The latest version of this document and related examples are found in http://myy.haaga-helia.fi/~taaak/q

Time series forecasting

The objective of time series methods is to discover a pattern in the historical data and then extrapolate this pattern into the future. The pattern may have several components. It is common to separate four components: trend, cyclical, seasonal and irregular.

Trend

- Persistent, overall upward or downward pattern
- Due to population increase, technology etc.
- Often several years duration but for example, in stock markets we may be interested in shorter term trends.

Cyclical

- Repeating up and down movements
- Due to multiyear cyclical movements in the economy
- Usually 2-10 years duration
- Exact duration of a cycle is not known beforehand -> difficult to consider when forecasting.

In this document the cyclical component is not considered. Forecasts related to the cyclical component are available from national and international research organizations and banks (e.g. Bank of Finland).

Seasonal

- Regular pattern of up and down fluctuations
- Due to weather, customs etc.
- Occurs usually within 1 year but for example, in stock markets we may be interested even in weekly seasonal variations.

Irregular

- Unsystematic, residual fluctuations>
- Due to random variation or unforeseen events like union strike or tornado
- Short duration and non repeating
- Impossible to forecast

Smoothing methods

The objective of smoothing methods is to smooth out the random fluctuations caused by the irregular component of the time series. Smoothing methods are appropriate for a stable time series - that is, one that exhibits no significant trend, cyclical, or seasonal effects. Smoothing methods are commonly used for short-range forecasts such as a forecast for the next time period. The two most common smoothing methods are the moving average method and the exponential smoothing method.

Moving average method

The moving average method uses the average of the most recent n data values in the time series as the forecast for the next period. In the following table sales data is available for 15 weeks. Five week moving averages have been calculated in the moving average column. The moving average for week 16 is the sales forecast for week 16.

	A	B	C	D
4	Week	Demand	5 week moving average	Absolute error
5	1	252922		
6	2	248559		
7	3	253342		
8	4	249532		
9	5	247693		
10	6	246618	250410	3792
11	7	261333	249149	12184
12	8	247447	251704	4257
13	9	251492	250525	967
14	10	249713	250917	1204
15	11	258563	251321	7242
16	12	249702	253710	4008
17	13	262667	251383	11284
18	14	258829	254427	4402
19	15	255631	255895	264
20	16		257078	
21	Mean absolute error			4960

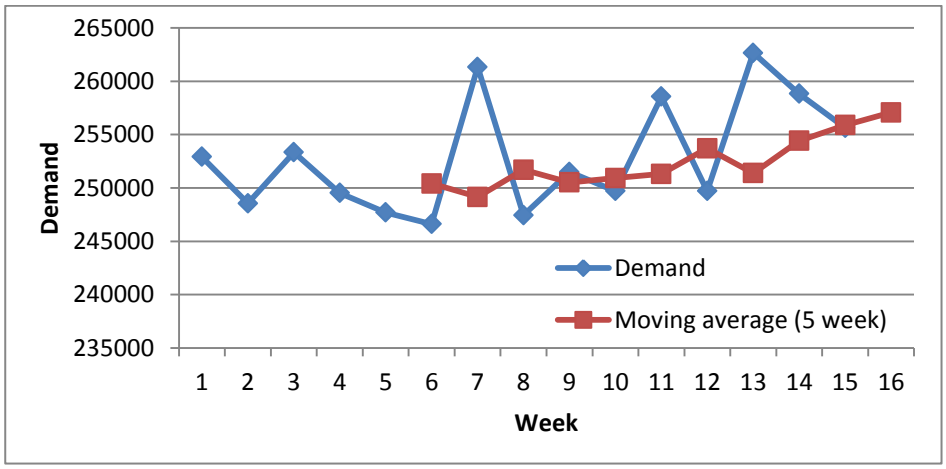
In the cell C10 there is a formula
=AVERAGE(B5:B9)

In the cell D10 there is a formula
=ABS(C10-B10)

In the cell D21 there is a formula
=AVERAGE(D10:D19)

When selecting the value of n you need to consider the accuracy of the forecast. For a particular time series, the selection of n will affect the accuracy of the forecast. To decide the value of n you can use trial and error to identify the n-value that minimizes the average error. In the table above the absolute forecast error for the past 10 weeks have been calculated. Finally the mean absolute error has been calculated as the mean of absolute forecast errors. By using different n-values you get different mean absolute error values. The n value giving the smallest average error is usually the best. At a later date, after new time series observations have been obtained, you may need to adjust the value of n.

You should always check the chart to see how the selected method seems to work.



If you assume that the older values are not as relevant as the most recent values you should consider using the exponential smoothing method instead of the moving average method.

Exponential smoothing

Exponential smoothing uses a weighted average of past time series values as the forecast. The weight for the most recent observation is selected by the forecaster. The weights for the other data values are automatically computed (built in the formula) and get smaller and smaller as the observations move farther into the past. The forecast is calculated with the following formula

alfa*previous observation + (1 - alfa)*previous forecast

Usually the first forecast is calculated as an average of known demands.

	A	B	C	D	E	F
4	Week	Demand	Forecast	Absolute error	alfa = 0,24	
5	1	252922	252936			
6	2	248559	252933	4374		
7	3	253342	251883	1459		
8	4	249532	252233	2701		
9	5	247693	251585	3892		
10	6	246618	250651	4033		
11	7	261333	249683	11650		
12	8	247447	252479	5032		
13	9	251492	251271	221		
14	10	249713	251324	1611		
15	11	258563	250938	7625		
16	12	249702	252768	3066		
17	13	262667	252032	10635		
18	14	258829	254584	4245		
19	15	255631	255603	28		
20	16		255610			
21	Mean absolute error			4327		

The first forecast in the cell C5 is calculated as **=AVERAGE(B5:B19)**

The second forecast in the cell C6 is calculated as **=$F$4*B5+(1-$F$4)*C5**

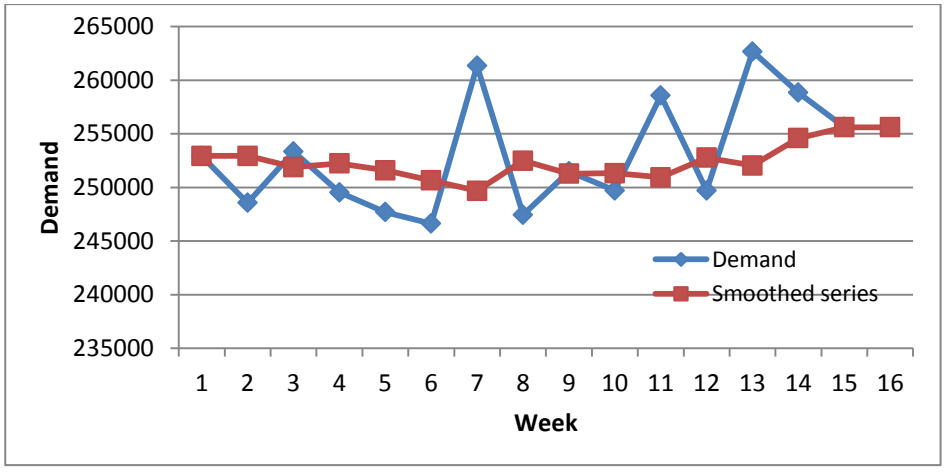
The absolute error in the cell D6 is calculated as **=ABS(C6-B6)**

The mean absolute error in the cell D21 is calculated as **=AVERAGE(D6:D19)**

The important issue is the selection of the smoothing constant alfa. The smoothing constant can be given any value between 0 and 1. Note, that the smoothing constant is the weight of the most recent value in the forecast. So, if the smoothing constant is 1 then the forecast is just the most recent value. The more weight you like to give to the older values the smaller smoothing constant you should use. To decide the best value for the smoothing constant you can use trial and error to identify the smoothing constant value that minimizes the mean absolute error. At a later date, after new time series observations have been obtained, you may need to adjust the smoothing constant.

You should try whether you can find a better smoothing constant for the example above.

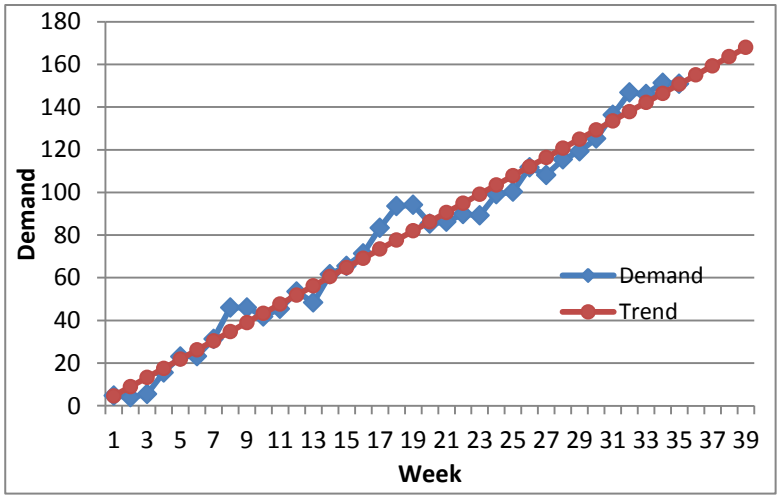
You should always check the chart to see how the selected method seems to work.



Trend projection

The trend projection method is applicable if a time series shows a consistent increase or decrease over time.

Although the next chart shows some up-and-down movement over the past 35 weeks, the time series for the demand seems to have an overall upward trend. The simple linear regression model gives an appropriate trend projection model for the example.



In Excel you get linear trend values by using the function **=FORECAST(x,known y's;known x's)**

- x=the number of the period (e.g. to get the forecast for week number 11 you should use x=11)
- known y's is the reference to the known time series values
- known x's is the reference to the known period numbers.

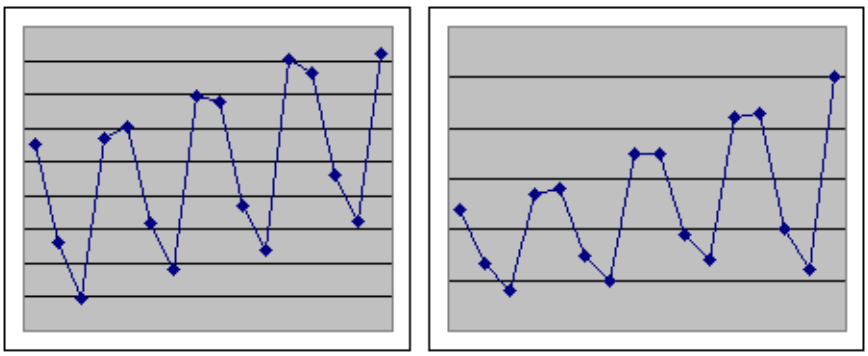
You may also find the **=SLOPE(known y's;known x's)** function useful. In the example above the slope equals 4,3. The slope of 4,3 indicates that over the past 35 weeks the firm has experienced an average growth in sales of about 4,3 units per week.

The use of a linear function to model the trend is common. However, sometimes time series exhibit a curvilinear (nonlinear) trend and curvilinear model should be used. Curvilinear models go beyond the scope of this document.

Considering both trend and seasonal variation

It takes several steps to calculate forecasts considering both the trend component and the seasonal component. At first you need to decide whether to use an additive or multiplicative model.

- In the additive model a time series value is considered to be the sum of the trend component, the seasonal component and the irregular component: $Y_t=T_t+S_t+I_t$
- In the multiplicative model a time series value is considered to be the product of the trend component, the seasonal component and the irregular component: $Y_t=T_t*S_t*I_t$



The additive model is appropriate when the magnitude of the seasonal variation is approximately constant (the first picture). On the other hand if the magnitude of the seasonal variation is dependent on the trend level then the multiplicative model is more appropriate (the second picture). In practice it is often difficult to know which model is better before doing the actual calculations (and comparing mean absolute errors).

The steps to calculate a forecast are explained in the following. You find more details in the corresponding Excel example.

Smooth seasonal variation in the time series

Use moving averages to smooth seasonal variation in the time series.

- In the case of quarterly data 4 period moving averages must be used to smooth the seasonal variation
- In the case of monthly data 12 month moving averages must be used to smooth the seasonal variation.

The positioning of the moving averages is not straightforward. In the case of paired number of seasons you need to calculate centered moving average in order to find appropriate location for the moving average:

- First moving average 19,50 belongs between 2. quarter and 3. quarter.
- Second moving average 20,00 belongs between 3. quarter and 4. quarter.
- Average of 19,50 and 20,00 (centered moving average) belongs to 3. quarter.

	A	B	C	D	E
4	Quarter		Demand	Moving average	Centered moving average
5	1	1	8		
6	2	2	13		
7	3	3	23	19,50	19,75
8	4	4	34	20,00	20,63
9	1	5	10	21,25	21,25
10	2	6	18	21,25	21,75
11	3	7	23	22,25	22,50
12	4	8	38	22,75	22,13
13	1	9	12	21,50	22,63
14	2	10	13	23,75	24,13
15	3	11	32	24,50	
16	4	12	41		

Calculate the trend

	A	B	C	D	E	F	G
4	Quarter		Demand	Moving average	Centered moving average	Trend	Seasonal
5	1	1	8			18,96	0,42
6	2	2	13			19,49	0,67
7	3	3	23	19,50	19,75	20,01	1,15
8	4	4	34	20,00	20,63	20,53	1,66
9	1	5	10	21,25	21,25	21,06	0,47
10	2	6	18	21,25	21,75	21,58	0,83
11	3	7	23	22,25	22,50	22,11	1,04
12	4	8	38	22,75	22,13	22,63	1,68
13	1	9	12	21,50	22,63	23,15	0,52
14	2	10	13	23,75	24,13	23,68	0,55
15	3	11	32	24,50		24,20	1,32
16	4	12	41			24,72	1,66
17	1	13				25,25	
18	2	14				25,77	
19	3	15				26,30	
20	4	16				26,82	

Calculate the trend forecast for each period you like to forecast.

The trend in the cell F5 has been calculated as **=FORECAST(B5;$E$7:$E$14;$B$7:$B$14)** and copied downwards.

Note, that the trend is based on centered moving average (time series without seasonal variation).

Note, that you need to use running quarter numbers 1, 2, 3, 4, 5, … in the column B.

Isolate the seasonal component

- In the additive model: Subtract the trend from the actual value to isolate the seasonal component.
- In the multiplicative model: Divide the actual value by the trend to isolate the seasonal component.

Calculate seasonal indexes

Calculate the average of the isolated seasonal components for each season to get the seasonal index for each season. Using the average removes irregular variation included in individual seasonal components. For example, the seasonal index for the first quarter **=AVERAGE(G5;G9;G13)**

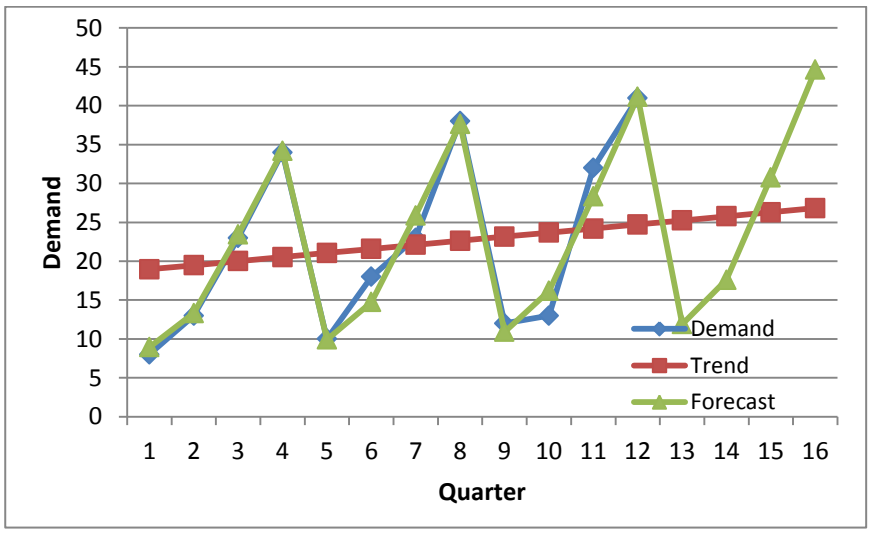
Quarter	Seasonal
1	0,47
2	0,68
3	1,17
4	1,66

Do seasonal adjustments to the trend forecasts

- In the case of the additive model add appropriate seasonal index to each trend forecast.
- In the case of the multiplicative model multiply each trend forecast by appropriate seasonal index.

	A	B	C	D	E	F	G	H	I
4	Quarter		Demand	Moving average	Centered moving average	Trend	Seasonal	Forecast	Absolute error
5	1	1	8			18,96	0,42	8,94	0,94
6	2	2	13			19,49	0,67	13,32	0,32
7	3	3	23	19,50	19,75	20,01	1,15	23,43	0,43
8	4	4	34	20,00	20,63	20,53	1,66	34,18	0,18
9	1	5	10	21,25	21,25	21,06	0,47	9,93	0,07
10	2	6	18	21,25	21,75	21,58	0,83	14,75	3,25
11	3	7	23	22,25	22,50	22,11	1,04	25,88	2,88
12	4	8	38	22,75	22,13	22,63	1,68	37,66	0,34
13	1	9	12	21,50	22,63	23,15	0,52	10,92	1,08
14	2	10	13	23,75	24,13	23,68	0,55	16,18	3,18
15	3	11	32	24,50		24,20	1,32	28,33	3,67
16	4	12	41			24,72	1,66	41,15	0,15
17	1	13				25,25		11,91	
18	2	14				25,77		17,61	
19	3	15				26,30		30,79	
20	4	16				26,82		44,64	
21	Absolute mean error								1,37

In the following chart you can find the original time series and forecasts for the next four periods. The trend line is also visible.



Final notes

Whatever forecasting method you use your forecasts are uncertain due to the uncertain nature of reality. So, your objective can't be to get exactly correct forecasts. Instead, your objective should be to get as accurate forecasts as possible. When you have to forecast something repeatedly (like weekly sales) then you get a great benefit, at least on the long run, from using an appropriate systematic method. Forecaster's experience may also give a valuable contribution to the forecasting.

The methods described in this document are not the only ones available. There are plenty of more sophisticated methods but the 'simple' methods explained in this document are accurate enough in many practical applications. This is due to irregular variation that cannot be forecasted no matter how sophisticated methods you use.