

14.3 EXAMPLE OF LINEAR TREND WITH SEASONALITY

Figure 14.5 Original Data

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Date	Claims			steph44haf								
2	04-Jan	17,414,897.94			If my data is cyclical in nature, what is the best function to use to								
3	04-Feb	10,699,109.47			forecast my next year or two years of sales volume? (I have higher volumes								
4	04-Mar	18,332,334.50			in certain times of the year and I need to account for this when projecting								
5	04-Apr	14,275,140.03			my sales.)								
6	04-May	12,305,352.33											
7	04-Jun	13,907,155.18			vezerid								
8	04-Jul	11,963,018.44			If your data is cyclical then you are probably best of to use a								
9	04-Aug	19,201,480.28			sinusoidal function. If there is an overall upward trend from period to								
10	04-Sep	15,623,457.98			period you might want to add a linear function. Thus I would recommend								
11	04-Oct	7,077,725.63			a function like:								
12	04-Nov	15,740,422.12			$f(t) = at + b + c \cdot \sin(dt+e)$								
13	04-Dec	13,761,418.33			Problem is the built-in functions for regression in Excel do not								
14	05-Jan	21,340,245.83			support such functions so you will need the Solver to perform the least								
15	05-Feb	9,409,514.83			squares method. For this you would need the column representing time to								
16	05-Mar	10,572,805.35			have numeric values or else you should provide an auxiliary column with								
17	05-Apr	12,339,659.95			consecutive numeric values (better off with 0, 1, ...). Say this is in								
18	05-May	11,986,746.47			column A:A starting from A2.								
19	05-Jun	10,252,392.46			You will need five cells for the five constants a-e. Say these are in								
20	05-Jul	12,416,685.61			F2:J2.								
21	05-Aug	17,892,569.26			Next to your dependent variable, say in D2, enter and copy down the								
22	05-Sep	26,618,694.92			formula:								
23	05-Oct	7,581,879.50			$=\$F\$2*A2+\$G\$2+\$H\$2*\text{SIN}(\$I\$2*A2+\$J\$2)$								
24	05-Nov	15,579,836.07			Next to it, in E2, enter the square difference of the dependent								
25	05-Dec	21,710,331.63			(assumed in column C:C) from the forecasted:								
26	06-Jan	21,665,556.58			$=(D2-C2)^2$								
27	06-Feb	13,653,795.27			Take the sum of column E:E and ask SOLver to minimize it by changing								
28	06-Mar	14,457,680.21			F2:J2. As this is a nonlinear problem and the built-in solver is not								
29	06-Apr	18,774,698.52			very industrial strength, your initial values in F2:J2 will have to be								
30	06-May	17,775,539.97			relatively close to the values you expect.								
31	06-Jun	16,774,408.35			Write back if you need further assistance.								
32					HTH								
33					Kostis Vezerides								
34													
35					steph44haf								
36					This stuff is great, but it might be a little over my head. Here is my data,								
37					unfortunately I didn't follow how to do the equations. I sort of figured out								
38					how to use Solver, but I wasn't sure what data I need in what columns, since								
39					I only have two rows right now. If you can't help me any more, I understand								
40					but I want to say thank you for your help already Kostis!								

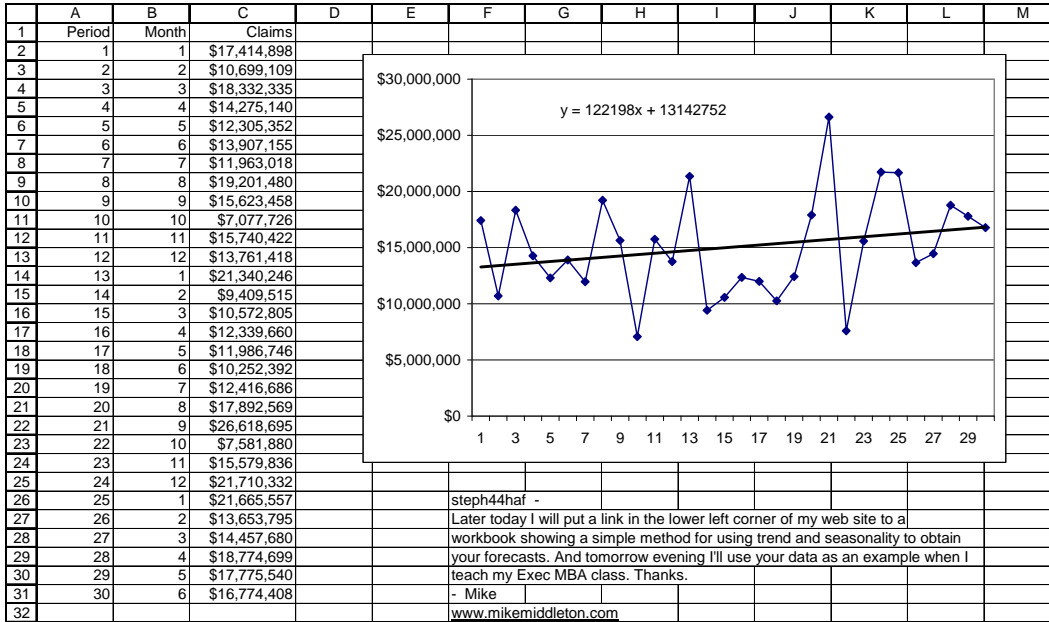
Stephanie posted on an Excel newsgroup. She described her data as "cyclical," and she wanted to make a forecast for the next year or two. Kostis replied with a method for using Solver to fit a sine function to data.

Figure 14.6 Formatted Data

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Period	Month	Claims			Stephanie,							
2	1	1	\$17,414,898			now I see... Well, I don't think you should call this data cyclical. At							
3	2	2	\$10,699,109			first I thought you were talking about a product with seasonal behavior							
4	3	3	\$18,332,335			but this is not the case. Judging from the headers and having charted							
5	4	4	\$14,275,140			the data: We have an overall growth pattern but large fluctuations from							
6	5	5	\$12,305,352			month to month, which is to be expected. Problem is, the fluctuations							
7	6	6	\$13,907,155			are rather large and they do not follow a specific pattern.							
8	7	7	\$11,963,018			In this case we have two choices: linear and exponential, unless there							
9	8	8	\$19,201,480			exist some other market-dependent conditions which would dictate a							
10	9	9	\$15,623,458			different type of function, e.g. quadratic. I give you two equations:							
11	10	10	\$7,077,726			Linear:							
12	11	11	\$15,740,422			13264930							
13	12	12	\$13,761,418			Exponential:							
14	13	1	\$21,340,246			12055160							
15	14	2	\$9,409,515			In both cases, K2 should contain the number of months between the start							
16	15	3	\$10,572,805			of your data and the month you want the projection for. You can use the							
17	16	4	\$12,339,660			function DATEDIFF(date2,DATE(2004,1,1),"m") to calculate this. For							
18	17	5	\$11,986,746			date2 you should use DATE(yr,month,day), i.e something like							
19	18	6	\$10,252,392			DATE(2007,5,1) for May 2007.							
20	19	7	\$12,416,686			However, I am afraid this is as far as my statistics will go. The							
21	20	8	\$17,892,569			number you will produce with these formulas is an estimate, however							
22	21	9	\$26,618,695			with low confidence. Maybe one of the resident experts, like Jerry							
23	22	10	\$7,581,880			Lewis, will jump in and direct you further so that you can also							
24	23	11	\$15,579,836			calculate the plus-or-minus expected fluctuation from the projection.							
25	24	12	\$21,710,332			HTH							
26	25	1	\$21,665,557			Kostis Vezerides							
27	26	2	\$13,653,795										
28	27	3	\$14,457,680										
29	28	4	\$18,774,699										
30	29	5	\$17,775,540										
31	30	6	\$16,774,408										

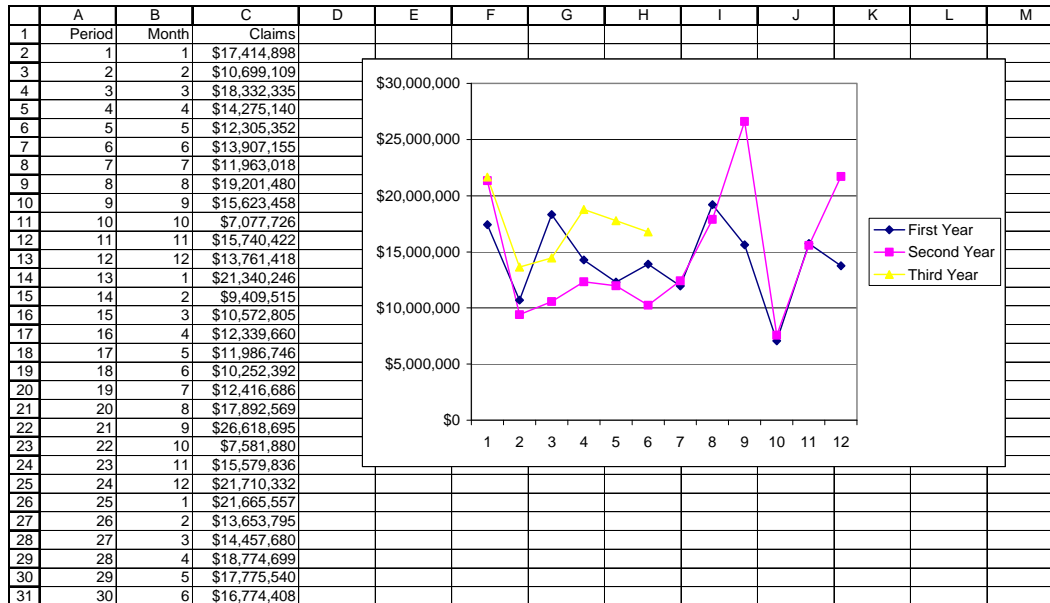
Kostis took a second look at the data and concluded that the data are not cyclical. He mentioned some other functions that he thought might be appropriate for fitting the data.

Figure 14.7 Visual Check Trend



Mike changed the dates to numbers, thereby avoiding any problems with the way that Excel interprets dates as specified in the operating system's Regional Options. He plotted the data, noticing a slight upward trend and possible slight seasonality.

Figure 14.8 Visual Check Seasonality



To visually check for seasonality, Mike plotted each year's data "on top of each other."

In Excel, he created a Line chart for one year of monthly data. Then he used the Source Data Series tab to add separate data series for each of the other two years.

This method of checking for seasonality works best when the data exhibits little trend. When there is significant trend or meandering with values for different years at very different levels, it is better to express the values as ratio-to-trend when making this check.

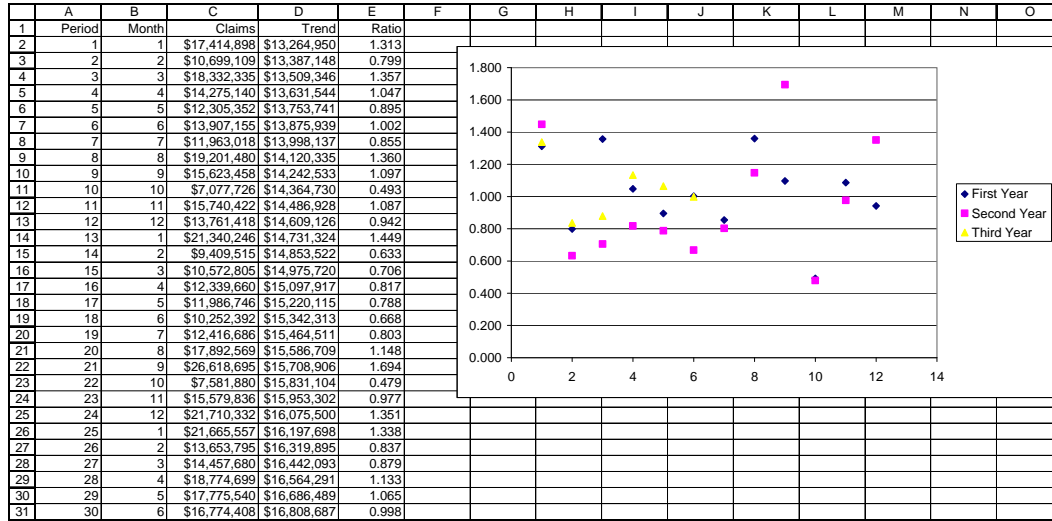
There does appear to be some seasonality.

Figure 14.9 Trends and Ratios

	A	B	C	D	E	F
1	Period	Month	Claims	Trend	Ratio	
2	1	1	\$17,414,898	\$13,264,950	1.313	
3	2	2	\$10,699,109	\$13,387,148	0.799	
4	3	3	\$18,332,335	\$13,509,346	1.357	
5	4	4	\$14,275,140	\$13,631,544	1.047	
6	5	5	\$12,305,352	\$13,753,741	0.895	
7	6	6	\$13,907,155	\$13,875,939	1.002	
8	7	7	\$11,963,018	\$13,998,137	0.855	
9	8	8	\$19,201,480	\$14,120,335	1.360	
10	9	9	\$15,623,458	\$14,242,533	1.097	
11	10	10	\$7,077,726	\$14,364,730	0.493	
12	11	11	\$15,740,422	\$14,486,928	1.087	
13	12	12	\$13,761,418	\$14,609,126	0.942	
14	13	1	\$21,340,246	\$14,731,324	1.449	
15	14	2	\$9,409,515	\$14,853,522	0.633	
16	15	3	\$10,572,805	\$14,975,720	0.706	
17	16	4	\$12,339,660	\$15,097,917	0.817	
18	17	5	\$11,986,746	\$15,220,115	0.788	
19	18	6	\$10,252,392	\$15,342,313	0.668	
20	19	7	\$12,416,686	\$15,464,511	0.803	
21	20	8	\$17,892,569	\$15,586,709	1.148	
22	21	9	\$26,618,695	\$15,708,906	1.694	
23	22	10	\$7,581,880	\$15,831,104	0.479	
24	23	11	\$15,579,836	\$15,953,302	0.977	
25	24	12	\$21,710,332	\$16,075,500	1.351	
26	25	1	\$21,665,557	\$16,197,698	1.338	
27	26	2	\$13,653,795	\$16,319,895	0.837	
28	27	3	\$14,457,680	\$16,442,093	0.879	
29	28	4	\$18,774,699	\$16,564,291	1.133	
30	29	5	\$17,775,540	\$16,686,489	1.065	
31	30	6	\$16,774,408	\$16,808,687	0.998	

Mike selected cells D2:D31, typed `=trend(c2:c31)`, and array-entered by holding down Control and Shift while pressing Enter. The TREND function assumes X values of 1,2,3,... The results in column D are the same as those shown on the Visual Check Trend chart. For ratio-to-trend, Mike entered the formula `=C2/D2` in cell E2 and copied it throughout the column. These ratios are tentative seasonal indexes, indicating how actual sales relate to trend in each month.

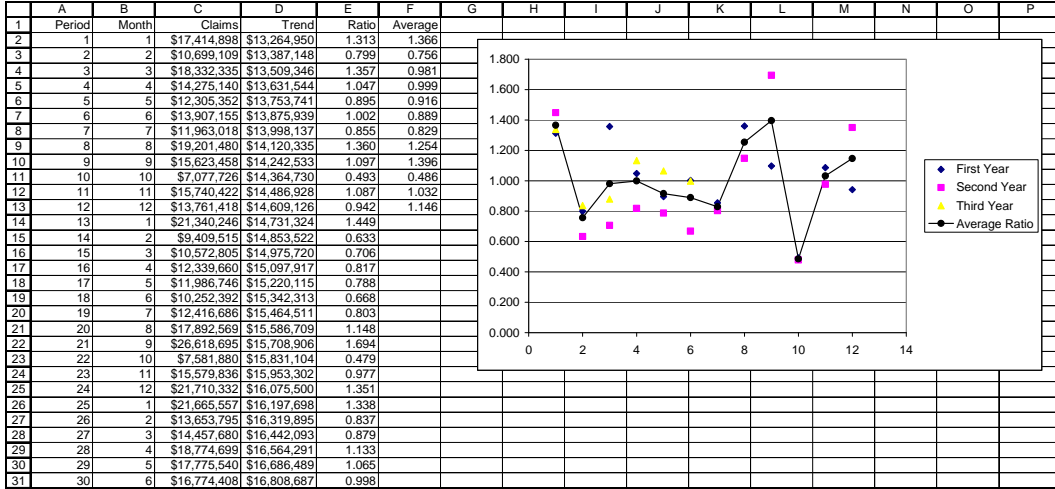
Figure 14.10 Visual Check Ratios



Mike created an XY (Scatter) chart using B2:B13 for X values and E2:E13 for Y values. Then he added the separate data series for the other two years. He observed some clustering of the ratios for months 1, 2, 7, and 10. There was considerable variation for months 3 and 9.

If there are data for five or more years, we could use a trimmed mean (Excel function TRIMMEAN) or median to get the "typical" ratio for each month. Here there are only two or three values for each month.

Figure 14.11 Average of Ratios

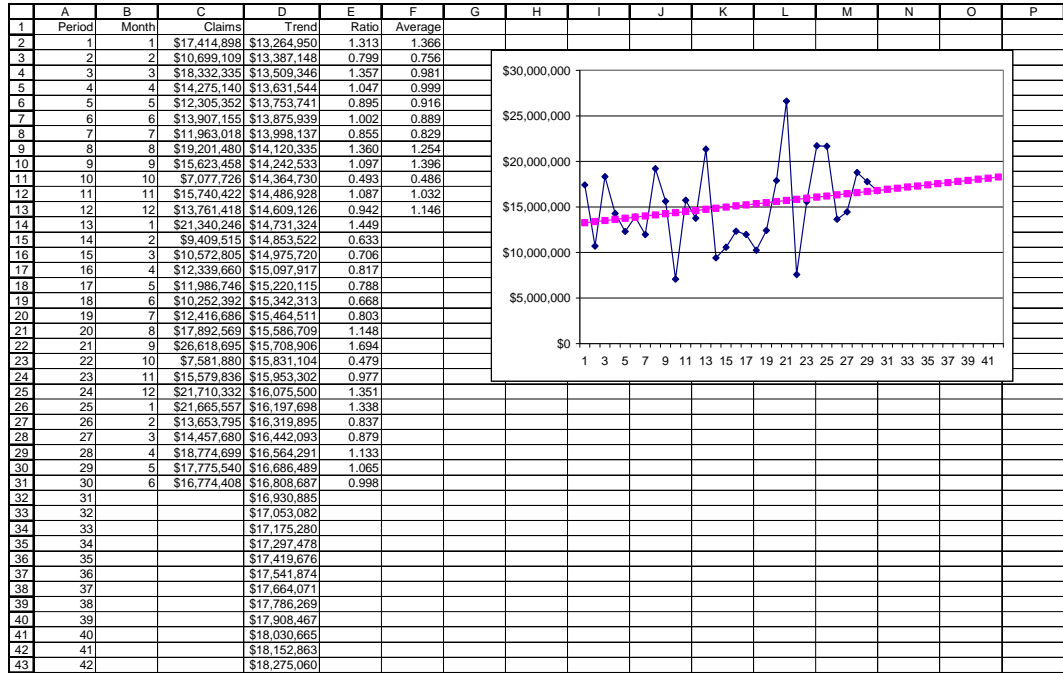


To get a "typical" seasonality ratio for each month, Mike entered =AVERAGE(E2,E14,E26) in cell F2 and copied it to cells F3:F13.

The result is that cell F13 contains =AVERAGE(E13,E25,E37). The function is appropriate because it ignores the contents of empty cells (the equivalent of Clear Contents), like cell E37

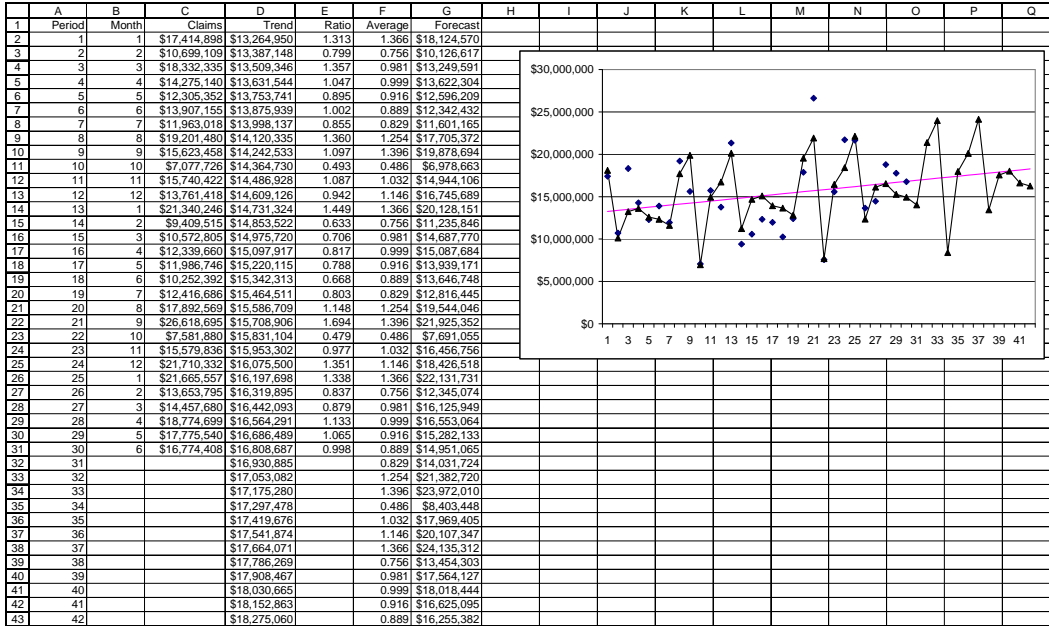
The resulting Average Ratios will be used to adjust the forecast of trend.

Figure 14.12 Extrapolate Trend



To extrapolate trend, Mike selected cells D32:D43 and array-entered `=TREND(C2:C31,A2:A31,A32:A43)`. Here it's necessary to include the X values as function arguments. The Line chart uses data series C2:C43 and D2:D43 with default X values. Stephanie should use her judgment regarding whether a linear extrapolation of trend is appropriate.

Figure 14.13 Include Seasonality



Mike selected the seasonal averages in F2:F13 and used Paste Special | Values into F14:F25, F26:F37, and F38:F49. In cell G2 he entered =D2*F2 and copied to G3:G43. The Line chart shows actual values C2:C43 as points, fitted and extrapolated trend values D2:D43 as a line without markers, and fitted and extrapolated forecast values (trend and seasonal components) G2:G43 with both markers and lines.

Figure 14.14 Residuals

