Are you reacting to outliers identified in your forecast model?

Posted on Sep 20, Posted by Tom Reilly Category Forecasting

You should be. There is information to be MINED in the model. Macro conclusions can be made from looking at commonalities across different series(ie 10% of the SKUs had an outlier four months ago---ask why this happened and investigate to learn what you are doing wrong or perhaps confirm what you are doing right!...and perhaps the other 90% SKUs also had some impact as well, but the model didn't detect it as it was borderline. You could then create a causal variable for all of the SKUs and rerun and now 100% of the SKUs have the intervention(maybe constrain all of the causals to stay in the model or lower the statistical test to accept them into the model) modeled to arrive at a better model and forecast. Let's explore more ways to use this valuable information:

LEVEL SHIFTS

When hurricane Sandy hit last October, it caused a big drop for a number of weeks. Your model might have identified a "level shift" to react to the new average. The forecast would reflect this new average, but we all know that things will return, but the model and forecast aren't smart enough to address that. It would make sense to introduce a causal variable that reflected the drop due to the hurricane, BUT the future values of the causal would NOT reflect the impact so the forecast would return to the original level. So, the causal would have a lot of leading zeroes, and 1's when the impact of Sandy was felt and 0's when the impact would disappear. You could actually transition the 1 to a 0 gradually with some ramping techniques we learned from the famous modeler/forecaster Peg Young of the US DOT . The 0 dummy variable might increment like this
10,0,0,0,0,0,0,,1,1,1,1,1,1,1,.9,.8,.7,.6,.5,.4,.3,.2,.1,0,0,0,0,0,0,etc.

OUTLIERS

When you see outliers you should be reviewiing them to see if there is any pattern to them. For example, if you don't properly model the "Super Bowl" impact, you might see an outlier on those days. It takes a little time and effort to review and think "why" does this happen. The benefits of taking the time to do this can have a powerful impact. You can then add a causal with a 1 in the history when the Supewr Bowls took place and then the provide a 1 for the next one. For monthly data, you might see a low June as an outlier. Don't adjust it to the mean as that is throwing the baby away with the bath water. This means you might not be modeling the seasonality correctly. You might need an AR12, seasonal differencing or seasonal dummies.

SEASONAL PULSES

Let's continue with the low June example. This doesn't necessarily mean all months have seasonality and assuming a model instead of modeling the data might lead to a false conclusion for the need of seasonality. We are talking about a "seasonal pulse" where only June has an impact and the other months are near the average. This is where your causal dummy variable has 0's and a 1 on the low Junes and also the future Junes(ie 1,0,0,0,0,0,0,0,0,0,0,0,1).

Tags: Tagged in: [outliers fox tiao tsay box box-jenkins](#)   [outliers sas](#)   [time series box-jenkins acf pacf level shift plot](#)

[time series forecasting trends level shifts seasonality outliers](#)