December because it is the average of September and October.

Month	Sales	Moving Average (2 periods) Last 2 periods	Moving Average (2 periods) 1 period back
Jan-12	£ 1,200,000.00		
Feb-12	£ 1,350,000.00		
Mar-12	£ 1,280,000.00	£ 1,275,000.00	
Apr-12	£ 1,300,000.00	£ 1,315,000.00	£ 1,275,000.00
May-12	£ 1,278,000.00	£ 1,290,000.00	£ 1,315,000.00
Jun-12	£ 1,300,000.00	£ 1,289,000.00	£ 1,290,000.00
Jul-12	£ 1,300,000.00	£ 1,289,000.00	£ 1,289,000.00
Aug-12	£ 1,400,000.00	£ 1,300,000.00	£ 1,289,000.00
Sep-12	£ 1,400,000.00	£ 1,350,000.00	£ 1,300,000.00
Oct-12	£ 1,500,000.00	£ 1,400,000.00	£ 1,350,000.00
Nov-12		£ 1,450,000.00	£ 1,400,000.00
Dec-12			£ 1,450,000.00

The three months Moving Average forecast is based on three months known values. You can use moving average forecast based on as many periods you need to get the most accurate forecast.

Procedures

1. Select the cell where the result is to be displayed.
2. Type =average(
3. Type the argument, or select the cells you want to insert in the function (the values for the periods).
4. Type) closing bracket.
5. Press enter.

WEIGHTED MOVING AVERAGE

Discussion

Another Excel forecasting option is the Weighted Moving Average. The Weighted Moving Average is nearly the same as the Moving Average except that it assumes that the most recent periods are the best predictors. Each period that is used for the forecast is given a weighting. The largest weighting is assigned to the most recent period. Each preceding period is successively less and less.

For example, during the calculation of a three-period Weighted Moving Average, the most recent three periods are used. The results from the previous period might be given the weight of three, the middle period might be assigned a weight of two, and the 3rd most distant period might be assigned a weight of one.

The formula for the Weighted Moving Average is as follows:

Forecast (period four) = ((weight of period three * Value of period three) + (weight of period two * Value of period two) + (weight of period one * Value of period one) / (weight of period three + weight of period two + weight of period one)

The example below:

Forecast April=((3*210)+(2*180)+(1*200)) / (1+2+3)=(630+360+200) / 6=1180 / 6=198.3

Month		Sales	Weighted Movingaverage 3 periods
Jan-12	£	200.00	1
Feb-12	£	180.00	2
Mar-12	£	210.00	3
Apr-12	£	170.00	
May-12	£	215.00	
Jun-12	£	190.00	
Jul-12	£	210.00	
Aug-12	£	180.00	
Sep-12	£	220.00	
Oct-12	£	210.00	
Nov-12			
Dec-12			

Procedures

1. Select the cell where the result is to be displayed.
2. Type =((weight value*period value)+(next weight value*next period value)) / sum of weight values.
3. Press enter.

D8 =((C7*$D$7)+(C6*$D$6)+(C5*$D$5))/6

	A	B	C	D
1	Forecasting			
2				
3				
4	Period	Month	Sales	Weighted Movingaverage 3 periods
5	1	Jan-12	£ 1,200,000.00	1
6	2	Feb-12	£ 1,350,000.00	2
7	3	Mar-12	£ 1,280,000.00	3
8	4	Apr-12	£ 1,300,000.00	£ 1,290,000.00
9	5	May-12	£ 1,278,000.00	£ 1,301,666.67
10	6	Jun-12	£ 1,300,000.00	£ 1,285,666.67
11	7	Jul-12	£ 1,300,000.00	£ 1,292,666.67
12	8	Aug-12	£ 1,400,000.00	£ 1,296,333.33
13	9	Sep-12	£ 1,400,000.00	£ 1,350,000.00
14	10	Oct-12	£ 1,500,000.00	£ 1,383,333.33
15	11	Nov-12		£ 1,450,000.00
16	12	Dec-12		

In the example above the weight for January, February, and March are typed in D5, D6, and D7, in D8 (the forecast for April) the formula.

*The weight is often a percentage. May be 10% for period 1, 30% for period 2, and 60% for period 3. The formula will now look like this. Forecast (period 4) = ((weight of period 3 * Value of period 3) + (weight of period 2 * Value of period 2) + (weight of period 1 * Value of period 1).*

THE DATA ANALYSIS TOOL MOVING AVERAGE

Discussion

In Excel you can also find a data analysis tool **Moving Average**. Select the **Data Analysis** command from the Analysis group at the far right end of the data ribbon. Select **Moving Average** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

The Data Analysis tool **Moving Average** is not a forecasting tool but a way to smooth data to easier spot the trends.

When you calculated moving average earlier in this book you did it for forecasting by using the equation:

$$\text{"April"} = \frac{\text{"January"} + \text{"February"} + \text{"March"}}{3}$$

or

$$F_{T+1} = \frac{A_t + A_{t-1} + A_{t-2}}{3}$$

The equation for the data analysis tool **Moving Average** tool is:

$$\text{"March"} = \frac{\text{"January"} + \text{"February"} + \text{"March"}}{3}$$

or

$$SD_T = \frac{A_t + A_{t-1} + A_{t-2}}{3}$$

After you have entered all necessary information into the **Moving Average** dialog box, you can get the result in the output range.

You can specify the following items in the **Moving Average** dialog box:

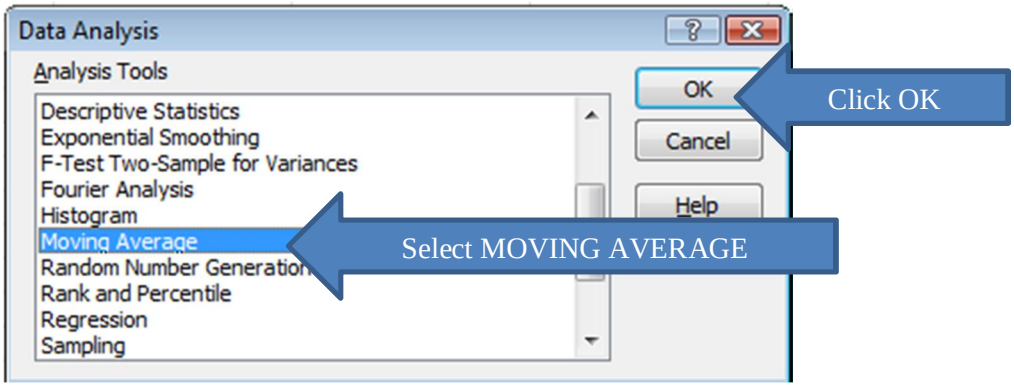
Parameter	Description
Input Range	The array where you have the data you want to examine
Labels	If labels are selected in the input range the box must be ticked
Interval	Enter how many periods you want to use
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook
Chart Output	Tick the box to get a chart output
Standard Errors	Tick the box to get calculated standard errors

When you have finished entering the information click OK and you will get the output.

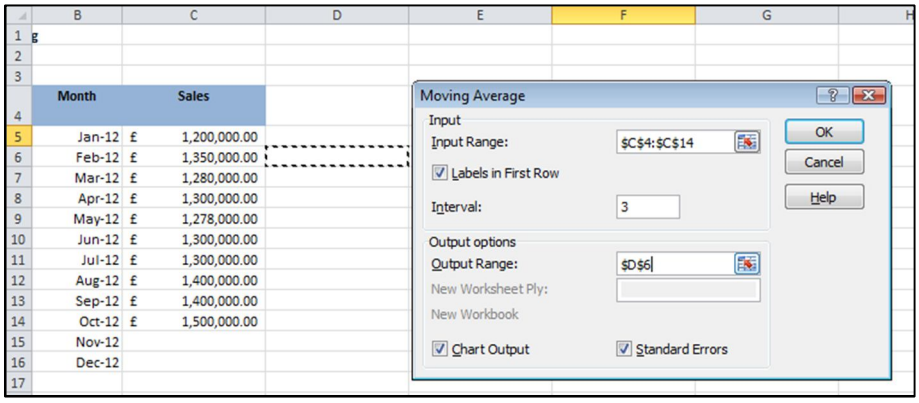
Procedures

1. To use the **Moving Average** tool, first click the Data tab.
2. In the Analysis group, click the Data Analysis button.

3. Select **Moving Average** from the list and click OK.

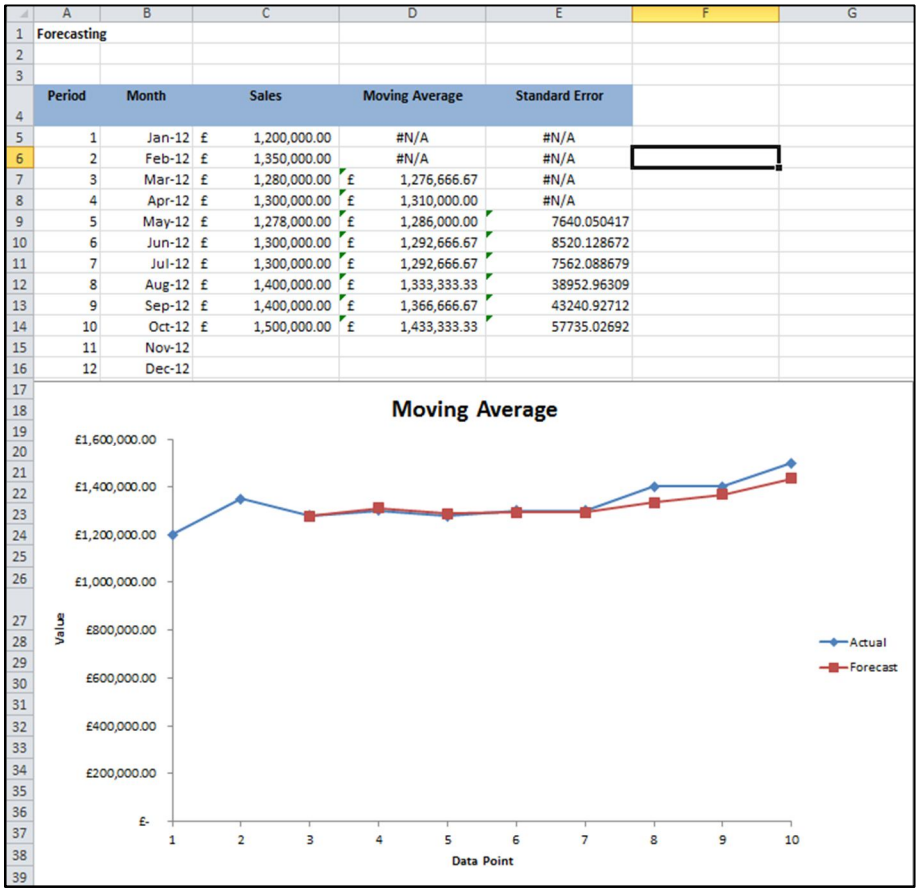


4. Select the Input Range and select the desired options.



5. Click OK

The output



Period	Month	Sales	Moving Average	Standard Error
1	Jan-12	£ 1,200,000.00	#N/A	#N/A
2	Feb-12	£ 1,350,000.00	#N/A	#N/A
3	Mar-12	£ 1,280,000.00	£ 1,276,666.67	#N/A
4	Apr-12	£ 1,300,000.00	£ 1,310,000.00	#N/A
5	May-12	£ 1,278,000.00	£ 1,286,000.00	7640.050417
6	Jun-12	£ 1,300,000.00	£ 1,292,666.67	8520.128672
7	Jul-12	£ 1,300,000.00	£ 1,292,666.67	7562.088679
8	Aug-12	£ 1,400,000.00	£ 1,333,333.33	38952.96309
9	Sep-12	£ 1,400,000.00	£ 1,366,666.67	43240.92712
10	Oct-12	£ 1,500,000.00	£ 1,433,333.33	57735.02692
11	Nov-12			
12	Dec-12			

Let us have a look at the output. You got the **Moving Average** output not for the first 2 periods as mentioned before you must have data from the previous 2 periods to work with 3 months moving average. Then you got standard error calculated based on the average from 3 moving average periods. That is why the standard error results starts from period 5. The error calculations are very important to see how accurate the moving average tool is. Later in this work book you will see examples calculating errors. The chart option was selected so you also got a chart in the output showing the factual known data and the moving average result.

SEASONAL FORECASTING

Discussion

If you have seasonal actual data (each year you are selling more in December and June or another seasonal pattern) you can use seasonal forecasting to forecast future data. A

weight can be added to the periods to increase the accuracy of the forecast. If you use the **Forecast** function to calculate a linear forecast you can see periods over forecast and under forecast. In the example below you have seasonal known actual data. The forecast for quarter 1 is above actual data, the forecast for quarter 2 is above actual data, the forecast for quarter 3 is below actual data, and the forecast for quarter 4 is above actual data.

Period	Quarter	Year	Data	Forecast	Seasonal Index	Seasonal Forecast
1	quarter 1	2006	100	108.84	0.9	97.96
2	quarter 2	2006	82	111.21	0.7	77.85
3	quarter 3	2006	180	113.58	1.5	170.37
4	quarter 4	2006	110	115.95	0.9	104.36
5	quarter 1	2007	110	118.32	0.9	106.49
6	quarter 2	2007	95	120.69	0.7	84.48
7	quarter 3	2007	173	123.06	1.5	184.59
8	quarter 4	2007	110	125.43	0.9	112.88
9	quarter 1	2008	110	127.80	0.9	115.02
10	quarter 2	2008	79	130.17	0.7	91.12
11	quarter 3	2008	200	132.53	1.5	198.80
12	quarter 4	2008	120	134.90	0.9	121.41
13	quarter 1	2009	120	137.27	0.9	123.55
14	quarter 2	2009	98	139.64	0.7	97.75
15	quarter 3	2009	215	142.01	1.5	213.02
16	quarter 4	2009	135	144.38	0.9	129.94
17	quarter 1	2010	140	146.75	0.9	132.07
18	quarter 2	2010	98	149.12	0.7	104.38
19	quarter 3	2010	217	151.49	1.5	227.23
20	quarter 4	2010	135	153.86	0.9	138.47
21	quarter 1	2011		156.23	0.9	140.60
22	quarter 2	2011		158.60	0.7	111.02
23	quarter 3	2011		160.96	1.5	241.45
24	quarter 4	2011		163.33	0.9	147.00

Add an index value (a weight) to each quarter to get a more accurate forecast. In this example the periods are quarters (4 periods groups). The sum of the index values must be 4 (12 if you are working with months etc.).

The index value multiplied by the forecast value should be very close to actual data (the index for period 1 (quarter 1 2006) must be a value multiplied by 108.84 gives a result close to 100 (the actual known value)). Later you will see how you can use error measurement calculations and the **SOLVER** tool to increase the accuracy when you are working with seasonal historical data. In the example above the seasonal index for quarter 1 is set to 0.9, set to 0.7 for quarter 2, set to 1.5 for quarter 3, and set to 0.9 for quarter 4.

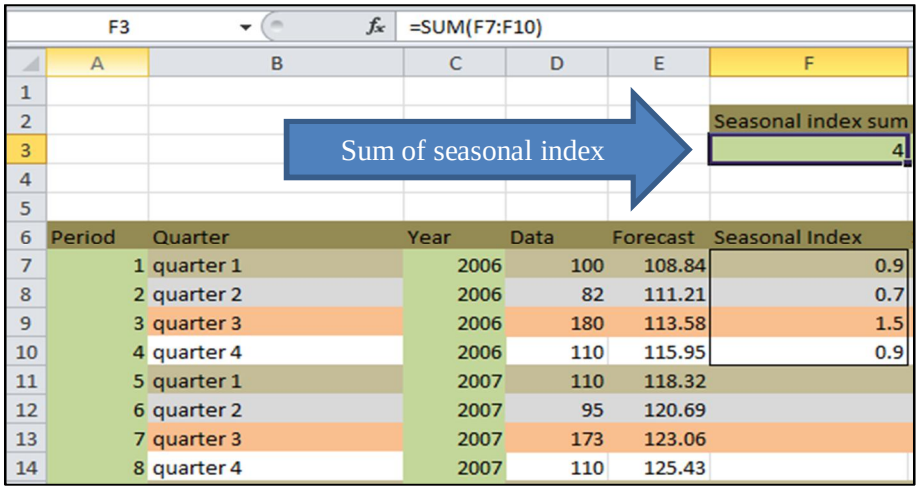
To get the seasonal forecast you just have to multiply the seasonal index value by the **Forecast** function's forecast and if you compare the actual data with the seasonal forecast you will see that the seasonal forecast is much more accurate than the **Forecast** function's forecast.

Procedures

1. Forecast the actual data using the **Forecast** function .

2. Compare actual data with the **Forecast** function's calculated forecast data period by period to spot any seasonal trends.

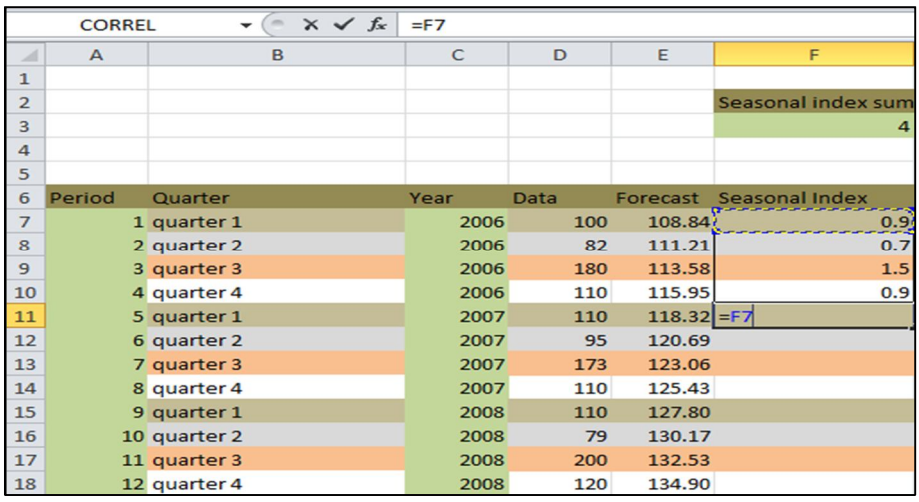
3. Add a seasonal index which sum should be the same as number of periods (4 if you are working with quarters, 12 if you are working with months etc.).



	A	B	C	D	E	F
1						
2						Seasonal index sum
3						4
4						
5						
6	Period	Quarter	Year	Data	Forecast	Seasonal Index
7	1	quarter 1	2006	100	108.84	0.9
8	2	quarter 2	2006	82	111.21	0.7
9	3	quarter 3	2006	180	113.58	1.5
10	4	quarter 4	2006	110	115.95	0.9
11	5	quarter 1	2007	110	118.32	
12	6	quarter 2	2007	95	120.69	
13	7	quarter 3	2007	173	123.06	
14	8	quarter 4	2007	110	125.43	

4. Add the seasonal index values to the rest of the quarters.

5. Select the cell where you want the next index value for quarter 1 and type = click on the cell with the quarter 1 seasonal index.



	A	B	C	D	E	F
1						
2						Seasonal index sum
3						4
4						
5						
6	Period	Quarter	Year	Data	Forecast	Seasonal Index
7	1	quarter 1	2006	100	108.84	0.9
8	2	quarter 2	2006	82	111.21	0.7
9	3	quarter 3	2006	180	113.58	1.5
10	4	quarter 4	2006	110	115.95	0.9
11	5	quarter 1	2007	110	118.32	=F7
12	6	quarter 2	2007	95	120.69	
13	7	quarter 3	2007	173	123.06	
14	8	quarter 4	2007	110	125.43	
15	9	quarter 1	2008	110	127.80	
16	10	quarter 2	2008	79	130.17	
17	11	quarter 3	2008	200	132.53	
18	12	quarter 4	2008	120	134.90	

6. Drag the fill-handle down to copy and paste the seasonal index value to the rest of the quarters.

7. Select the cell where the result is to be displayed.

8. Multiply the **Forecast** function's forecast with the seasonal index.

9. Copy down the calculation.

G7 =E7*F7

	A	B	C	D	E	F	G
1							
2						Seasonal index sum	
3						4	
4							
5							
6	Period	Quarter	Year	Data	Forecast	Seasonal Index	Seasonal Forecast
7	1	quarter 1	2006	100	108.84	0.9	97.96
8	2	quarter 2	2006	82	111.21	0.7	77.85
9	3	quarter 3	2006	180	113.58	1.5	170.37
10	4	quarter 4	2006	110	115.95	0.9	104.36
11	5	quarter 1	2007	110	118.32	0.9	106.49
12	6	quarter 2	2007	95	120.69	0.7	84.48
13	7	quarter 3	2007	173	123.06	1.5	184.59
14	8	quarter 4	2007	110	125.43	0.9	112.88
15	9	quarter 1	2008	110	127.80	0.9	115.02
16	10	quarter 2	2008	79	130.17	0.7	91.12
17	11	quarter 3	2008	200	132.53	1.5	198.80
18	12	quarter 4	2008	120	134.90	0.9	121.41
19	13	quarter 1	2009	120	137.27	0.9	123.55
20	14	quarter 2	2009	98	139.64	0.7	97.75
21	15	quarter 3	2009	215	142.01	1.5	213.02
22	16	quarter 4	2009	135	144.38	0.9	129.94
23	17	quarter 1	2010	140	146.75	0.9	132.07
24	18	quarter 2	2010	98	149.12	0.7	104.38
25	19	quarter 3	2010	217	151.49	1.5	227.23
26	20	quarter 4	2010	135	153.86	0.9	138.47
27	21	quarter 1	2011		156.23	0.9	140.60
28	22	quarter 2	2011		158.60	0.7	111.02
29	23	quarter 3	2011		160.96	1.5	241.45
30	24	quarter 4	2011		163.33	0.9	147.00

It can be difficult to spot a seasonal trend in actual data but even that you cannot it is a forecasting model which in many situations is more accurate than other models.

Exercises

Exercise 1 Linear Regression

1. Open the file **Forecasting and data analysis**.
2. Open the **Forecast Exercises** sheet.
3. Use the **Forecast** function in F12 to calculate the forecast for the first quarter 2007.
4. Copy down the **Forecast** function to get the forecast for the rest of the quarters.
5. Select the range G12:G35 and use the **Trend** function to calculate the trend for known actual values (when entered the function arguments remember it is an array function and you must press Ctrl Shift Enter).
6. Select the range G36:G39 and use the **Trend** function to forecast the 4 quarters in 2013. Remember to enter the range B36:B39 as the **Trend** function argument **New**_x's (when entered the function arguments remember it is an array function and you must press Ctrl Shift Enter).
7. Select G10 and use the **Slope** function to find the slope.
8. Select H10 and use the **Intercept** function to find Y-interceptor.
9. Select H12 and calculate the equation for trend and forecast using the slope and Y-interceptor (the equation: Y = mX + c, m is the slope, c is the Y-interceptor, and X is the period).
10. Save the file.

Exercise 2 Exponential Regression

1. Open the file **Forecasting and data analysis**.

2. Open the **Forecast Exercises** sheet.

3. Select the range Q12:Q27 and use the **Growth** function to calculate the growth for the known actual data for year 2007 to 2010 (when entered the function arguments remember it is an array function and you must press Ctrl Shift Enter).

4. Select the range Q28:Q31 and use the **Growth** function to forecast the 4 quarters in 2011. Remember to enter the range L28:L31 as the **Growth** function argument **New**_x's (when entered the function arguments remember it is an array function and you must press Ctrl Shift Enter).

5. Compare the difference between column P (the **Forecast** function) and column Q (the **Growth** function).

6. Save the file.

Exercise 3 Exponential Smoothing forecasting

1. Open the file **Forecasting and data analysis**.
2. Open the **Forecast Exercises** sheet.
3. Enter an alpha value of 50% in G41.
4. Enter 120,000 in G45 (you must guess the first period's forecast).
5. Select G46 and enter the **Exponential Smoothing** forecast equation =$G$41*E45+(1-$G$41)*G45 (Ft+1=alpha * Dt + (1-alpha) * Ft).
6. Copy down the equation to forecast period 25 (quarter 1 2013).
7. Change the alpha value in G41 and observe the changes in the range G45:G69.
8. Click the **Data** tab and click **Data Analysis** in the **Analysis** Group.
9. Select **Exponential Smoothing.**
10. Enter the input range $E$44:$E$68. Damping factor of 0.7 to get an Alpha value of 30%. Tick labels and output range should be $H$45. Tick Chart output and standard errors.
11. Change Alpha in G41 to 30% and compare the two outputs.
12. Save the file.

Exercise 4 Naïve forecasting and Moving Average

1. Open the file **Forecasting and data analysis**.
2. Open the **Forecast Exercises** sheet.
3. Select Q46 and type =O45.
4. Drag the fill-handle down to row 69 to copy down the cell content. This is the Naïve forecast which is based on the previous period.
5. Select R48 and use the **Average** function to calculate the average of the previous 3 periods known actual figures (=average(O45:O47) and copy down the **Average** function to R69. You now have the 3 periods moving average forecast.
6. Enter 10% in S41, enter 20% in S42, and 70% in S43. The sum of the weights must be 100%. You assume that the previous period's actual known value is more important to the forecast than two and three periods back.
7. Select S48 and type =$S$41*O45+$S$42*O46++$S$43*O47.
8. Copy down the calculation to S69.
9. Examine the result. Change the weights in the range S41:S43, but remember that the sum must be 100%. You now have a 3 periods weighted moving average forecast.
10. Save the file.

Exercise 5 Seasonal forecasting

1. Open the file **Forecasting and data analysis**.
2. Open the **Forecast Exercises** sheet.
3. Examine the values in the range E84:E107. Is there any seasonal patterns? Yes there is. Quarter 4 has the highest sales figures and quarter 3 the lowest.
4. Select G84 and enter the weight for the first period. You have to compare the sale in E84 and the forecast in F84. If the forecast is higher than the actual value the weight should be less than 1. The weight = Sales / Forecast function or G84=E84 / F84.
5. Enter the weights in G85, G86 and G87 (the sum of the weights must equals 4 the number of periods (12 if you are working with months)). Work with 2 decimals and round up or down to make sure that the sum equals number of periods here 4.
6. Select G88 and type =G84 and copy down the cell to G111.
7. Select H84 and type =G84*F84.
8. Copy down the calculation to H111 and you now have a seasonal forecast.
9. Save the file.

LESSON 2 - MEASURING FORECAST ACCURACY

In this lesson, you will learn how to:

- Measure forecast accuracy
- Calculate error / deviation
- Calculate absolute error / deviation
- Calculate percentage error / deviation
- Calculate absolute percentage error / deviation
- Calculate square error
- Calculate standard error
- Calculate MAD or MAE (Mean Absolute Deviation or Error)
- Calculate MSQ (Mean Square Error)
- Calculate MPE (Mean Percentage Error)
- Calculate MAPE (Mean Absolute Percentage Error)
- Calculate TSE (Tracking Signal Error)

CONCEPTS AND TERMS

Discussion

A key question in any forecasting is how to measure performance. Such measures are very important when you select a forecasting method and create a forecasting model, since you may compare alternatives and choose the method with the highest level of accuracy.

Then, once the method is being used on a regular basis, you need similar measures to tell you whether the forecasts are maintaining their accuracy. If not you will need to get things back on track by putting improvements in place such as more timely data, better statistical methods and may be change forecasting method. To keep up getting more accuracy forecasts you should work with more than one forecasting method so you always can compare the accuracy.

To measure how accurate the forecast is you can use different methods. To be able to compare different forecasting methods you will look at the mean errors but first you will calculate the errors. Later you compare the different forecast methods to find out which one is the most accurate for your data right now, because this can change over time so it is important to use more than one method and keeping up calculate the errors.

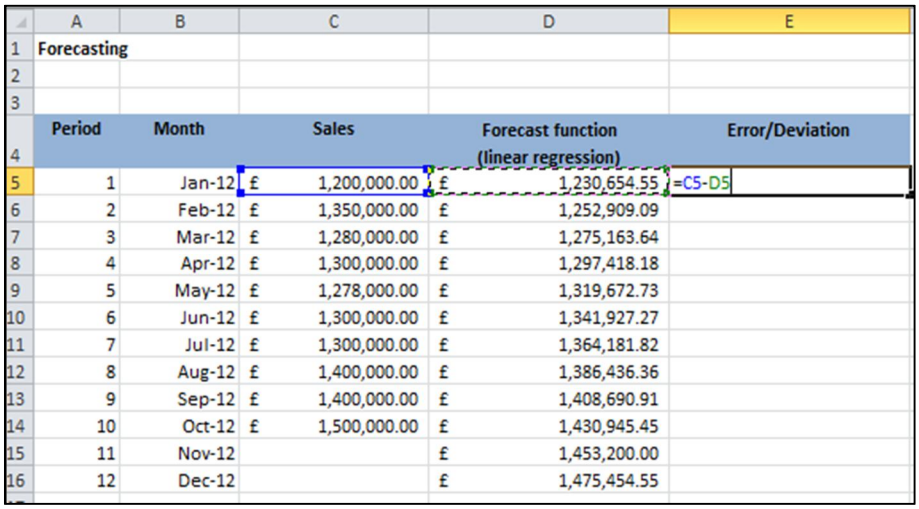
CALCULATE ERROR/DEVIATION

Discussion

The error or the deviation is the difference between actual value and forecasted value. You can calculate the error or deviation by simple subtract the forecast from the actual. This will give a positive or negative result for each period or tell us if the forecast for each period is over or under actual value.

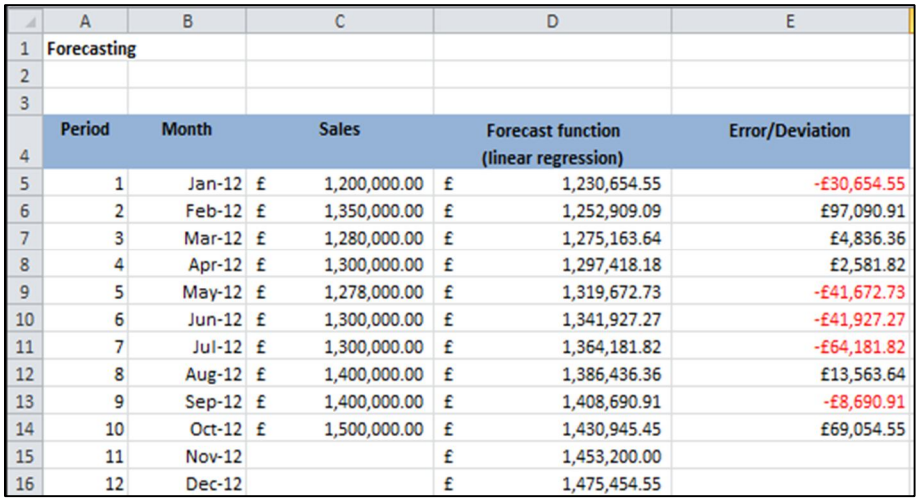
Procedures

1. Select the cell where the result is to be displayed.
2. Type = select the cell with the actual value type – (subtract) select the cell with the forecast value.
3. Press enter.



	A	B	C	D	E
1	Forecasting				
2					
3					
4	Period	Month	Sales	Forecast function (linear regression)	Error/Deviation
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	=C5-D5
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	

The calculation will tell us if our forecast is over or under actual value for each period.



	A	B	C	D	E
1	Forecasting				
2					
3					
4	Period	Month	Sales	Forecast function (linear regression)	Error/Deviation
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	-£30,654.55
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	£97,090.91
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	£4,836.36
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	£2,581.82
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	-£41,672.73
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	-£41,927.27
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	-£64,181.82
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	£13,563.64
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	-£8,690.91
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	£69,054.55
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	

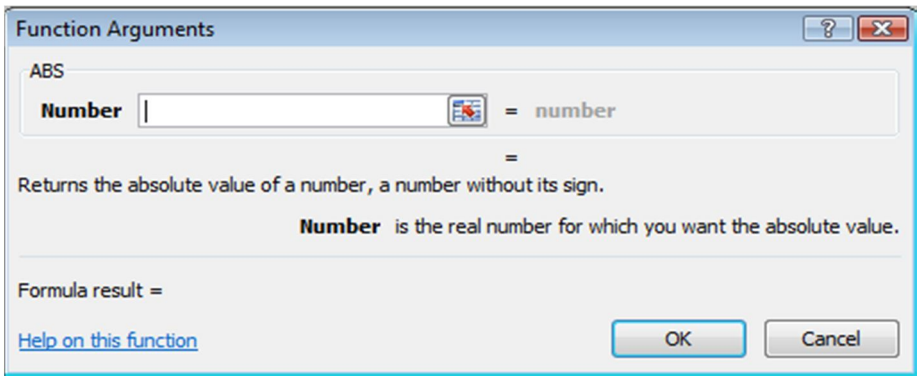
CALCULATE ABSOLUTE ERROR/DEVIATION

Discussion

It is a good idea to calculate absolute error / deviation to get the right picture of the errors. To compare forecast methods the average error / deviation is very important. If you calculate the average from positive and negative errors the result will not be very useful. To calculate the average you need to summarise all the values and then divide the sum with the number of data. You can have negative and positive errors which will neutralise or even out each other and the average result will not tell us anything about the forecast accuracy.

You need to calculate the absolute error / deviation to get a useful result in an average calculation. To calculate absolute values in Excel you can use the **Abs** function. The **Abs** function has only one argument and it is **number**. If you type a number negative or positive the function will always return is as an absolute (positive) value.

To calculate the Absolute Error you need to type =abs(the actual value – the forecast value).



Procedures

1. Select the cell where the result is to be displayed.
2. Type =abs(select the cell with the actual value type – (subtract) select the cell with the forecast value).
3. Press enter.

	A	B	C	D	E
1	Forecasting				
2					
3					
4	Period	Month	Sales	Forecast function (linear regression)	Absolute Error/Deviation
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	=abs(C5-D5)
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	

The result of the absolute error / deviation calculation will show the error as a positive value not if the forecast is above or below factual data.

	A	B	C	D	E
1	Forecasting				
2					
3					
4	Period	Month	Sales	Forecast function (linear regression)	Absolute Error/Deviation
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	£ 30,654.55
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	£ 97,090.91
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	£ 4,836.36
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	£ 2,581.82
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	£ 41,672.73
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	£ 41,927.27
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	£ 64,181.82
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	£ 13,563.64
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	£ 8,690.91
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	£ 69,054.55
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	

CALCULATE PERCENTAGE ERROR/DEVIATION

Discussion

It is always useful to get something like forecast accuracy shown as percentage and the percentage error can also help you to decide which forecasting method to use for your data.

Procedures

1. Select the cell where the result is to be displayed.
2. Type = select the cell with the error / deviation and divide it with actual value.
3. Press enter.

	A	B	C	D	E	F
1	Forecasting					
2						
3						
4	Period	Month	Sales	Forecast function (linear regression)	Error/Deviation	Percentage Error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	-30,654.55	=E5/C5
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	97,090.91	
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	4,836.36	
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	2,581.82	
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	-41,672.73	
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	-41,927.27	
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	-64,181.82	
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	13,563.64	
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	-8,690.91	
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	69,054.55	
15	11	Nov-12		£ 1,453,200.00		
16	12	Dec-12		£ 1,475,454.55		

	A	B	C	D	E	F
1	Forecasting					
2						
3						
4	Period	Month	Sales	Forecast function (linear regression)	Error/Deviation	Percentage Error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	-30,654.55	-2.55%
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	97,090.91	7.19%
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	4,836.36	0.38%
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	2,581.82	0.20%
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	-41,672.73	-3.26%
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	-41,927.27	-3.23%
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	-64,181.82	-4.94%
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	13,563.64	0.97%
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	-8,690.91	-0.62%
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	69,054.55	4.60%
15	11	Nov-12		£ 1,453,200.00		
16	12	Dec-12		£ 1,475,454.55		

CALCULATE ABSOLUTE PERCENTAGE ERROR/DEVIATION

Discussion

You need the Absolute Percentage Error to be able to calculate Mean Absolute Percentage Error, which are very useful to compare forecasting methods.

Procedures

1. Select the cell where the result is to be displayed.
2. Type = select the cell with the absolute error / deviation and divide it with actual value
3. Press enter.

	A	B	C	D	E	F
1	Forecasting					
2						
3						
4	Period	Month	Sales	Forecast function (linear regression)	Absolute Error/Deviation	Absolute Percentage Error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	30,654.55	=E5/C5
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	97,090.91	
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	4,836.36	
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	2,581.82	
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	41,672.73	
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	41,927.27	
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	64,181.82	
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	13,563.64	
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	8,690.91	
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	69,054.55	
15	11	Nov-12		£ 1,453,200.00		
16	12	Dec-12		£ 1,475,454.55		

	A	B	C	D	E	F
1	Forecasting					
2						
3						
4	Period	Month	Sales	Forecast function (linear regression)	Absolute Error/Deviation	Absolute Percentage Error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	30,654.55	2.55%
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	97,090.91	7.19%
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	4,836.36	0.38%
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	2,581.82	0.20%
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	41,672.73	3.26%
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	41,927.27	3.23%
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	64,181.82	4.94%
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	13,563.64	0.97%
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	8,690.91	0.62%
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	69,054.55	4.60%
15	11	Nov-12		£ 1,453,200.00		
16	12	Dec-12		£ 1,475,454.55		

CALCULATE SQUARE ERROR

Discussion

A popular way of looking at forecast accuracy is calculating square error. By squaring the deviation the errors will be more obvious and easy to spot. By squaring the deviation the result will always be absolute.

Procedures

1. Select the cell where the result is to be displayed.
2. Type =(select the cell with the actual value – (subtract) the forecast and type) then type ^ (shift 6) and type 2.
3. Press enter.

	A	B	C	D	E
4	Period	Month	Sales	Forecast function (linear regression)	Squared Error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	=(C5-D5)^2
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	

	A	B	C	D	E
4	Period	Month	Sales	Forecast function (linear regression)	Squared Error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	939701157
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	9426644628
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	23390413.22
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	6665785.124
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	1736616198
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	1757896198
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	4119305785
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	183972231.4
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	75531900.83
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	4768530248
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	

The SUMSQ function can also be used to calculate Square Error.
Type =sumsq(select the actual data cell – the forecast cell and type).
In the example above =sumsq(C5-D5).

CALCULATE STANDARD ERROR

Discussion

Standard error is a statistical standard for error calculations. You need to summarise the squares of the difference between the three previous known actual periods and the three previous periods forecast and divide the result by three (number of periods) and calculate the square root of the result. You can use the **SQRT** function together with the **SUMXMY2** function to calculate the standard error.

The SQRT function:

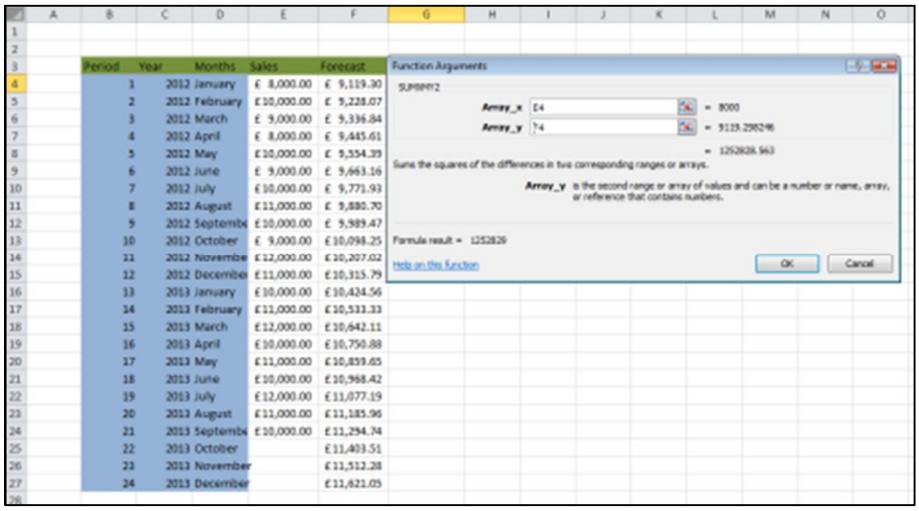
The **SQRT** function returns the square root of a number.

Procedures

1. Select the cell where the result is to be displayed.
2. Type =SQRT(type a number or select a cell).
3. Press **OK** and the function will return the square root.

The SUMXMY2 function:

The **Sumxmy2** function returns the sum of the square of the difference of two arrays. The function has two arguments **Array_x** and **Array_y**. Enter the range for the first array in **Array_x** and enter the range for the second array in **Array_y**.

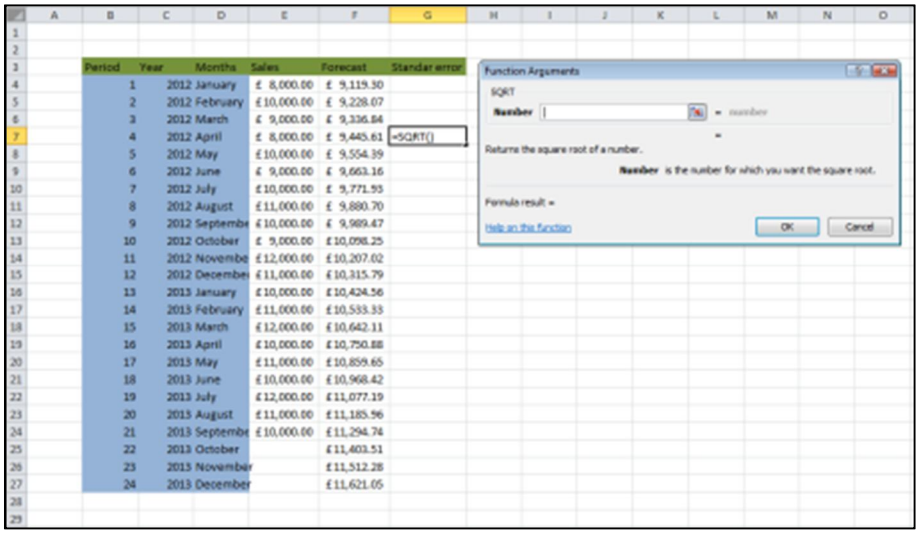


Procedures

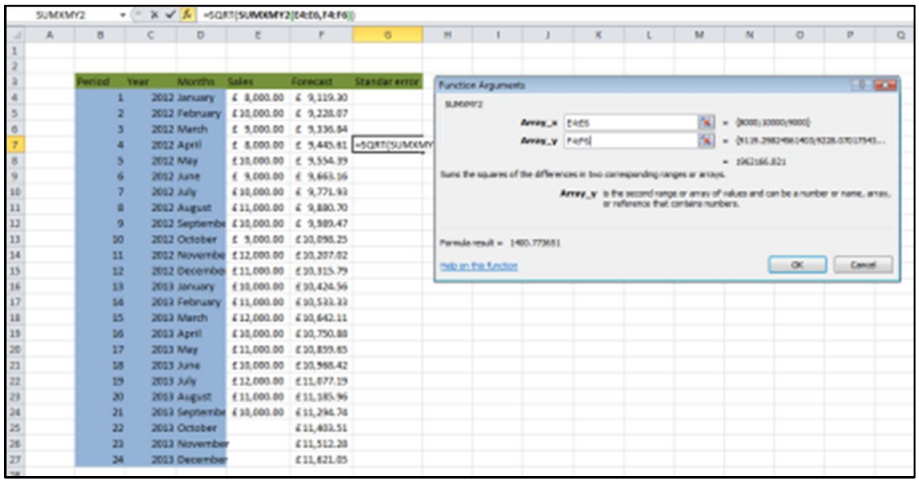
1. Select the cell where the result is to be displayed.
2. Type =SUMXMY2(in **Array_x** enter the range of the first array. In **Array_y** enter the range of the second range).
3. Press OK.

Use the SQRT function and the SUMXMY2 to calculate standard error

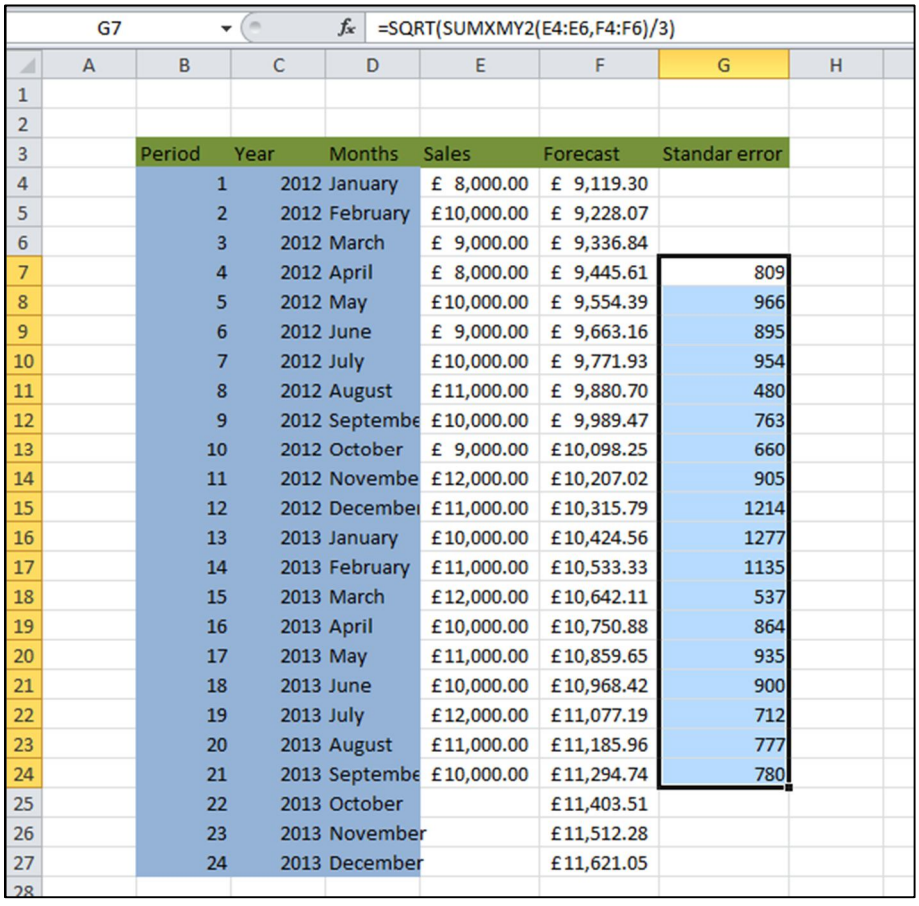
Select the cell where you want the result to be displayed. Start with the **SQRT** function.



Nest the **SUMXMY2** function and select the two arrays (the previous three known values in **Array_x** and the previous three forecast values in **Array_y**).



Click **OK** and type " /3" after the **SUMXMY2** function inside the **SQRT** function. Press enter.



	Period	Year	Months	Sales	Forecast	Standar error
4	1	2012	January	£ 8,000.00	£ 9,119.30	
5	2	2012	February	£10,000.00	£ 9,228.07	
6	3	2012	March	£ 9,000.00	£ 9,336.84	
7	4	2012	April	£ 8,000.00	£ 9,445.61	809
8	5	2012	May	£10,000.00	£ 9,554.39	966
9	6	2012	June	£ 9,000.00	£ 9,663.16	895
10	7	2012	July	£10,000.00	£ 9,771.93	954
11	8	2012	August	£11,000.00	£ 9,880.70	480
12	9	2012	Septembe	£10,000.00	£ 9,989.47	763
13	10	2012	October	£ 9,000.00	£10,098.25	660
14	11	2012	Novembe	£12,000.00	£10,207.02	905
15	12	2012	Decembe	£11,000.00	£10,315.79	1214
16	13	2013	January	£10,000.00	£10,424.56	1277
17	14	2013	February	£11,000.00	£10,533.33	1135
18	15	2013	March	£12,000.00	£10,642.11	537
19	16	2013	April	£10,000.00	£10,750.88	864
20	17	2013	May	£11,000.00	£10,859.65	935
21	18	2013	June	£10,000.00	£10,968.42	900
22	19	2013	July	£12,000.00	£11,077.19	712
23	20	2013	August	£11,000.00	£11,185.96	777
24	21	2013	Septembe	£10,000.00	£11,294.74	780
25	22	2013	October		£11,403.51	
26	23	2013	November		£11,512.28	
27	24	2013	December		£11,621.05	

Procedures

1. Select the cell where the result is to be displayed.

2. Type =SQRT(SUMXMY2(select the range with the actual values. Type”,” select the range with the forecast) / 3)
3. Click **OK.**

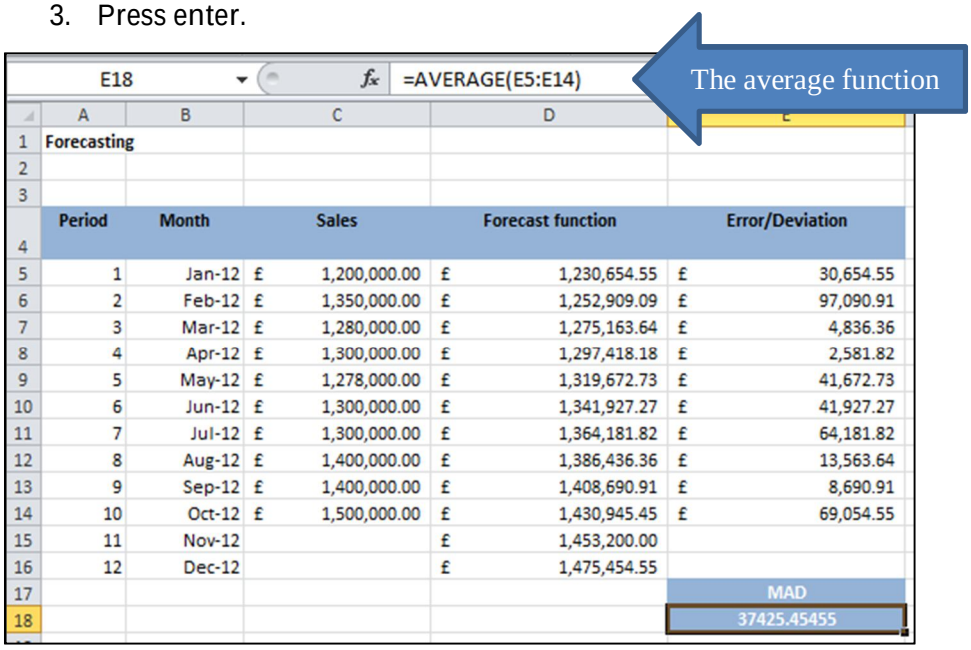
CALCULATE MAD OR MAE (MEAN ABSOLUTE DEVIATION OR MEAN ABSOLUTE ERROR)

Discussion

The MAD / MAE is a very important measurement to compare forecast methods. You need an overall result to compare methods. MAD / MAE is the average of the absolute deviation / error.

You can calculate MAD (Mean Absolute Deviation) or MAE (Mean Absolute Error) by using the average function.

1. Select the cell where the result is to be displayed.
2. Type =average(select the range or enter the cell references for the range containing absolute error / deviation and type)
3. Press enter.



	A	B	C	D	E
1	Forecasting				
2					
3					
4	Period	Month	Sales	Forecast function	Error/Deviation
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	£ 30,654.55
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	£ 97,090.91
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	£ 4,836.36
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	£ 2,581.82
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	£ 41,672.73
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	£ 41,927.27
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	£ 64,181.82
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	£ 13,563.64
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	£ 8,690.91
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	£ 69,054.55
15	11	Nov-12		£ 1,453,200.00	
16	12	Dec-12		£ 1,475,454.55	
17					MAD
18					37425.45455

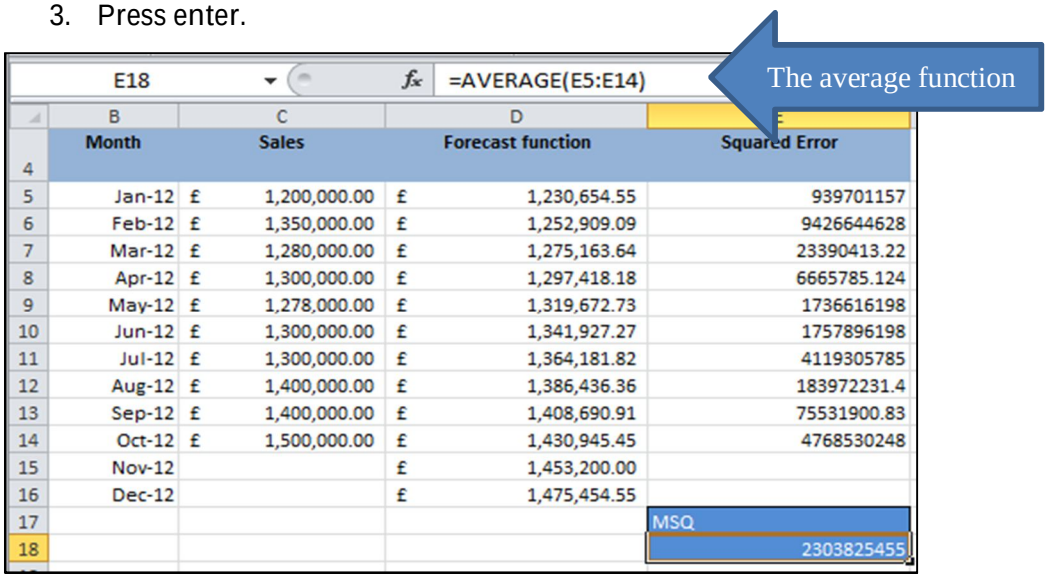
CALCULATE MSQ (MEAN SQUARE ERROR)

Discussion

MSQ (Mean Square Error) is a method to calculate forecast accuracy and to compare forecasting methods. MSQ is the average of the calculated Square Error.

Procedures

1. Select the cell where the result is to be displayed.
2. Type =average(select the range or enter the cell references for the range containing absolute error / deviation and type)
3. Press enter.



	B	C	D	E
4	Month	Sales	Forecast function	Squared Error
5	Jan-12	£ 1,200,000.00	£ 1,230,654.55	939701157
6	Feb-12	£ 1,350,000.00	£ 1,252,909.09	9426644628
7	Mar-12	£ 1,280,000.00	£ 1,275,163.64	23390413.22
8	Apr-12	£ 1,300,000.00	£ 1,297,418.18	6665785.124
9	May-12	£ 1,278,000.00	£ 1,319,672.73	1736616198
10	Jun-12	£ 1,300,000.00	£ 1,341,927.27	1757896198
11	Jul-12	£ 1,300,000.00	£ 1,364,181.82	4119305785
12	Aug-12	£ 1,400,000.00	£ 1,386,436.36	183972231.4
13	Sep-12	£ 1,400,000.00	£ 1,408,690.91	75531900.83
14	Oct-12	£ 1,500,000.00	£ 1,430,945.45	4768530248
15	Nov-12		£ 1,453,200.00	
16	Dec-12		£ 1,475,454.55	
17				MSQ
18				2303825455

CALCULATE MPE (MEAN PERCENTAGE ERROR)

Discussion

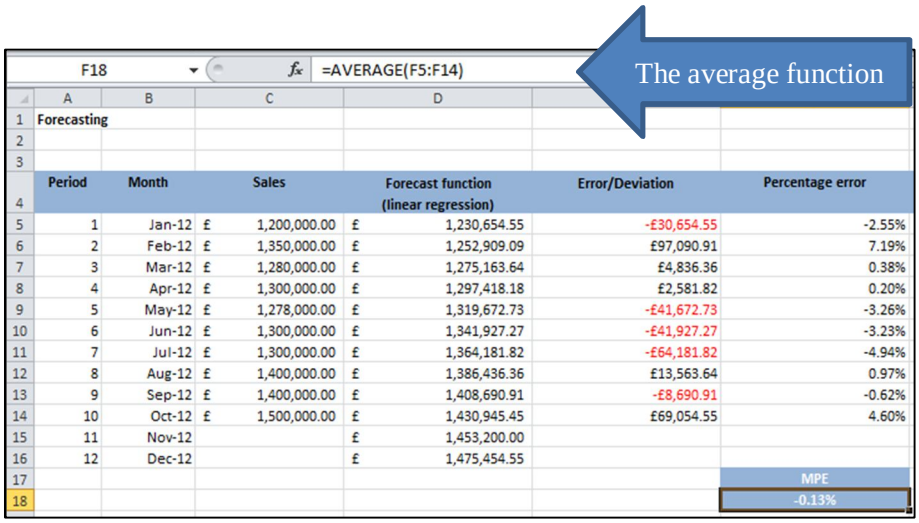
The MPE (Mean Percentage Error) is the average of the Percentage Error. Use the AVERAGE function to calculate MPE.

Procedures

1. Select the cell where the result is to be displayed.

2. Type =average(select the range or enter the cell references for the range containing absolute error / deviation and type)

3. Press enter.



F18 =AVERAGE(F5:F14)

	A	B	C	D		
1	Forecasting					
2						
3						
4	Period	Month	Sales	Forecast function (linear regression)	Error/Deviation	Percentage error
5	1	Jan-12	£ 1,200,000.00	£ 1,230,654.55	-£30,654.55	-2.55%
6	2	Feb-12	£ 1,350,000.00	£ 1,252,909.09	£97,090.91	7.19%
7	3	Mar-12	£ 1,280,000.00	£ 1,275,163.64	£4,836.36	0.38%
8	4	Apr-12	£ 1,300,000.00	£ 1,297,418.18	£2,581.82	0.20%
9	5	May-12	£ 1,278,000.00	£ 1,319,672.73	-£41,672.73	-3.26%
10	6	Jun-12	£ 1,300,000.00	£ 1,341,927.27	-£41,927.27	-3.23%
11	7	Jul-12	£ 1,300,000.00	£ 1,364,181.82	-£64,181.82	-4.94%
12	8	Aug-12	£ 1,400,000.00	£ 1,386,436.36	£13,563.64	0.97%
13	9	Sep-12	£ 1,400,000.00	£ 1,408,690.91	-£8,690.91	-0.62%
14	10	Oct-12	£ 1,500,000.00	£ 1,430,945.45	£69,054.55	4.60%
15	11	Nov-12		£ 1,453,200.00		
16	12	Dec-12		£ 1,475,454.55		
17						MPE
18						-0.13%

CALCULATE MAPE (MEAN ABSOLUTE PERCENTAGE ERROR):

Discussion

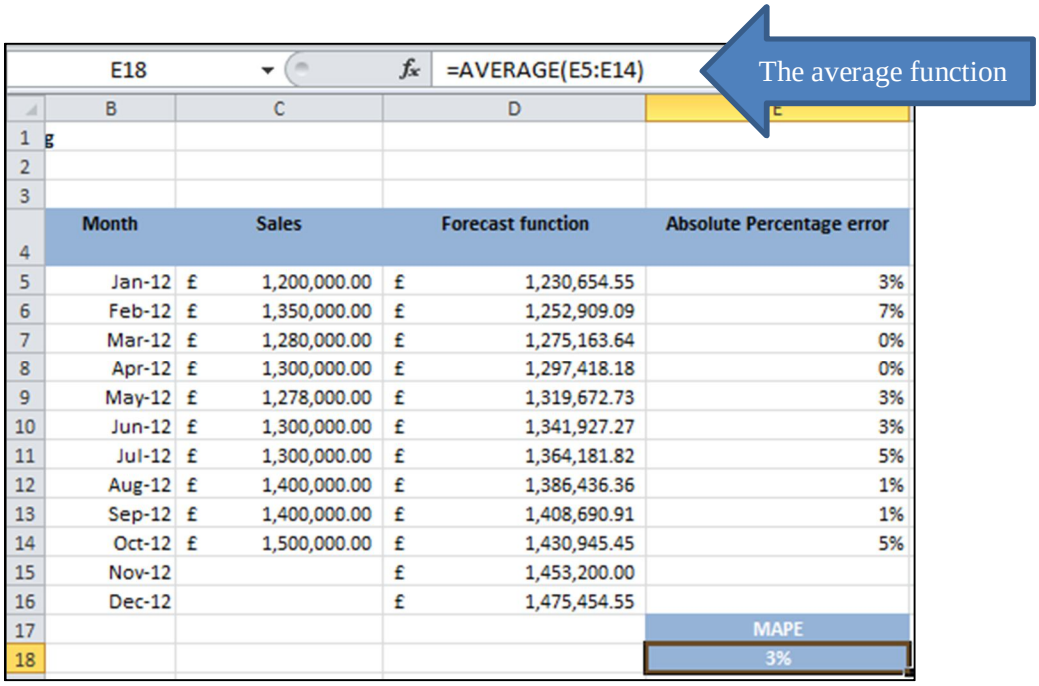
The MAPE (Mean Absolute Percentage Error) is the average of the Absolute Percentage Error. Use the AVERAGE function to calculate MAPE.

Procedures

1. Select the cell where the result is to be displayed.

2. Type =average(select the range or enter the cell references for the range containing absolute error / deviation)

3. Press enter.



	B	C	D	E
1	g			
2				
3				
4	Month	Sales	Forecast function	Absolute Percentage error
5	Jan-12	£ 1,200,000.00	£ 1,230,654.55	3%
6	Feb-12	£ 1,350,000.00	£ 1,252,909.09	7%
7	Mar-12	£ 1,280,000.00	£ 1,275,163.64	0%
8	Apr-12	£ 1,300,000.00	£ 1,297,418.18	0%
9	May-12	£ 1,278,000.00	£ 1,319,672.73	3%
10	Jun-12	£ 1,300,000.00	£ 1,341,927.27	3%
11	Jul-12	£ 1,300,000.00	£ 1,364,181.82	5%
12	Aug-12	£ 1,400,000.00	£ 1,386,436.36	1%
13	Sep-12	£ 1,400,000.00	£ 1,408,690.91	1%
14	Oct-12	£ 1,500,000.00	£ 1,430,945.45	5%
15	Nov-12		£ 1,453,200.00	
16	Dec-12		£ 1,475,454.55	
17				MAPE
18				3%

CALCULATE TSE (TRACKING SIGNAL ERROR)

Discussion

Once a forecast model is developed it should indicate if the actual values following the forecast, and indicate if any deviation in order to correct the model. TSE (Tracking Signal Error) is used to pinpoint forecasting models that need adjustment. As long as the tracking signal is between –4 and 4, assume the model is working correctly.

To calculate TSE you need to calculate MAD (Mean Absolute Deviation). Summarise the error / deviation and divide it by MAD and you will have TSE.

Procedures

1. Select the cell where the result is to be displayed.
2. Type =sum(select the Error range and type) type / select the cell with the MAD.
3. Press enter.

E18 =SUM(E6:E14)/F18

	A	B	C	D	E	F
1	Forecasting					
2						
3						
4	Period	Month	Sales	Naïve Forecast	Error/Deviation	Absolute Error/Deviation
5	1	Jan-12	£ 1,200,000.00			
6	2	Feb-12	£ 1,350,000.00	£ 1,200,000.00	£150,000.00	£150,000.00
7	3	Mar-12	£ 1,280,000.00	£ 1,350,000.00	-£70,000.00	£70,000.00
8	4	Apr-12	£ 1,300,000.00	£ 1,280,000.00	£20,000.00	£20,000.00
9	5	May-12	£ 1,278,000.00	£ 1,300,000.00	-£22,000.00	£22,000.00
10	6	Jun-12	£ 1,300,000.00	£ 1,278,000.00	£22,000.00	£22,000.00
11	7	Jul-12	£ 1,300,000.00	£ 1,300,000.00	£0.00	£0.00
12	8	Aug-12	£ 1,400,000.00	£ 1,300,000.00	£100,000.00	£100,000.00
13	9	Sep-12	£ 1,400,000.00	£ 1,400,000.00	£0.00	£0.00
14	10	Oct-12	£ 1,500,000.00	£ 1,400,000.00	£100,000.00	£100,000.00
15	11	Nov-12		£ 1,500,000.00		
16	12	Dec-12				
17					Tracing Signal Error	MAD
18					5.5785	£53,777.78

Exercises

Exercise 1 Error/deviation

1. Open the file **Forecasting and data analysis**.
2. Open the **Error-deviation exercises** sheet.
3. Select G2 and calculate the error (=E2-F2).
4. Copy down the calculation to G25.
5. Select H2 and calculate the absolute error (=abs(E2-F2)) and copy down the calculation to H25.
6. Select I2 and calculate the percentage error (=G2/E2) and copy down the calculation to I25. Add percentage style and 2 decimals.
7. Select J2 and calculate the absolute percentage error (=H2/E2) and copy down the calculation to J25. Add percentage style and 2 decimals.
8. Select K2 and calculate the square error (=(E2-F2)^2) and copy down the calculation to K25.
9. Save the file.

Exercise 2 Mean errors/deviation

1. Open the file **Forecasting and data analysis**.
2. Open the **Error-deviation exercises** sheet.
3. Select H32 and calculate the mean absolute deviation (MAD) (=average(H2:H25)).
4. Select I32 and calculate mean percentage error (MPE) (=average(I2:I25)).
5. Select J32 and calculate mean absolute percentage error (MAPE) (=average(J2:J25)).
6. Select K32 and calculate mean square error (MSQ) (=average(K2:K25)).
7. Select J32 and calculate mean absolute percentage error (MAPE) (=average(J2:J25)).
8. Select G32 and calculate tracking signal error (TSE) (=sum(G2:G25) / H32)
9. Save the file.

LESSON 3 – USING THE SOLVER TO OPTIMISE FORECASTS

In this lesson, you will learn how to:

- Install the **Solver**
- Use the **Solver** to optimise exponential smoothing forecasts
- Use the **Solver** to optimise weighted moving average forecasts
- Use the **Solver** to optimise seasonal forecasts

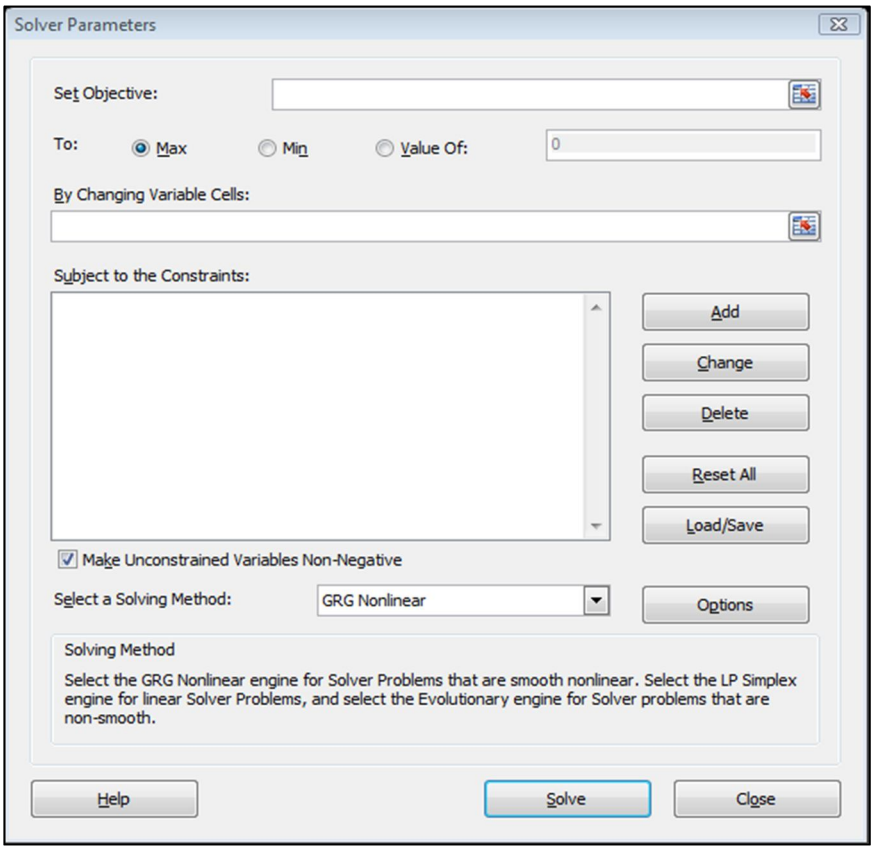
CONCEPT AND TERMS

Discussion

The **Solver** is an advanced Excel tool you can use in a lot of different situations. In this book the **Solver** will be used to optimise Weighted Moving Average forecasts, Seasonal forecasts, and Exponential Smoothing forecasts.

You can specify the following items in the **Solver** Parameters dialog box:

Parameter	Description
Set Target cell	The cell containing the formula you want **Solver** to use
To (equal to Excel 2007)	Whether to set the target cell to its maximum value, its minimum value, or a specified value
By Changing cells	The variable cells that **Solver** can change in order to reach the desired target cell value
Constraints	You can set constraints or limitations on the changing cells, the target cell, or other cells in the worksheet



(This is a screenshot of the Excel 2010 **Solver**. The Excel 2007 **Solver** looks a little different, but the parameters and how to use it for optimising forecasts will be the same**).**

INSTALLING THE SOLVER

Discussion

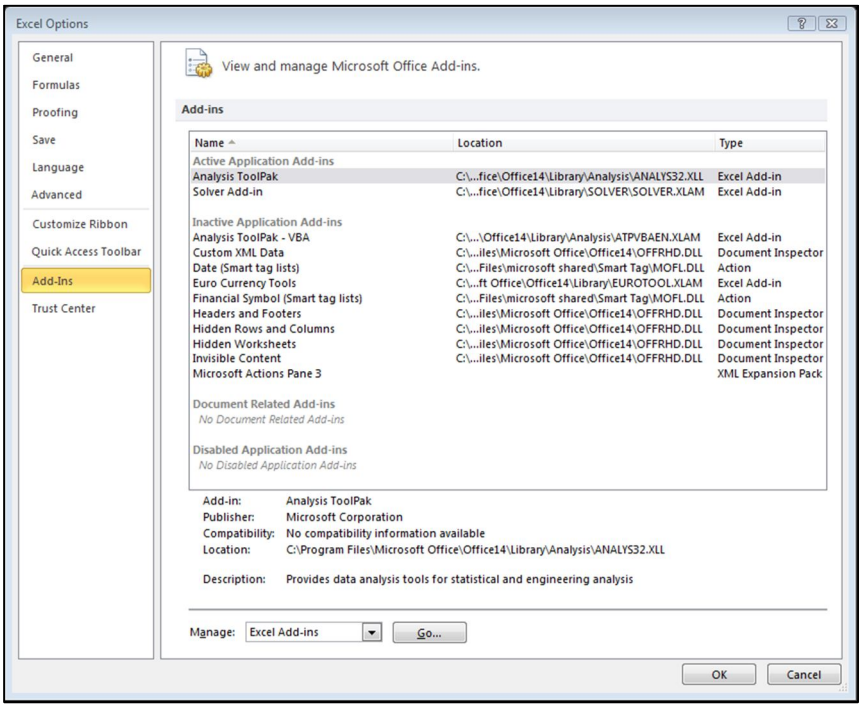
Solver is not part of the standard Excel application as normally installed, it is an Add-In (often referred to as a "plug-in."). Add-Ins are additional Excel components that are not commonly used and as a result, have to be installed on your computer. There are many third party (non-Microsoft) Add-Ins that help Excel do useful and clever things that it cannot normally do, or that enhance some of the things that it can do.

These can be found and downloaded from the Internet although it is wise to consult with your IT department first in case they have objections on grounds of security and computer software safety. It is unlikely that in the working environment you will be able to install these programs yourself anyway.

There are, however, selections of Add-Ins that come with Excel and that are perfectly safe and suitable for use; one of these is **Solver**.

Procedure

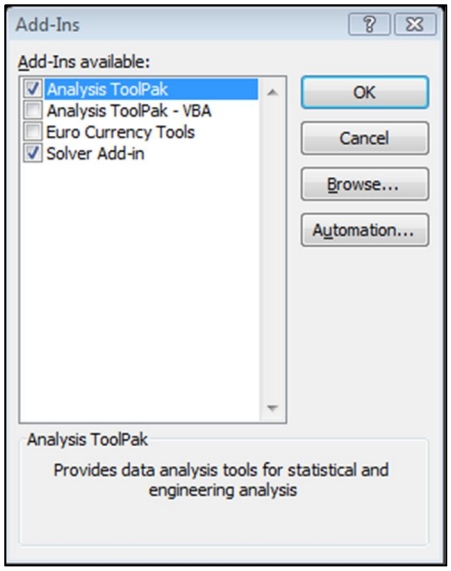
1. Click the Office Button (Excel 2007) or the File tab (Excel 2010).
2. Click Excel Options.
3. Click Add-Ins at the left.
4. Click Add-Ins at the left.



5. Select Excel Add-ins from the Manage: list at the bottom of the dialog box.

6. Click Go….

7. Click in box next to **Solver** Add-in.



8. Click in box next to **Solver** Add-in.

USE THE SOLVER TO OPTIMISE EXPONENTIAL SMOOTHING FORECASTS

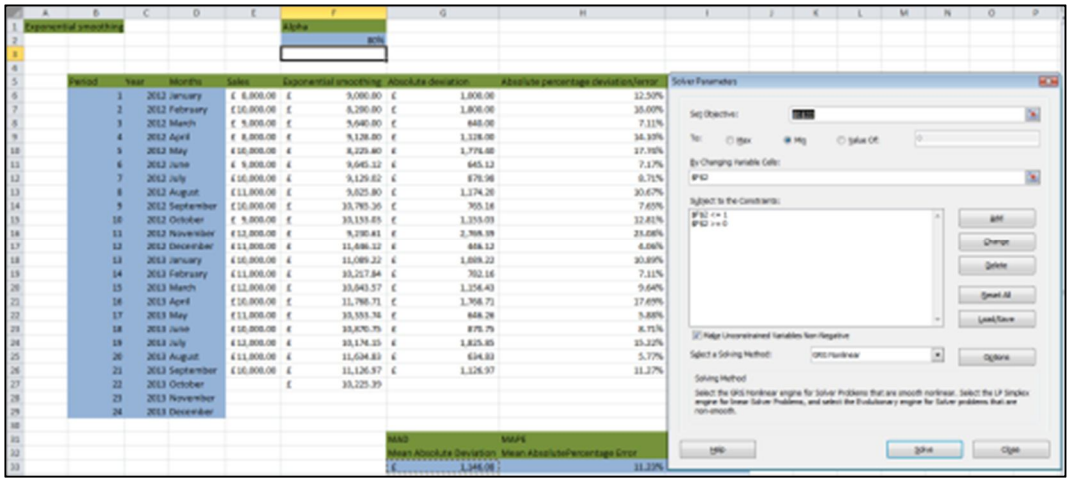
Discussion

The **Exponential Smoothing** forecasting model needs an **Alpha** value. A percentage of the previous known actual value and a percentage of the previous forecast. You can analyse the data and calculate the percentage but you can also ask the **Solver** to find the right percentage to get the most accurate forecast.

In the example below the **Exponential Smoothing** model uses an **Alpha** of 80%. The forecast is based on 80% of the previous actual known sales figure and 20% of the previous period's forecast. This percentage gives a **MAD** of 1146.08.

Solver parameters:

The **Set Objective:** $G$33 is the cell with **MAD. To: Min** is selected because you want the lowest possible **MAD** to get the most accurate forecast. **By Changing Variable Cells:** $F$2 is selected. It is **Alpha** the **Solver** must change to bring down **MAD. Subject to the Constraints:** 2 constrains are entered **Alpha** must be less than or equal to 1 (100%) or greater than or equal to 0 (0%).



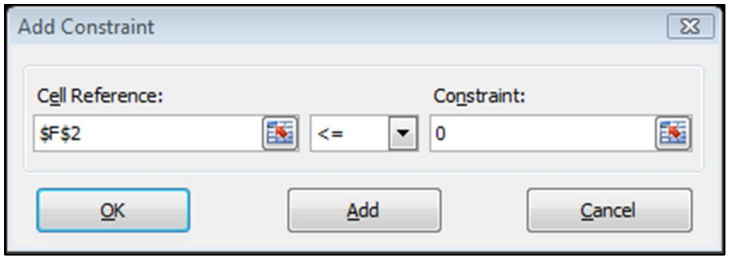
The **Solver** found the best percentage to be 28% and this changed the **MAD** from 1146.08 to 885.58.

Exponential smoothing				Alpha				
				28%				
Period	Year	Months	Sales	Exponential smoothing	Absolute deviation	Absolute percentage deviation/error	Square error	
1	2012	January	£ 8,000.00	£ 9,000.00	£ 1,000.00	12.50%		
2	2012	February	£ 10,000.00	£ 8,715.32	£ 1,284.68	12.85%		
3	2012	March	£ 9,000.00	£ 9,081.04	£ 81.04	0.90%		
4	2012	April	£ 8,000.00	£ 9,057.97	£ 1,057.97	13.22%		
5	2012	May	£ 10,000.00	£ 8,756.78	£ 1,243.22	12.43%		
6	2012	June	£ 9,000.00	£ 9,110.71	£ 110.71	1.23%		
7	2012	July	£ 10,000.00	£ 9,079.19	£ 920.81	9.21%		
8	2012	August	£ 11,000.00	£ 9,341.33	£ 1,658.67	15.08%		
9	2012	September	£ 10,000.00	£ 9,813.53	£ 186.47	1.86%		
10	2012	October	£ 9,000.00	£ 9,866.61	£ 866.61	9.63%		
11	2012	November	£ 12,000.00	£ 9,619.90	£ 2,380.10	19.83%		
12	2012	December	£ 11,000.00	£ 10,297.48	£ 702.52	6.39%		
13	2013	January	£ 10,000.00	£ 10,497.47	£ 497.47	4.97%		
14	2013	February	£ 11,000.00	£ 10,355.85	£ 644.15	5.86%		
15	2013	March	£ 12,000.00	£ 10,539.23	£ 1,460.77	12.17%		
16	2013	April	£ 10,000.00	£ 10,955.09	£ 955.09	9.55%		
17	2013	May	£ 11,000.00	£ 10,683.19	£ 316.81	2.88%		
18	2013	June	£ 10,000.00	£ 10,773.38	£ 773.38	7.73%		
19	2013	July	£ 12,000.00	£ 10,553.21	£ 1,446.79	12.06%		
20	2013	August	£ 11,000.00	£ 10,965.09	£ 34.91	0.32%		
21	2013	September	£ 10,000.00	£ 10,975.03	£ 975.03	9.75%		
22	2013	October		£ 10,697.45				
23	2013	November						
24	2013	December						
					MAD	MAPE	MSQ	
					Mean Absolute Deviation	Mean AbsolutePercentage Error	Mean Square Error	
					£ 885.58	8.59%		

Now the model can be optimised every time you get new actual known data in this example, when you get the October figures. Just use the **Solver** each time new data is entered.

Procedure

1. To use **Solver**, first click the Data tab.
2. In the Analysis group, click the **Solver** button.
3. Select the target cell.
4. Set the Equal To: option to the one you require.
5. In the By Changing Cells: box, select, or type, the range of cells the **Solver** must use to compare against the target cell value.
6. Click the Add button.
7. In the Add Constraint dialog box, specify the first cell constraint.



8. If necessary, click on Add to specify more cell constraints.
9. Click the OK button.

10. In the **Solver** Parameters dialog box, click the Solve button. The **Solver** Results

The **Solver** Results dialog box opens with the results.

USE THE SOLVER TO OPTIMISE WEIGHTED MOVING AVERAGE FORECASTS

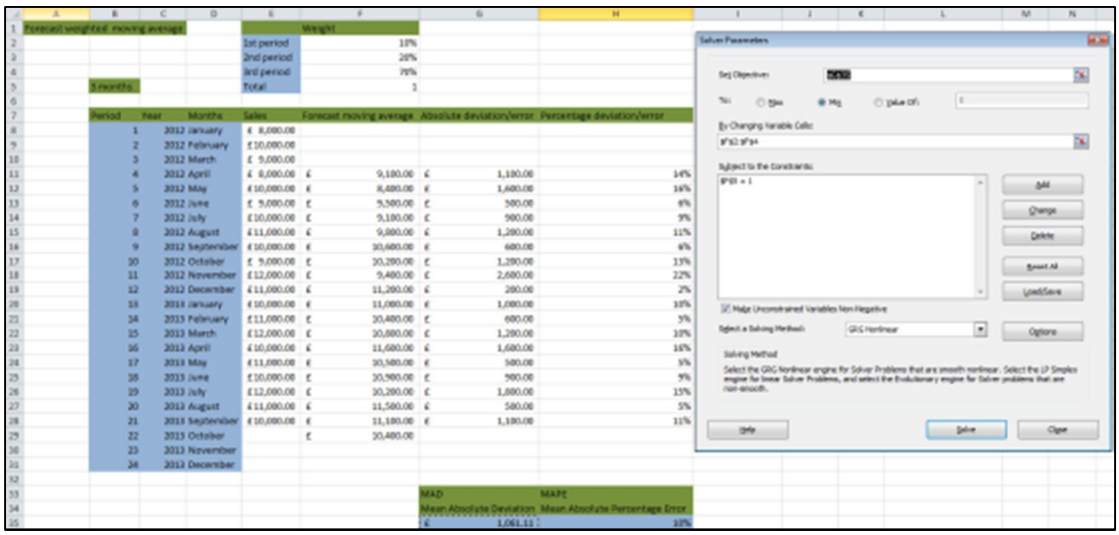
Discussion

The **Weighted Moving Average** forecasting model needs a weight / percentage value for previous periods. Instead analysing the data the **Solver** can find the right weights to get the most accurate forecast.

In the example below the **Weighted Moving Average** model uses a weight for the 1st period of 10%, 20% for the 2nd period, and 70% for the 3rd period. These percentages return a **MAD** 1061.11.

Solver parameters:

The **Set Objective:** $G$35 is the cell with **MAD. To: Min** is selected because you want the lowest possible **MAD** to get the most accurate forecast. **By Changing Variable Cells:** the range $F$2:$F$4 is selected. It is weights the **Solver** must change to bring down **MAD. Subject to the Constraints:** 1 constrain is entered the sum of the weights must be equal to 1 (100%).

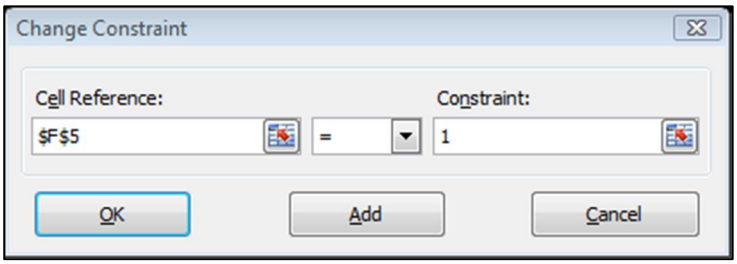


The **Solver** found the best percentage to be 71%, 14%, and 14% and this changed the **MAD** from 1061.11 to 769.80.

	A	B	C	D	E	F	G	H
1	Forecast weighted moving average					Weight		
2					1st period	71%		
3					2nd period	14%		
4					3rd period	14%		
5		3 months			Total	1		
6								
7		Period	Year	Months	Sales	Forecast moving average	Absolute deviation/error	Percentage deviation/error
8		1	2012	January	£ 8,000.00			
9		2	2012	February	£10,000.00			
10		3	2012	March	£ 9,000.00			
11		4	2012	April	£ 8,000.00	£ 8,431.03	£ 431.03	5%
12		5	2012	May	£10,000.00	£ 9,568.97	£ 431.03	4%
13		6	2012	June	£ 9,000.00	£ 9,000.00	£ 0.00	0%
14		7	2012	July	£10,000.00	£ 8,431.03	£ 1,568.97	16%
15		8	2012	August	£11,000.00	£ 9,856.32	£ 1,143.68	10%
16		9	2012	September	£10,000.00	£ 9,431.03	£ 568.97	6%
17		10	2012	October	£ 9,000.00	£ 10,143.68	£ 1,143.68	13%
18		11	2012	November	£12,000.00	£ 10,568.97	£ 1,431.03	12%
19		12	2012	December	£11,000.00	£ 10,143.68	£ 856.32	8%
20		13	2013	January	£10,000.00	£ 9,718.38	£ 281.62	3%
21		14	2013	February	£11,000.00	£ 11,568.97	£ 568.97	5%
22		15	2013	March	£12,000.00	£ 10,856.32	£ 1,143.68	10%
23		16	2013	April	£10,000.00	£ 10,431.03	£ 431.03	4%
24		17	2013	May	£11,000.00	£ 11,000.00	£ 0.00	0%
25		18	2013	June	£10,000.00	£ 11,568.97	£ 1,568.97	16%
26		19	2013	July	£12,000.00	£ 10,143.68	£ 1,856.32	15%
27		20	2013	August	£11,000.00	£ 11,000.00	£ 0.00	0%
28		21	2013	September	£10,000.00	£ 10,431.03	£ 431.03	4%
29		22	2013	October		£ 11,568.97		
30		23	2013	November				
31		24	2013	December				
32								
33							MAD	MAPE
34							Mean Absolute Deviation	Mean Absolute Percentage Error
35							£ 769.80	7%

Procedure

11. To use **Solver**, first click the Data tab.
12. In the Analysis group, click the **Solver** button.
13. Select the target cell.
14. Set the Equal To: option to the one you require.
15. In the By Changing Cells: box, select, or type, the range of cells the **Solver** must use to compare against the target cell value.
16. Click the Add button.
17. In the Add Constraint dialog box, specify the first cell constraint.



18. If necessary, click on Add to specify more cell constraints.
19. Click the OK button.

20. In the **Solver** Parameters dialog box, click the Solve button. The **Solver** Results

The **Solver** Results dialog box opens with the results.

USE THE SOLVER TO OPTIMISE SEASONAL FORECASTS

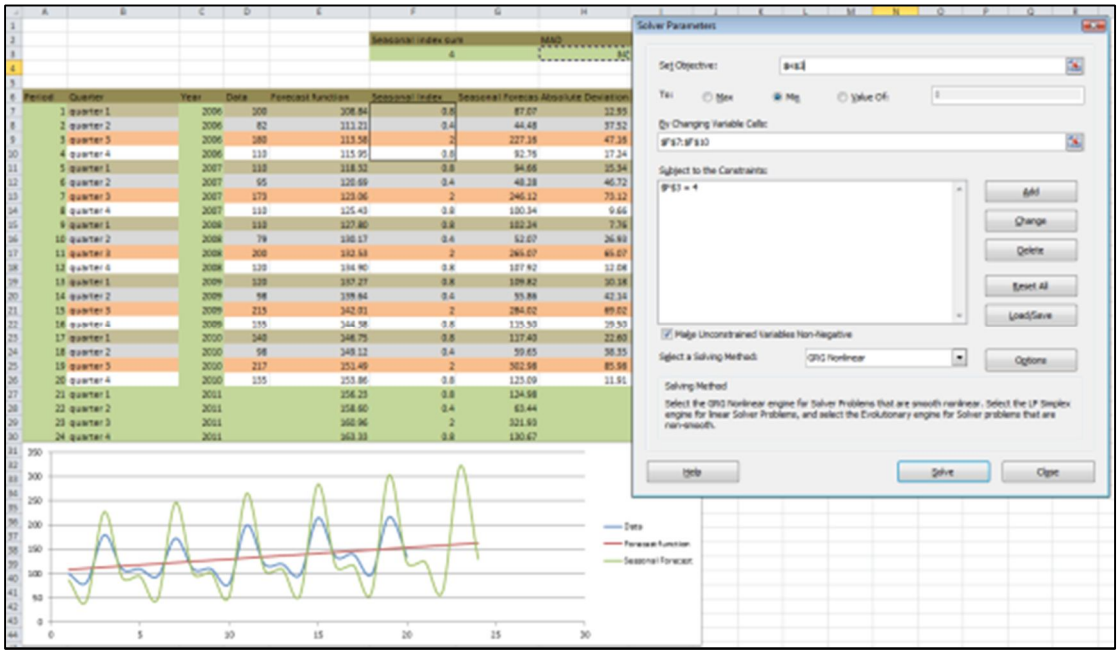
Discussion

The **Seasonal** forecasting model needs a weighted value for each period. The weights must equal the number of periods (if you are working with quarters the sum of the weights must equal 4, Months 12 etc.). Instead analysing the data the **Solver** can find the right weights to get the most accurate forecast.

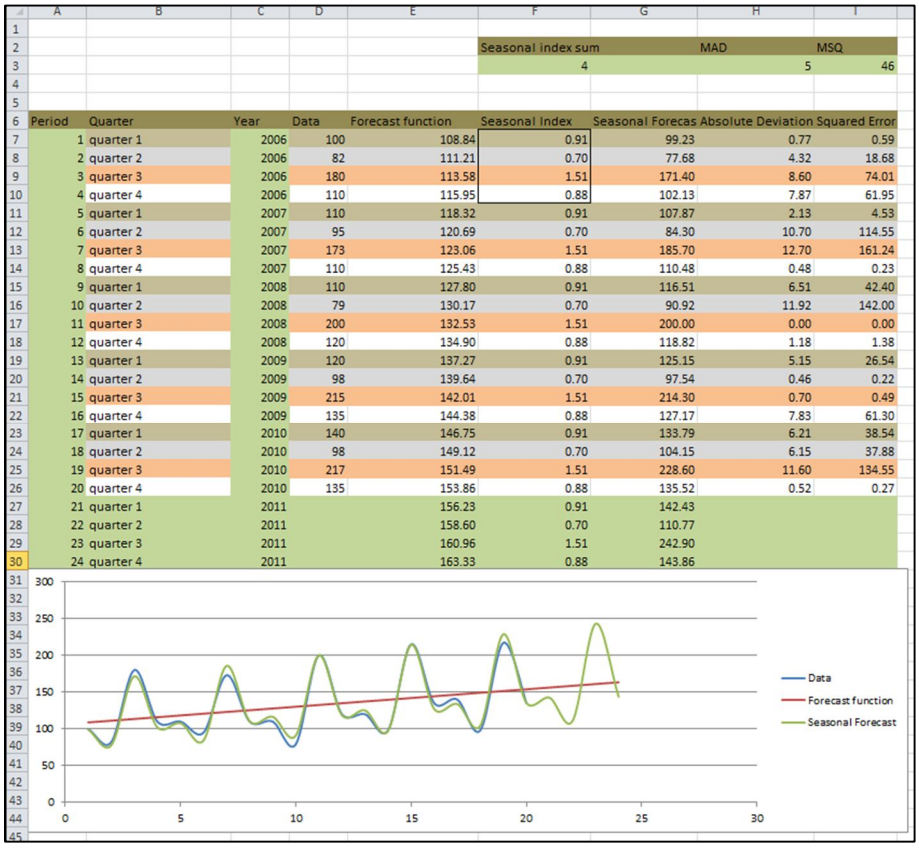
In the example below the **Seasonal** forecast model uses a weight for the 1st period of 0.8, 0.4 for the 2nd period, 2 for the 3rd period, and 0.8 for the 4th period. These percentages return a **MAD** 34.

Solver parameters:

The **Set Objective:** $H$3 is the cell with **MAD. To: Min** is selected because you want the lowest possible **MAD** to get the most accurate forecast. **By Changing Variable Cells:** the range $F$7:$F$10 is selected. It is weights the **Solver** must change to bring down **MAD. Subject to the Constraints:** 1 constrains are entered the sum of the weights must be equal to 4 (the sum of the weights and number of periods).



The **Solver** found the best suitable weights for the Seasonal forecast model and the forecast is much more accurate now with a MAD of 5 and please look at the chart. It is very clear that the forecast is very accurate.



					Seasonal index sum		MAD	MSQ
					4		5	46
Period	Quarter	Year	Data	Forecast function	Seasonal Index	Seasonal Forecas	Absolute Deviation	Squared Error
1	quarter 1	2006	100	108.84	0.91	99.23	0.77	0.59
2	quarter 2	2006	82	111.21	0.70	77.68	4.32	18.68
3	quarter 3	2006	180	113.58	1.51	171.40	8.60	74.01
4	quarter 4	2006	110	115.95	0.88	102.13	7.87	61.95
5	quarter 1	2007	110	118.32	0.91	107.87	2.13	4.53
6	quarter 2	2007	95	120.69	0.70	84.30	10.70	114.55
7	quarter 3	2007	173	123.06	1.51	185.70	12.70	161.24
8	quarter 4	2007	110	125.43	0.88	110.48	0.48	0.23
9	quarter 1	2008	110	127.80	0.91	116.51	6.51	42.40
10	quarter 2	2008	79	130.17	0.70	90.92	11.92	142.00
11	quarter 3	2008	200	132.53	1.51	200.00	0.00	0.00
12	quarter 4	2008	120	134.90	0.88	118.82	1.18	1.38
13	quarter 1	2009	120	137.27	0.91	125.15	5.15	26.54
14	quarter 2	2009	98	139.64	0.70	97.54	0.46	0.22
15	quarter 3	2009	215	142.01	1.51	214.30	0.70	0.49
16	quarter 4	2009	135	144.38	0.88	127.17	7.83	61.30
17	quarter 1	2010	140	146.75	0.91	133.79	6.21	38.54
18	quarter 2	2010	98	149.12	0.70	104.15	6.15	37.88
19	quarter 3	2010	217	151.49	1.51	228.60	11.60	134.55
20	quarter 4	2010	135	153.86	0.88	135.52	0.52	0.27
21	quarter 1	2011		156.23	0.91	142.43		
22	quarter 2	2011		158.60	0.70	110.77		
23	quarter 3	2011		160.96	1.51	242.90		
24	quarter 4	2011		163.33	0.88	143.86		

Procedure

1. To use **Solver**, first click the Data tab.
2. In the Analysis group, click the **Solver** button.
3. Select the target cell.
4. Set the Equal To: option to the one you require.
5. In the By Changing Cells: box, select, or type, the range of cells the **Solver** must use to compare against the target cell value.
6. Click the Add button.
7. In the Add Constraint dialog box, specify the first cell constraint.
8. If necessary, click on Add to specify more cell constraints.

9. Click the OK button.

10. In the **Solver** Parameters dialog box, click the Solve button. The **Solver** Results

The **Solver** Results dialog box opens with the results.

Exercises

Exercise 1 Optimising exponential smoothing forecasts using the Solver

1. Open the file **Forecasting and data analysis**.
2. Open the **Optimise using Solver exercises** sheet.
3. Open the **Solver** parameters dialog box.
4. The **Solver** will be used to bring down the MAD (Mean Absolute Deviation). **Set Objective:** $G$35 **To:** select **Min.** to get the lowest possible value.
5. **By Changing Variable Cells:** the cell with the **Alpha** ($F$2) value is the cell the **Solver** can change to bring down the error.
6. Add two constraints the **Alpha** value must be greater than or equal to 0 and less than or equal to 1 (between 0% and 100%).
7. Click **Solve** and examine the MAD and the chart.
8. Save the file.

Exercise 2 Optimising weighted 3 periods moving average using the Solver

1. Open the file **Forecasting and data analysis**.
2. Open the **Optimise using Solver exercises** sheet.
3. Open the **Solver** parameters dialog box.
4. The **Solver** will be used to bring down the MAD (Mean Absolute Deviation). **Set Objective:** $G$71 **To:** select **Min.** to get the lowest possible value.
5. **By Changing Variable Cells:** the cells with **weights** $F$38:$F$40 values are the cells the **Solver** can change to bring down the error.
6. Add one constraint the **Sum of weights** ($G$38) must equals 100%.
7. Click **Solve** and examine the MAD and the chart.
8. Save the file.

Exercise 3 Optimising seasonal forecasts using the Solver

1. Open the file **Forecasting and data analysis**.
2. Open the **Optimise using Solver exercises** sheet.
3. Open the **Solver** parameters dialog box.
4. The **Solver** will be used to bring down the MAD (Mean Absolute Deviation). **Set Objective:** $J$106 **To:** select **Min.** to get the lowest possible value.
5. **By Changing Variable Cells:** the cell with the seasonal weights or index values ($G$77:$G$80) are the cells the **Solver** can change to bring down the error.
6. Add one constraint the sum of the weights ($G$74) must equals 4.
7. Click **Solve** and examine the MAD and the chart.
8. Save the file.

LESSON 4 - SHOWING TRENDS AND FORECASTS USING CHARTS

In this lesson, you will learn how to:

- Choose the right chart type
- Create trendlines
- Choose the best trendline for your data
- Visualise forecasts and forecast errors in a chart

CONCEPTS AND TERMS

Discussion

Excel provides us with a wide selection of charts. Charts are very powerful tools to visualise data and can make it much easier to understand data analysis. In this book the primary idea is to look at the best charts for visualising forecasts and data analysis. How trends easy can be spotted, how forecasts easily can be done in a chart, how you can visualise forecast errors.

CHOOSE THE RIGHT CHART TYPE

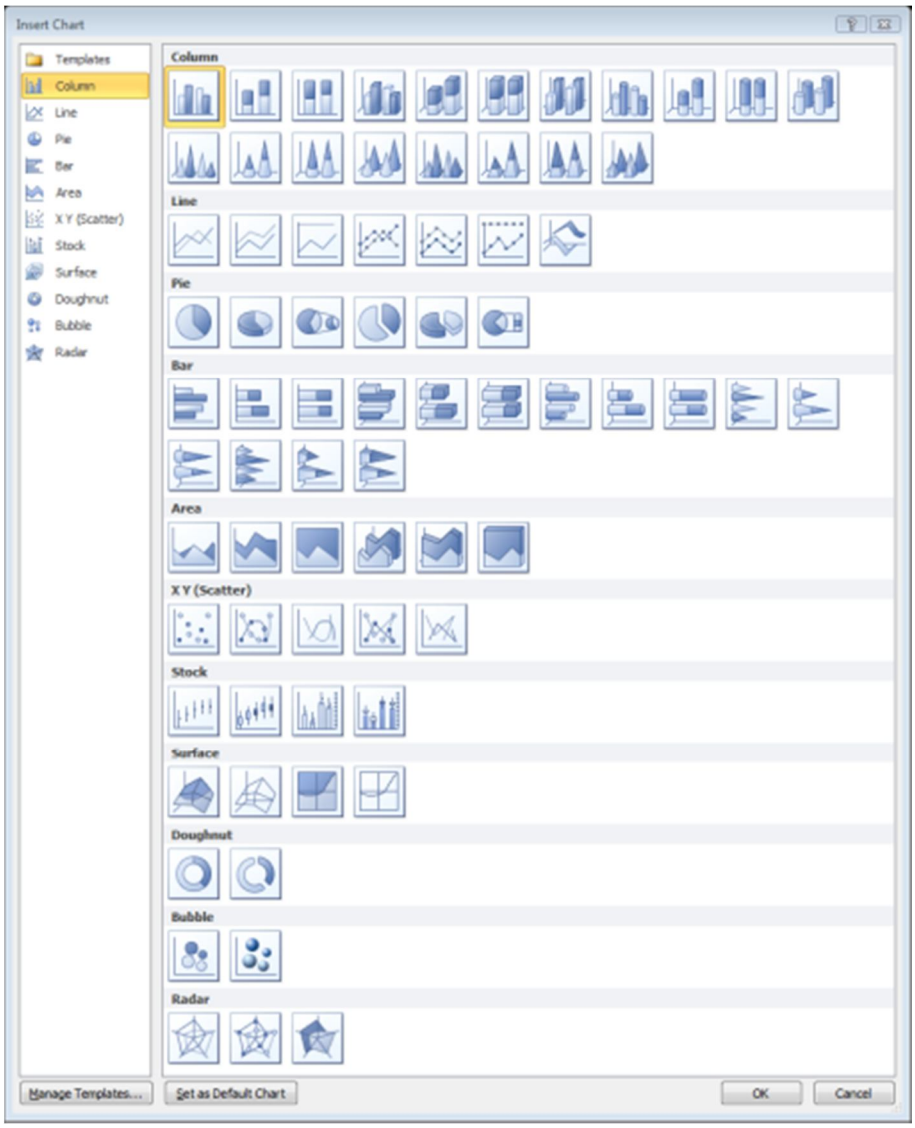
Discussion

The most common charts used to visualise data analysis are **Column** charts, **Line** charts, **Scatter** charts, and **Pie** charts, but all the charts provided by Excel can be used.

Chart type	Description
Pie charts:	Pie charts are superb for displaying data points as a percentage of the whole, but as with all chart types if the value for the data points getting to small (the gap between the values to big) it is difficult to distinguish each data point in the chart. If a data point is less than 5% of the pie, it becomes hard to distinguish the slice. Pie charts can only display one data series, which means you can include only one column or row of values in your selection when you create a pie chart.
Column charts:	Data that is arranged in columns or rows on a worksheet can be plotted in a column chart. Column charts are useful for showing data changes over a period of time or for illustrating comparisons among items. In column charts, categories are typically organized along the horizontal axis and values along the vertical axis. Consider using a column chart when: • You have one or more data series that you want to plot. • Your data contains positive, negative, and zero (0) values. • You want to compare the data for numerous categories side by side.

Line charts:	Line charts are great tools to track and trend data. Line charts can display continuous data over time, set against a common scale, and are therefore ideal for showing trends over time. As a general rule, you may want to use a Line chart if your data has non-numeric x-values. (For numeric x-values, it's usually better to use a Scatter chart).
Scatter charts:	Scatter charts are commonly used for displaying and comparing numeric values, such as scientific, statistical, and engineering data. You should use a Scatter chart if you want to change the scale of the x-axis, or make it a logarithmic scale. You can use this chart type to effectively: Display worksheet data that includes pairs or grouped sets of values. You can adjust the independent scales of a Scatter chart to reveal more information about the grouped values. Show similarities between large sets of data. Rather than showing the differences between data points, a Scatter chart can point out interesting similarities. Compare large numbers of data points without regard to time. The more data that you include in a Scatter chart, the better the comparisons that you can make. Scatter charts can be displayed with or without lines to connect the data points, and connecting lines can be displayed with or without data markers.

Below you can see the Excel chart dialog box where you can find the entire chart types in Excel.



CREATE TRENDLINES

Discussion

Before you start creating charts and trendlines make sure that:

You have enough data to show a meaningful chart and enough data for Excel to be able to calculate a reliable trend. Insufficient data might twist the results. For example, seasonal fluctuations might be mistaken for long-term trends if the data is from only one year. The chart must be a 2-D chart. Excel cannot create trendlines on 3-D charts.

- The data is ordered from earliest to most recent.
- No data is missing. If data is unavailable for a period, enter an estimate.

You can forecast with trendlines you will see how later in this book.

So what are trendlines, really?

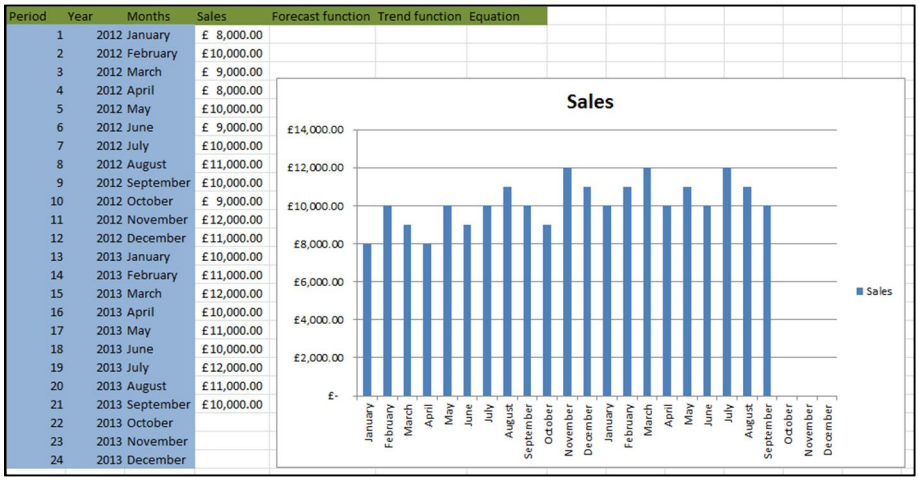
Trendlines are graphical representations of trends in data that you can use to analyse problems of prediction. Such analysis is also called regression analysis. By using regression analysis, you can extend a trendline in a chart beyond the actual data to predict future values. For example, the preceding chart uses a simple logarithmic trendline that is forecast ahead four quarters to clearly show a trend toward rising revenue.

R^2 (R-squared)

So you might be wondering: How reliable are these trendlines, anyway? The answer has to do with something called R-squared — or, more specifically, the R-squared value of the trendline (this is where math comes in). Think of the R-squared value as a magic number — in this case, a number between 0 and 1.

You don't have to understand all about R-squared values — just remember: A trendline is most reliable when its R-squared value is at or near 1. When you fit a trendline to your existing data, Excel automatically calculates its R-squared value based on a formula. The more closely a trendline match your existing data, the more accurate a forecast that uses this trendline is likely to be. If you want, you can display this value on your chart.

First you need to create a chart.

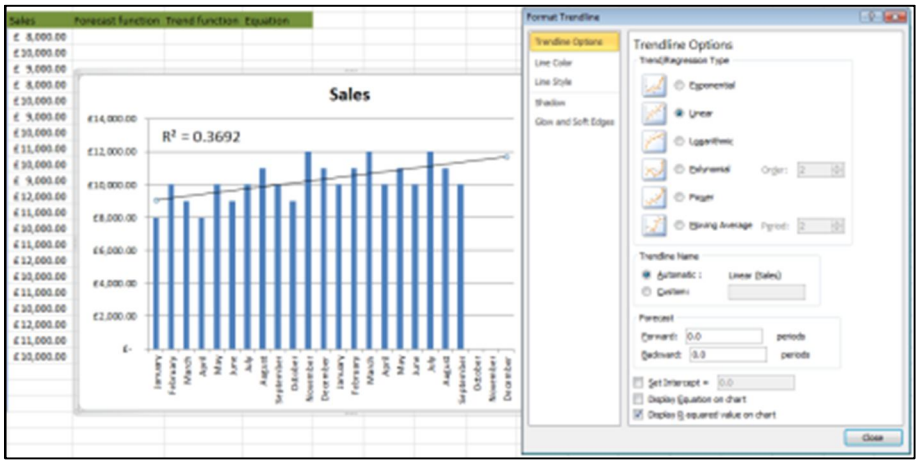


Period	Year	Months	Sales	Forecast function	Trend function	Equation
1	2012	January	£ 8,000.00			
2	2012	February	£10,000.00			
3	2012	March	£ 9,000.00			
4	2012	April	£ 8,000.00			
5	2012	May	£10,000.00			
6	2012	June	£ 9,000.00			
7	2012	July	£10,000.00			
8	2012	August	£11,000.00			
9	2012	September	£10,000.00			
10	2012	October	£ 9,000.00			
11	2012	November	£12,000.00			
12	2012	December	£11,000.00			
13	2013	January	£10,000.00			
14	2013	February	£11,000.00			
15	2013	March	£12,000.00			
16	2013	April	£10,000.00			
17	2013	May	£11,000.00			
18	2013	June	£10,000.00			
19	2013	July	£12,000.00			
20	2013	August	£11,000.00			
21	2013	September	£10,000.00			
22	2013	October				
23	2013	November				
24	2013	December				

It can be difficult to see a trend just looking at the chart. To add a trendline you need to select the chart and select the contextual **Chart Tool** tab **layout.** Select **Linear Trendline** from the trendline list.

You just created a trendline!

Right click the trendline and click Format Trendline. Click **Display R-squared value on chart** to display R^2.



The R-squared value, which tells you how closely your trendline follows your data, appears on the chart. With a linear trendline in this example above, the value is 0.3692. As you can see the trendline doesn't follow the data very closely because the sales goes up and down. A linear trendline with an R-squared value of 0.3692 is not a very reliable trendline for a forecast. The linear trendline in this example will may be not be able to give us an accurate forecast but it is easy to see that the trend is going up so you have increasing sales trend and that is of course nice.

You can change the trendline to another trendline type and compare the R-square values. May be there is a trendline which is a better match to the data. In the next chapter you will see a description of all the different trendlines and when to use the different types.

Procedure

1. Create a chart.
2. Make sure that the chart is selected. Click **Trendline** on the contextual **Chart Tools** tab.
3. Click on the trend line type you want to create.
4. Right click the trendline and choose **Format Trendline.**
5. Click **Display R-squared value on chart** to display R^2.

*Instead of right click the trendline to format the trendline you can click **More Trendline Options** in the trendline type dialog list.*

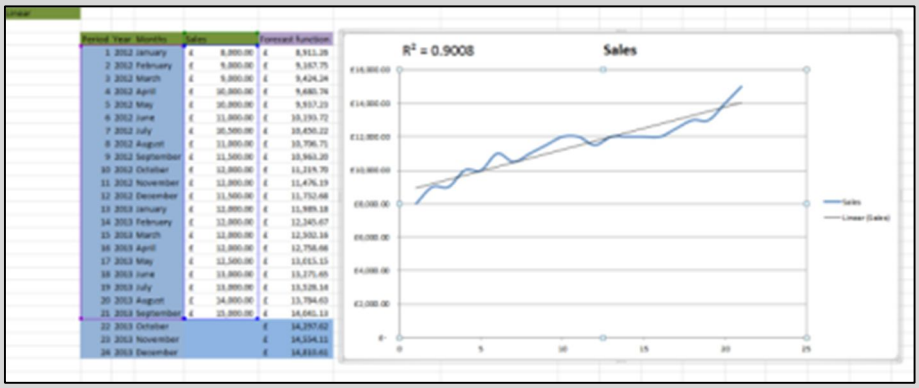
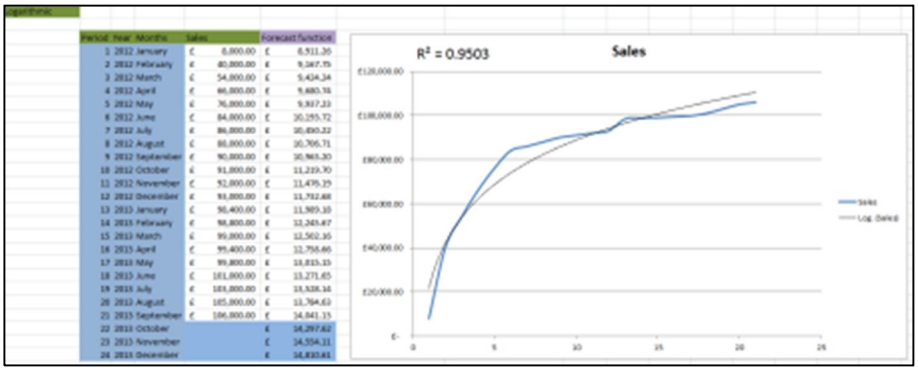
*You can create a trendline by right click the data series and click **Add Trendline.***

CHOOSING THE BEST TRENDLINE FOR YOUR DATA

Discussion

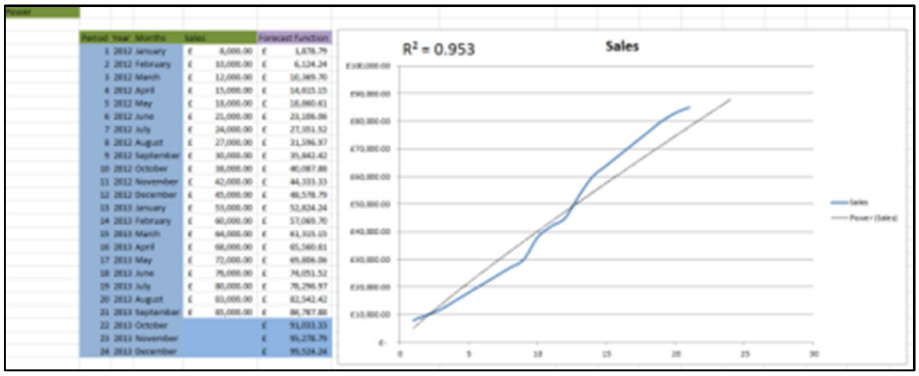
When you want to add a trendline to a chart in Excel, you can choose between six different trend / regression types. You should use the type which fit best to your data.

Choosing the right trendline for your data is very important. Checking the R-squared value can help you choose the best trendline for your data. It also helps to understand the types of trendlines that are likely to fit different scenarios.

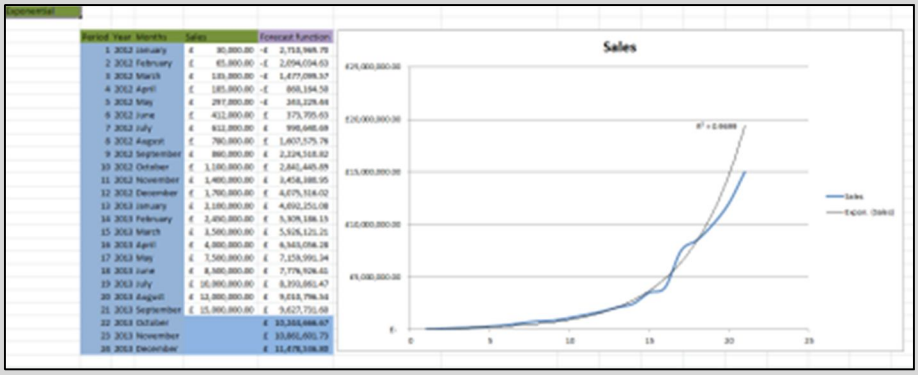
Type	Description
Linear	A linear trendline is a best-fit straight line that is used with simple linear data sets. Your data is linear if the pattern in its data points resembles a line. A linear trendline usually shows that something is increasing or decreasing. 
Logarithmic	A logarithmic trendline is a best-fit curved line that is most useful when the rate of change in the data increases or decreases quickly and then levels out. A logarithmic trendline can use negative and / or positive values. 

Polynomial	A polynomial trendline is a curved line that is used when data fluctuates. It is useful, for example, for analysing gains and losses over a large data set. The order of the polynomial can be determined by the number of fluctuations in the data or by how many bends (hills and valleys) appear in the curve. An Order 2 polynomial trendline generally has only one hill or valley. Order 3 generally has one or two hills or valleys. Order 4 generally has up to three.

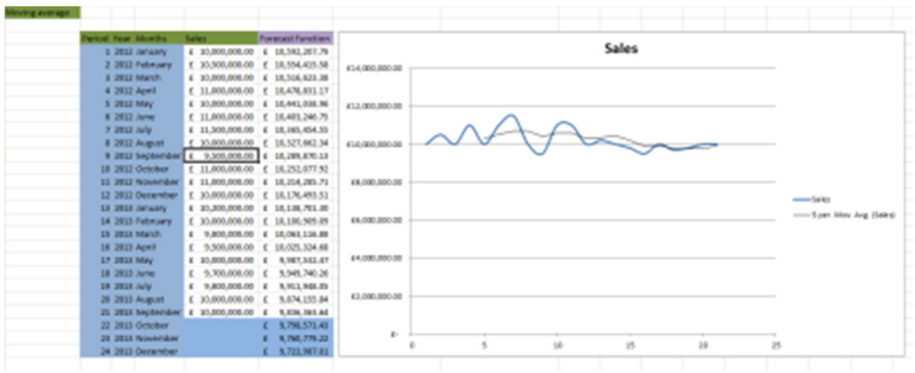
Power	A power trendline is a curved line that is best used with data sets that compare measurements that increase at a specific rate — for example, the acceleration of a race car at one-second intervals. You cannot create a power trendline if your data contains zero or negative values.



Exponential	An exponential trendline is a curved line that is most useful when data values rise or fall at increasingly higher rates. You cannot create an exponential trendline if your data contains zero or negative values.



Moving average	A moving average trendline smooth out fluctuations in data to show a pattern or trend more clearly. A moving average trendline uses a specific number of data points (set by the Period option), averages them, and uses the average value as a point in the trendline. If Period is set to 2, for example, then the average of the first two data points is used as the first point in the moving average trendline. The average of the second and third data points is used as the second point in the trendline, and so on.

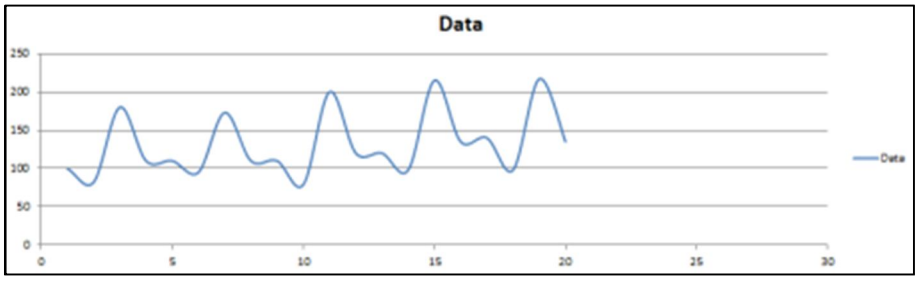


VISUALISE FORECASTS AND FORECAST ERRORS IN A CHART

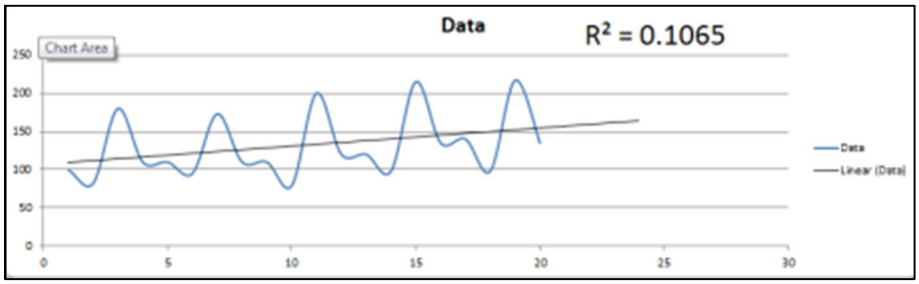
Discussion

Forecast charts are exceptional to show the result in a very user friendly easy understandable manner. Line charts, scatter chart, columns charts, or a combination of column charts and line charts should be the charts used for forecasts.

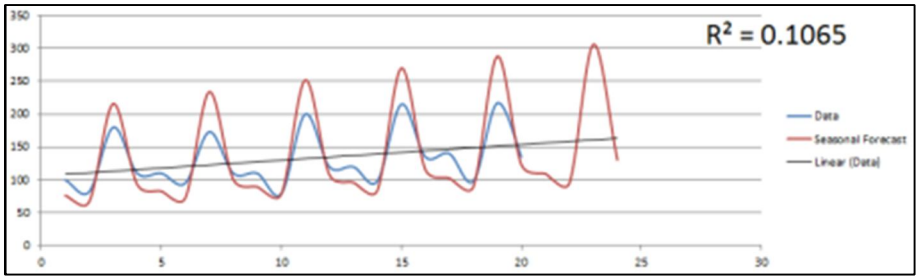
In the scatter chart below the blue line is data over 20 quarters. The vertical axis shows the data and the horizontal axis shows the quarters. It is easy to see that it is seasonal data.



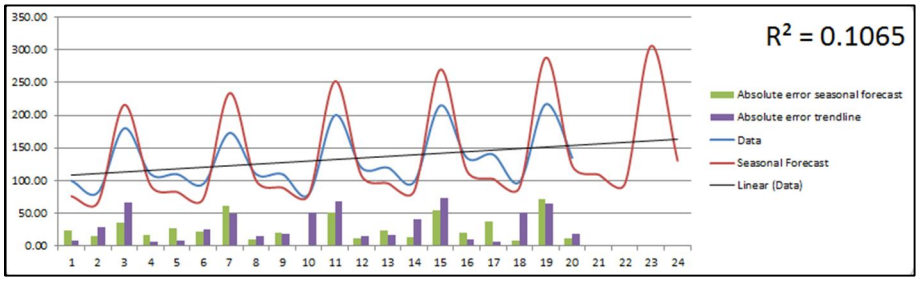
By adding a 4 periods forecast trendline it is obvious that the values increase over time but also that the data is not linear data (the R^2=0.1065). It tells us that the trendline's prediction for period 21 to 24 is not reliable.



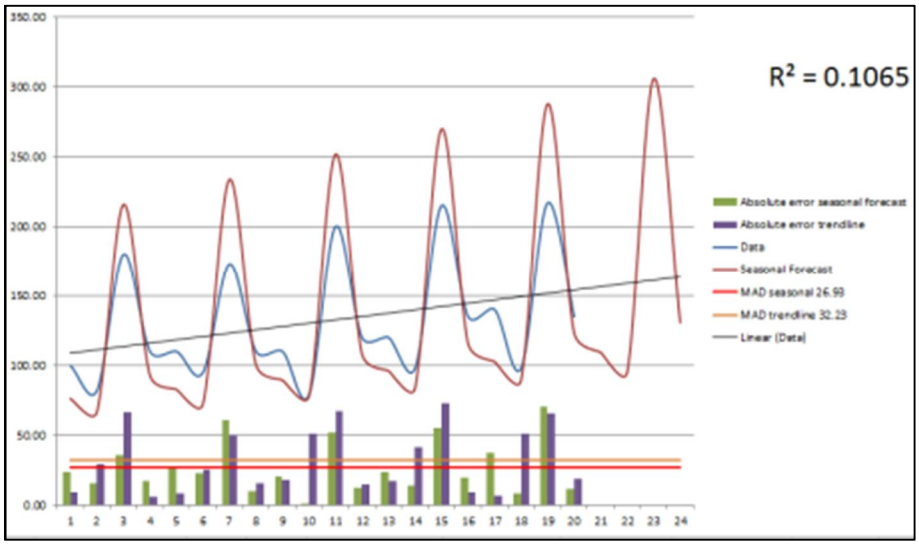
For seasonal data a seasonal forecast should be better to predict the future. In the chart below the result of a seasonal forecast is added and you can see that it is much better to "find" the ups and downs in the data.



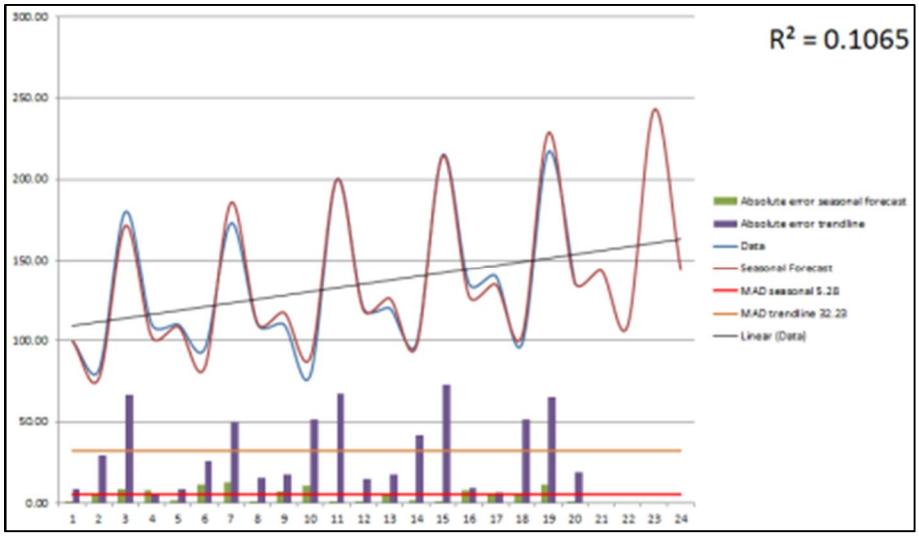
By adding the absolute error for the two forecast methods the forecast accuracy can be shown in the chart period by period, but it is not obvious which one is the most accurate.



By adding the MAD (Mean Absolute Deviation) to the chart you can now see that the seasonal forecast is the most accurate forecast, but it is not because it is much more accurate than the trendline forecast.

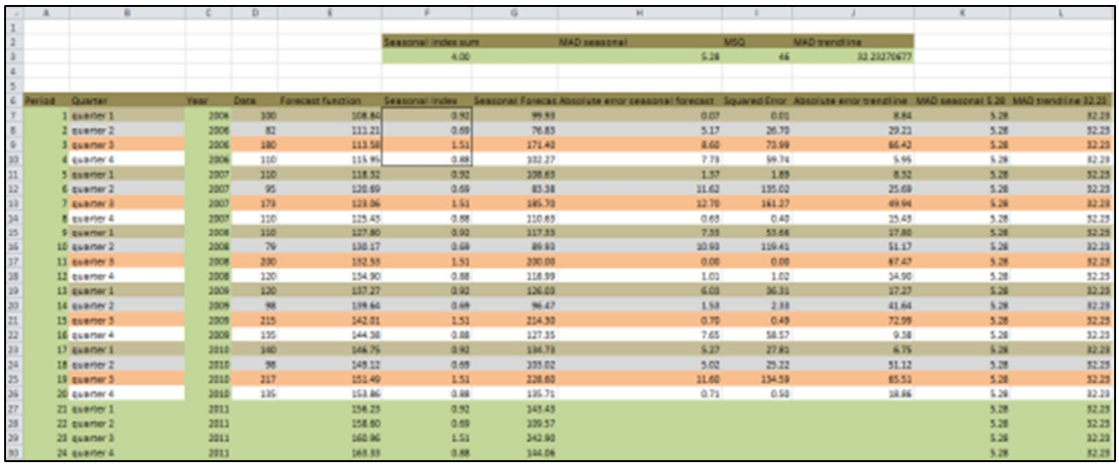


In the example below the **Solver** has been used to optimise the seasonal forecast and it is easy to see that now the seasonal forecast is much more accurate than the trendline forecast. The **Solver** decreased MAD from 26.93 to 5.28.

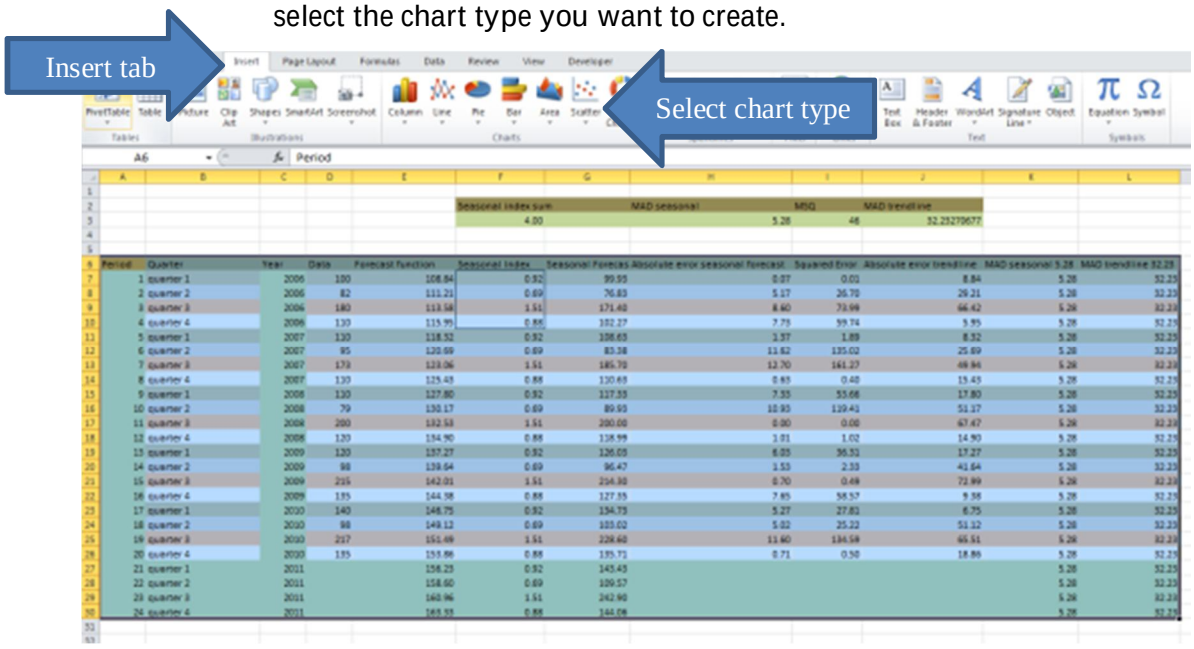


Procedures

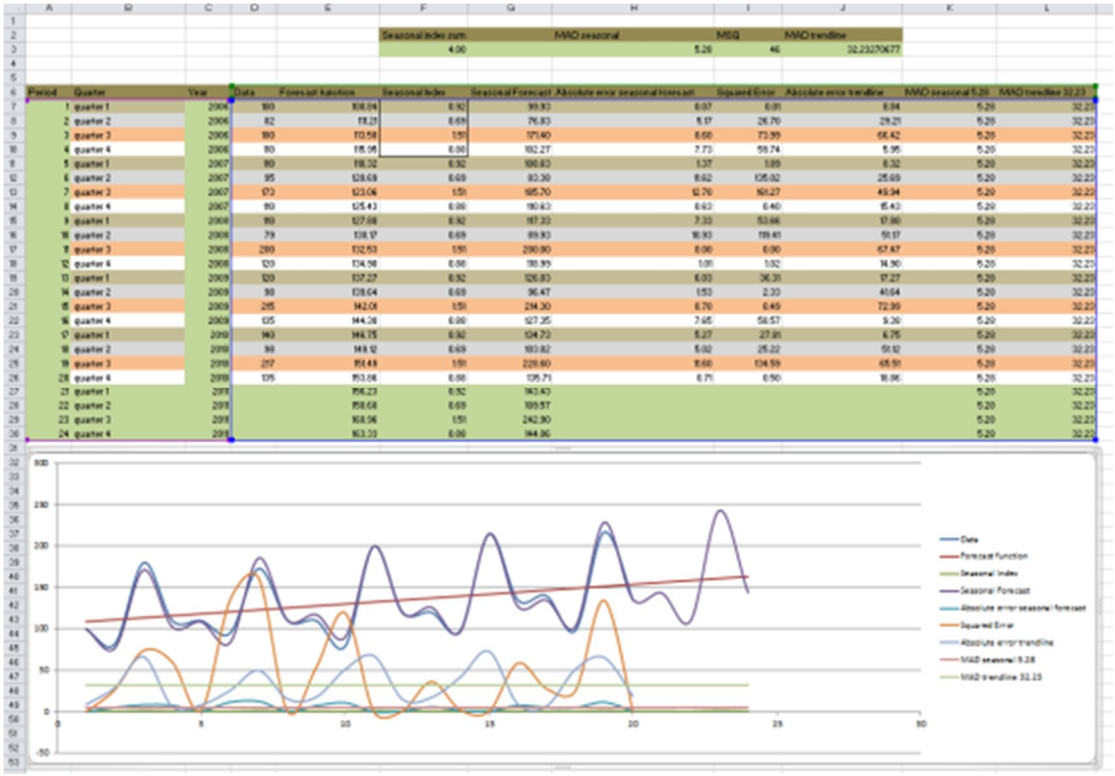
1. Create the data you want to visualise in a chart.



2. Select the data, click **Insert** tab and on the charts group select the chart type you want to create.



3. Move and resize the chart.



Exercises

Exercise 1 Adding trendlines

1. Open the file **Forecasting and data analysis**.
2. Open the **Visualise forecasts exercise** sheet.
3. Right click the data series **Sales** on the chart in **Exercise 1**.
4. Click **Add Trendline**.
5. Create a linear trendline and tick **Display Equation on chart** and **Display R-squared value on chart**.
6. Click **Line Color** top left in the **Format Trendline** dialog box.
7. Select **Solid line** and the colour red.
8. Examine the R^2 value. Do you think this is the right trendline for this chart?
9. Right click the trendline and click **Format Trendline.**
10. Type 4 in the **Forward:** box under **Forecast.**
11. Click **Close.**
12. Save the file.

Exercise 2 Trendline types

1. Open the file **Forecasting and data analysis**.
2. Open the **Visualise forecasts exercise** sheet.
3. Right click the data series **Sales** on the chart in **Exercise 2**.
4. Click **Add Trendline**.
5. Create a linear trendline and tick **Display Equation on chart** and **Display R-squared value on chart**.
6. Examine the R^2 value. Do you think this is the right trendline for this chart? The **R-squared** value is relatively high but the trendline does not reflect the data series very well.
7. Write down the **R-squared** value. Right click the trendline and click **Format Trendline.**
8. Try the other trendline types and compare the **R-squared** value.
9. Use the trendline with the highest **R-squared** value
10. Save the file.

Exercise 3 Visualising forecasts and errors

In the exercise 3 area (starting from row 40) two different methods have been used to forecast the sales the **Forecast** function and seasonal forecasting. Use the chart to compare the two methods and to decide which one you should use to get the most accurate forecast.

1. Open the file **Forecasting and data analysis**.
2. Open the **Error-deviation exercises** sheet.
3. Add the **Forecast** function data to the chart (copy the range F40:F68 click on the chart area and paste the data).
4. Do the same with the seasonal forecast data (copy the range O40:O68 click on the chart area and paste the data).
5. Add the forecast function's Absolute error (column H) and the seasonal forecast's Absolute error (column Q) to the second chart (the blank chart).
6. Right click the two error series and change chart type to column (If they are not in columns already).
7. Add the forecast function's MAD (range L40:L68) and the seasonal forecast's MAD (U40:U68) to the second chart.
8. Right click the two MAD series and change chart type to line.
9. Use the **Solver** to optimise the seasonal forecast. The right parameters are already added to the **Solver** parameter dialog box.
10. Compare the two methods which one is the best and most accurate?
11. Save the file.

LESSON 5 – COMPARING FORECASTING METHODS AND MODELS

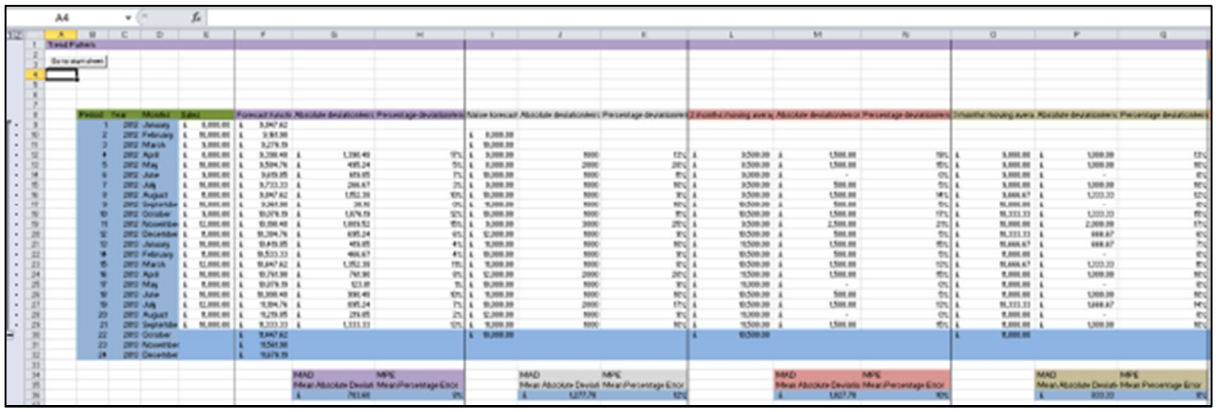
In this lesson, you will learn how to:

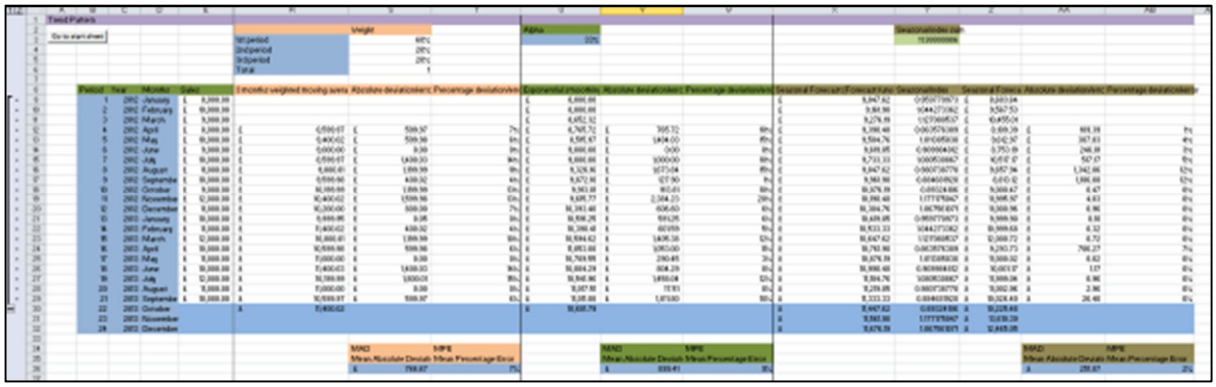
- Compare different forecasting models
- Understand which forecasting model to use

CONCEPT AND TERMS

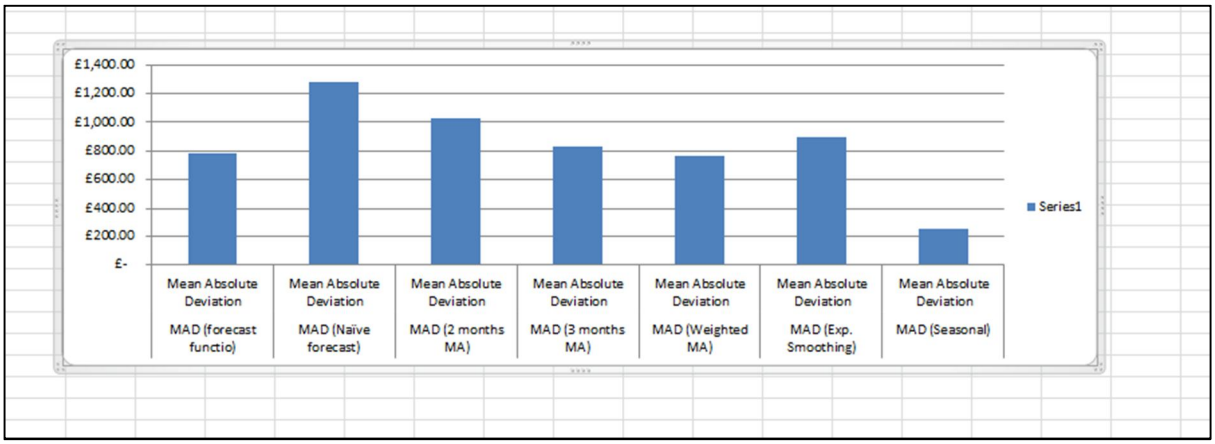
Discussion

To keep your forecast on track it is important to continuously developing your forecast model. You have seen different forecasting models and how to calculate the errors. You should use more than one method and every time you add more data calculate the errors and optimise your model. It can change over time which forecast method generates the most accurate forecasts. If you by comparison of the different method can see a change in accuracy between the methods you can keep your forecasts on track by always knowing which forecast method generates the best result for your data.





Create a column chart displaying all MAD results from all the forecasting method can be a nice tool. Keep an eye on the chart when you are adding new data to you forecasting model and you will very easy visualise which model is the most reliable.



You can also after trying the different methods decide which one to use in the future and then use TSE (Tracking Signal Error) to make sure that the model is on track. The TSE value must be between -4 to 4.

LESSON 6 – FORECASTING USING WHAT-IF ANALYSIS

In this lesson, you will learn how to:

- Use the **Scenarios Manager**
- Use the **Goal Seek** tool
- Use the **Data Table** tool

CONCEPTS AND TERMS

Discussion

The What-If analysis tools are very powerful tools when you need to forecast and you have some knowledge which Excel cannot calculate from the historical data. Your company will add 5 new products to the product line. The costs will be reduced because of an investment in new manufacturing equipment. You expect a marketing investment to pay off. You can manipulate your forecasts using What-If analysis.

THE SCENARIOS MANAGER

Discussion

The **Scenarios** tool can be a very nice supply to forecasting. You may need to forecast December. You have the data for each month from January 2011 to September 2012. You have the units you have sold, the unit price, the variable costs, the fixed costs, the sales, the profit, the total market units, and your company's market shares. The **Forecast** function has been used to calculate the forecast for Market shares and total market units for December. You know that the forecast will not be accurate.

First of all you know that the sales for December will be very good. You also know that your company will spend a lot of money getting a bigger market share. Your company is also working on reducing the fixed and variable costs and you also have plans changing the price policy for your products.

You want now to work with the forecast to see what will happen with all these changes. You do not know if all the changes will happen so you want to create different scenarios to be able to compare the different situations.

You can create a scenario summary report, which lists all the scenarios you have created in a side-by-side format so that you can compare them. When you create a scenario summary report, the Scenario Manager automatically inserts a Scenario Summary sheet in the workbook and places the report on it. This sheet allows you to easily view and print the scenario summary report. A scenario summary report appears in an outline format.

Period	Year	Months	Units	Unit Price	Variable Costs	Fixed costs	Sales	Profit	Total Market (units)	Total Market (units) Forecast	Market Shares	Market Shares Forecast
1	2012	January	1000	£ 8.00	£ 2.00	£ 500.00	£ 8,000.00	£ 5,500.00	100000	107273	1.00%	1.03%
2	2012	February	1250	£ 8.00	£ 2.00	£ 500.00	£10,000.00	£ 7,000.00	100000	107545	1.25%	1.03%
3	2012	March	1125	£ 8.00	£ 2.00	£ 500.00	£ 9,000.00	£ 6,250.00	100000	107818	1.13%	1.03%
4	2012	April	1000	£ 8.00	£ 2.00	£ 500.00	£ 8,000.00	£ 5,500.00	120000	108091	0.83%	1.02%
5	2012	May	1250	£ 8.00	£ 2.00	£ 500.00	£10,000.00	£ 7,000.00	120000	108364	1.04%	1.02%
6	2012	June	1125	£ 8.00	£ 2.00	£ 500.00	£ 9,000.00	£ 6,250.00	120000	108636	0.94%	1.01%
7	2012	July	1000	£ 10.00	£ 2.00	£ 500.00	£10,000.00	£ 7,500.00	120000	108909	0.83%	1.01%
8	2012	August	1100	£ 10.00	£ 2.00	£ 500.00	£11,000.00	£ 8,300.00	120000	109182	0.92%	1.01%
9	2012	September	1000	£ 10.00	£ 2.00	£ 500.00	£10,000.00	£ 7,500.00	100000	109455	1.00%	1.00%
10	2012	October	900	£ 10.00	£ 2.00	£ 500.00	£ 9,000.00	£ 6,700.00	100000	109727	0.90%	1.00%
11	2012	November	1200	£ 10.00	£ 2.00	£ 500.00	£12,000.00	£ 9,100.00	100000	110000	1.20%	0.99%
12	2012	December	1500	£ 10.00	£ 2.00	£ 500.00	£15,000.00	£ 11,500.00	130000	110273	0.85%	0.99%
13	2013	January	1000	£ 10.00	£ 1.80	£ 500.00	£10,000.00	£ 7,700.00	90000	110545	1.11%	0.99%
14	2013	February	1100	£ 10.00	£ 1.80	£ 500.00	£11,000.00	£ 8,520.00	100000	110818	1.10%	0.98%
15	2013	March	1200	£ 10.00	£ 1.80	£ 500.00	£12,000.00	£ 9,340.00	100000	111091	1.20%	0.98%
16	2013	April	1000	£ 10.00	£ 1.80	£ 500.00	£10,000.00	£ 7,700.00	100000	111364	1.00%	0.97%
17	2013	May	1100	£ 10.00	£ 1.80	£ 500.00	£11,000.00	£ 8,520.00	120000	111636	0.92%	0.97%
18	2013	June	1000	£ 10.00	£ 1.80	£ 500.00	£10,000.00	£ 7,700.00	120000	111909	0.83%	0.97%
19	2013	July	1200	£ 10.00	£ 1.80	£ 500.00	£12,000.00	£ 9,340.00	120000	112182	1.00%	0.96%
20	2013	August	1100	£ 10.00	£ 1.80	£ 500.00	£11,000.00	£ 8,520.00	120000	112455	0.92%	0.96%
21	2013	September	1000	£ 10.00	£ 1.80	£ 500.00	£10,000.00	£ 7,700.00	110000	112727	0.91%	0.95%
22	2013	October	1072	£ 10.00	£ 1.80	£ 500.00	£10,720.00	£ 8,290.40	113000	113000	0.95%	0.95%
23	2013	November	1070	£ 10.00	£ 1.80	£ 500.00	£10,700.00	£ 8,274.00	113273	113273	0.94%	0.94%
24	2013	December	1068	£ 10.00	£ 1.80	£ 500.00	£10,680.00	£ 8,257.60	113545	113545	0.94%	0.94%

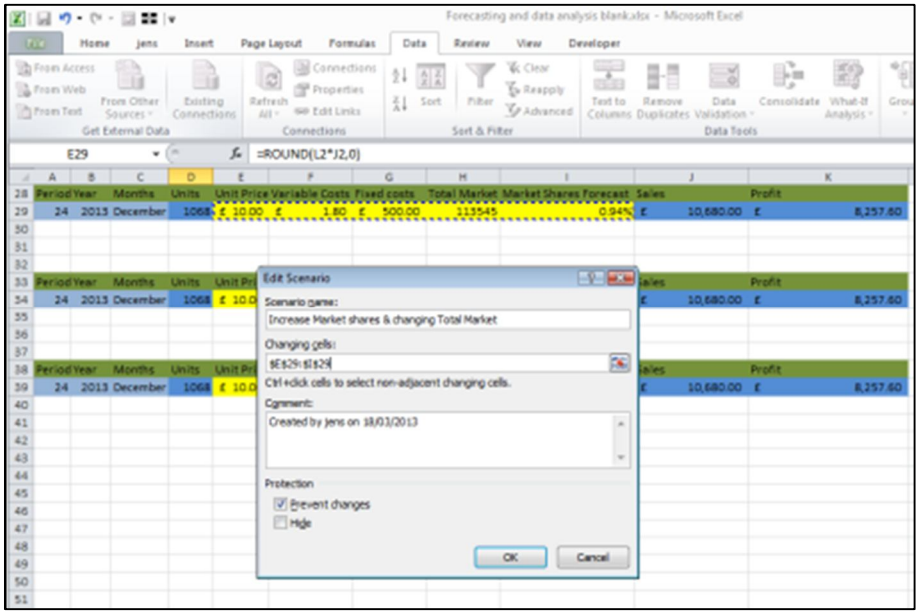
Period	Year	Months	Units	Unit Price	Variable Costs	Fixed costs	Total Market	Market Shares Forecast	Sales	Profit
24	2013	December	1068	£ 10.00	£ 1.80	£ 500.00	113545	0.94%	£ 10,680.00	£ 8,257.60

Period	Year	Months	Units	Unit Price	Variable Costs	Fixed costs	Total Market	Market Shares Forecast	Sales	Profit
24	2013	December	1068	£ 10.00	£ 1.80	£ 500.00	113545	0.94%	£ 10,680.00	£ 8,257.60

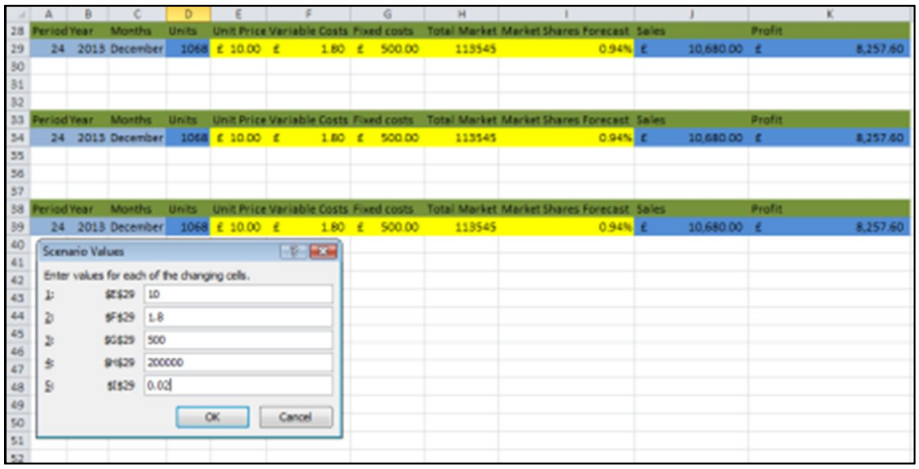
Period	Year	Months	Units	Unit Price	Variable Costs	Fixed costs	Total Market	Market Shares Forecast	Sales	Profit
24	2013	December	1068	£ 10.00	£ 1.80	£ 500.00	113545	0.94%	£ 10,680.00	£ 8,257.60

Open up the scenarios tool by clicking the **Data** tab and click the **What-if Analysis** button in the **Data Tools** group. Click **Scenarios Manager** and the **Edit Scenario** dialog box will open. You first want to see how the forecast will change if the total market will increase because of Christmas and if your company's investment will pay off by giving you more shares of the market.

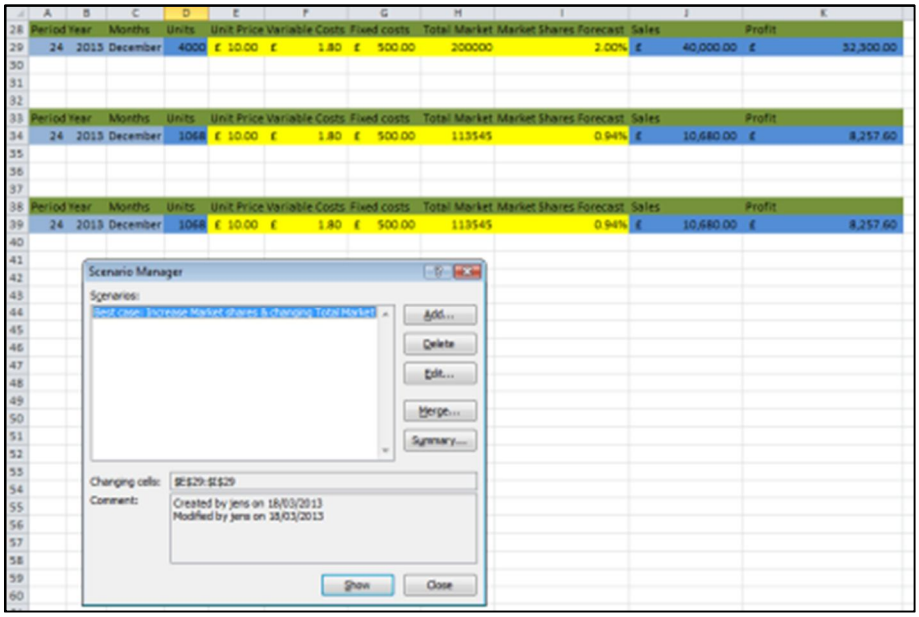
In the **Edit Scenario** dialog box first enter a name for the scenario. In the **Changing cells:** box you must select the changeable cells. Click **OK.**



Based on Christmas forecast from previous years your best guess for Total Market for December will be 200.000 units. Your hope that the investment will increase the Market share to 2%. Enter the values in the **Scenario Values** dialog box and click **OK.**



If you click **Show** you can see the result of your scenario in the worksheet. If the increase of market shares and total market units will be as predicted the sales increase from £ 10,680.00 to £ 40.000,00. To add more scenarios click add and just go through the steps again.



Procedures

Creating scenarios

1. Click the **Data** tab.
2. Click the **What-If Analysis** button in the **Data Tools** group.
3. Click the **Scenario Manager** command.

4. Click the **Add**… button.
5. In the **Scenario name:** box, type a name that identifies the scenario.
6. In the **Changing cells:** box, select or type the cells you want to modify in the scenario.
7. In the **Comments:** box, type an explanatory note regarding that particular scenario.
8. Click the **OK** button.
9. In the **Scenario Values** box, type values for each of the displayed changing cells.
10. Click the **OK** button.
11. The **Scenario Manager** box now displays the scenario.
12. . Click the **Close** button.

Showing a scenario

1. To display a scenario, first click the **Data tab**.
2. Click the **What-If Analysis** button.
3. Click the **Scenario Manager**… button.
4. Select the scenario you want to display.
5. Click the **Show** button.
6. The chosen scenario data is now displayed.
7. Click the **Close** button.

Deleting a scenario

1. To delete a scenario, first click the **Data** tab.
2. Click the **What-If Analysis** button.
3. Click the **Scenario Manager**… button.
4. Select the scenario you want to delete.
5. Click the **Delete** button.

6. The chosen scenario data is deleted.
7. Click the Close button.

Edit values in a scenario

1. To edit values in a scenario, first click the **Data** tab.
2. In the **Data Tools** group, click the **What-If analysis** arrow.
3. Click the **Scenario Manager**… button.
4. Select the name of the scenario you want to edit in the Scenarios list box.
5. Click the **Edit**… button.
6. In the Edit Scenario dialog box, reselect the Changing cells that you want to edit, if necessary.
7. Click the **OK** button.
8. In the **Scenario Values** dialog box, type a new value for the cell you want to change.
9. Click the **OK** button.
10. To display the effect, click the **Show** button.
11. Click the **Close** button.
12. The effect for that scenario is now displayed.

Creating a scenario summary report

1. To create a scenario summary report, first click the **Data** tab.
2. Click the **What-If Analysis** button.
3. Click the **Scenario Manager**… button.
4. In the Scenario Manager dialog box, click the scenario you want to view.
5. Click the **Summary**… button.
6. In the **Scenario Summary** dialog box, click the **OK** button.

7. A Scenario Summary worksheet is displayed with the report.

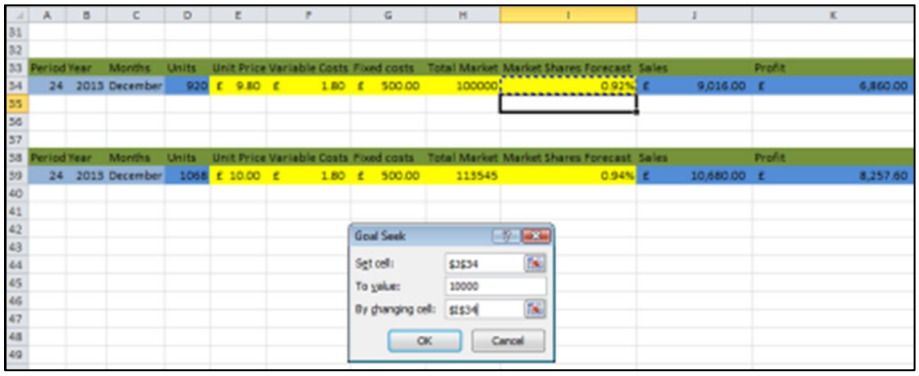
8. The effect for that scenario is now displayed.

THE GOAL SEEK TOOL

Discussion

The **Goal Seek** tool can be very useful in a forecasting situation. Your forecast predicts that you in December will reach sales of £ 9,000.00 but you really need the sales to be £ 10,000.00.

The **Goal Seek** tool can now tell you what to do to get what you want. Open up the **Goal Seek** tool by clicking the **Data** tab and click the **What-if Analysis** button in the **Data Tools** group. Click **Goal Seek** and the **Goal Seek** dialog box will open. Click in the **Set Cell:** box and select the cell with the value you want to change. Click in the **To Value:** box and type the value you want the **Set cell:** cell to have. Click in **By Changing Cell:** box and click the cell you want Excel to change to get the result you want in the **Set Cell:** cell.



Click **OK** and you will get the result.

Procedures

1. To create a scenario summary report, first click the **Data** tab.

2. Click the **What-If Analysis** button.

3. Click the **Scenario Manager**… button.

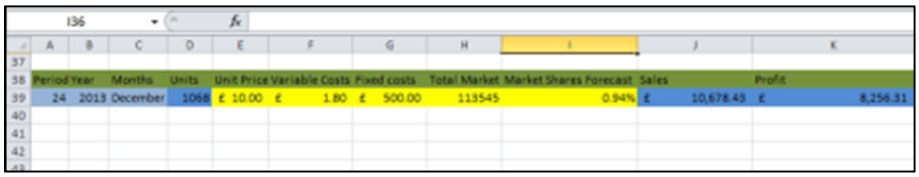
4. In the Scenario Manager dialog box, click the scenario you want to view.
5. Click the **Summary**… button.
6. In the **Scenario Summary** dialog box, click the **OK** button.
7. A Scenario Summary worksheet is displayed with the report.
8. The effect for that scenario is now displayed.

THE DATA TABLE TOOL

Discussion

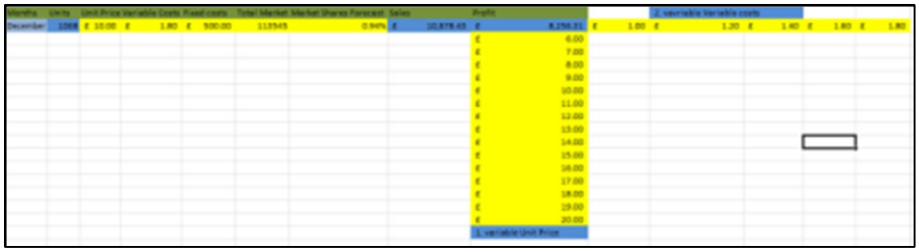
The **Data Table** tool is another **What-if Analysis** tool which is working very well with forecasts. A lot of forecasters must have asked themselves what if this or this happens. What will happen to my forecast? The **Data Table** tool can work with one or two variables. The **Data Table** tool can return a lot of results very fast.

You have a forecast for December but you want to see what will happen to the forecast if you change the price for the product and reduce the costs.

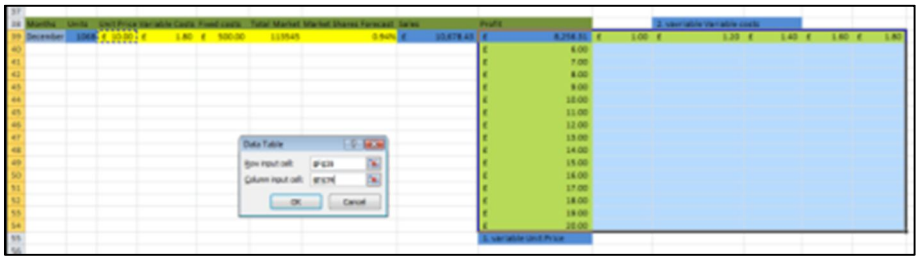


Period Year	Months	Units	Unit Price	Variable Costs	Fixed costs	Total Market	Market Shares Forecast	Sales	Profit
24	2013 December	1068	£ 10.00	£ 1.80	£ 500.00	113545	0.94%	£ 10,678.43	£ 8,256.31

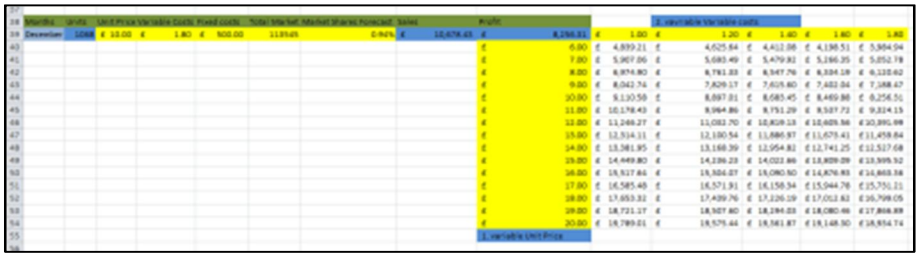
If you want to work with two variables you must have the variables in the same column and row as the formula.



Select the formula, the variable rows and columns and click **Data** tab, click **What-If Analysis** in the **Data Tools** group. Click **Data Table** and in the **Data Table** dialog box **Row input cell:** box enter the cell reference for the cell the formula use to calculate the variable in the row of the selection. In the **Column input cell:** box the **Data Table** tool need the same for the column input.



Click **OK.**



Procedures

One-variable

1. To create a one-variable (column) data table, first select the cell immediately above the output column.
2. Enter the formula you want to use.
3. Press the [Enter] key to confirm the entry.
4. Select the range containing the formula, the input column, and the output column.
5. Click the Data tab.
6. In the Data Tools group, click the What-If Analysis button.
7. Click the Data Table… button.
8. In the Data Table dialog box, click in the Column input cell: text box.
9. Enter, or select, the input cell.
10. Click the OK button.
11. The results are displayed in your one column data table.

Two-variable

1. To create a two-variable data table, first select the cell above the column input values and to the left of the row input values.
2. Enter the formula you want to use.
3. Select the complete data table.
4. Click the Data tab.
5. In the Data Tools group, click the What-If Analysis button.
6. Click the Data Table… button.
7. In the Data Table dialog box, click in the Row input cell: text box.
8. Enter, or select, the input cell.
9. In the Data Table dialog box, click in the Column input cell: text box.
10. Enter, or select, the input cell.
11. Enter, or select, the input cell.
12. The results are displayed in your two-variable data table.

Exercises

Exercise 1 Forecasting using scenarios

You have the data for each quarter from 2007 to 2011. You have the units you have sold, the unit price, the variable costs, the fixed costs, the sales, the profit, the total market units, and your company's market shares. The **Forecast** function has been used to calculate the forecast for Market shares and a seasonal forecast has been used to forecast Total Market (units) for 2012.

What will happen with the forecast for quarter 4 2012 if you can increase your market shares with 2%, decrease fixed and variable costs and if the total market units will increase?

1. Open the file **Forecasting and data analysis**.
2. Open the **What-If analysis** sheet.
3. Open up the **Scenarios** tool's parameters dialog (**Data tab**, **Data Tools** group, **What-If Analysis**).
4. Click **Add** and call the scenario "Best case forecast".
5. In **Changing cells:** type **F32:I32** or click in the box and select the range **F32:I32**.
6. In **Comment:** type "Decrease variable and fixed costs. Increase market shares and total market".
7. Click **OK** button.
8. Change **$F$32** (variable costs) to 1.6, change **$G$32** (fixed costs) to 16000, change **$H$32** (total market units forecast) to 180000, and **$I$32** (market shares forecast) to 0.22.
9. Click **OK** button.
10. Click **Add** and call the scenario "Worst case forecast" do not change **Changing cells:**.
11. In **Comment:** type "Decrease market shares and total market units".
12. Click **OK** button.
13. Change **$H$32** (total market units forecast) to 140000, and **$I$32** (market shares forecast) to 0.18.
14. Click **OK** button.
15. Select **Best case forecast** and click **Show** button. Examine the result.
16. Select **Worst case forecast** and click **Show** button. Examine the result.

Exercise 2 Forecasting using goal seek

You have the data for each quarter from 2007 to 2011. You have the units you have sold, the unit price, the variable costs, the fixed costs, the sales, the profit, the total market units, and your company's market shares. The **Forecast** function has been used to calculate the forecast for Market shares and a seasonal forecast has been used to forecast Total Market (units) for 2012.

In the best case scenario from the previous exercise you reached a profit of £ 376,040.00, but you will like to find out how you can get a profit of £ 400,000.00.

1. Open the file **Forecasting and data analysis**.
2. Open the **What-If analysis** sheet.
3. Open up the **Goal Seek** tool's parameters dialog (**Data tab**, **Data Tools** group, **What-If Analysis**).
4. In **Set cell:** type **$K$37** or click in the box and select **K37**.
5. In **To value:** type 400000.
6. In **By changing cell:** type **$I$37** or click in the box and select **I37**.
7. Click **OK** button and examine the result (you need market shares of 26.67% to get £400,000.00 in profit).
8. Click **Cancel** button to reset the values and to close down the **Goal Seek** dialog box.
9. Go through step 3 to 8 again but this time please find out how much you need to increase the unit price to get £ 400,000.00 in profit.
10. Go through step 3 to 8 again but this time please find out how much you need to decrease the variable costs to get £ 400,000.00 in profit. You will find out that it is not possible to get a profit of £ 400,000.00 by only reducing the variable costs.
11. Save the file.

Exercise 3 Forecasting using data tables

You have the data for each quarter from 2007 to 2011. You have the units you have sold, the unit price, the variable costs, the fixed costs, the sales, the profit, the total market units, and your company's market shares. The **Forecast** function has been used to calculate the forecast for Market shares and a seasonal forecast has been used to forecast Total Market (units) for 2012.

You want to find out what will happen to the forecast if you can bring down the variable and fixed cost.

1. Open the file **Forecasting and data analysis**.
2. Open the **What-If analysis** sheet.
3. For the first variable **fixed costs** use column **K**. Select **K43** and type 10000. Type 12000 in **K44,** 14000 in **K45**, 16000 in **K46**, 18000 in **K47**, 20000 in **K48,** and 22000 in **K49.**
4. For the second variable **Variable costs** use row **42**. Select **L42** and type 1. Type 1.2 in **M42,** 1.4 in **N42**, 1.6 in **O42**, 1.8 in **P42**, 2 in **Q42,** and 2,2 in **R42.**
5. Select the range **K42:R49.**
6. Open up the **Data Table** tool's parameters dialog (**Data tab**, **Data Tools** group, **What-If Analysis**).
7. In **Row input cell:** type **F42** or click in the box and select the range **F42**. The row variable is **Variable Costs.**
8. In **Column input cell:** type **G42** or click in the box and select the range **G42**. The column variable is **Fixed Costs.**
9. Click **OK** button.
10. Please do step 3 to 9 again. This time use the profit formula in **K55** and **Unit Price** (use the range 8 to 20 pounds increased by 2 each time) in column **K** and **Market Shares Forecast** (use the range 16% to 30% increased by 2 each time) in row **55** as the two variables.
11. Compare the two tables. To increase the profit is it best to increase the costs or to try to get a bigger part of the market shares?
12. Save the file

LESSON 7 - CORRELATION COEFFICIENT

In this lesson, you will learn how to:

- Use the **Correl** function
- Use the Data Analysis Tool **Correlation**
- Create a scatter chart to visualise Correlation Coefficient

CONCEPTS AND TERMS

Discussion

The **Correl** function determines the relationship (the correlation coefficient) between two arrays of data.

You may need to find out if your company's increased advertising costs also increase your company's sales figures or if higher temperatures in summer increase sales of ice cream. Or if you increase the price for your products will it affect the quantity you are selling.

The **Correl** function will return a result between -1 and 1.

Between 1 and 0:

If the result is close to 1 then it is because there is a close positive relationship between the two arrays (If the temperature rises the ice cream sales increase). Close to 0 then there is not a close relationship between the two arrays (the sales do not increase because of increased advertising costs).

Between 0 and -1:

If the result is close to -1 then there is a close negative relationship between the two arrays (if the price goes up the sales goes down). Close to 0 but a negative value then there is not a close relationship between the two arrays (the sales does not decrease if the price goes up).

THE CORREL FUNCTION

Discussion

The **Correl** function consists of two required arguments, in the following order: **Array1**, **Array2**. **Array1** and **Array2** are the arrays of data you want to examine to see if there is any relationship. In the **Array1** box enter the first range of data. In the **Array2** box enter the second range of data.

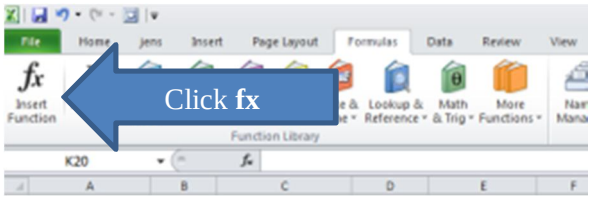
Array1 and array2 must have the same number of data points. If not the **Correl** function returns the #N/A error value.

If either array1 or array2 is empty the **Correl** function returns the #DIV/0! Error.

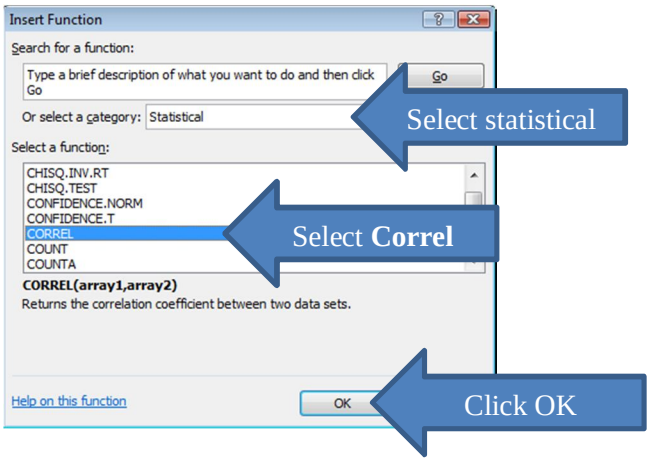
Month	Advertising costs	Sales
January	20000	100000
February	30000	80000
March	20000	100000
April	10000	150000
May	20000	150000
June	20000	300000
July	30000	200000
August	20000	100000
September	10000	100000
October	20000	100000
November	10000	200000
December	30000	300000

Procedures

1. To use a **Correl** function, first create a data range with two columns of values.
2. Click in the cell where you want to place the function.
3. Click on the **Formulas** tab.
4. In the **Function Library** group, click on the **Insert Function** button.



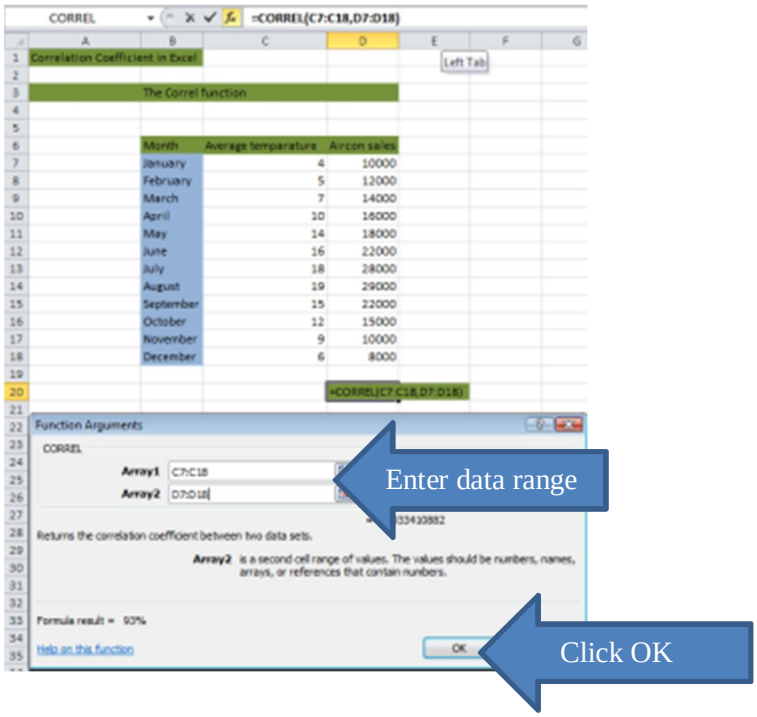
5. In the Insert function dialog box, locate **Statistical** category in the **Or select a category:** box.
6. Click on **Correl**.
7. Click on the **OK** button.



8. In the **Function Arguments** dialog box, click in the **Array1** box.
9. Select the cells containing the value from the first column in the data range.

10. In the **Array2** box, select the cells containing the value from the second column in the data range.

11. Click on the OK button.



*The **COVARIANCE.P** function is another useful function to compare data from two different arrays. The **COVARIANCE.P** function has two arguments **Array1** and **Array2**. Enter the ranges for the two data sets and the function will return a positive or negative number. If it is positive close to zero there is a close positive relationship between the 2 arrays. Negative and close to zero there is a close negative relationship. If the value increases the two arrays are less related.*

THE DATA ANALYSIS TOOL CORRELATION

Discussion

The Data Analysis Tool **Correlation** is working the same way as the **Correl** function.

The difference between the two tools is with the **Correl** function you can find the relationship between two arrays of data, with the Data Analysis Tool **Correlation** you can find the relationship between any numbers of arrays.

As the **Correl** function the Data Analysis Tool **Correlation** will return a result between -1 and 1.

Between 1 and 0:

If the result is close to 1 then it is because there is a close positive relationship between the two arrays (If the temperature rise the ice cream sales increase). Close to 0 then there is not a close relationship between the two arrays (the sales do not raise because of increased advertising costs).

Between 0 and -1:

If the result is close to -1 then there is a close negative relationship between the two arrays (if the price goes up the sales goes down). Close to 0 but a negative value then there is not a close relationship between the two arrays (the sales does not decrease if the price goes up).

Using the Correlation tool

You can use **Correlation** tool by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Correlation** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Correlation** dialog box, you can get the result in the output range.

You can specify the following items in the **Correlation** parameters dialog box:

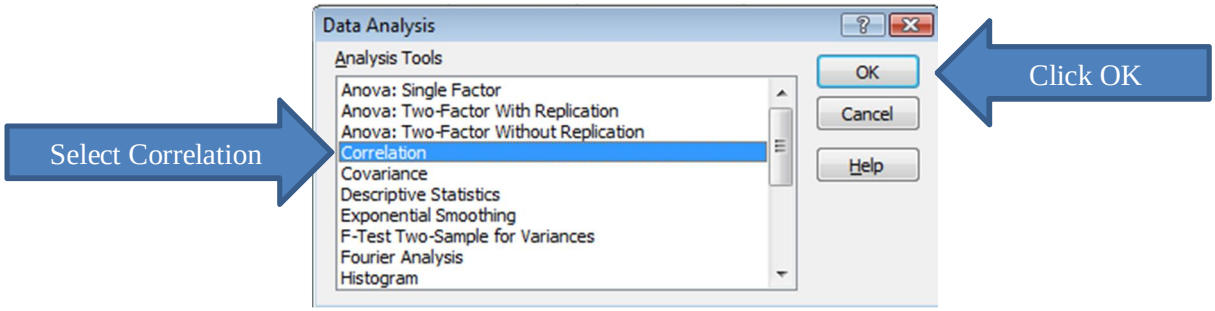
Parameter	Description
Input	The array where you have the data you want to examine
Grouped By	input data grouped by columns or rows
Labels	If labels are selected in the input range the box must be ticked
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook

When you have finished entering the information click OK and you will get the output.

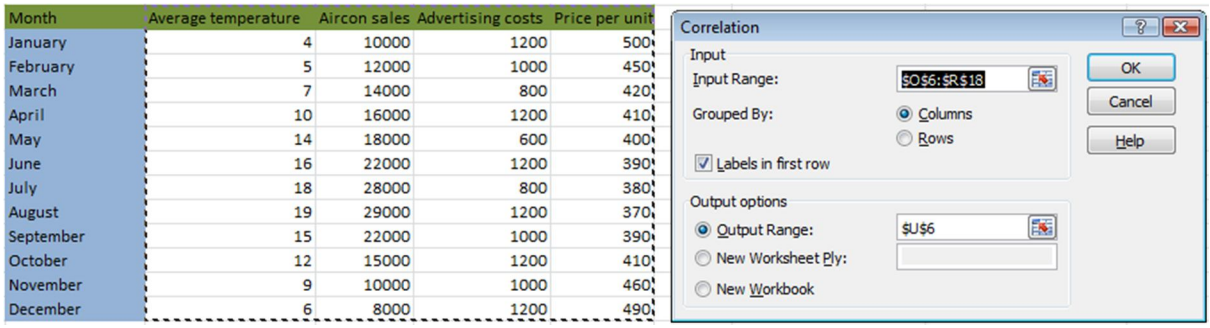
Procedures

1. To use the **Correlation** tool, first click the Data tab.
2. In the **Analysis** group, click the **Data Analysis** button.

3. Select **Correlation** from the list and click OK.



4. Select the Input Range.



Month	Average temperature	Aircon sales	Advertising costs	Price per unit
January	4	10000	1200	500
February	5	12000	1000	450
March	7	14000	800	420
April	10	16000	1200	410
May	14	18000	600	400
June	16	22000	1200	390
July	18	28000	800	380
August	19	29000	1200	370
September	15	22000	1000	390
October	12	15000	1200	410
November	9	10000	1000	460
December	6	8000	1200	490

5. Select the desired options for **Grouped By**, **Labels**, and **Output**.

6. Click OK

The output

Month	Average temperature	Aircon sales	Advertising costs	Price per unit
January	4	10000	1200	500
February	5	12000	1000	450
March	7	14000	800	420
April	10	16000	1200	410
May	14	18000	600	400
June	16	22000	1200	390
July	18	28000	800	380
August	19	29000	1200	370
September	15	22000	1000	390
October	12	15000	1200	410
November	9	10000	1000	460
December	6	8000	1200	490

	Average temperature	*Aircon sales*	*Advertising costs*	*Price per unit*
Average temperature	1			
Aircon sales	0.933410882	1		
Advertising costs	-0.145073235	-0.1520328	1	
Price per unit	-0.896696678	-0.8976502	0.257331145	1

Let us have a look at the output. You got a table with the column labels from the input array, because you selected the column labels when you selected the Input Range. In the output table you have the labels both as row and columns labels.

You can see that the relationship between Aircon sales and Average temperature is close, because the result is close to 1. It means that if the average temperature rises you are selling more air-condition units.

If you look at the relationship between Aircon sales and Advertising costs you got a negative result close to 0. It means that it is not a good idea spending money on advertising. The sales of air-condition units do not increase if you spend more money on advertising. In this example the sales actually decrease a little bit if you increase the advertising budget.

You got a result close to minus 1 for the relationship between Aircon sales and Price per unit. It means that you have a close negative relationship between the two columns. If the price goes up the sales go down. If the price goes down the sales go up.

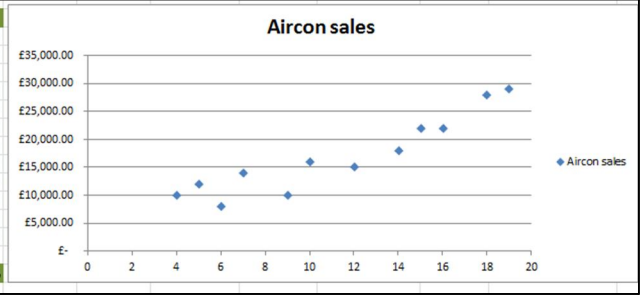
CREATE A SCATTER CHART TO DISPLAY CORRELATION COEFFICIENT

Discussion

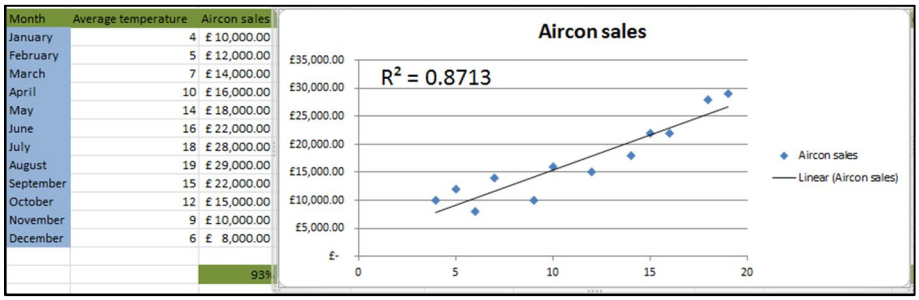
Charts are very useful to visualise data analysis. The **Scatter** chart is the most beneficial when working with data analysis.

In the example below it is easy to spot the trend in the chart. There is a close relationship between average temperature and aircon sales. Of course you can also see that the **Correl** function returns 93% wich also emphasise that there is a close relationship between the two arrays.

Month	Average temperature	Aircon sales
January	4	£ 10,000.00
February	5	£ 12,000.00
March	7	£ 14,000.00
April	10	£ 16,000.00
May	14	£ 18,000.00
June	16	£ 22,000.00
July	18	£ 28,000.00
August	19	£ 29,000.00
September	15	£ 22,000.00
October	12	£ 15,000.00
November	9	£ 10,000.00
December	6	£ 8,000.00
		93%

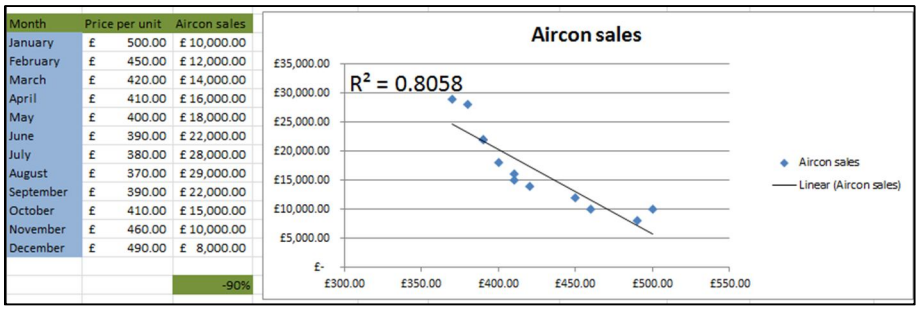


To make it even easier to spot the trend you can add a linear **Trendline** to the chart and if you add the R^2 (R square error) to the **Trendline** you can see how accurate your **Trendline** is. A value close to 1 shows that the **Trendline** is very accurate. Close to 0 there is not really any trend.



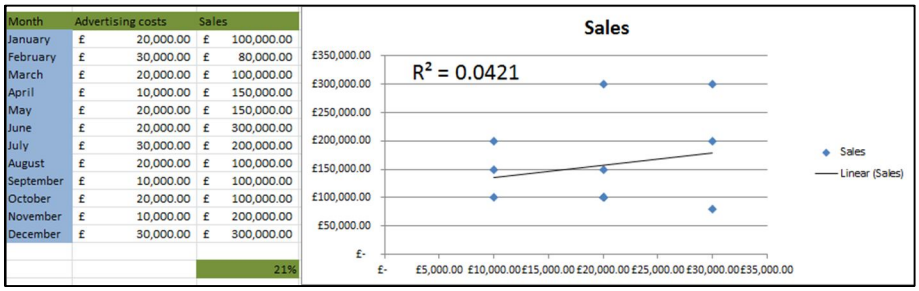
Month	Average temperature	Aircon sales
January	4	£ 10,000.00
February	5	£ 12,000.00
March	7	£ 14,000.00
April	10	£ 16,000.00
May	14	£ 18,000.00
June	16	£ 22,000.00
July	18	£ 28,000.00
August	19	£ 29,000.00
September	15	£ 22,000.00
October	12	£ 15,000.00
November	9	£ 10,000.00
December	6	£ 8,000.00
		93%

In the example below there is a close negative relationship between the Price per unit and Aircon sales. Again it is easy to see in a scatter chart.



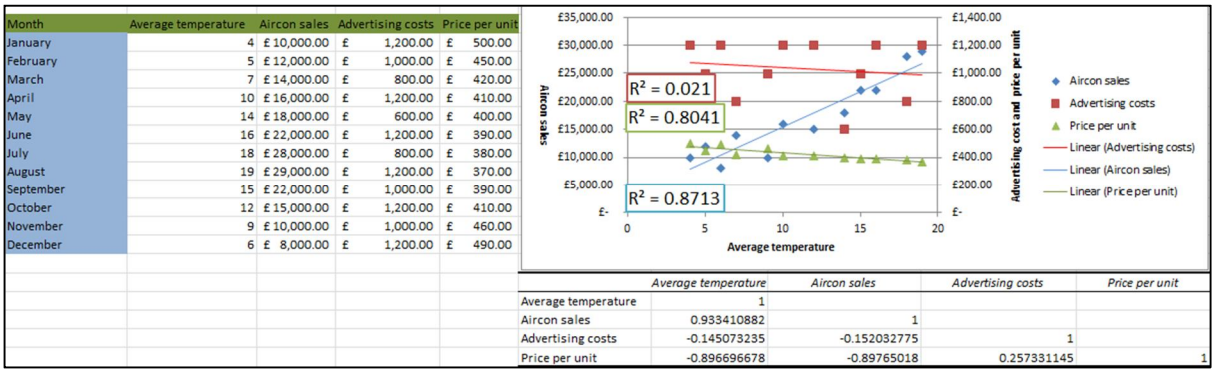
Month	Price per unit	Aircon sales
January	£ 500.00	£ 10,000.00
February	£ 450.00	£ 12,000.00
March	£ 420.00	£ 14,000.00
April	£ 410.00	£ 16,000.00
May	£ 400.00	£ 18,000.00
June	£ 390.00	£ 22,000.00
July	£ 380.00	£ 28,000.00
August	£ 370.00	£ 29,000.00
September	£ 390.00	£ 22,000.00
October	£ 410.00	£ 15,000.00
November	£ 460.00	£ 10,000.00
December	£ 490.00	£ 8,000.00
		-90%

In the example below there is no close relationship between Advertising costs and Sales. The **Correl** function returns 21% and the R^2 value is 0.0421 which tells you that there is not really a trend in the data.



Month	Advertising costs	Sales
January	£ 20,000.00	£ 100,000.00
February	£ 30,000.00	£ 80,000.00
March	£ 20,000.00	£ 100,000.00
April	£ 10,000.00	£ 150,000.00
May	£ 20,000.00	£ 150,000.00
June	£ 20,000.00	£ 300,000.00
July	£ 30,000.00	£ 200,000.00
August	£ 20,000.00	£ 100,000.00
September	£ 10,000.00	£ 100,000.00
October	£ 20,000.00	£ 100,000.00
November	£ 10,000.00	£ 200,000.00
December	£ 30,000.00	£ 300,000.00
		21%

In the example below the **Data Analysis** tool **Correlation** has returned the correlation for three arrays and again a chart is very useful to examine the outcome.



Month	Average temperature	Aircon sales	Advertising costs	Price per unit
January	4	£ 10,000.00	£ 1,200.00	£ 500.00
February	5	£ 12,000.00	£ 1,000.00	£ 450.00
March	7	£ 14,000.00	£ 800.00	£ 420.00
April	10	£ 16,000.00	£ 1,200.00	£ 410.00
May	14	£ 18,000.00	£ 600.00	£ 400.00
June	16	£ 22,000.00	£ 1,200.00	£ 390.00
July	18	£ 28,000.00	£ 800.00	£ 380.00
August	19	£ 29,000.00	£ 1,200.00	£ 370.00
September	15	£ 22,000.00	£ 1,000.00	£ 390.00
October	12	£ 15,000.00	£ 1,200.00	£ 410.00
November	9	£ 10,000.00	£ 1,000.00	£ 460.00
December	6	£ 8,000.00	£ 1,200.00	£ 490.00

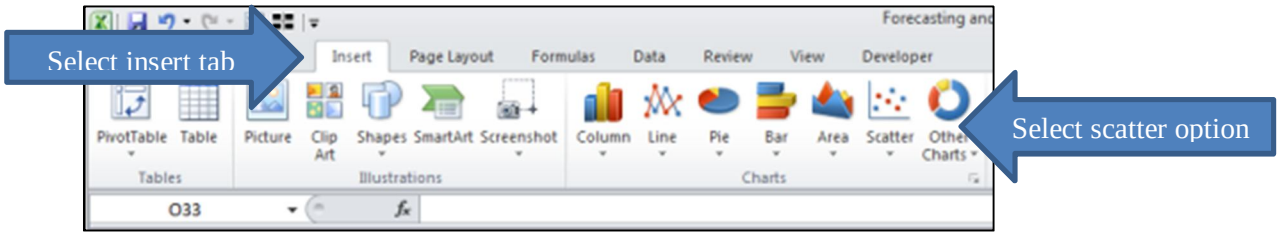
	Average temperature	*Aircon sales*	*Advertising costs*	*Price per unit*
Average temperature	1			
Aircon sales	0.933410882	1		
Advertising costs	-0.145073235	-0.152032775	1	
Price per unit	-0.896696678	-0.89765018	0.257331145	1

Procedures

1. To create a chart to visualise correlation select the arrays with the data you want to examine.

2. Click **Insert** tab.

3. Click the scatter chart option you want to use.

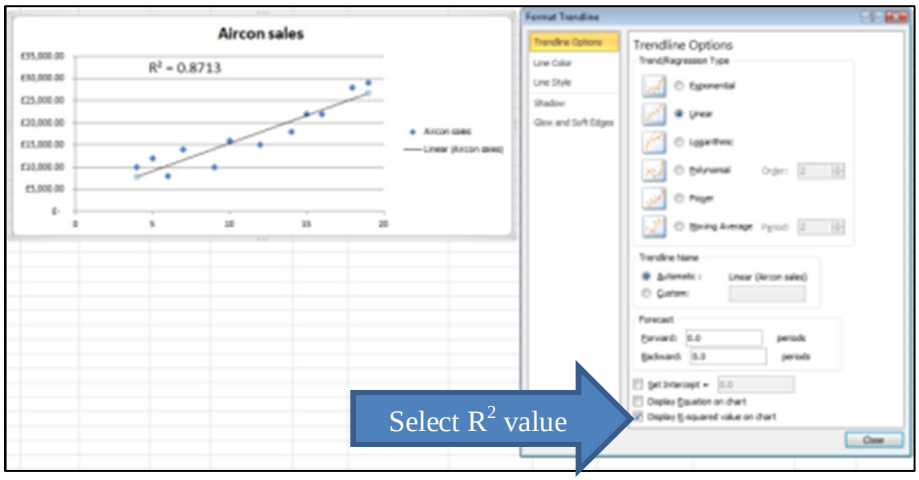


4. Click **Layout** tab (the chart most be selected to have access to the **Layout** tab).

5. In the **Analysis** group click **Trendline** and **Linear Trendline**.

6. Right click the **Trendline** and click **Format Trendline**.

7. Select the option **Display R Squared Value on Chart**



8. Click **Close**.

EXERCISE

Using Correl function and the data analysis tool Correlation

1. Open the file **Forecasting and data analysis**.
2. Open the **Correlation Coefficient** sheet.
3. Use the **Correl** function in D60 to get the correlation coefficient for the two arrays C47:C58 and D47:D58 (Average temperature and Ice sales).
4. The result should be 0.9566537. What does this tell us about the relationship between the two arrays?
5. Create a scatter chart showing the Correlation Coefficient between the average temperature and the ice sales.
6. Click the **Data** tab. Click **Data analysis** in the Analysis group. Open the **Correlation** tool.
7. Use the Data Analysis Tool **Correlation** to find the Correlation Coefficient between the average temperature, the ice sales, and advertising costs (the input should be O46:Q58).
8. The result should be:

	Average temperature	*Ice sales*	*Advertising costs*
Average temperature	1		
Ice sales	0.956653704	1	
Advertising costs	-0.204621004	-0.196872578	1

9. What does this tell us about the relationships between the two arrays?
10. Create a scatter chart showing the Correlation Coefficient between the average temperatures, the ice sales, and advertising costs.
11. Save and close the file

LESSON 8 – BREAK-EVEN ANALYSIS

In this lesson, you will learn how to:

- Calculate Break-even
- Visualise break-even using scatter chart
- Use the **goal seek** tool to calculate break-even
- Use the **Solver** tool to calculate break-even
- Exercise

CONCEPTS AND TERMS

Discussion

When you are considering an investment or launching a new product you want to know at what point you will break-even. When will you start earning money? How many units do I need to sell to Break-even if you invest £ 100.000 on television commercials? How should the products be prised to break-even this year?

For forecasting break-even analysis is very important to forecast new product's financial success.

In this workbook you will see the math behind break-even and the break-even will be manually calculated. The Goal Seek tool in Excel can do these calculations and is a very valuable tool for break-even analysis.

CALCULATE BREAK-EVEN

The formula for break-even looks like this:

0 (profit) =Units *(Sales Price – Variable Costs) – Fixed Costs

Example:

- 100 Units
- Sales Price = £10
- Variable Costs = £4
- Fixed Costs = £600

With above values the formula will be true a profit of 0. You need to sell 100 units for a price of £10 if the variable costs are £4 and fixed costs are £600

If you want to find out how many units you need to sell to break-even you must change the formula using the math rules for equations.

0 (profit) =Units *(Sales Price – Variable Costs) – Fixed Costs

First add Fixed Costs on both sides of the equal sign:

Fixed Costs=Units *(Sales Price – Variable Costs)

This equation can now be used to calculate the Fixed Costs to break-even.

By using the values from the example: 600=100*(10-4)

Divide with (Sales Price – Variable Costs) on both sides of the equal sign:

Units = Fixed Costs / (Sales Price – Variable Costs)

This equation can calculate how many units you need to break-even.

By using the values from the example: 100=600 / (10-4)

The equation to find the sales price is:

Sales Price = Fixed Costs / Units + Variable Costs

To get this equation start from the break-even equation again:

0 (profit) =Units *(Sales Price – Variable Costs) – Fixed Costs

First add Fixed Costs and both sides of the equal sign:

Fixed Costs=Units *(Sales Price – Variable Costs)

Then divide Units on both sides of the equal sign:

Fixed Costs / Units=Sales Price – Variable Costs

Then add Variable Costs on both sides of the equal sign:

Sales Price = Fixed Costs / Units + Variable Costs

This equation can calculate sales price you need to break-even.

By using the values from the example: 10=600 / 100 + 4

Procedures (calculate units to break-even)

Units = Fixed Costs / (Sales Price – Variable Costs)

- 100 Units
- Sales Price = £10
- Variable Costs = £4
- Fixed Costs = £600

1. Type =600 / (10-4).

2. Press enter.

Procedures (calculate sales price to break-even)

Sales Price = Fixed Costs / Units + Variable Costs

- 100 Units
- Sales Price = £10
- Variable Costs = £4
- Fixed Costs = £600

1. Type =600 / 100+4
2. Press enter.

Procedures (calculate fixed costs to break-even)

Fixed Costs=Units *(Sales Price – Variable Costs)

- 100 Units
- Sales Price = £10
- Variable Costs = £4
- Fixed Costs = £600

1. Type =100*(10-4)
2. Press enter.

VISUALISE BREAK-EVEN USING SCATTER CHART

Discussion

Charts can be very useful to understand and visualise data analysis. Scatter charts is the chart type, which should be used to visualise break-even analysis.

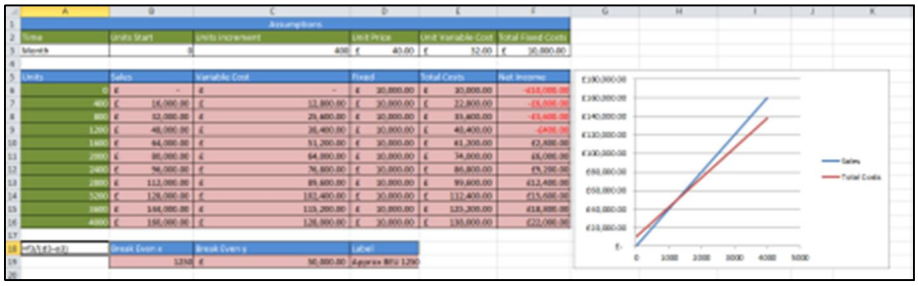
In the example below in column A there are units going from 0 to 4000 with an increment of 400. In column B the sales (units times unit price), in column C variable cost (units times unit variable cost), in column D the fixed cost, in

column E total costs (variable cost + fixed cost), and in column F the net income (sales – total costs).

If you look at the table it is not easy to see the break-even. It is somewhere between 1200 and 1600 units.

	A	B	C	D	E	F
1	Assumptions					
2	Time	Units Start	Units increment	Unit Price	Unit Variable Cost	Total Fixed Costs
3	Month	0	400	£ 40.00	£ 32.00	£ 10,000.00
4						
5	Units	Sales	Variable Cost	Fixed	Total Costs	Net Income
6	0	£ -	£ -	£ 10,000.00	£ 10,000.00	-£10,000.00
7	400	£ 16,000.00	£ 12,800.00	£ 10,000.00	£ 22,800.00	-£6,800.00
8	800	£ 32,000.00	£ 25,600.00	£ 10,000.00	£ 35,600.00	-£3,600.00
9	1200	£ 48,000.00	£ 38,400.00	£ 10,000.00	£ 48,400.00	-£400.00
10	1600	£ 64,000.00	£ 51,200.00	£ 10,000.00	£ 61,200.00	£2,800.00
11	2000	£ 80,000.00	£ 64,000.00	£ 10,000.00	£ 74,000.00	£6,000.00
12	2400	£ 96,000.00	£ 76,800.00	£ 10,000.00	£ 86,800.00	£9,200.00
13	2800	£ 112,000.00	£ 89,600.00	£ 10,000.00	£ 99,600.00	£12,400.00
14	3200	£ 128,000.00	£ 102,400.00	£ 10,000.00	£ 112,400.00	£15,600.00
15	3600	£ 144,000.00	£ 115,200.00	£ 10,000.00	£ 125,200.00	£18,800.00
16	4000	£ 160,000.00	£ 128,000.00	£ 10,000.00	£ 138,000.00	£22,000.00
17						
18	=f3/(d3-e3)	Break Even x	Break Even y	Label		
19		1250	£ 50,000.00	Approx BEU 1250		

To visualise break-even in a scatter chart select the units (A5:A16), the sales (B5:B16), and total costs (E5:E16), and create the straight line scatter chart.



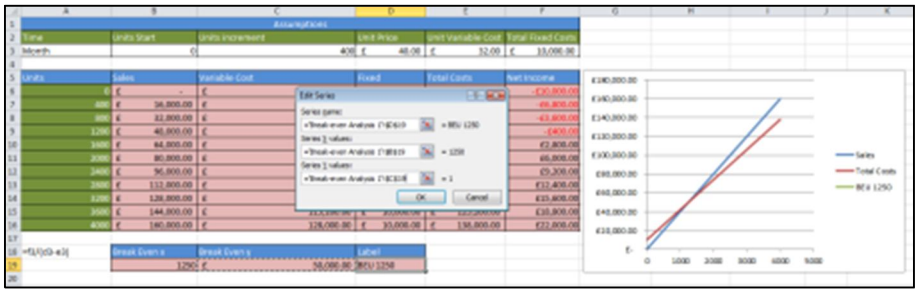
The break-even point is where the sales line cross the total costs line.

To make it easier to see the break-even in the scatter chart you can calculate the break-even using the break-even equation you have seen earlier in this chapter;

Units = Fixed Costs / (Sales Price – Variable Costs)

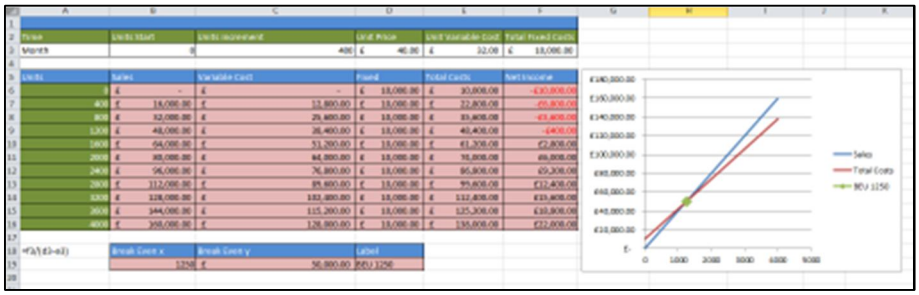
In this example the calculation is; =f3 / (d3-e3) and the result is 1250. You will also need to calculate the Y value to plot it on the chart. The Y value is break-even units multiplied by unit price. In this example; =B19*D3 (the break-even equation is calculated in B19). The result is calculated in C19 and is 50000. To make the exact break-even point very clear in the chart you can add the exact value to the legend label in the chart. In cell D19; ="BEU "&B19 has been entered. This will return “BEU 1250” on the legend label.

To add the break-even point and the legend label you need to select the chart and click the contextual Design tab. Click Select Data in the Data group. Click Add in Series Name click D19 in Series X values click B19 and in Series Y values click C19.

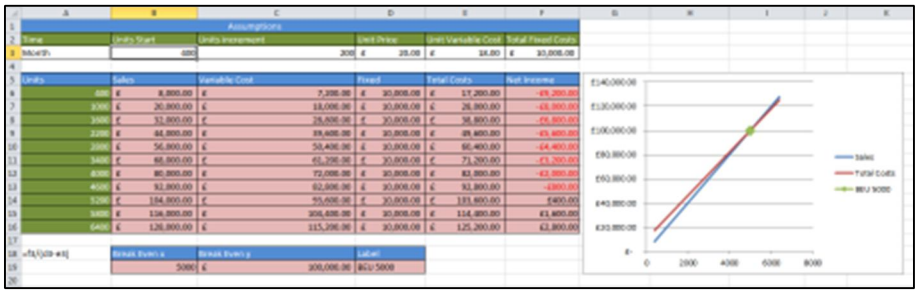


The break-even point is still not visible on the chart. Well the label is visible in the legend. To make it visible you must change the data series' formatting. Select the data series and right click. Choose **Format Data** Series click **Marker Options** select **Build-in** and select a type and a size. Click Close.

Now it is very easy to see the break-even point.



If you have created the table as the one in this example you can change the variables in row three and everything will be updated immediately.

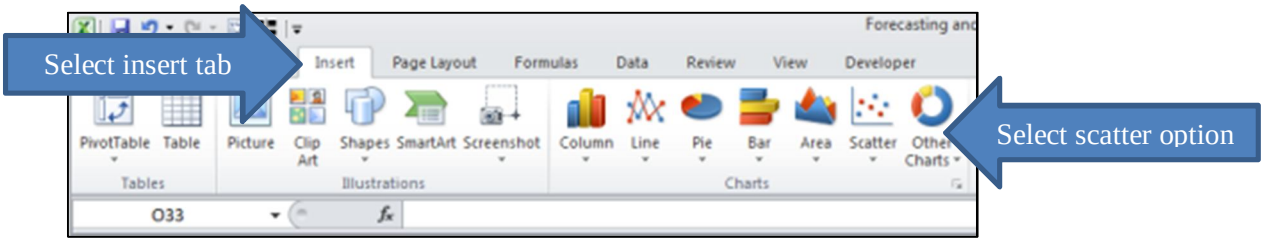


Procedures

1. To create a scatter chart to visualise Break-Even for units sold select the arrays with the data you want to visualise.

2. Click **Insert** tab.

3. Click the scatter chart option you want to use and create the chart.



4. Calculate Break-Even using the break-even equation. Units = Fixed Costs / (Sales Price – Variable Costs)

5. Calculate the Y value for break-even (break-even multiplied by unit price).

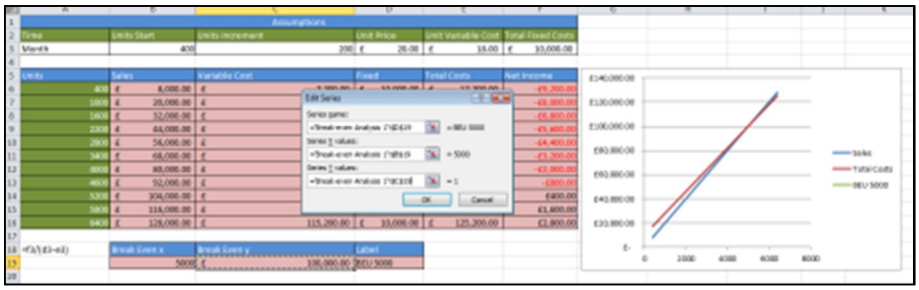
6. Select the chart and click the contextual **Design** tab and click **Select Data** in the **Data** group.

7. Click **Add** and in the **Edit Series** parameters dialog box in the **Series Name** box type a name for the data series.

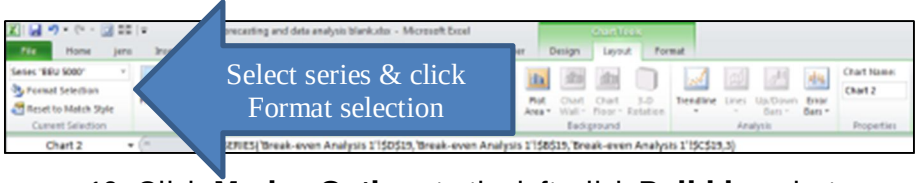
8. In the **Series X values** box enter the cell reference for the cell with the break-even equation / calculation.

9. In the **Series Y values** box enter the cell reference for the cell with Y value calculation.

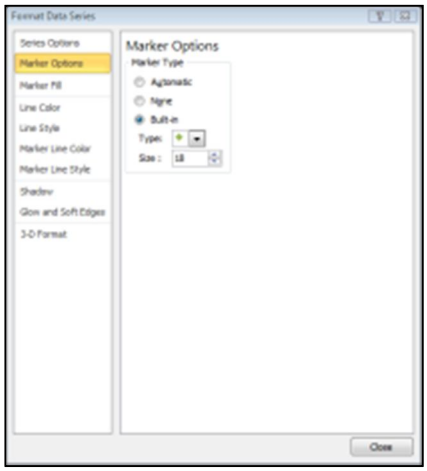
10. Click the **OK** button.



11. To make the break-even data series visible in the chart right click the series and choose **Format Data Series** or click **Layout** tab and select the data series from the chart object list in the **Current Selection** group and then in the same group click **Format Selection.**



12. Click **Marker Options** to the left, click **Build-in,** select a marker type and size. Click the **Close** button.



USE SCENARIOS AND THE GOAL SEEK TOOL TO CALCULATE BREAK-EVEN

Discussion

The **What-if** analysis tools in Excel are very useful also when you are working with Break-Even analysis. How can I change the break-even point? Which parameters will affect my break-even analysis most if I change them?

EXERCISE

Calculating break-even

1. Open the file **Forecasting and data analysis**.

2. Open the **Break-even Analysis 1** sheet.

3. Select C68 and calculate break-even. How many units must be sold to get 0 in profit (use the equation "Units=FixedCosts/(SalesPrice-VariableCosts)").

4. Select F68 and calculate the selling price to break-even selling 10000 units (use the equation "SalesPrice = FixedCosts/Units + VariableCosts").

5. Select H68 and calculate the Fixed Costs to break-even selling 10000 units (use the equation "FixedCosts= Units*(SalesPrice - VariableCosts)").

6. Select K68 and calculate the Variable Costs to break-even selling 10000 units (use the equation "VariableCosts=SalesPrice-FixedCosts/Units").

8. Use the Goal Seek tool to Break-even Units, Sales Price, Fixed Costs, and Variable Cost in the same 4 tables.

7. Save and close the file.

LESSON 9 – DATA ANALYSIS TOOLS

In this lesson, you will learn how to:

- Use Descriptive Statistics
- Use Histogram
- Use Regression
- Use Sampling

CONCEPTS AND TERMS

Discussion

You have already seen some of the tools from the **Data Analysis TOOLPAK** add-in. In this chapter some of the other tools will be briefly demonstrated. All the statically terms will not be explained.

DESCRIPTIVE STATISTICS

Discussion

The **Descriptive Statistics** tool can return statistical information from data in Excel.

Usually when you analyse data it is important examine the distribution and to discover information such as the minimum and maximum values and to determine if there are outliers. This is an important step in any analysis since it helps you understand if your data meet rules required by other analyses such as t-tests and regression and if the data are needed for forecasting.

The **Descriptive Statistics** tool can provide you with these information and much more.

Sales	
Mean	2708.180556
Standard Error	39.70767757
Median	2679
Mode	2600
Standard Deviation	336.9308169
Sample Variance	113522.3754
Kurtosis	0.873835822
Skewness	0.642971537
Range	1551
Minimum	2147
Maximum	3698
Sum	194989
Count	72
Largest(3)	3600
Smallest(3)	2155
Confidence Level(95.0%)	79.17486035

Information you should notice includes:

1. Search for outliers: Look at the Minimum and Maximum values to see if these values fall within your expected range for these data. If a value is unexpectedly small or large, you should examine your data to see if they were wrong. If there are corrections that need to be made, make them before continuing. If you have values that are unexpectedly large or small, but are actual values, it may indicate that your data are not normally distributed. This knowledge will help you to decide forecasting method. It may also indicate that the average is not the best value to report to describe the trend of this data set.

2. Symmetry: Another measure that helps you decide normality is Skewness and Kurtosis. The Skewness measure indicates the level of non-symmetry. If the distribution of the data is symmetric then skewness will be close to 0 (zero). The further from 0, the more skewed the data are.

How do you tell if the skewness is large enough to cause concern? Excel doesn't give you this value, but a measure of the standard error of skewness can be calculated as =SQRT(6 / N) where N is the sample size. If the skewness is more than twice this amount, then it indicates that the distribution of the data is non-symmetric.

3. Kurtosis characterizes the relative peakedness or flatness of a distribution compared with the normal distribution. Positive kurtosis indicates a relatively peaked distribution. Negative kurtosis indicates a relatively flat distribution. Again, for normally distributed data the kurtosis is 0 (zero). As with skewness, if the value of kurtosis is too big or too small, there is concern about the normality of the distribution. In this case, a rough formula for the standard error for kurtosis is =SQRT(24 / N) where N is the sample size. If the kurtosis is more than twice this amount the data may be considered not to meet the criteria for normality by this measure.

4. Evaluation of central trend: For normally distributed data the average is the typical value to measure. However, the average is not interesting without some evaluation of the inconsistency of the data. Therefore you should at least report three values – the average, the standard error of the average, and the sample size.

The median is another measure of central tendency and is usually reported when the data are not normally distributed. The mode, or the most frequent value, is a third measure of central tendency.

5. Measures of inconsistency: The measures of inconsistency reported in Excel's descriptive statistics include the standard error, the standard deviation, the variance and the range.

Standard Deviation = SQRT(Variance)

Standard Error = Standard Deviation / SQRT(N)

You can find The **Descriptive Statistics** tool by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Descriptive Statistics** from the list and click OK.

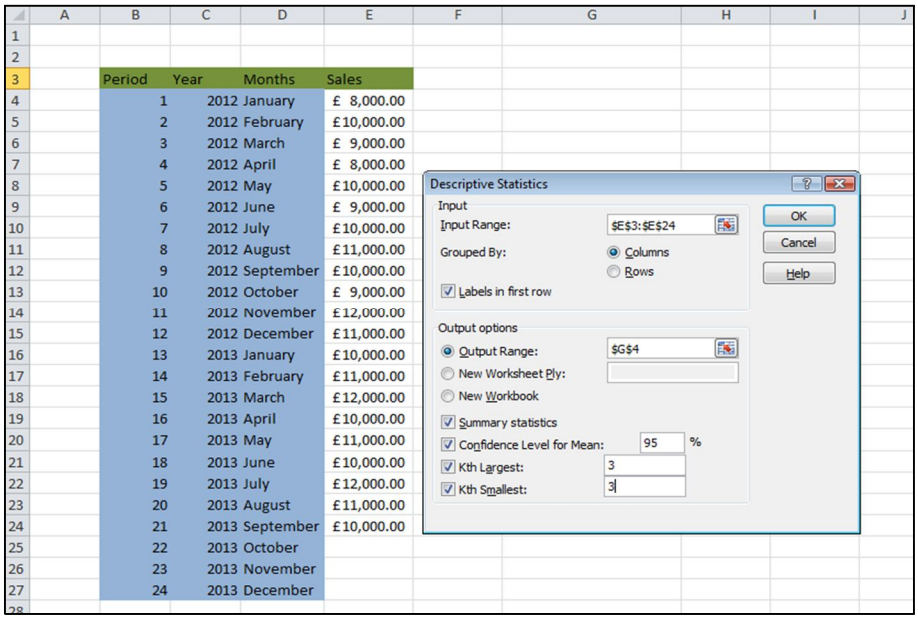
If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Descriptive Statistics** dialog box, you can get the result in the output range.

You can specify the following items in the **Descriptive Statistics** parameters dialog box:

Parameter	Description
Input Range	The range where you have the data you want to examine
Grouped By	Do you have the data grouped by columns or rows
Labels	If labels are selected in the input range the box must be ticked
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook
Summary statistics	Tick the box if you want the output statistics summarised
Confidence Level for Mean	Tick the box if you want the confidence calculated for the average
Kth Largest	Tick the box if you want not only the largest value in the data range shown in the output. By typing 3 in the box the third largest value will be shown in the output
Kth Smallest	Tick the box if you want not only the smallest value in the data range shown in the output. By typing 3 in the box the third smallest value will be shown in the output

When you have finished entering the information click OK and you will get the output.



Period	Year	Months	Sales
1	2012	January	£ 8,000.00
2	2012	February	£10,000.00
3	2012	March	£ 9,000.00
4	2012	April	£ 8,000.00
5	2012	May	£10,000.00
6	2012	June	£ 9,000.00
7	2012	July	£10,000.00
8	2012	August	£11,000.00
9	2012	September	£10,000.00
10	2012	October	£ 9,000.00
11	2012	November	£12,000.00
12	2012	December	£11,000.00
13	2013	January	£10,000.00
14	2013	February	£11,000.00
15	2013	March	£12,000.00
16	2013	April	£10,000.00
17	2013	May	£11,000.00
18	2013	June	£10,000.00
19	2013	July	£12,000.00
20	2013	August	£11,000.00
21	2013	September	£10,000.00
22	2013	October	
23	2013	November	
24	2013	December	

The output:

Period	Year	Months	Sales
1	2012	January	£ 8,000.00
2	2012	February	£10,000.00
3	2012	March	£ 9,000.00
4	2012	April	£ 8,000.00
5	2012	May	£10,000.00
6	2012	June	£ 9,000.00
7	2012	July	£10,000.00
8	2012	August	£11,000.00
9	2012	September	£10,000.00
10	2012	October	£ 9,000.00
11	2012	November	£12,000.00
12	2012	December	£11,000.00
13	2013	January	£10,000.00
14	2013	February	£11,000.00
15	2013	March	£12,000.00
16	2013	April	£10,000.00
17	2013	May	£11,000.00
18	2013	June	£10,000.00
19	2013	July	£12,000.00
20	2013	August	£11,000.00
21	2013	September	£10,000.00
22	2013	October	
23	2013	November	
24	2013	December	

Sales	
Mean	10190.4762
Standard Error	254.661752
Median	10000
Mode	10000
Standard Deviation	1167.00675
Sample Variance	1361904.76
Kurtosis	-0.36267376
Skewness	-0.19853301
Range	4000
Minimum	8000
Maximum	12000
Sum	214000
Count	21
Largest(3)	12000
Smallest(3)	9000
Confidence Level(95.0%)	531.215105

Procedures

1. To use a **Forecast** function, first create a data range with data values you know

2. Click in the cell where you want to place the function.

3. Click on the **Formulas** tab.

4. In the **Function Library** group, click on the **Insert Function** button.

HISTOGRAM

Discussion

The **Histogram** tool can return statistical information from data in Excel and group the output data. You must first create a **bin** to tell the **Histogram** tool how you want the output data grouped.

	A	B	C	D	E	F	G	H
1								
2								
3		Period	Year	Months	Sales		Bind	
4		1	2008	January	£ 3,000.00		2000	
5		2	2008	February	£ 2,000.00		4000	
6		3	2008	March	£ 1,000.00		6000	
7		4	2008	April	£ 500.00		8000	
8		5	2008	May	£ 3,000.00		10000	
9		6	2008	June	£ 1,000.00		12000	
10		7	2008	July	£ 2,000.00		14000	
11		8	2008	August	£ 1,000.00		16000	
12		9	2008	September	£ 2,000.00		18000	
13		10	2008	October	£ 1,000.00			
14		11	2008	November	£ 5,000.00			
15		12	2008	December	£10,000.00			
16		13	2009	January	£ 1,000.00			
17		14	2009	February	£ 5,000.00			
18		15	2009	March	£ 2,000.00			
19		16	2009	April	£ 3,000.00			
20		17	2009	May	£ 500.00			
21		18	2009	June	£ 1,000.00			
22		19	2009	July	£ 2,000.00			
23		20	2009	August	£ 5,000.00			
24		21	2009	September	£ 6,000.00			

In the example above the bind starts with 2000 and goes to 18000 with an increment of 2000.

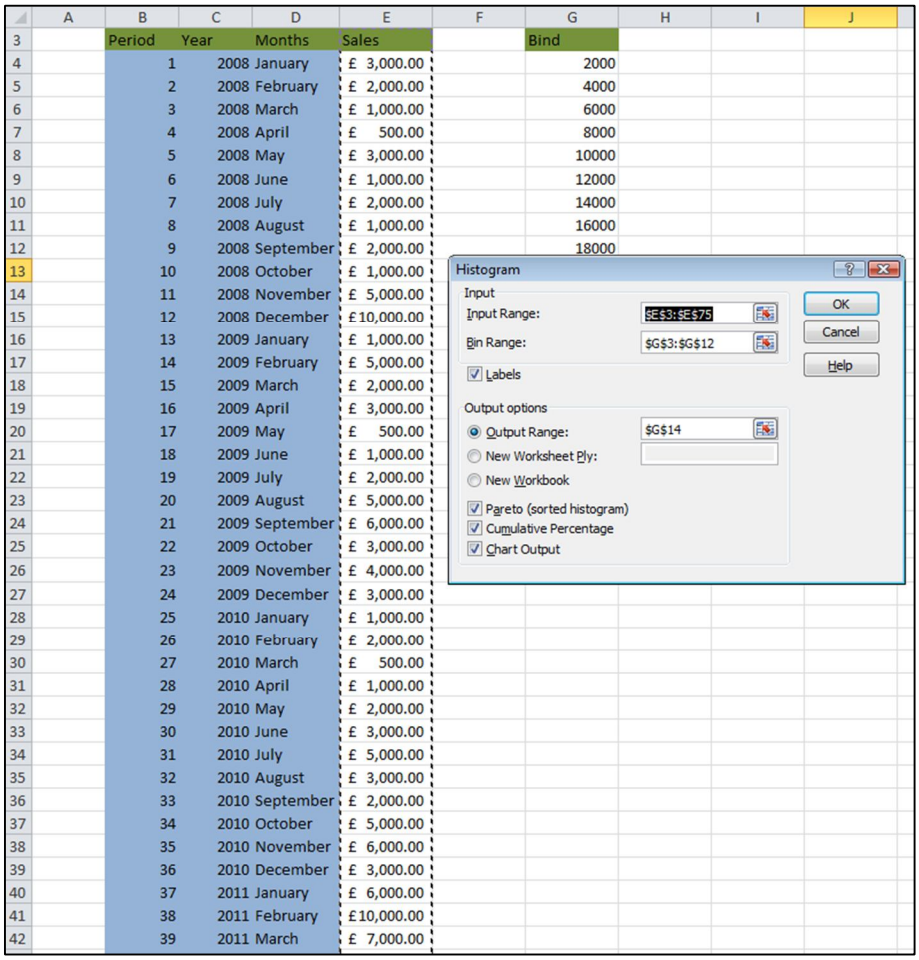
You can find the **Histogram** tool by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Histogram** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Histogram** dialog box, you can get the result in the output range.

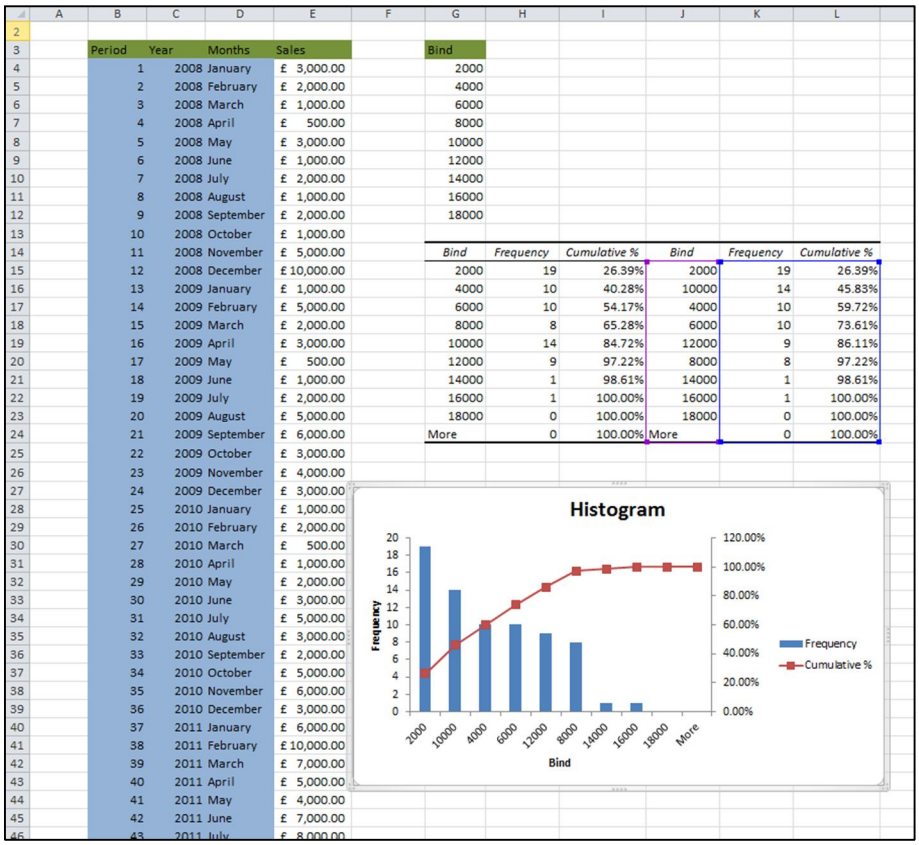
You can specify the following items in the **Histogram** parameters dialog box:

Parameter	Description
Input Range	The range where you have the data you want to examine
Bin Range	Select the range where the bin range has been defined
Labels	If labels are selected in the input range the box must be ticked
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook
Pareto (sorted histogram)	Tick the box if you want the output statistics sorted
Cumulative Percentage	Tick the box if you want cumulative percentage to be a part of the output
Chart output	Tick the box if you want to create a chart to visualise the output



	A	B	C	D	E	F	G	H	I	J
3		Period	Year	Months	Sales		Bind			
4		1	2008	January	£ 3,000.00		2000			
5		2	2008	February	£ 2,000.00		4000			
6		3	2008	March	£ 1,000.00		6000			
7		4	2008	April	£ 500.00		8000			
8		5	2008	May	£ 3,000.00		10000			
9		6	2008	June	£ 1,000.00		12000			
10		7	2008	July	£ 2,000.00		14000			
11		8	2008	August	£ 1,000.00		16000			
12		9	2008	September	£ 2,000.00		18000			
13		10	2008	October	£ 1,000.00					
14		11	2008	November	£ 5,000.00					
15		12	2008	December	£10,000.00					
16		13	2009	January	£ 1,000.00					
17		14	2009	February	£ 5,000.00					
18		15	2009	March	£ 2,000.00					
19		16	2009	April	£ 3,000.00					
20		17	2009	May	£ 500.00					
21		18	2009	June	£ 1,000.00					
22		19	2009	July	£ 2,000.00					
23		20	2009	August	£ 5,000.00					
24		21	2009	September	£ 6,000.00					
25		22	2009	October	£ 3,000.00					
26		23	2009	November	£ 4,000.00					
27		24	2009	December	£ 3,000.00					
28		25	2010	January	£ 1,000.00					
29		26	2010	February	£ 2,000.00					
30		27	2010	March	£ 500.00					
31		28	2010	April	£ 1,000.00					
32		29	2010	May	£ 2,000.00					
33		30	2010	June	£ 3,000.00					
34		31	2010	July	£ 5,000.00					
35		32	2010	August	£ 3,000.00					
36		33	2010	September	£ 2,000.00					
37		34	2010	October	£ 5,000.00					
38		35	2010	November	£ 6,000.00					
39		36	2010	December	£ 3,000.00					
40		37	2011	January	£ 6,000.00					
41		38	2011	February	£10,000.00					
42		39	2011	March	£ 7,000.00					

The output.



Period	Year	Months	Sales
1	2008	January	£ 3,000.00
2	2008	February	£ 2,000.00
3	2008	March	£ 1,000.00
4	2008	April	£ 500.00
5	2008	May	£ 3,000.00
6	2008	June	£ 1,000.00
7	2008	July	£ 2,000.00
8	2008	August	£ 1,000.00
9	2008	September	£ 2,000.00
10	2008	October	£ 1,000.00
11	2008	November	£ 5,000.00
12	2008	December	£ 10,000.00
13	2009	January	£ 1,000.00
14	2009	February	£ 5,000.00
15	2009	March	£ 2,000.00
16	2009	April	£ 3,000.00
17	2009	May	£ 500.00
18	2009	June	£ 1,000.00
19	2009	July	£ 2,000.00
20	2009	August	£ 5,000.00
21	2009	September	£ 6,000.00
22	2009	October	£ 3,000.00
23	2009	November	£ 4,000.00
24	2009	December	£ 3,000.00
25	2010	January	£ 1,000.00
26	2010	February	£ 2,000.00
27	2010	March	£ 500.00
28	2010	April	£ 1,000.00
29	2010	May	£ 2,000.00
30	2010	June	£ 3,000.00
31	2010	July	£ 5,000.00
32	2010	August	£ 3,000.00
33	2010	September	£ 2,000.00
34	2010	October	£ 5,000.00
35	2010	November	£ 6,000.00
36	2010	December	£ 3,000.00
37	2011	January	£ 6,000.00
38	2011	February	£ 10,000.00
39	2011	March	£ 7,000.00
40	2011	April	£ 5,000.00
41	2011	May	£ 4,000.00
42	2011	June	£ 7,000.00
43	2011	July	£ 8,000.00

Bind
2000
4000
6000
8000
10000
12000
14000
16000
18000

Bind	Frequency	Cumulative %	Bind	Frequency	Cumulative %
2000	19	26.39%	2000	19	26.39%
4000	10	40.28%	10000	14	45.83%
6000	10	54.17%	4000	10	59.72%
8000	8	65.28%	6000	10	73.61%
10000	14	84.72%	12000	9	86.11%
12000	9	97.22%	8000	8	97.22%
14000	1	98.61%	14000	1	98.61%
16000	1	100.00%	16000	1	100.00%
18000	0	100.00%	18000	0	100.00%
More	0	100.00%	More	0	100.00%

REGRESSION

Discussion

The **Regression** tool can return statistical information from data in Excel. The tool can make a linear regression forecast returning the slope and Y-interceptor. It can also return the error or deviation and much more statistical informations.

Simple linear regression analysis can be easily completed using Excel.

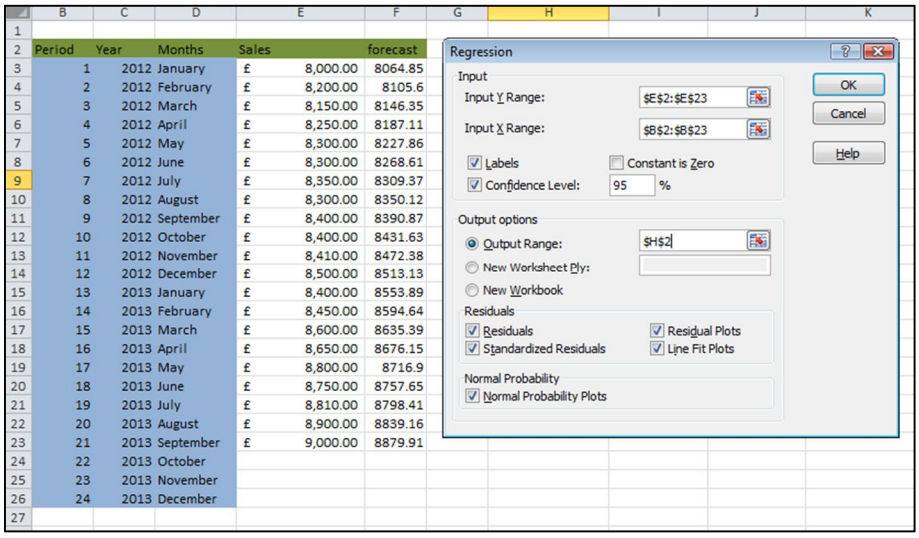
You can find the **Regression** tool by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Regression** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Regression** dialog box, you can get the result in the output range.

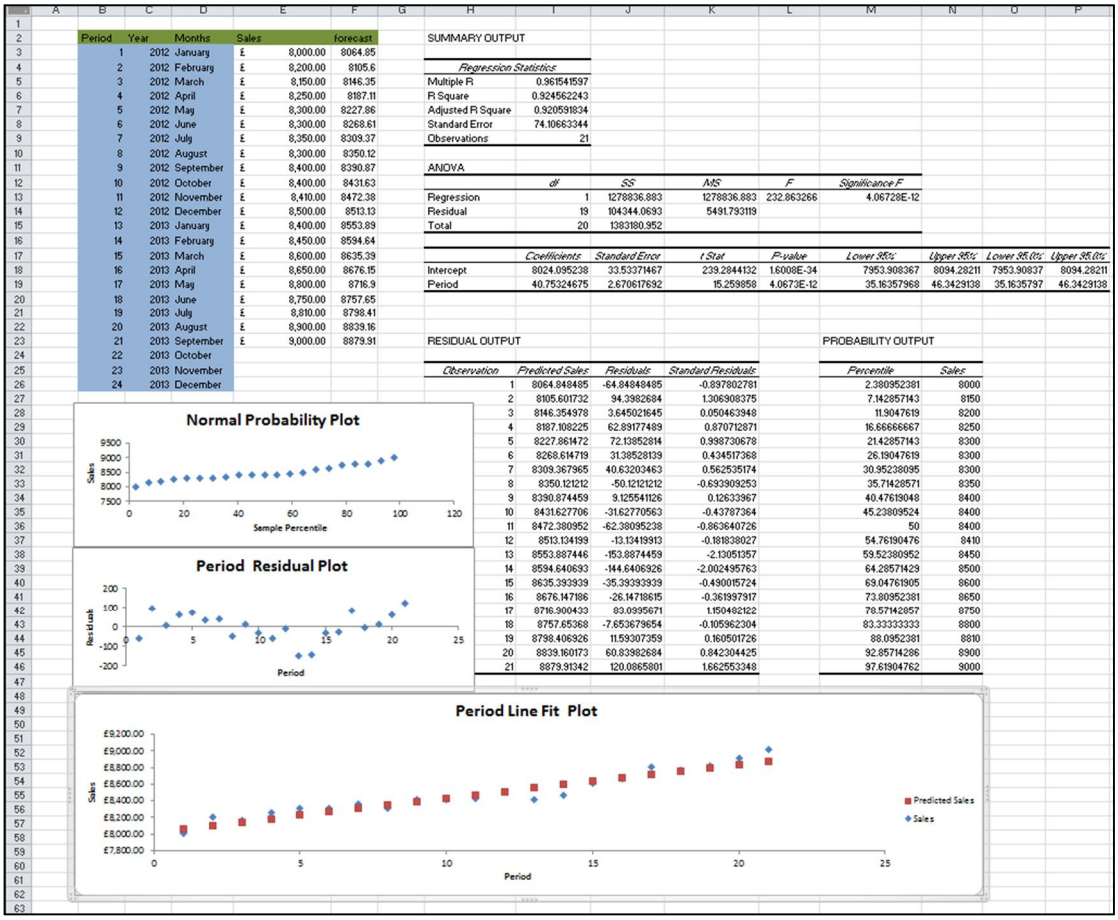
You can specify the following items in the **Regression** parameters dialog box:

Parameter	Description
Input Y Range	The range where you have the Y data you want to examine (values)
Input X Range	The range where you have the X data you want to examine (periods)
Labels	If labels are selected in the input range the box must be ticked
Constant is Zero	Tick the box if you want to work with a constant of zero.
Confidence Level	Enter a percentage from 0 to 100 per cent if you want to work with confidence level
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook
Residuals	Tick the box if you want the residuals (errors, deviation) as a part of the output
Standardized Residuals	Tick the box if you want the standardized residuals (standard errors, standard deviation) as a part of the output
Residuals Plots	Tick the box if you want the residuals (errors, deviation) visualised in a chart as a part of the output
Line Fit Plots	Tick the box if you want a chart showing the trends as a part of the output
Normal Probability	Tick the box if you want a chart showing the probability as a part of the output



Period	Year	Months	Sales	forecast
1	2012	January	£ 8,000.00	8064.85
2	2012	February	£ 8,200.00	8105.6
3	2012	March	£ 8,150.00	8146.35
4	2012	April	£ 8,250.00	8187.11
5	2012	May	£ 8,300.00	8227.86
6	2012	June	£ 8,300.00	8268.61
7	2012	July	£ 8,350.00	8309.37
8	2012	August	£ 8,300.00	8350.12
9	2012	September	£ 8,400.00	8390.87
10	2012	October	£ 8,400.00	8431.63
11	2012	November	£ 8,410.00	8472.38
12	2012	December	£ 8,500.00	8513.13
13	2013	January	£ 8,400.00	8553.89
14	2013	February	£ 8,450.00	8594.64
15	2013	March	£ 8,600.00	8635.39
16	2013	April	£ 8,650.00	8676.15
17	2013	May	£ 8,800.00	8716.9
18	2013	June	£ 8,750.00	8757.65
19	2013	July	£ 8,810.00	8798.41
20	2013	August	£ 8,900.00	8839.16
21	2013	September	£ 9,000.00	8879.91
22	2013	October		
23	2013	November		
24	2013	December		

The output:



SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.961541597
R Square	0.924562243
Adjusted R Square	0.920591834
Standard Error	74.10663344
Observations	21

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1278836.883	1278836.883	232.863266	4.06728E-12
Residual	19	104344.0693	5491.793119		
Total	20	1383180.952			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	8024.095238	33.53371467	239.2844132	1.6008E-34	7953.908367	8094.28211	7953.90837	8094.28211
Period	40.75324675	2.670617692	15.259858	4.0673E-12	35.16357968	46.3429138	35.1635797	46.3429138

RESIDUAL OUTPUT

Observation	*Predicted Sales*	*Residuals*	*Standard Residuals*
1	8064.848485	-64.84848485	-0.897802781
2	8105.601732	94.3982684	1.306908375
3	8146.354978	3.645021645	0.050463948
4	8187.108225	62.89177489	0.870712871
5	8227.861472	72.13852814	0.998730678
6	8268.614719	31.38528139	0.434517368
7	8309.367965	40.63203463	0.562535174
8	8350.121212	-50.1212121	-0.693909253
9	8390.874459	9.125541126	0.12633967
10	8431.627706	-31.62770563	-0.43787364
11	8472.380952	-62.38095238	-0.863640726
12	8513.134199	-13.13419913	-0.18183827
13	8553.887446	-153.8874459	-2.13051357
14	8594.640693	-144.6406926	-2.002495763
15	8635.393939	-35.39393939	-0.490015724
16	8676.147186	-26.14718615	-0.361997317
17	8716.900433	93.0995671	1.150482122
18	8757.65368	-7.653679654	-0.105962304
19	8798.406926	11.59307359	0.160501726
20	8839.160173	60.83982684	0.842304425
21	8879.91342	120.0865801	1.662553348

PROBABILITY OUTPUT

Percentile	*Sales*
2.380952381	8000
7.142857143	8150
11.9047619	8200
16.66666667	8250
21.42857143	8300
26.19047619	8300
30.95238095	8300
35.71428571	8350
40.47619048	8400
45.23809524	8400
50	8400
54.76190476	8410
59.52380952	8450
64.28571429	8500
69.04761905	8600
73.80952381	8650
78.57142857	8750
83.33333333	8800
88.0952381	8810
92.85714286	8900
97.61904762	9000

SAMPLING

Discussion

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 visualise in a chart, you can use a representative sample. You can also create a sample that contains only values from a particular part of a cycle if you believe that the input data is periodic.

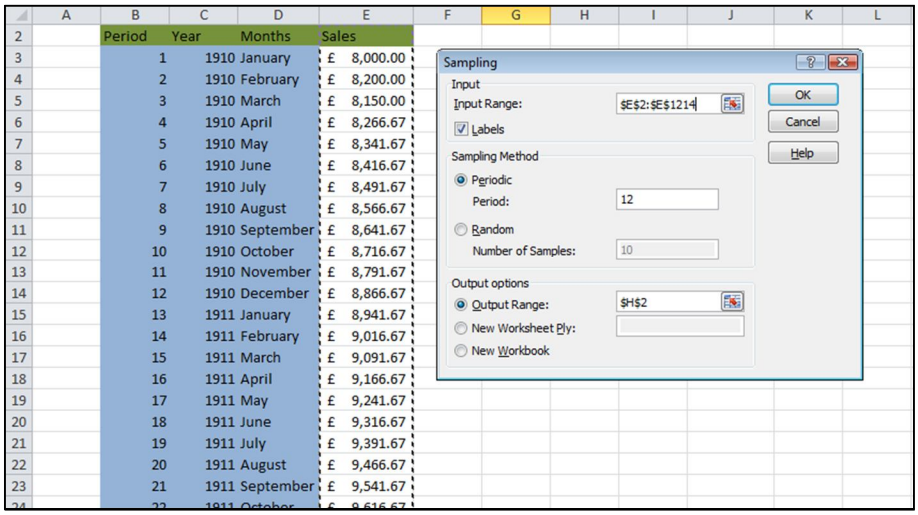
You can find the **Sampling** tool by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Sampling** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Sampling** dialog box, you can get the result in the output range.

You can specify the following items in the **Sampling** parameters dialog box:

Parameter	Description
Input Range	The array where you have the data you want to examine
Labels	If labels are selected in the input range the box must be ticked
Sampling Method Period	If there is a periodic trend in the data select this option. Enter a period number.
Sampling Method Random	If you want the tool to select randomly select this option. Enter a number of samples you need.
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook



The output:

Period	Year	Months	Sales	
1	1910	January	£ 8,000.00	8866.667
2	1910	February	£ 8,200.00	9766.667
3	1910	March	£ 8,150.00	10666.67
4	1910	April	£ 8,266.67	11566.67
5	1910	May	£ 8,341.67	12466.67
6	1910	June	£ 8,416.67	13366.67
7	1910	July	£ 8,491.67	14266.67
8	1910	August	£ 8,566.67	15166.67
9	1910	September	£ 8,641.67	16066.67
10	1910	October	£ 8,716.67	16966.67
11	1910	November	£ 8,791.67	17866.67
12	1910	December	£ 8,866.67	18766.67
13	1911	January	£ 8,941.67	19666.67
14	1911	February	£ 9,016.67	20566.67
15	1911	March	£ 9,091.67	21466.67
16	1911	April	£ 9,166.67	22366.67
17	1911	May	£ 9,241.67	23266.67
18	1911	June	£ 9,316.67	24166.67
19	1911	July	£ 9,391.67	25066.67
20	1911	August	£ 9,466.67	25966.67
21	1911	September	£ 9,541.67	26866.67
22	1911	October	£ 9,616.67	27766.67
23	1911	November	£ 9,691.67	28666.67
24	1911	December	£ 9,766.67	29566.67
25	1912	January	£ 9,841.67	30466.67
26	1912	February	£ 9,916.67	31366.67
27	1912	March	£ 9,991.67	32266.67
28	1912	April	£ 10,066.67	33166.67

RANK & PERCENTILE

Discussion

The **Rank and Percentile** tool ranks data. For each rank the tool return a percentage for data below the specific ranking.

Finding and displaying the percentile of specific results can be valuable to your audience. Sometimes stand alone numbers do not tell the story, but, if I know that 80% of the other results are lower than me, that tells a more complete story.

The percentile is defined as the value of a variable below which a certain percent of observations fall. For example, if you score in the 80th percentile on a test, 80% of the other test takers did worse than you.

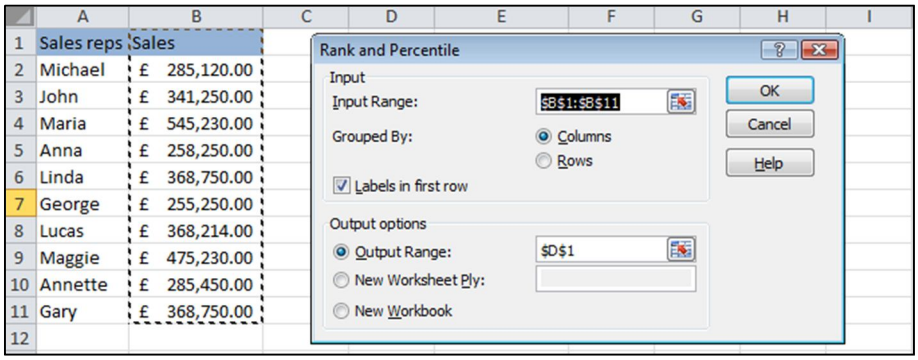
You can find the **Rank and Percentile** tool by selecting the Data Analysis command from the Analysis group at the far right end of the data ribbon. Select **Rank and Percentile** from the list and click OK.

If there is no Analysis group or **Data Analysis** command on the Ribbon, you must install **Data Analysis TOOLPAK** (please go to **Appendix A** in this workbook to see how to install **Data Analysis TOOLPAK)**.

After you have entered all necessary information into the **Rank and Percentile** dialog box, you can get the result in the output range.

You can specify the following items in the **Rank and Percentile** parameters dialog box:

Parameter	Description
Input Range	The array where you have the data you want to examine
Grouped By	Do you have the data grouped by columns or rows
Labels	If labels are selected in the input range the box must be ticked
Output Range	Enter a cell reference for the output or New Worksheet Ply to get the output in a new worksheet in the workbook with the input data or New Workbook to get the output in a new workbook



The output:

	A	B	C	D	E	F	G
1	Sales reps	Sales		*Point*	*Sales*	*Rank*	*Percent*
2	Michael	£ 285,120.00		3	£ 545,230.00	1	100.00%
3	John	£ 341,250.00		8	£ 475,230.00	2	88.80%
4	Maria	£ 545,230.00		5	£ 368,750.00	3	66.60%
5	Anna	£ 258,250.00		10	£ 368,750.00	3	66.60%
6	Linda	£ 368,750.00		7	£ 368,214.00	5	55.50%
7	George	£ 255,250.00		2	£ 341,250.00	6	44.40%
8	Lucas	£ 368,214.00		9	£ 285,450.00	7	33.30%
9	Maggie	£ 475,230.00		1	£ 285,120.00	8	22.20%
10	Annette	£ 285,450.00		4	£ 258,250.00	9	11.10%
11	Gary	£ 368,750.00		6	£ 255,250.00	10	0.00%
12							

APPENDIX A – INSTALLING DATA ANALYSIS TOOLPAK TO EXCEL

INSTALLING DATA ANALYSIS TOOLPAK

Discussion

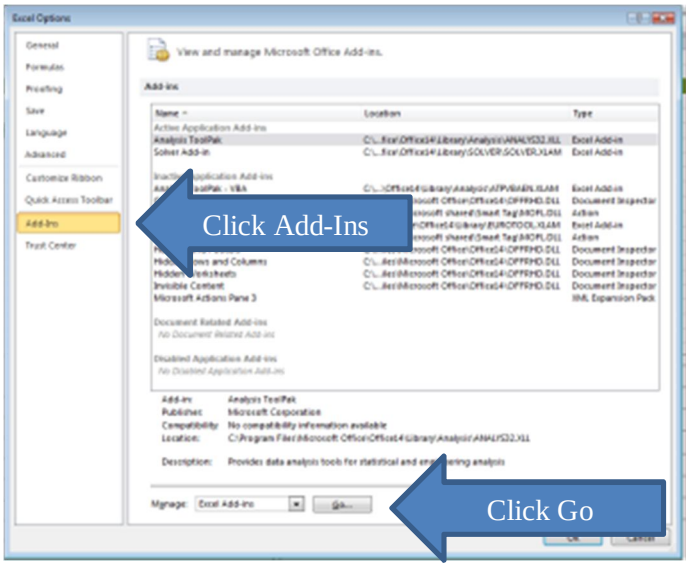
The Data Analysis Toolpak is not part of the standard Excel application as normally installed, it is an Add-In (often referred to as a "plug-in."). Add-Ins are additional Excel components that are not commonly used and as a result, have to be installed on your computer. There are many third party (non-Microsoft) Add-Ins that help Excel do useful and clever things that it cannot normally do, or that enhance some of the things that it can do.

These can be found and downloaded from the Internet although it is wise to consult with your IT department first in case they have objections on grounds of security and computer software safety. It is unlikely that in the working environment you will be able to install these programs yourself anyway.

There are, however, selections of Add-Ins that come with Excel and that are perfectly safe and suitable for use; one of these is **ANALYSIS TOOLPAK**. This Add-in will add 19 analysis tools to Excel. After installation you will be able to find the **ANALYSIS TOOLPAK** tools in the data ribbon.

Procedure

1. Click the Office Button (Excel 2007) or the File tab (Excel 2010).
2. Click Excel Options.

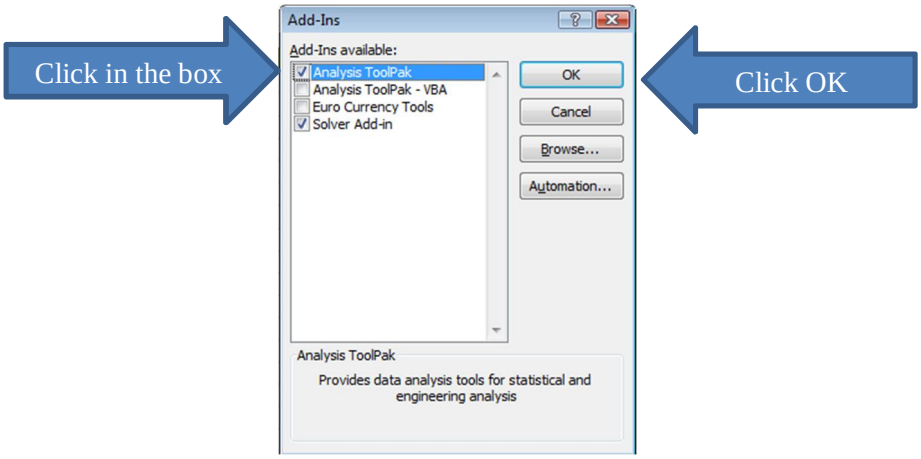


3. Click Add-Ins at the left.

4. Select Excel Add-ins from the Manage: list at the bottom of the dialog box.

5. Click Go….

6. Click in box next to **ANALYSIS TOOLPAK** Add-in.



7. Click OK.

APPENDIX B – TEXT FUNCTIONS

In this appendix, you will learn how to:

- Use the **Right, Left**, and **Mid** functions
- Use the **Concatenate** function
- Use the **Len** function
- Use the **Find** function
- Use nested text functions

CONCEPT AND TERMS

Discussion

Text functions are very useful if you need to clean up text strings, concatenate text, or split a text string to many cells. Specially if you import data from a database for analysing in Excel it is a advantage to know the text functions in Excel.

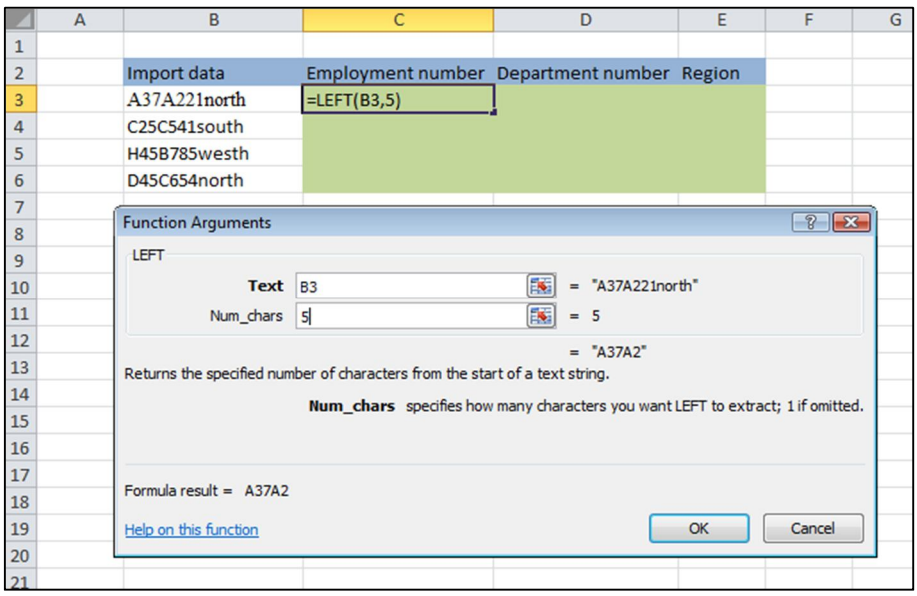
USE THE RIGHT, LEFT, AND MID FUNCTIONS

Discussion

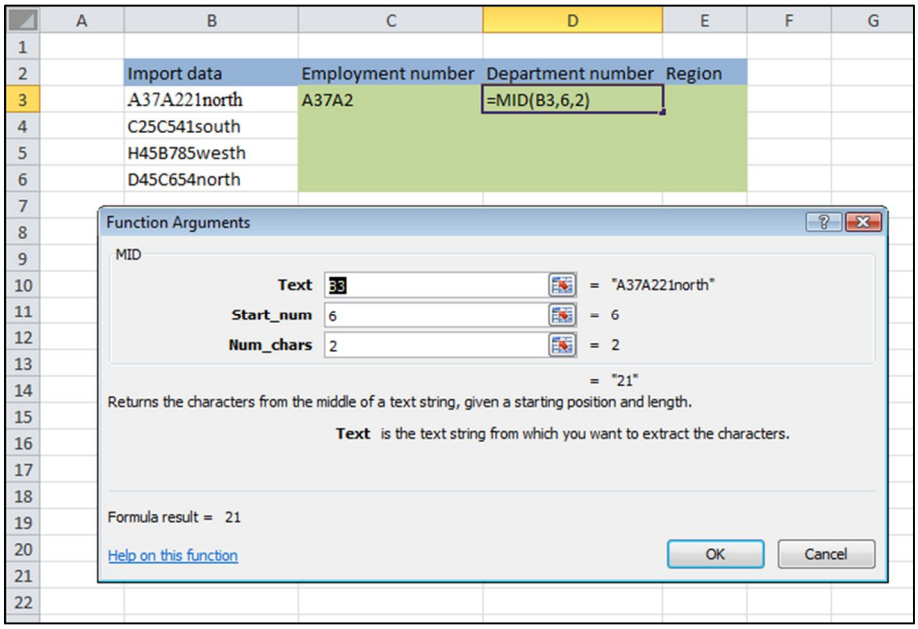
If you want to extract characters from a text string you can use the **Left, Right,** and **Mid** function. The **Left** function can extract characters from the left of the text string, the **Right** function can extract characters from the right of a text string, and the **Mid** function can extract characters from the midle of a text string.

The first five characters of the text string "A37A221regionnorth"are the imployment number for a staff member, the next two characters are department number and the last five characters are region information. To extract the employment number you need the **Left** function.

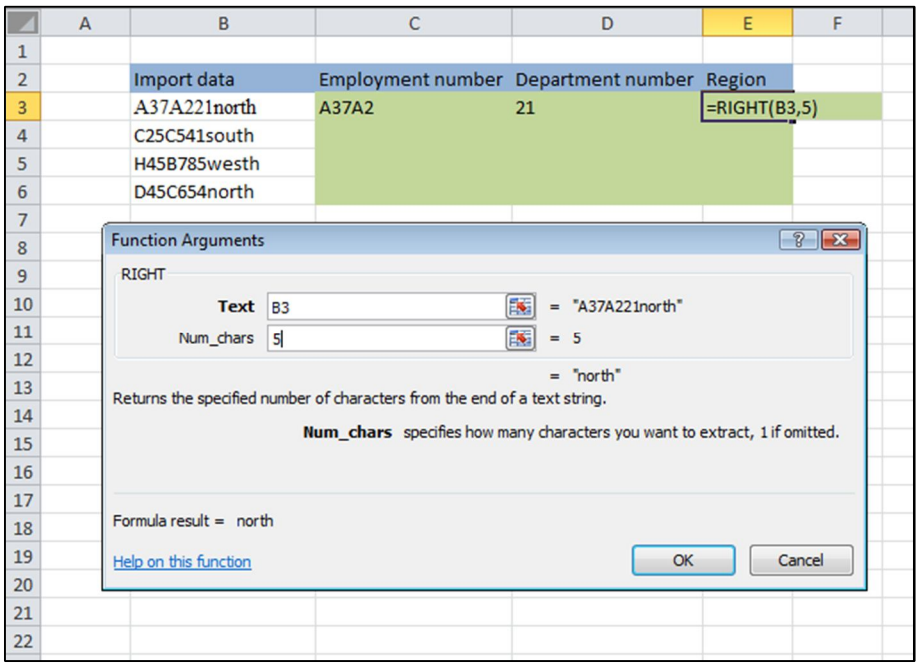
The **Left** function consists of one required argument and one optional, in the following order: **Text**, **Num_chars. Text** is the box in which the cell reference for the cell with the text string must be entered. In the **Num_chars** box enter the number of characters you want to extract from left.



The **Mid** function consists of three required arguments, in the following order: **Text**, **Start_num,** and **Num_chars. Text** is the box in which the cell reference for the cell with the text string must be entered. In the **Start_num** box enter the start number of the first character you want to extract from the middle of the text string. **Num_chars** is the number of characters you want to extract.



The **Right** function consists of one required argument and one optional, in the following order: **Text**, **Num_chars. Text** is the box in which the cell reference for the cell with the text string must be entered. In the **Num_chars** box enter the number of characters you want to extract from right.

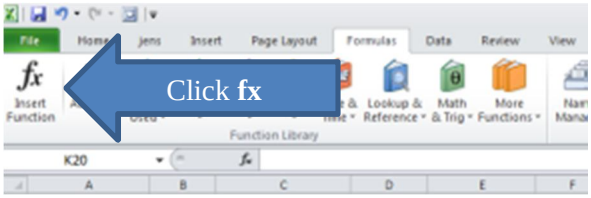


Procedures

1. To use a **Left** function, select a cell where you want the function to extract characters from the left of a text string.

2. Click on the **Formulas** tab.

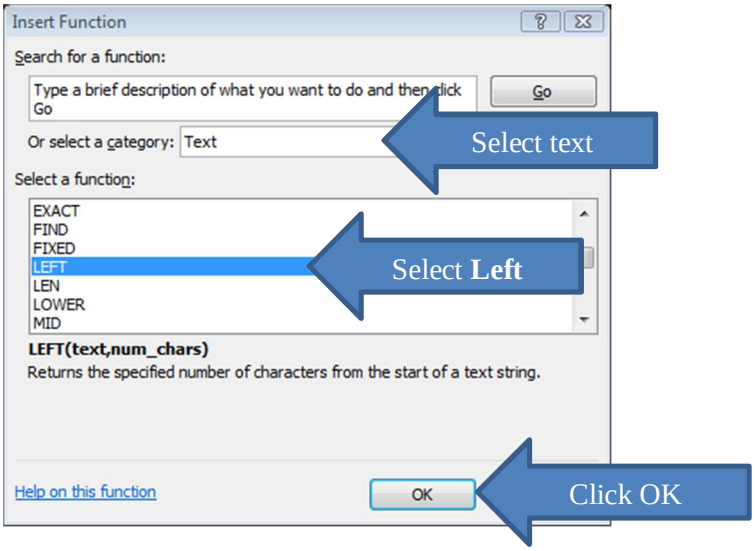
3. In the **Function Library** group, click on the **Insert Function** button.



4. In the Insert function dialog box, locate **Text** category in the **Or select a category:** box.

5. Click on **Left**.

6. Click on the **OK** button.



7. In the **Function Arguments** dialog box, click in the **Text** box.

8. Click the cell containing the text string from which you want to extract the characters.

9. In the **Num_chars** box, type the number of characters you want to extract.

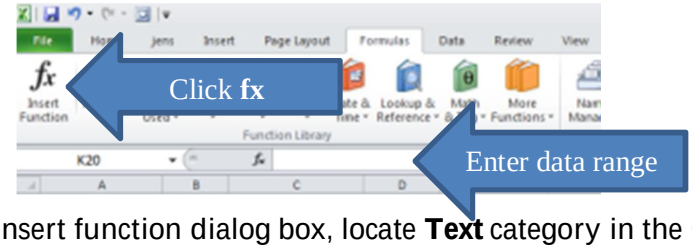
10. Click on the OK button.

Procedures

1. To use a **Mid** function, select a cell where you want the function to extract characters from the middle of a text string.

2. Click on the **Formulas** tab.

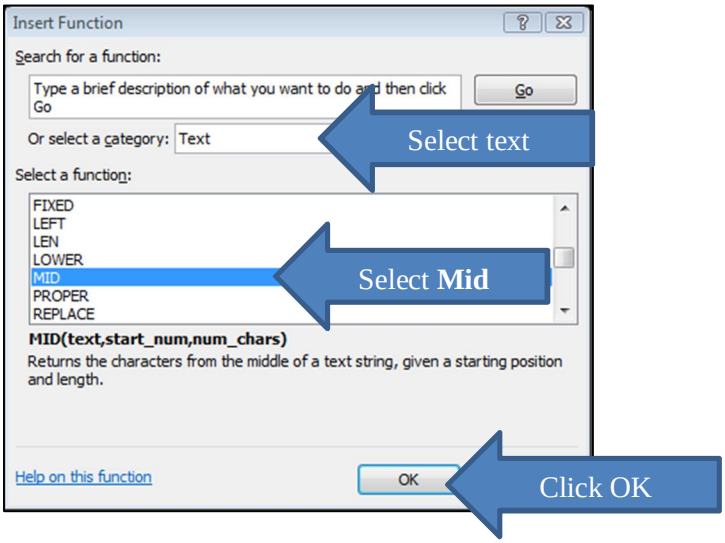
3. In the **Function Library** group, click on the **Insert Function** button.



4. In the Insert function dialog box, locate **Text** category in the **Or select a category:** box.

5. Click on **Mid**.

6. Click on the **OK** button.



7. In the **Function Arguments** dialog box, click in the **Text** box.

8. Click the cell containing the text string from which you want to extract the characters.

9. In the **Start_num** box, type the number of the first character you want to extract from the middle of you text string.

10. In the **Num_chars** box, type the number of characters you want to extract.

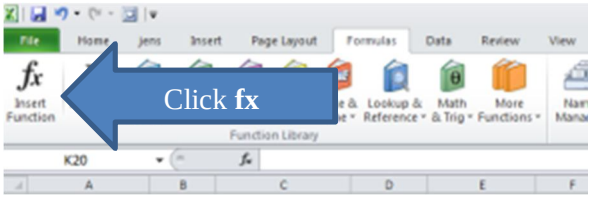
11. Click on the OK button.

Procedures

1. To use a **Right** function, select a cell where you want the function to extract characters from the right of a text string.

2. Click on the **Formulas** tab.

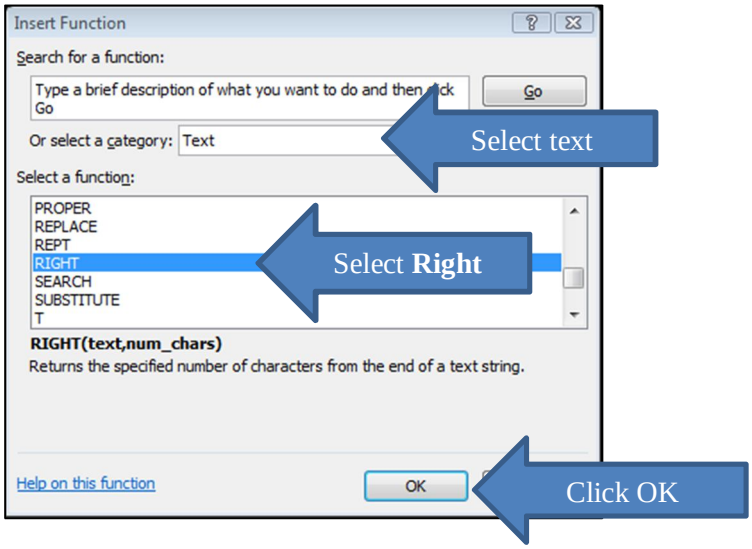
3. In the **Function Library** group, click on the **Insert Function** button.



4. In the Insert function dialog box, locate **Text** category in the **Or select a category:** box.

5. Click on **Right**.

6. Click on the **OK** button.



7. In the **Function Arguments** dialog box, click in the **Text** box.

8. Click the cell containing the text string from which you want to extract the characters.

9. In the **Num_chars** box, type the number of characters you want to extract.

10. Click on the OK button.

USE THE CONCATENATE FUNCTION

Discussion

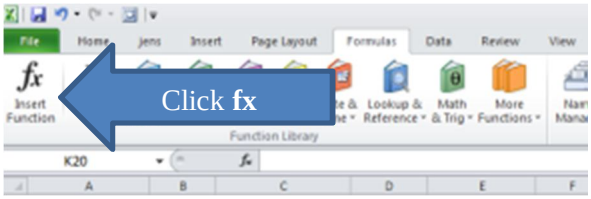
The **Concatenate** function joins up to 255 text strings into one text string. The joined items can be text, numbers, cell references, or a combination of those items. For example, if your worksheet contains a person's first name in cell A1 and the person's previous name in cell B1, you can combine the two values in another cell by using the following formula:

=concatenate(A1," ",B1)

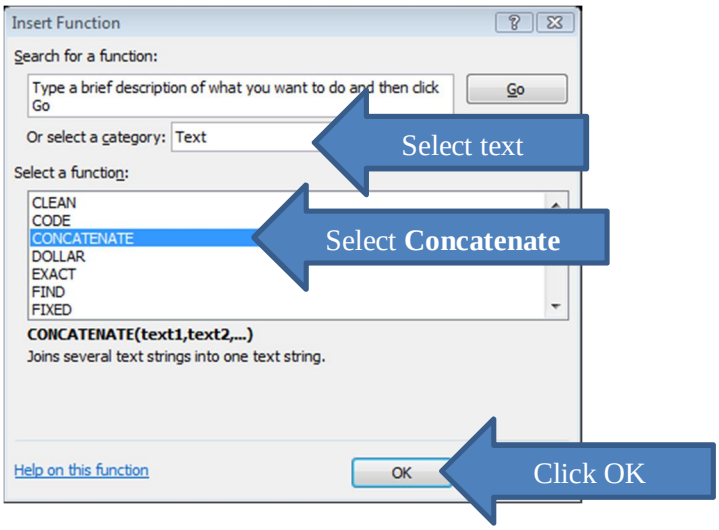
The **Concatenate** function consists of one required argument and as many optional arguments you need up to 254, in the following order: **Text1**, **Text2, Text3, Text4** ….to **Text255. Text1** is the box in which the cell reference for the first cell with the text string must be entered. In the **Text2** box enter the cell reference of the cell containing the second text string you want to join. In the **Text3** box enter the cell reference of the cell containing the third text string you want to join etc... If you want to have a space between the text strings you can enter " " ("space") in every second argument.

Procedures

1. To use a **Concatenate** function, select a cell where you want the function to join text strings.
2. Click on the **Formulas** tab.
3. In the **Function Library** group, click on the **Insert Function** button.



4. In the Insert function dialog box, locate **Text** category in the **Or select a category:** box.
5. Click on **Concatenate**.
6. Click on the **OK** button.



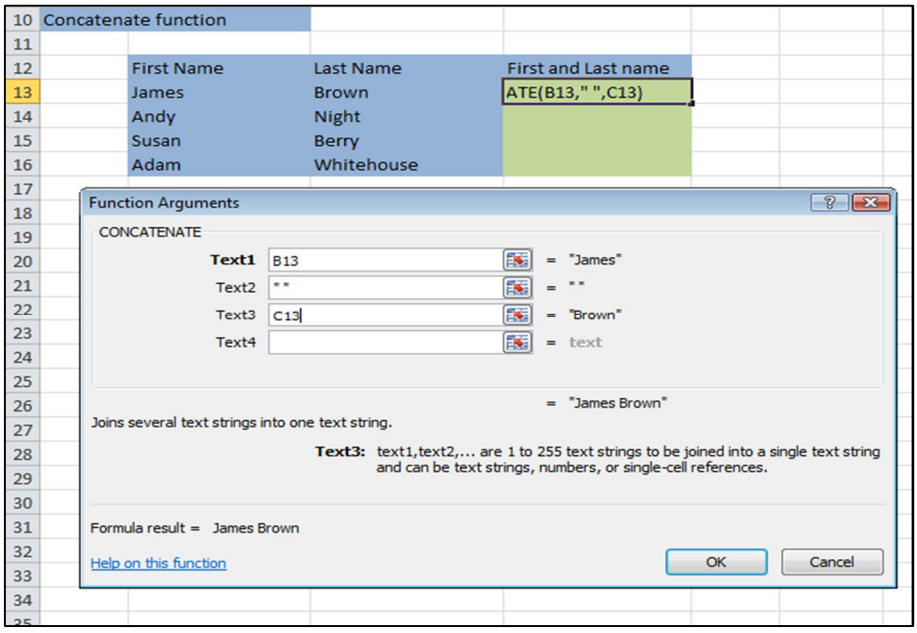
7. In the **Function Arguments** dialog box, click in the **Text1** box.

8. Click the cell containing the first text string you want to join.

9. In the **Text2** box type " " to get a space between the first and second text string.

10. In the **Text2** box click the cell containing the second text string you want to join.

11. Click on the OK button.



USE THE LEN FUNCTION

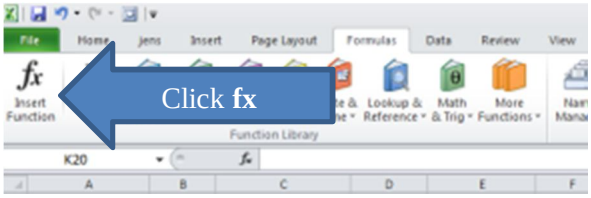
Discussion

The **Len** function returns the number of characters in a text string. A space will also be counted as a character.

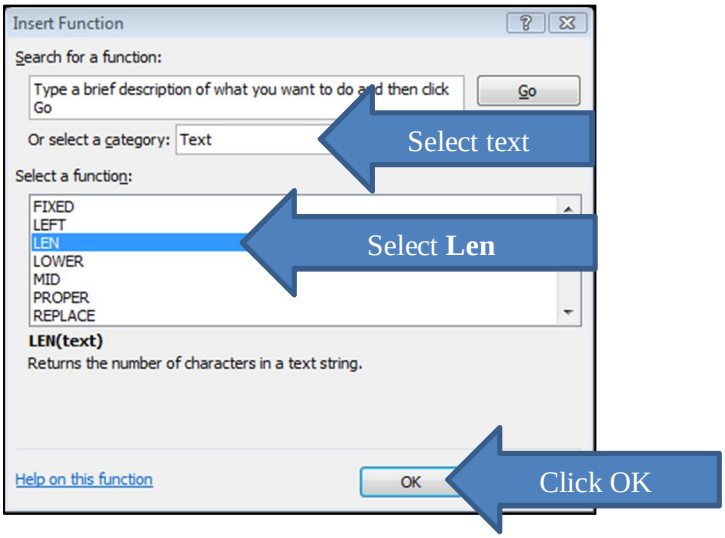
The **Len** function consists of one required argument: **Text1.** In **Text1** enter the cell reference for the cell with the text string and the function will return how many characters there is in text string.

Procedures

1. To use a **Len** function, select a cell where you want the function to display the result.
2. Click on the **Formulas** tab.
3. In the **Function Library** group, click on the **Insert Function** button.



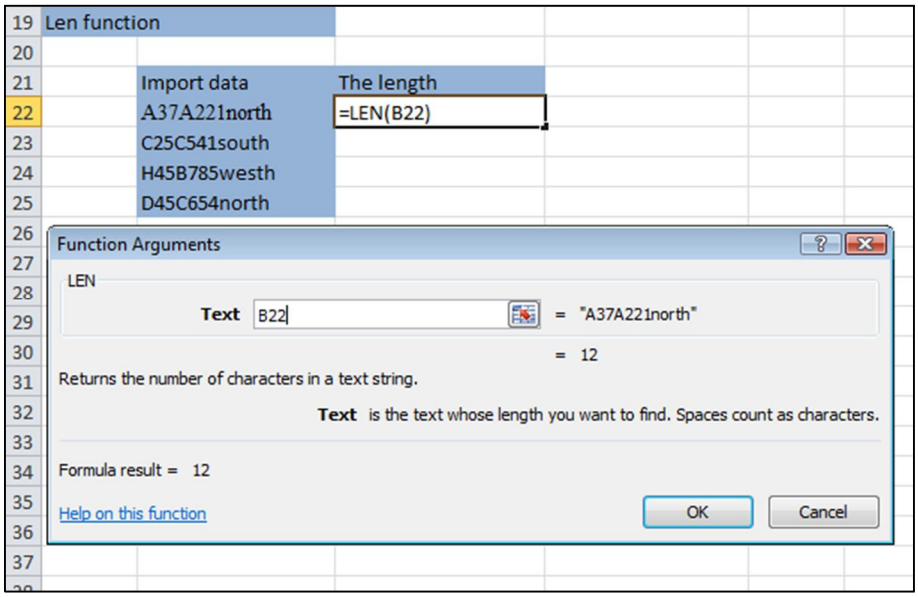
4. In the Insert function dialog box, locate **Text** category in the **Or select a category:** box.
5. Click on **Len**.
6. Click on the **OK** button.



7. In the **Function Arguments** dialog box, click in the **Text** box.

8. Click the cell containing the text string you want to examine.

9. Click on the OK button.



USE THE FIND FUNCTION

Discussion

The **Find** function can find a character or a text string within a text string and return the position of the character or text string. The **Find** function is very often used together with other text function such as the **Mid** function.

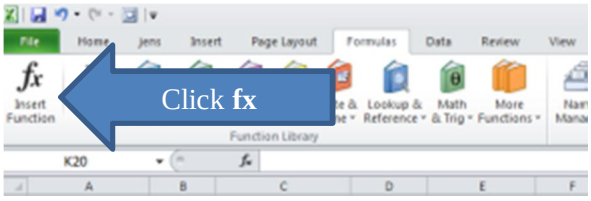
The **Find** function consists of two required arguments and one optional argument, in following order: **Find_text, Within_text, Start_num.** In **Find_text** you must enter the text string or character you want the function to find (the argument is case sensitive). In **Within_text** enter the cell reference for the cell with the text string where you want to find the character or text string. In **Start_num** you can enter the character number you want the function to start the search from. If you in a text string want to find the character "A" and there is two characters "A" then the function will return the position number for the first "A".

Procedures

1. To use a **Find** function, select a cell where you want the function to display the result.

2. Click on the **Formulas** tab.

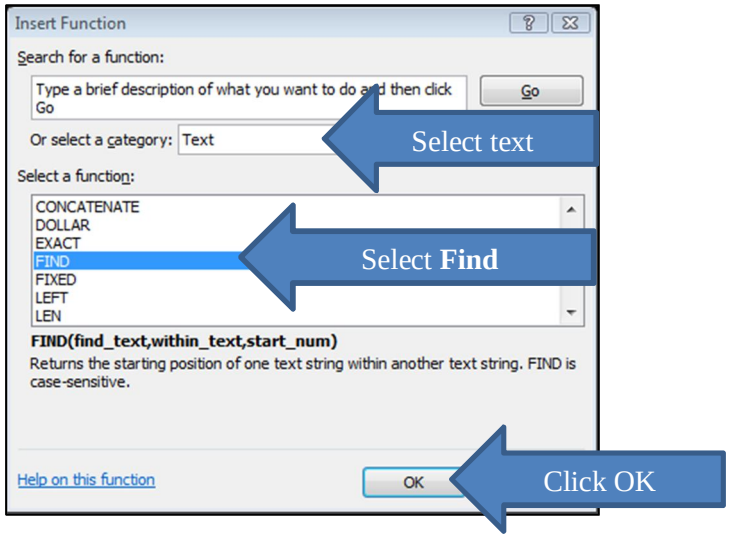
3. In the **Function Library** group, click on the **Insert Function** button.



4. In the Insert function dialog box, locate **Text** category in the **Or select a category:** box.

5. Click on **Find**.

6. Click on the **OK** button.



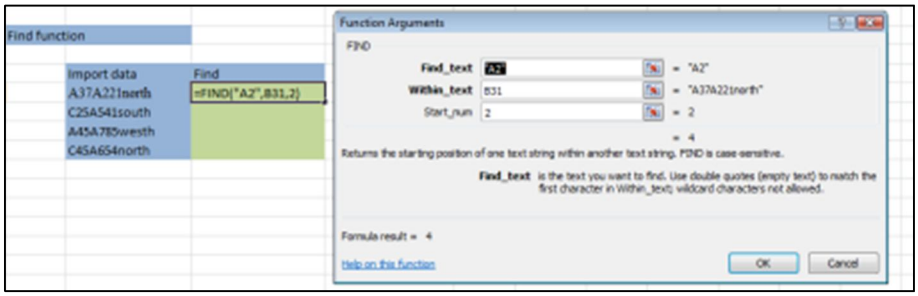
7. In the **Function Arguments** dialog box, click in the **Find_text** box.

8. Type the character or text string you want the function to find.

9. Enter the cell reference for the cell you want to examine in the **Within_text** box.

10. In the **Start_num** box enter a number you want the function to start the search from.

11. Click on the OK button.



USE NESTED TEXT FUNCTIONS

Discussion

You can nest the text functions to get more dynamic options to extract text strings. May be you have imported text strings from a database and you want to extract different part of the text string to different cells in Excel but you do not know how many characters you want in each cell. You can use a combination of the **Right, Left, Mid, Len,** and **Find** functions to solve the problem. There is many solutions in Excel to deal with this but you will see some examples here.

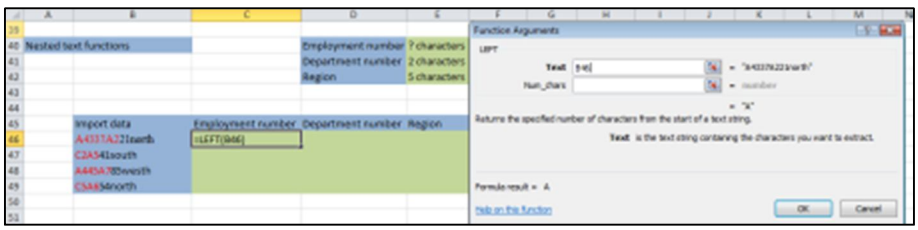
Example 1

You have imported a number of text strings "A4337A221north", "C2A541south","A445A785westh", and "C5A654north". You know that the first part of the text string is the employment number but you do not know how many characters. You know that the second part of the text string is department number and the department number is 2 characters. The last 5 characters of the text string is the region.

	A	B	C	D	E
39					
40	Nested text functions			Employment number	? characters
41				Department number	2 characters
42				Region	5 characters
43					
44					
45		Import data	Employment number	Department number	Region
46		A4337A221north			
47		C2A541south			
48		A445A785westh			
49		C5A654north			
50					

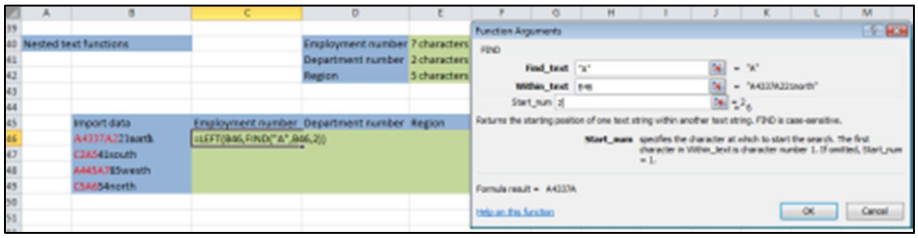
You need the employment number in column C, department number in column D, and region in column E. The employment number is a mix of letters and digits but always there is the letter "A" and one digit at the end of the employment number. With this knowledge you can extract the employment number from the text string by using the **Left** and **Find** function.

Select the cell where you want the characters extracted. Start with the **Left** function. In the **Text** box enter the cell reference for the cell with the text string.

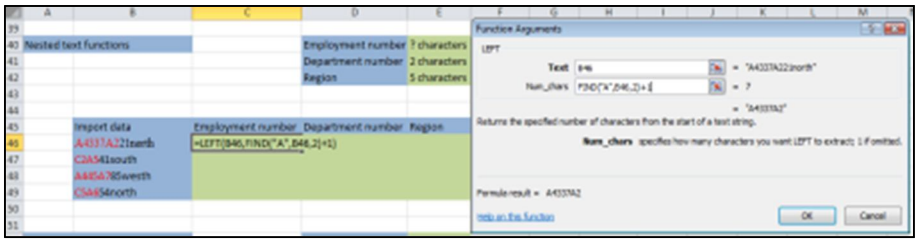


Click in the **Num_char** box and nest the **Find** function. The **Find** function must find the letter "A" the second last character in the employment number. In the

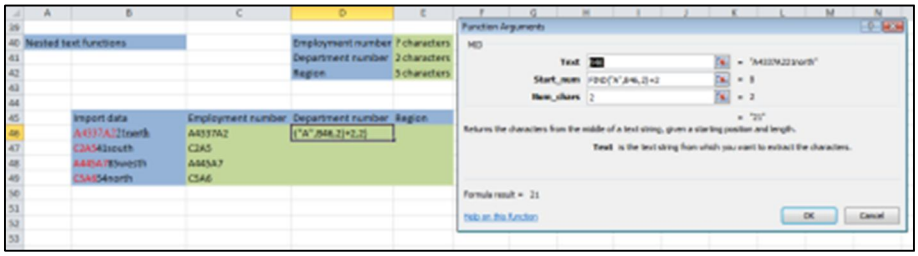
Find_text box enter the character you want to find. In the **Within_text** box enter the cell reference for the cell with the text string. In the **Start_num** box enter the start number for the search. You don't want to find "A" if it is the first character in the text string as it is in this example that is why 2 is entered in the box. You want to start the search from character 2.



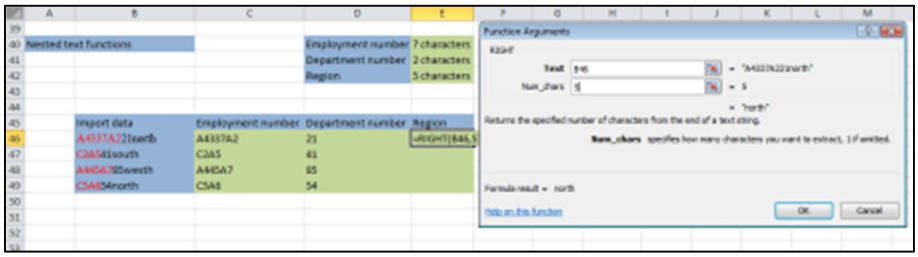
This will find the character number for "A" in the text string but the digit after "A" is also a part of the employment number. What you need to do is to click after the **Find** function in the **Left** function dialog box and type "+1". Click **OK** button and you have extracted the employment number.



To extract the department number you need the **Mid** function. In the **Text** box enter the cell reference for the cell with the text string. In **Start_num** you need to nest the **Find** function. Again the **Find** function must find the character "A". This time you need to type "+2" after the **Find** function in the **Mid** function dialog box because the department number starts 2 characters from "A". In **Number_char** box type 2 and click **OK.**



To extract the region you do not need any nested function just use the **Right** function.



Example 2

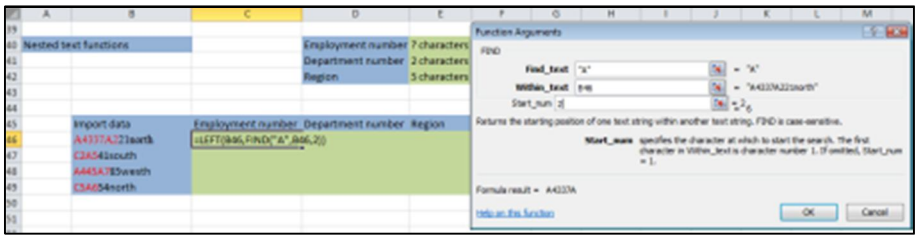
You have imported a number of text strings "A4337A221northwest", "C2A545east","A445A785southwesth", and "C5A654north". You know that the first part of the text string is the employment number but you do not know how many characters. You know that the second part of the text string is department number and the department number is 2 characters. The last part of the text string is the region but you do not know how many characters.

	A	B	C	D	E	F
63						
64	Nested text functions			Employment number	? characters	
65				Department number	2 characters	
66				Region	? characters	
67						
68						
69		Import data	Employment number	Department number	Region	
70		A4337A221northwest	A4337A2	21	northwest	
71		C2A545east	C2A5	45	east	
72		A445A785southwesth	A445A7	85	southwesth	
73		C5A654north	C5A6	54	north	
74						
75						

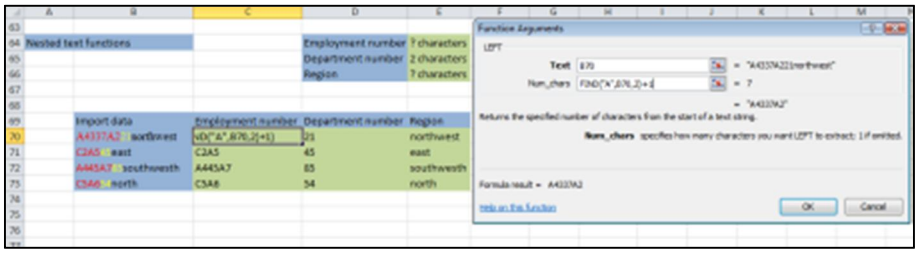
You need the employment number in column C, department number in column D, and region in column E. The employment number is a mix of letters and digits but always there is the letter "A" and one digit at the end of the employment number. With this knowledge you can extract the employment number from the text string by using the **Left** and **Find** function.

Select the cell where you want the characters extracted. Start with the **Left** function. In the **Text** box enter the cell reference for the cell with the text string.

Click in the **Num_char** box and nest the **Find** function. The **Find** function must find the letter "A" the second last character in the employment number. In the **Find_text** box enter the character you want to find. In the **Within_text** box enter the cell reference for the cell with the text string. In the **Start_num** box enter the start number for the search. You don't want to find "A" if it is the first character in the text string as it is in this example that is why 2 is entered in the box.

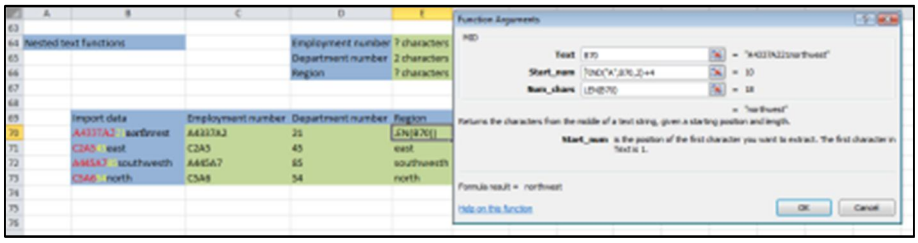


This will find the character number for "A" in the text string but the digit after "A" is also a part of the employment number. What you need to do is to click after the **Find** function in the **Left** function dialog box and type "+1". Click **OK** button and you have extracted the employment number.



To extract the department number you need the **Mid** function. In the **Text** box enter the cell reference for the cell with the text string. In **Start_num** you need to nest the **Find** function. The **Find** function must find the character "A". This time you need to type "+2" after the **Find** function in the **Mid** function dialog box because the department number starts 2 characters from "A". In **Number_char** box type 2 and click **OK.**

To extract the region without knowing the number of characters you can not use the **Right** function. you must use a combination of the **Mid, Find** and **Len** functions. Start with the **Mid** function. You need to nest the **Find** function in **Start_num** again to find the character "A". This time you need to type "+4" after the **Find** function in the **Mid** function dialog box to find the position of the first character in the region information. Nest the **Len** function in **Num_char** this will reurn the number of characters for the whole text string. Click **OK** and the **Mid** function will return the region information.



INDEX

P

R

S

T

V

W