

Modeling of Cyclical (Seasonal) Data

Many times in finance the historical data that needs to be modeled and/or used as a basis for a forecast contains cyclical patterns. Examples of this would be the seasonal sales cycle of a retail business, or the longer-term business cycle of an industry, or perhaps an operations-based cycle. These cycles can vary from a partial year to spanning double digit years, but whatever the duration of the cycle, provided that it is consistent in its repetition, the cycle can be modeled using a standardized set of procedures. This post references an Excel workbook model (attached below in this post), and the two should be used in unison to learn this procedure.

In the attached Workbook please goto the "Created Data Set" worksheet. In this worksheet I created a three year set of data (Column K) by applying a periodic (monthly) growth rate of 0.6% and an assumed annual cycle (Column F) to a base sales value of \$100. Notice that the dataset created looks quite random at first glance.

The created dataset is then analyzed in the Worksheet called "Develop Forecast Model". There are three steps used in developing the model:

1. Graph all of the data and identify if a cycle exists, and if one does exist, then characterize what the typical cycle looks like in terms of start, pattern and duration. When plotting the dataset created, an annual 12 month seasonal pattern was observed.
2. Normalize the dataset by removing the growth rate experienced over the life of the data. This results in a flattening of the growth trend of the cycles, but the cycles themselves remain fully intact. The growth that was removed will be a linear growth curve, and this will serve as the baseline of sales, with the slope matching the CAGR of sales growth. This growth curve (rate) can later be adjusted due to a top-down forecast based on economic conditions, or perhaps a bottom-up forecast based on future marketing campaigns or projections from your sales pipeline.
3. Cut the remaining cycles into their appropriate segments based upon the normally observed durations, and then average each period across the various cycle segments. In this example, and for an annual seasonal cycle, you will end up with 12 averaged period values that comprise the annual seasonal cycle. Next normalize the averaged period cycle values by dividing them by the first averaged period value in the cycle. The first normalized value will always be 1 (unity), with the averaged period cycle values representing multiplication factors to be applied against the linear growth curve derived in Step 2.

You now have the elements characterized to build a model that can then be extended out for a forecast. The worksheet called "Run Forecast Model" builds the forecast by starting with a seed sales value, growing this out with the linear growth rate (CAGR) obtained in Step 2 above, and then multiplying these values with the seasonal multiplication factors derived in Step 3 above. Another way to model the seasonal multiplication factors is to curve-fit the factors into a polynomial equation, and then using this polynomial equation to derive the multiplication factors. Both methods have been outlined in the "Develop Forecast Model" and "Run Forecast Model" tabs. A hidden benefit of this model is that any CAGR can now be supplied, and the seasonality then applied after the growth rate - it is a very flexible model.

Error rates were calculated by running the forecasts using both modeling techniques (i.e., ratio and polynomial), and these forecast values were then compared against the original sales model. Absolute average error rates of 2% and 8.5% were experienced for the ratio-based and polynomial-based models respectively. These errors rates were traced back to a slight error in the base sales growth, with the majority of the error being caused by offsets in the seasonal multiplication factors. An error rate of only 2% is excellent when considering that this is for a rather volatile cyclical model. The reason that a polynomial-based model may be preferred to a ratio-based model

despite the higher error rates would be if inter-period values are required. In that case a spline-model or linear interpolation off the ratio-based model would be a more accurate and preferable method of estimating inter-period multiplication factors. Future posts will address how to create spline functions and linear-interpolation in Excel.

©2018 Ben Etzkorn