Forecasting with Seasonality

Dr. Ron Lembke

Sept 25, 2015

Forecasting with seasonality and a trend is obviously more difficult than forecasting for a trend or for seasonality by itself, because compensating for both of them is more difficult than either one alone.

There are other methods a person could find to use for taking into account both a trend and seasonality, but the approach we will follow is the following:

1. Estimate the amount of seasonality - the seasonal relatives (or factors or indices)
2. Estimate the trend (the rate demand is growing at)
3. Make a straight-line prediction of future demand
4. Adjust straight-line projection for seasonality to get a seasonalized forecast

Unfortunately, as we will see, we can't just throw all the data into linear regression and see what comes out. Linear Regression finds a line of best fit based on minimizing the sum of squared errors. With seasonal data, some points will be very far away from the trend line, which can skew the trend line, and give a low R^2 value, which will result in us not having a lot of confidence in the Linear Regression values that we would get. Since we are concerned with seasonal demand here, we need to take it into account it in our considerations.

We will use the following terminology:

F_i	Forecast of demand in period i.
A_i	Actual demand in period i.
forecast was made at time t.	

1 Estimating Seasonal Relatives

To get an estimate of the seasonal relative for each month (or quarter, week, etc., depending on the data), we need to first talk about seasonality. Seasonal demand has a pattern that repeats. Demand for clothing has a seasonal pattern that repeats every 12 months. Some companies may analyze annual seasonal patterns quarterly. The months (or quarters or weeks, etc.) we will refer to as **periods**. Demand with an annual seasonal pattern has a **cycle** that is 12 periods long if the periods are months, or 4 periods long if the periods are quarters. There could also be seasonality on a smaller time scale, like per week. If you think about the sales at restaurants on campus, Fridays are probably slower days that Monday through Thursday. And there is even a recurring pattern to sales throughout the day.

A seasonal relative (also known as a seasonal index or seasonal factor) is how much the demand for that particular period tends to be above (or below) the average demand. So to get an accurate estimate of this, we have to get some kind of average for the demand in the first period of the cycle, and the second period, etc., and then compare these average demands per period to some kind of overall average demand. This might be the average demand for the past cycle, or over multiple cycles.

Suppose that your demand varies over the course of the year, but that there is no long term growth in demand. Suppose we are concerned about how demand goes up or down by quarter. The way we will look at it, we will ask how much the first quarter's demand is higher or lower than the average, and ask the same about the other quarters. To do that, we will compute something called a seasonal relative. An value of 1.00 means that the demand for that period is exactly the same as the average. An value of 1.10 means the demand is 10% higher than the average, and an value of 0.85 means this period's demand is 15% lower than the average.

2 Seasonality and a Trend

We will use the data below, shown in Graph 1 as an example. Tthe data clearly has a cycle that is 4 periods long.

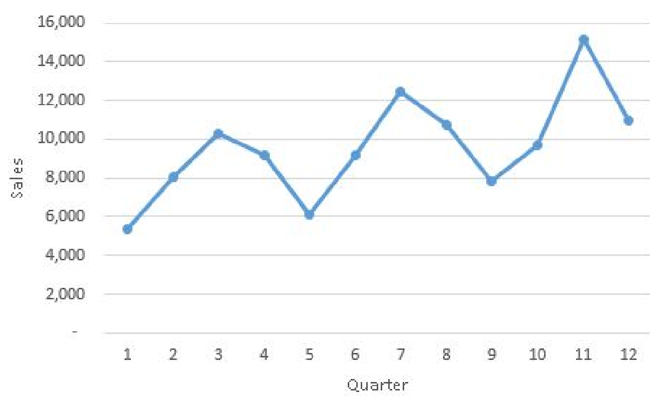

Figure 1: Seasonal Data

The exact sales numbers are

Month	Sales
1	5,384
2	8,081
3	10,282
4	9,156
5	6,118
6	9,139
7	12,460
8	10,717
9	7,825
10	9,693
11	15,177
12	10,990

2.1 Average Method

To estimate the seasonal relatives, we are going to do it by averaging the demands each period, and dividing by the overall average.

Period	Cycle 1	Cycle 2	Cycle 3	Average
1	5,384	6,118	7,825	6,442.3
2	8,081	9,139	9,693	8,971.0
3	10,282	12,460	15,177	12,639.7
4	9,156	10,717	10,990	10,287.7
Overall Average				9,585.17

So we have found what the average demand is for period 1 of a cycle, for period 2, etc. If we divide these averages by the overall average, we get the following seasonal indices:

Period	Period Avg	Overall Avg	Index
1	6,442.3	÷ 9,585.17	=0.672
2	8,971.0	÷ 9,585.17	=0.936
3	12,639.7	÷ 9,585.17	=1.319
4	10,287.7	÷ 9,585.17	=1.073

3 Estimating the Growth Rate

To estimate the growth rate, we will deseasonalize the individual demand numbers to see how demands per period go up or down.

3.1 Deseasonalizing Method

The deseasonalizing method uses the seasonal relatives we already created to try to get a picture of how much we've been growing over time. What is deseasonalizing? After we make a straight-line forecast of the future, we are going to multiply it by the seasonal indices to get a seasonalized forecast of the future.

Deseasonalizing is basically the opposite: we are going to take the actual, seasonal data, and divide it by the seasonal factors to get something that looks more like a straight line. Then we are going to do a linear regression through this line. The deseasonalized demands should be a lot more like a straight line than the original data is, that is, it should generally show a more consistent growth rate than we see with seasonality in it. When we do a linear regression through these deseasonalized points, the linear regression should give us a pretty good fit through the points.

Month	Seasonal Index	Deseas.	
1	5,384	÷ 0.672	= 8,011.90
2	8,081	÷ 0.936	= 8,633.55
3	10,282	÷ 1.319	= 7,795.30
4	9,156	÷ 1.073	= 8,533.08
5	6,118	÷ 0.672	= 9,104.17
6	9,139	÷ 0.936	= 9,763.89
7	12,460	÷ 1.319	= 9,446.55
8	10,717	÷ 1.073	=9,987.88
9	7,825	÷ 0.672	=11,644.35
10	9,693	÷ 0.936	= 10,355.77
11	15,177	÷ 1.319	=11,506.44
12	10,990	÷ 1.073	=10,242.31

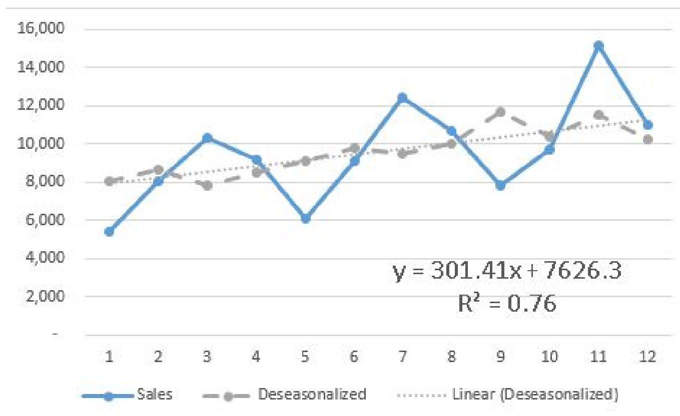

Figure 2: Seasonal Data

If we do a linear regression through these deseasonalized numbers, we get an intercept of 7626.25 and a slope of 301.41, and an R^2 value of 0.76. For comparison, if we did a Linear Regression on the original data, we get $R^2 = 0.39$

4 Straight-Line Projection of Future Demand

Use the linear regression calculated previously to extrapolate out into the future.

Period	Linear Forecast
1	7,927.66
2	8,229.08
3	8,530.49
4	8,831.90
5	9,133.31
6	9,434.73
7	9,736.14
8	10,037.55
9	10,338.96
10	10,640.38
11	10,941.79
12	11,243.20
13	11,544.61
14	11,846.03
15	12,147.44
16	12,448.85

5 Making Seasonalized Forecasts

Multiply the linear forecast you made by the seasonal factors to get a seasonalized forecast.

Period	Linear Forecast	Index	Forecast
1	7,927.66	* 0.672	= 5,327.4
2	8,229.08	* 0.931	= 7,702.4
3	8,530.49	* 1.319	= 11,251.7
4	8,831.90	* 1.073	= 9,476.6
5	9,133.31	* 0.672	= 6,137.6
6	9,434.73	* 0.931	= 8,830.9
7	9,736.14	* 1.319	= 12,842.0
8	10,037.55	* 1.073	= 10,770.3
9	10,338.96	* 0.672	= 6,947.8
10	10,640.38	* 0.931	= 9,959.4
11	10,941.79	* 1.319	= 14,432.2
12	11,243.20	* 1.073	= 12,064.0
13	11,544.61	* 0.672	= 7,758.0
14	11,846.03	* 0.931	= 11,087.9
15	12,147.44	* 1.319	= 16,022.5
16	12,448.85	* 1.073	= 13,357.6

6 Triple Exponential Smoothing

When there is a trend and no seasonality, we used double exponential smoothing, in which we smoothed our estimates of the trend and the intercept in every period. It is possible to take this approach one step further, by smoothing the estimates of the seasonal relatives every time. This is known as Winter's Method. To do this, you first smooth the seasonal factor of the most recent period of demands, then you have to make sure that all of the seasonal indexes add up to the right thing. If there are 4 periods in a cycle, all the factors need add up to 4. If they all add up to 4.2, then each factor needs to be multiplied by 4/4.2.

Then, the estimates of the intercept and the trend need to be revised. However, to properly take the seasonal factors into account, the formulas are more complicated than the ones for double exponential smoothing. As you would guess, this gets pretty complicated pretty fast.

Statistical studies show that there is, in theory, some value to be gained from using such a complicated method. However, the actual gains are likely to be small for any real world usage.