Paper SAS1833-2015

Strengthening Diverse Retail Business Processes with Forecasting: Practical Application of Forecasting Across the Retail Enterprise

Alex Chien, Beth Cubbage, Wanda Shive, SAS Institute Inc., Cary, NC

ABSTRACT

In today's omni-channel world, consumers expect retailers to deliver the product they want, where they want it, when they want it, at a price they accept. A major challenge many retailers face in delighting their customers is successfully predicting consumer demand. Business decisions across the enterprise are affected by these demand estimates. Forecasts used to inform high-level strategic planning, merchandising decisions (planning assortments, buying products, and pricing, allocation and replenishment of inventory) and operational execution (labor planning) are similar in many respects. However, each business process requires careful consideration of specific input data, modeling strategies, and output requirements.

Learn how leading retailers are increasing sales and profitability by operationalizing forecasts that improve decisions across their enterprise.

INTRODUCTION

Achieving excellent customer satisfaction while maximizing sales and profits is a universal goal of retailers. For decades, retailers have managed their businesses by relying on estimates of future sales.

Key operational decisions influenced by these future sales estimates or forecasts include:

- Inventory – goal of maximizing inventory turns
  - How much inventory should be bought or manufactured?
  - When should inventory be purchased?
  - How much inventory does each selling location need?
  - When should additional orders be submitted?
- Pricing – goal of maximizing sales and gross margins
  - How should products and offerings be priced?
  - When should promotions occur and at what price?
  - When should products be marked down and by how much in order to clear out excess inventory?
- Distribution & Labor Planning – goal of minimizing operating costs
  - To which selling locations should products be shipped and when?
  - How many associates should be scheduled to work during each operating hour?

While very related in many ways, the frequency at which these business decisions are made and the level of detail required can be dramatically different. Thus, there is not a one-size-fits-all forecasting approach for these various business decisions. Determining the most effective forecasting approach or modeling strategy depends heavily upon the output requirements and the nature of the historical input data.

Start with the end in mind. First, one needs to describe the data elements of the forecast output. What does the required forecast output need to look like? The answer lies in understanding three key factors related to the output requirements:

- Forecast Granularity – What level of operational detail is required?
  - Product (for example, All products, Department, Category or Class, Sub-Category or Sub-Class, SKU, or Item)
  - Location (for example, All locations, Country, Channel, Promotional or Price Zone, Distribution Center, or Store)
  - Time interval (for example, Year, Quarter, Period, Week, Day, Hour, or Minutes)
- Forecast Horizon – How far into the future should forecasts be generated?
  - Year(s), Quarter(s), Period(s), Week(s), Day(s)

- Forecast Features – What major demand influences should be considered?
    - Price sensitivity
        - Effect of everyday price changes
        - Effect of promotions
        - Effect of markdowns
    - Key events
        - Company-wide (for example, anniversary sales, Back-to-School)
        - Location specific (for example, events associated with local charities)
        - Product category specific (for example, white sale, seafood extravaganza)

Once the above data output questions are answered, the next critical consideration is understanding historical sales results or input data. Key characteristics of the input data that influence which forecasting approach or modeling strategy works best include:

- Amount of historical sales data – short or long
- Pattern – consistent or intermittent (sparse) sales
- Seasonality – seasonal or non-seasonal products
- Volume – high, average, low
- Data Quality – outliers, zero price, accuracy of historical promotions

This paper presents two examples of how different modeling techniques are used to achieve fit for purpose forecasts for inventory replenishment and for enhancing clearance pricing or markdown decisions. Leading retailers who leverage these best practice approaches enjoy sales increases and improved profitability.

REPLENISHMENT FORECASTING – A MULTI-STAGE APPROACH

Store or location replenishment forecasts are typically one of the lowest level product forecasts needed by retailers. The retailer's overall replenishment goal is to minimize over stocking and limit out-of-stocks of every product for each selling location (maintain adequate inventory) while maximizing sales and profits. To do so, a reasonably accurate replenishment forecast is required.

Sales for a specific product can greatly vary from one selling location to another. Sporadic sales across time, including weeks of zero sales, within the same selling location further complicates forecasting weekly sales of a particular product for a given store. These challenges require special attention in order to achieve a reasonable replenishment forecast.

Specifications of the required forecast output of the example in this paper are as follows:

- Forecast Granularity: by Product or SKU, by Store, by Week
- Forecast Horizon: 12 weeks of future forecasts
- Forecast Features: include price sensitivity and events

As previously mentioned, the granularity of a forecast can significantly affect the best overall modeling approach. The following SKU-Store-Week forecast example explains how two forecasting methods are combined to provide a more accurate replenishment forecast.

INPUT DATA

Typically, at least two years of historical input data is needed for replenishment forecasting. In this example, the input data consists of two years of unit sales history by product by store by week.

To explain this replenishment forecasting approach, we draw attention to the following portion of data:

- 1 Product Category: 100
- 4 Products or SKUs: A, B, C and D
- 2 Selling Locations: X and Y (of 40 locations included in the input data)

VISUALIZING HISTORICAL UNIT SALES (INPUT DATA)

To determine the best forecast modeling strategy, we need to understand the demand patterns of our historical data. We first explore the unit sales patterns for all products within category 100 across all stores. Total unit sales volume for Category 100 for the 104 weeks of history totaled almost 168,000 units. The 40 stores included in this example average selling just over 1,600 units of Category 100 during this historical period or slightly over 40 units per store per week.

In Figure 1 below, we see a significant seasonal pattern with weekly sales peaking at around 3,000 units across all stores.

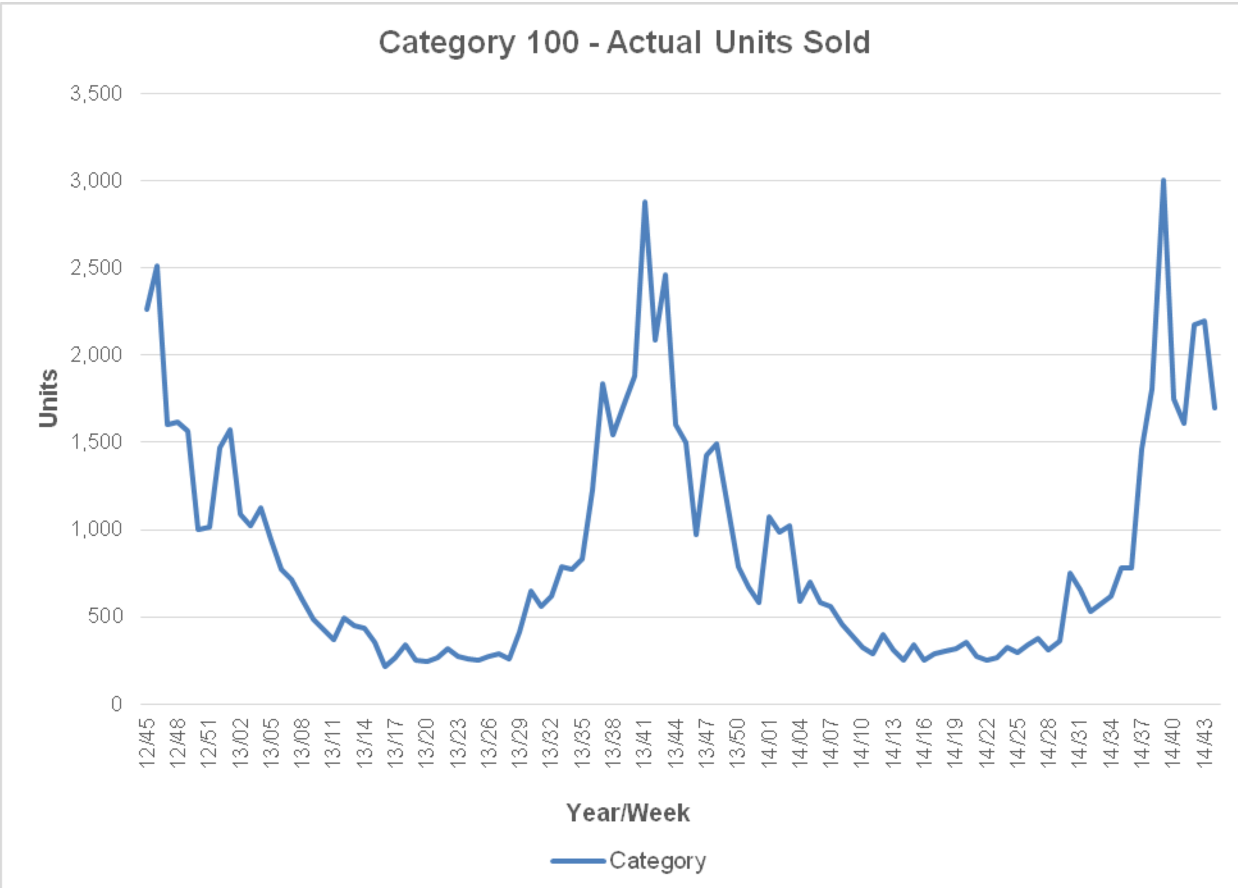

Figure 1 – Product Category 100 (Two Years of Historical or Actual Unit Sales for the Total Category for All Stores)

Next we analyze the unit sales pattern for each of the four products included in Category 100. In Figures 2 – 5, we see substantial variations in selling patterns among the four products.

Product A accounts for just over 38,000 of Category 100's unit volume, or 22.5% of total category unit sales, during the 104 weeks of history. In Figure 2, we see some variation in the unit sales per week. However, the volume peaks for Product A are shifted to later weeks than the high volume weeks for the overall category. In Figure 1, we see the total category units sold peaks in week 41 of 2013 and week 39 of 2014. Conversely, Figure 2 below shows that Product A's units sold are highest in weeks 2 and 3 for both 2013 and 2014.

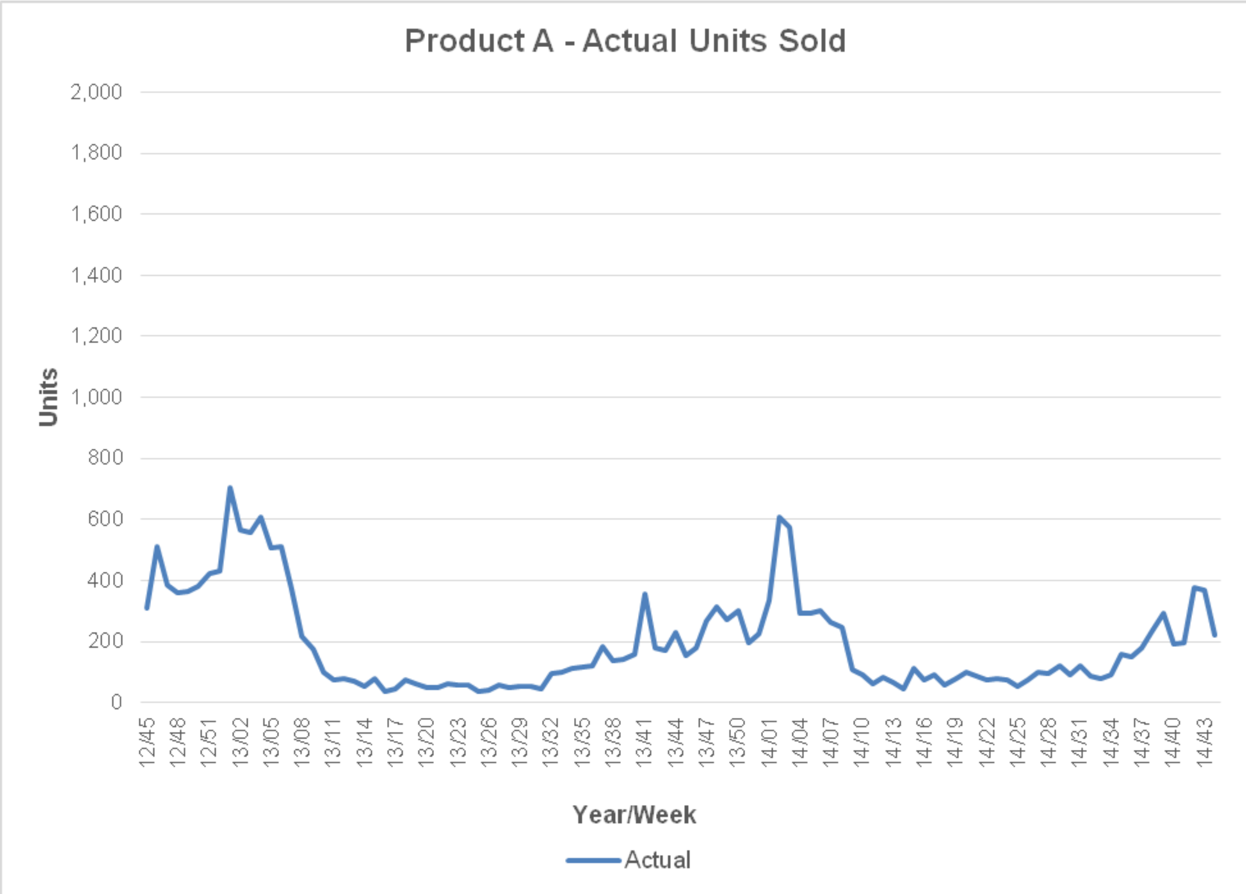

Figure 2 – Product A (104 Weeks of Historical or Actual Unit Sales for All Stores)

As we continue to review the historical unit sales of each of the four products within Category 100, we see other patterns emerge. In Figure 3 below, we see sales by week for Product B. This product contributes the lowest overall unit volume within the category. Product B accounts for about 35,000 units which is 21% of total Category 100 unit sold. The sales pattern across weeks is even less defined for Product B than the pattern we saw for Product A above.

Figure 3 – Product B (104 Weeks of Historical or Actual Unit Sales for All Stores)

Moving to the third product of Category 100, in Figure 4, we see that Product C has a similar seasonal sales pattern to the total category unit volume by week seen in Figure 1. Product C contributed the highest unit volume sales (almost 58,000 units or 34.5% of the total category volume) among the four products included in Category 100.

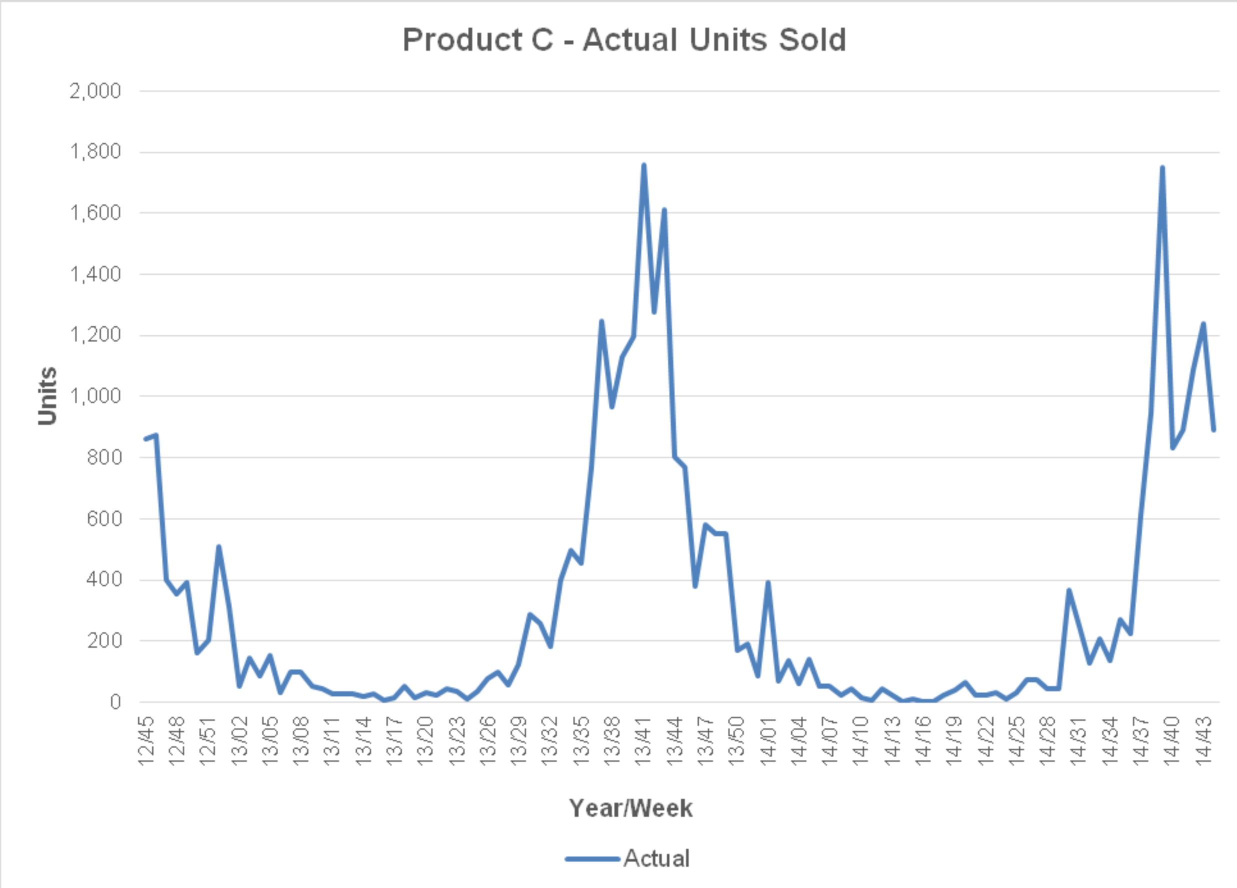

Figure 4 – Product C (104 Weeks of Historical or Actual Unit Sales for All Stores)

Finally, for Product D we see yet another sales behavior displayed. A seasonal pattern that aligns with Product C and the overall category, but the sales is the third lowest of the four products observed. Product D accounts for just over 37,000 of Category A's unit volume, or 22% of total category unit sales, during the 104 weeks of history.

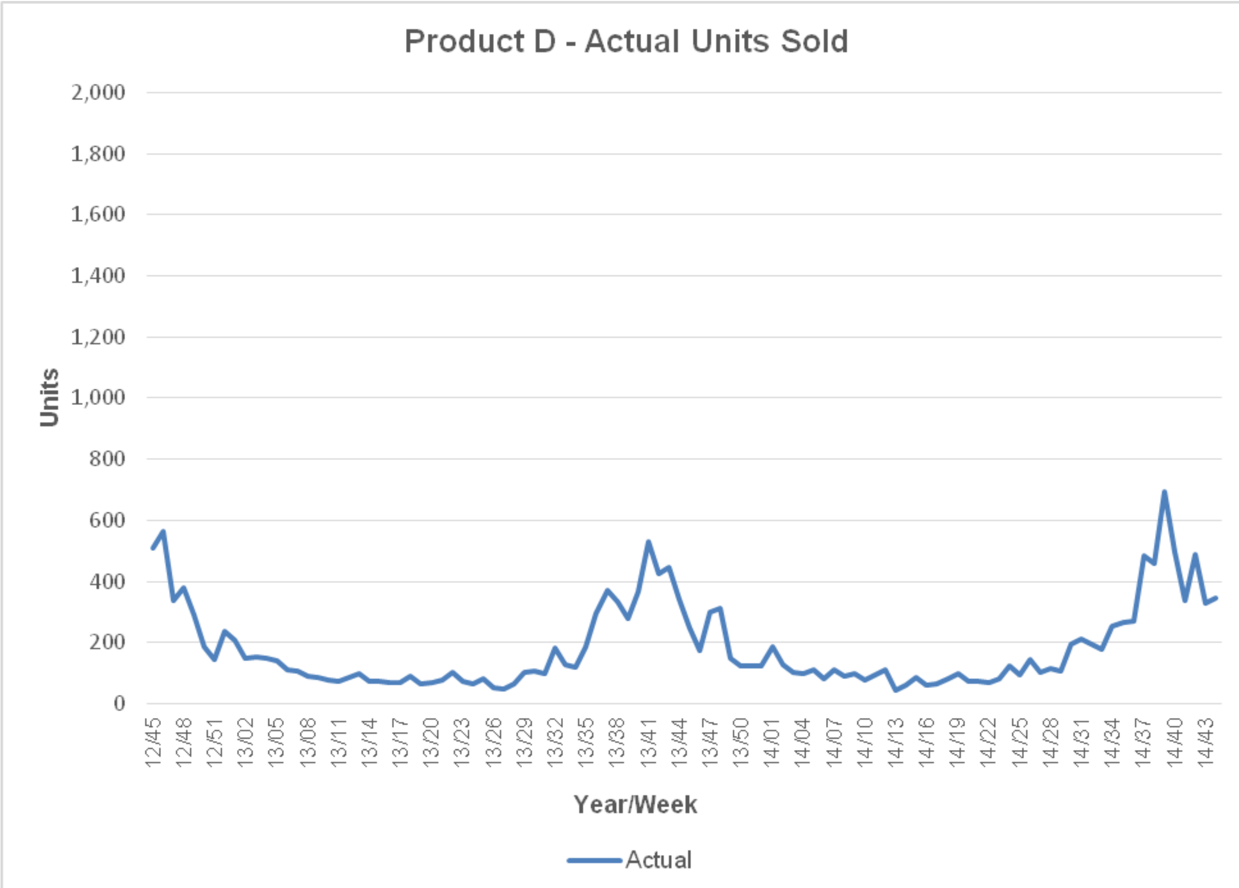

Figure 5 – Product D (104 Weeks of Historical or Actual Unit Sales for All Stores)

Because our required output for this replenishment forecast example is for each SKU-Store-Week combination, understanding store level demand is critical. Thus, after reviewing units sold for each of the four products within Category 100 across all stores, we now turn our attention to understanding sales variations among stores. As mentioned previously, our example input data contains 40 stores' unit sales by product, store, and week. In Figures 6 – 9, we will observe individual store sales patterns for Products B and C for two stores to demonstrate the varying demand patterns by store.

Remember from our earlier discussion about Figure 3, Product B has the least defined unit sales pattern of the four products within Category 100. When we dig further into the details of how Product B sells at individual stores, we see even more erratic behavior. We compare the sales of Product B in Store X to the sales of the same product in Store Y. Store X sold 921 units of Product B during the 104 weeks of history – that's almost 9 units per week on average. Store Y sold 122 units of Product B during the same time period – that's just over 1 unit per week. We start our comparison by looking at Product B unit sales at Store X, in Figure 6.

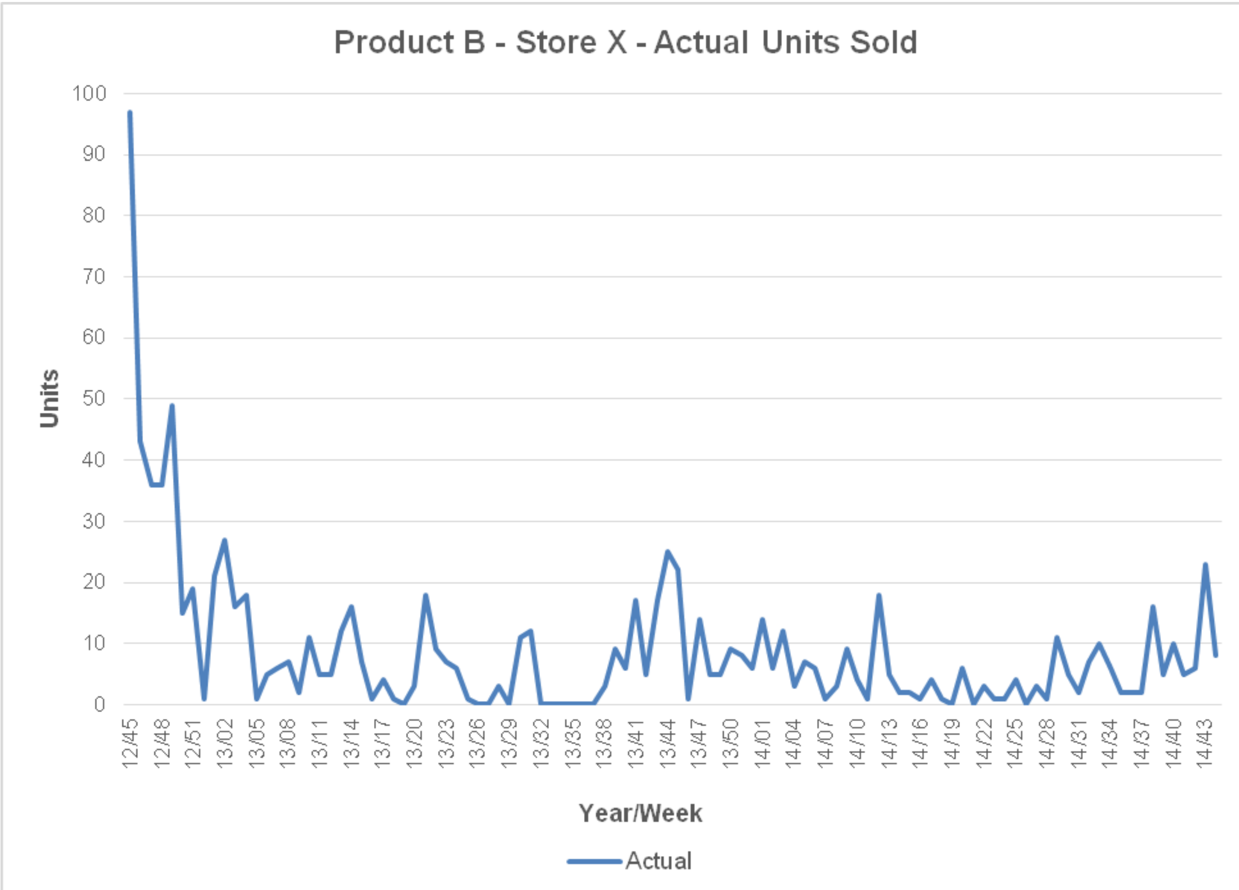

Figure 6 – Product B for Store X only (104 Weeks of Historical or Actual Unit Sales)

In Figure 6, we see the units sold of Product B at Store X by week. Notice that units sold at Store X peaked in the early weeks of the 104 week history. In fact, 41% of Store X's volume occurred in the first twelve weeks of its history. The following 92 weeks' demand pattern is rather low and inconsistent.

In Figure 7, we see even more sparse data in the same Product B for the different selling location, Store Y. This store sold only 122 units during the historical period. Store Y sold zero units during 43 of the 104 weeks and only one (1) unit during 30 other weeks. That means Store Y sold one or less units in 70% of total historical period. We consider this demand pattern very intermittent.

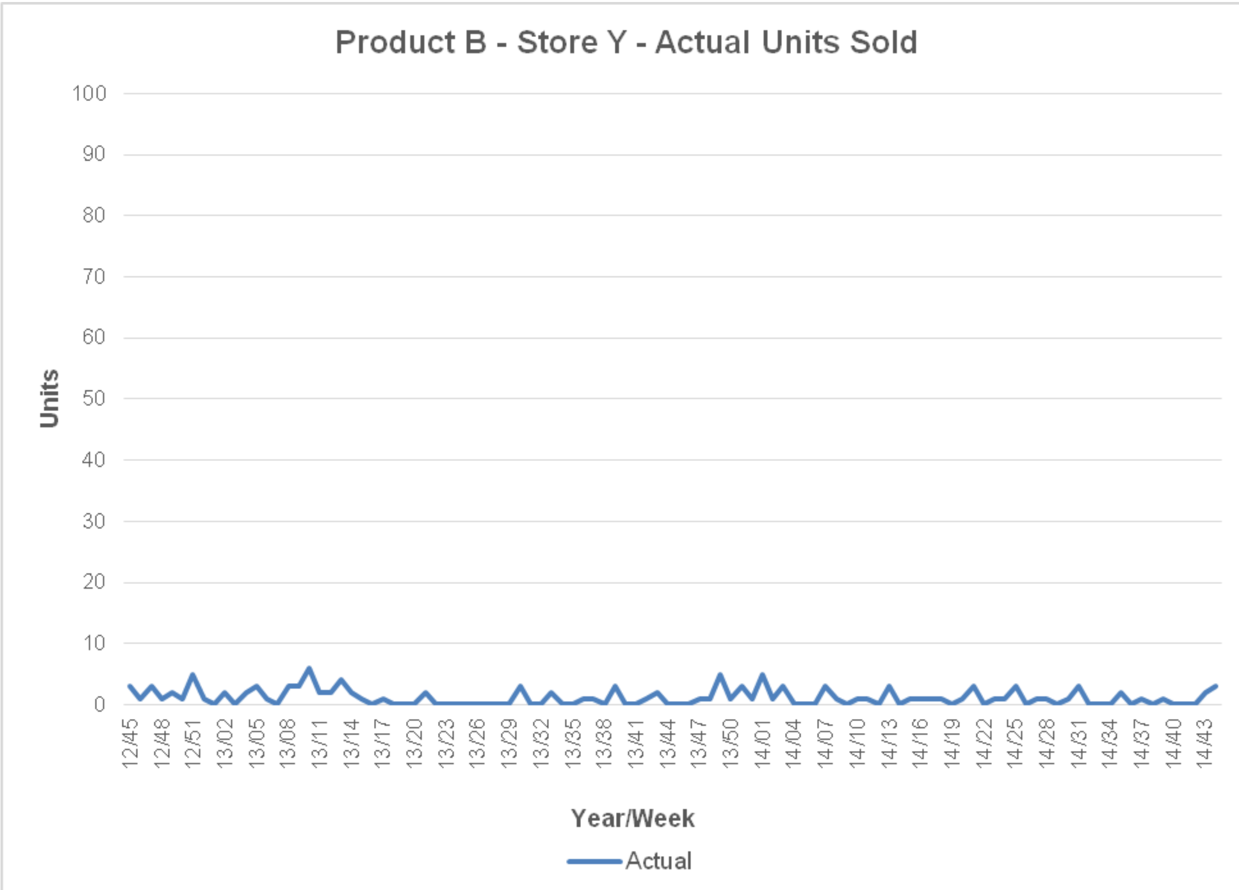

Figure 7 – Product B for Store Y only (104 Weeks of Historical or Actual Unit Sales)

Next we analyze two stores' unit sales for Product C, the highest selling volume product in Category 100.

In Figure 8, we observe sales activity of Product C for Store X. In this example, we see dramatic variations in sales demand across the weeks. Strong promotional activity is evident near the end of the historical period. The Product C unit volume per week for Store X ranges from none (0) in 30% of the 104 weeks to a peak of 530 during the week of the promotion.

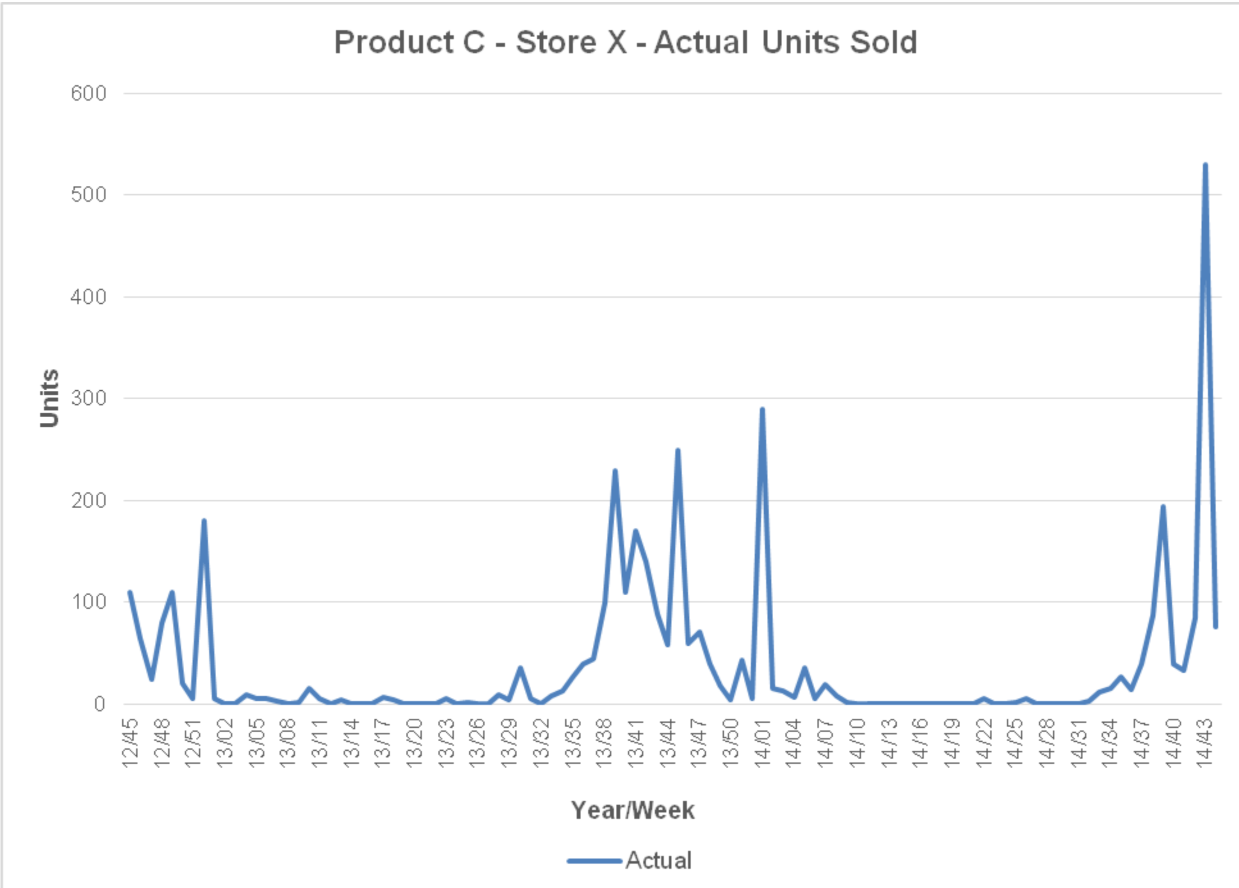

Figure 8 – Product C for Store X only (104 Weeks of Historical or Actual Unit Sales)

Units sold at Store Y paint yet another picture. In Figure 9, we find seasonal intermittency. Store Y sold no units of Product C in 64 or 61.5% of the historical weeks and the intermittent demand that exists appears to be seasonal in nature.

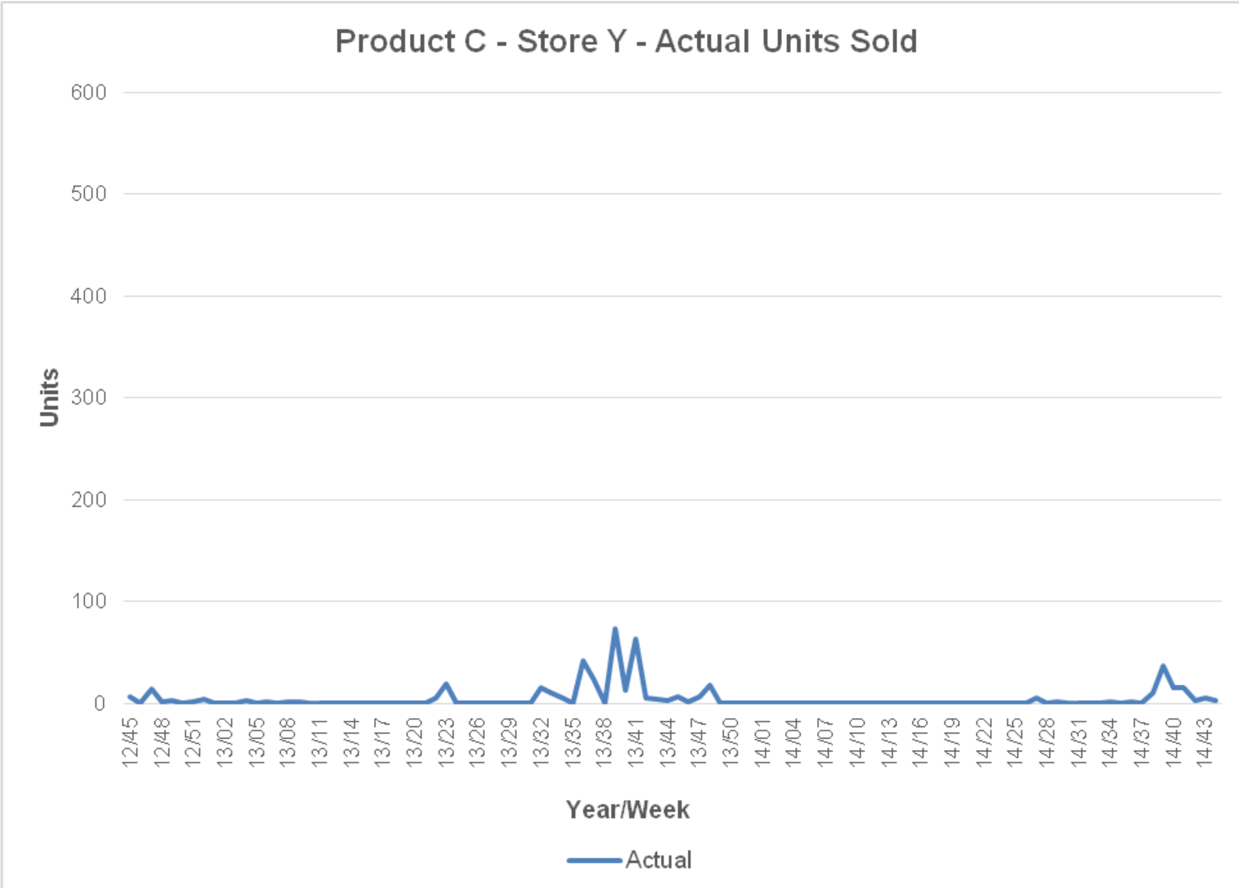

Figure 9 – Product C for Store Y only (104 Weeks of Historical or Actual Unit Sales)

MODELING STRATEGY – A MULTI-STAGE APPROACH

Based on the visualizations previously reviewed, we next consider which forecast modeling strategies to apply in order to achieve the most robust sales forecasts. We must evaluate which approach handles the most variations between products and stores observed in Figures 1-9. While doing so, we keep in mind all output requirements, including estimating price sensitivity of future promotions.

At the aggregate level of Category 100 (see Figure 1), the demand pattern is rather well behaved. We expect a time series model to operate most effectively in achieving reasonable future sales estimates for the overall category.

Conversely, at the lower level of specific product (SKU) and individual store (see Figures 6-9), we found intermittency or sparse data. We also found vast differences in price sensitivity features needed for estimating the effect of future promotions. These low-level data do not on their own provide a strong enough signal to estimate price sensitivity. Because of the intermittency and sparse data at the lower level, a typical time series model will not be effective. We should consider cross-sectional models to pool information from multiple products, multiple stores, or both in order to obtain robust estimates of the individual SKU-Store features. One example of such models is regression.

The process below describes a multi-stage approach to generate a solid SKU-Store-Week replenishment forecast that handles all of the output requirements in our example. This multi-stage approach incorporates both time series and regression models.

Stage 1

In Stage 1 of our multi-stage approach our goal is to generate the best forecast possible for each of the four products across all stores.

As our first step, we apply time series forecasting models to generate a forecast at the category level – Category 100 which includes all four products across all 40 stores. A time series forecast model works well for this aggregated, well behaved data (see Figure 1). This time series model captures the general trend, seasonality, and price sensitivity across the entire category.

Next, we generate forecasts for each of the four products for all stores that reflect the varying demand patterns of Products A, B, C and D. Once again, we use times series models to produce these forecasted unit sales by product aggregated across all stores.

In Figure 10, we see the results from our first two steps. The time series forecast for Category 100 estimates selling 13,702 units over the 12 week forecast horizon, while the sum of the individual product forecasts suggests 14,116 will be sold.

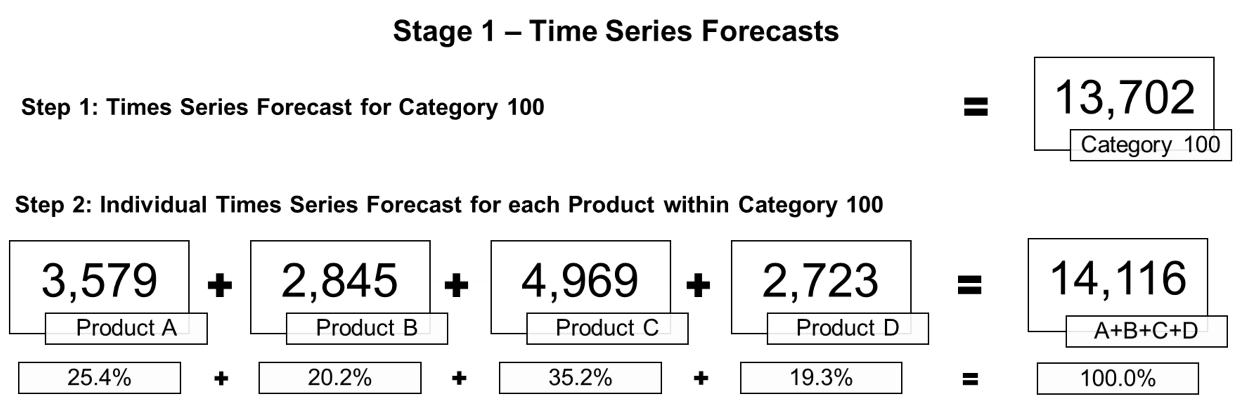

Figure 10 – Time Series Forecasts for Total Category 100 and Each Individual Product

In the third step of our process, we perform top-down reconciliation from the category level forecast (Step 1) to calculate a reconciled product level forecast for each of the four products (Figure 11). For example, the Product A reconciled forecast of 3,474 is generated by multiplying its 25.4% share of the total category forecast (from Step 2) times the total category forecast of 13,702 (from Step 1).

Stage 1 – Top-Down Reconciliation (Step 3)

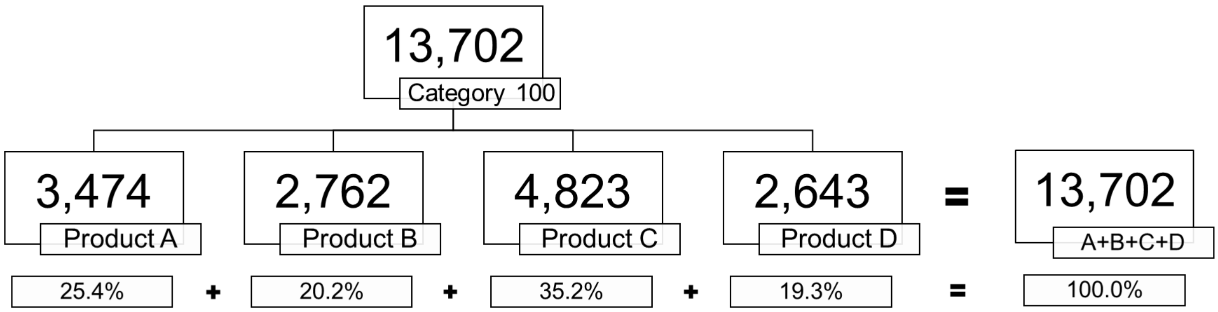

Figure 11 – Top-Down Reconciliation (Step 3)

Next, we compare the independently generated time series forecast for each product (Step 2) to the reconciled forecast by product (Step 3) to determine which forecast results in the lowest forecast error during the holdout period. The best performing forecasts are highlighted in Figure 12. We see that the individual time series forecast for Product C is selected because it out performs the reconciled forecast for Product C. The reconciled forecast for the other three products performed best during the holdout period.

Stage 1 – Product Forecast Comparison

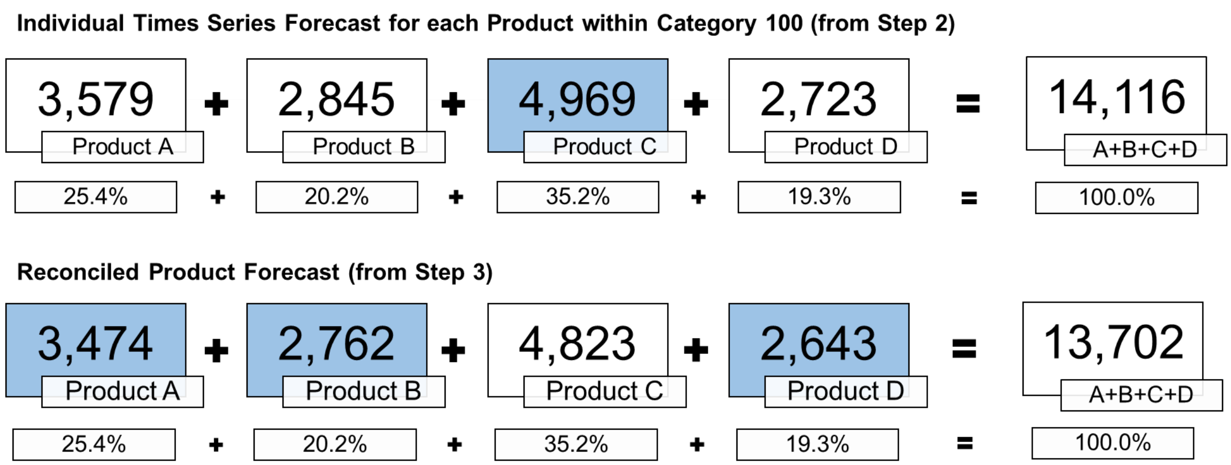

Figure 12 – Product Forecast Comparison

The final output of Stage 1 is the best forecast for each of the four products within Category 100 aggregated across all stores as shown in Figure 13. Notice that the final forecast for all four products is now 13,848 units.

Stage 1 – Final Product Forecast

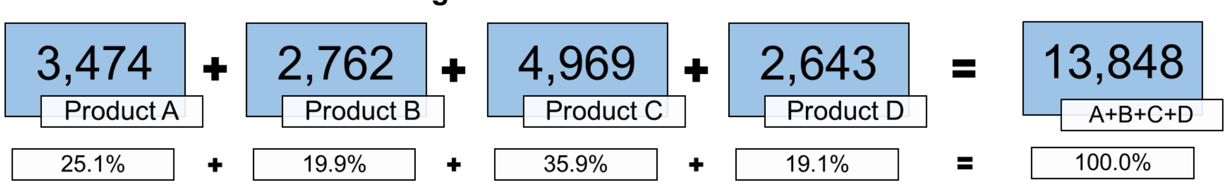

Figure 13 – Final Product Forecast

Stage 2

In Stage 2, our goal is to generate the best forecast possible for each SKU-Store combination.

With highly intermittent data, price and promotion lift estimates at the lowest level are nearly impossible without data pooling.

Historical data of Products A, B & D contained little to no evidence of price or promotion effects in the past. However, we can infer the price sensitivities for Products A, B & D using historical price and promotion effects from Product C.

We apply cross-sectional models to the granular level of by Product, Store, and Week historical data. During this process, data is pooled or accumulated across SKU-Store-Week combinations in order to extract features that could not be recognized within each individual series. We use regression techniques to capture the price and promotional effects by product and by store. Figure 14, shows regression forecast results for Store X and Store Y, which are two of the 40 stores in our example.

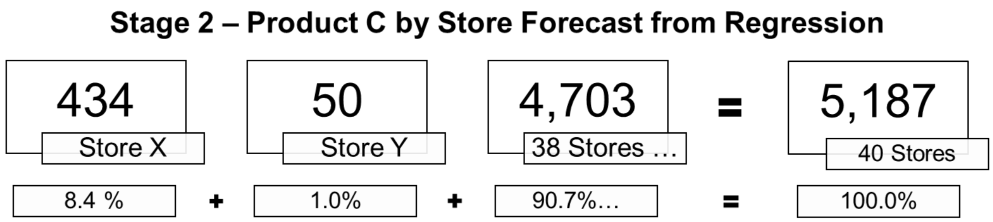

Figure 14 – Product C by Store Forecast from Regression

Next, we perform another top-down reconciliation by using the final product level forecast of Product C from Stage 1 and the regression forecast from Stage 2 to generate the final SKU-Store-Week replenishment forecast. Figure 15 shows the final reconciled forecast for Stores X and Y. For example, the reconciled forecast of 416 for Store X is generated by multiplying its 8.4% share of the total Product C forecast from the regression results above. That value is multiplied by the total final Product C forecast of 4,969 for all 40 stores from Stage 1 to get 416.

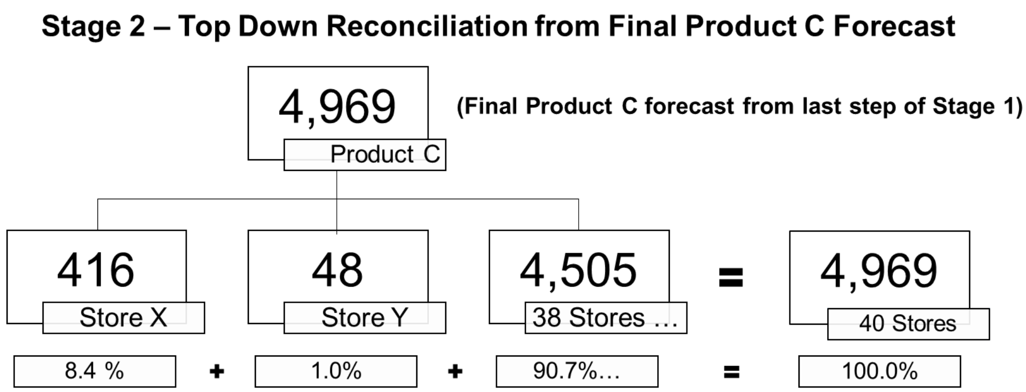

Figure 15 – Top Down Reconciliation from Final Product C Forecast to Final SKU-Store Forecast

In summary, weekly SKU-Store specific price and promotional features are nearly impossible to extract with time series procedures due to intermittency and sparse data. We use cross-sectional models, regression in this case, to capture price and promotional effects across pooled SKU-Store-Week data. In this way, we benefit from both the strength of a solid product forecast using time series procedures and the power of understanding those effects at the lowest level from regression. These in turn generate a robust final forecast that satisfies the output requirements.

The multi-stage modeling strategy demonstrated above can be achieved by leveraging SAS® analytics such as SAS® Forecast Server, SAS® Enterprise Miner, and SAS® High Performance Analytics.

In replenishment forecasting, it is important to get both reasonably accurate forecasts and strong estimates of price sensitivity. It is even more critical to obtain a precise price sensitivity measure for clearance pricing. In the following section, we use examples and simulations to demonstrate the importance of the price sensitivity estimate.

FORECASTING TO SUPPORT CLEARANCE PRICE DECISIONS

Price optimization systems utilize demand forecasts and price elasticity estimates to recommend optimal pricing strategies. In such a system, the forecast is an input to an optimization problem. Other inputs to the problem include price elasticity, price rules, and price optimization goals (such as Minimize Inventory at Risk or Maximize Revenue and Margin). Forecasts are often evaluated by reviewing holdout forecast results, which compare forecasts for the most recent six- to eight-week period to the actual outcomes for that period. Demand models are then tuned based on this evaluation. If forecast error is very high, the models may be tweaked. Forecast accuracy is easy to understand and measure, so some retailers may focus on it almost exclusively. However, optimization simulations indicate that there are far more benefits from estimating the right price elasticity than from getting a very low forecast error. In this section of the paper, we use SAS Markdown Optimization to show improvements from accurate price elasticities. We also discuss several approaches used by SAS to obtain the best possible elasticity estimates.

PRICE ELASTICITY OVERVIEW

Before detailing the analysis described, this section defines price elasticity and provides an example of demand changes using three fictional products, each with a different estimated price sensitivity. *Price elasticity* measures the rate of response of quantity demanded when price is changed (Chvosta et. al. 2013). To estimate price elasticity rates, an analyst typically uses regression techniques on a data set that contains past prices and quantity sold. Because demand usually increases when price decreases, the elasticity rates are negative. The more negative the elasticity rate, the more price sensitive the product. For example, consider three products that are each expected to sell 100 units at regular price in a given week, but have different price sensitivities. As shown in Figure 16, if a 25% markdown is offered for each product in that week, we estimate that each product would sell a different number of additional units.

Product	Price Sensitivity	Elasticity Estimate	Additional Units Sold with 25% Markdown	Total Estimated Units Sold
1	Low	-1.5	45	145
2	Medium	-2.5	87	187
3	High	-3.5	137	237

Figure 16 – Estimated Additional Units Sold by Price Elasticity Estimate

ESTIMATED MARKDOWN OUTCOMES – A THREE PRODUCT EXAMPLE

As demonstrated in the previous section, the estimated price sensitivity rate can have a major impact on the effect of a markdown on total unit demand. Because of this impact, the optimal markdown cadence will differ between products with low, medium and high elasticity. In this section, we show the markdown paths and final revenue results for three different products with low (-1.5 elasticity estimate), medium (-2.5 elasticity estimate), and high (-3.5 elasticity estimate) price sensitivity. In this section, we assume that the price elasticity estimates are correct. The goal of the markdown optimization is to minimize inventory at risk, with a secondary goal of maximizing revenue. To simplify the analysis, we assume that the base forecast is the same for each product. We also assume that the starting inventory is 200,000 and the regular price is $50 for all three products. The price sensitivity is the only variable between the products. Each product is eligible to be marked down for 12 consecutive weeks and the allowable markdowns are 30%, 50% and 75%. There are no other rules. In a true implementation, there would very likely be other optimization constraints related to business processes. However, the application of additional constraints will only slightly dampen the importance of price elasticity.

Markdown Depth, Timing & Sell-Through

Markdown depth, markdown timing, estimated units sold, revenue, and ending sell-through are all relevant metrics. Sell-through is defined as the total units sold over the 12-week markdown period divided by the inventory units at the start of the 12-week period. The three products each have unique markdown depth, timing, revenue, and sell-through. The product that has low price sensitivity is not able to meet the sell-through goal of 90%, even if the maximum markdown were taken in the first week. Therefore, we simply maximize revenue by marking the product to 30% off in the first week and not making any additional price reductions. The medium elasticity product is able to meet the sell-through target by taking its first markdown to 50% in week 1 and an additional mark to 75% in week 8. The highly price sensitive product can meet the sell-through goal by taking just one markdown to 50% in week 1. The markdown depth, timing, and sell-through values can be seen in Figure 17.

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Estimated Sell-Through
Low Elasticity	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	34%
Medium Elasticity	50%	50%	50%	50%	50%	50%	50%	75%	75%	75%	75%	75%	97%
High Elasticity	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	97%

Figure 17 – Markdown Depth, Timing, and Sell-Through (Three Different Products with Three Different Elasticity Estimates)

Revenue Earned During Markdown Period

Revenue also differs depending on each product's price sensitivity. Figure 18 shows that the product with low price sensitivity earns the least revenue during the markdown period ($2,369,325). In contrast, the product with high price sensitivity earns significantly more during the 12-week markdown period ($4,831,575).

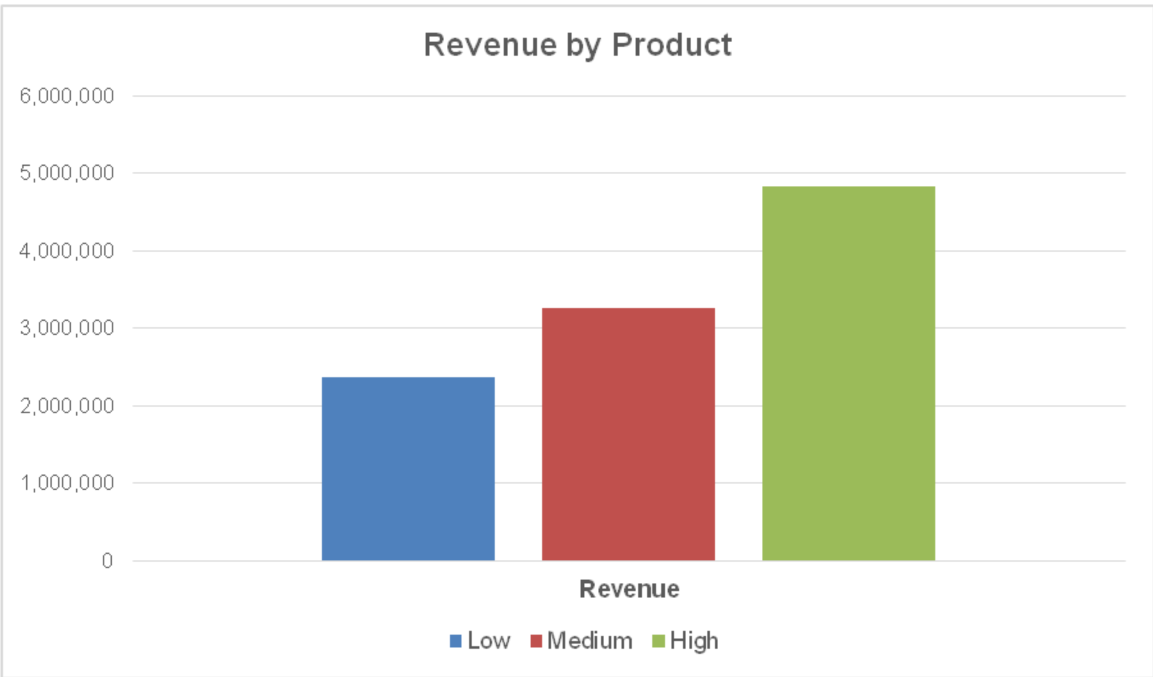

Figure 18 – Revenue by Product during the Markdown Period

Inventory Path through Markdown Period

In addition, the remaining unit inventory at the end of each week is different for each product. As seen in Figure 19, both the high and medium elasticity product plans achieve the sell-through target near the end of the markdown period (Week 11) while the low elasticity product does not reach the sell-through target.

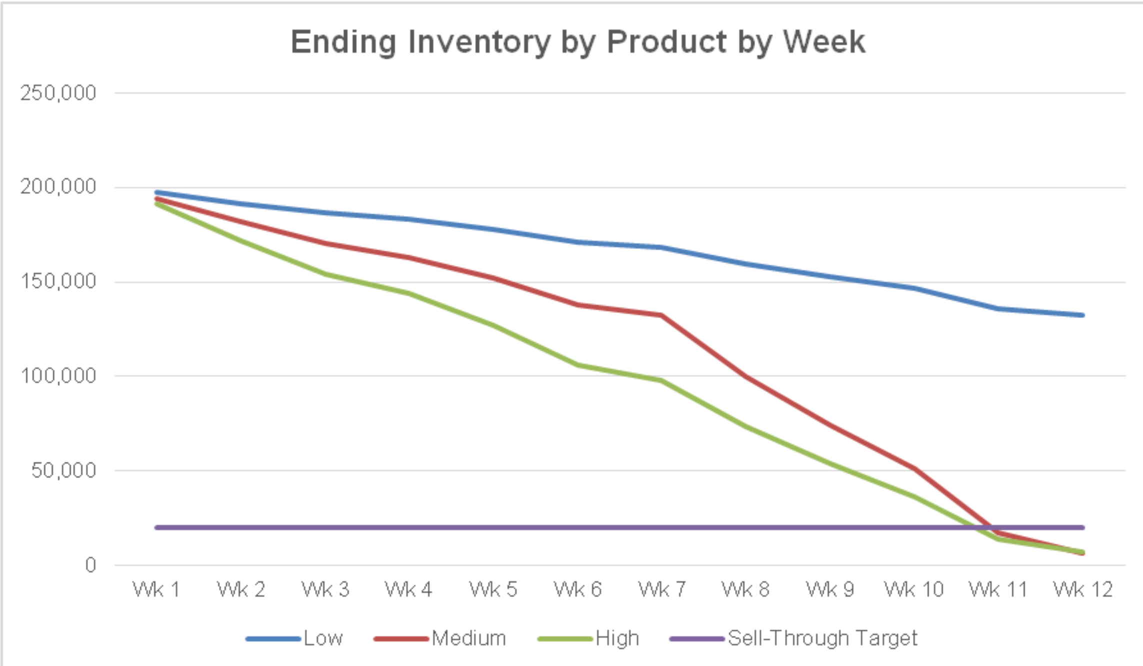

Figure 19 – Ending Inventory by Product by Week

SIMULATING MARKDOWN RECOMMENDATIONS WITH ELASTICITY ERROR

In this section, we focus on the product from the previous section that has medium price sensitivity. We will simulate three optimal markdown paths for this product. In the first simulation, we under-forecast the price sensitivity. That is, the estimated lift from a markdown is smaller than the actual lift from a markdown. In the second simulation, we over-forecast the price sensitivity. In this case, the estimated lift from a markdown is larger than the actual lift from a markdown. In the third simulation, the price sensitivity is estimated correctly and, thus, the estimated lift from a markdown is correct. The recommended markdown paths for the three simulations can be seen in Figure 20.

	Week 1	**Week 2**	**Week 3**	**Week 4**	**Week 5**	**Week 6**	**Week 7**	**Week 8**	**Week 9**	**Week 10**	**Week 11**	**Week 12**	**Estimated Sell-Through**
Under-estimate Price Sensitivity	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	49%
Over-estimate Price Sensitivity	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	59%
Correctly Estimated Price Sensitivity	50%	50%	50%	50%	50%	50%	50%	75%	75%	75%	75%	75%	97%

Figure 20 – Markdown Depth, Timing, and Sell-Through (One Product with Three Different Elasticity Assumptions)

The simulation with the correct estimate reaches the sell-through goal, but the others do not. In the simulation where the price sensitivity is under-estimated, the optimization does not show that the sell-through target can be reached, so it maximizes revenue instead. In the case of the over-estimated price sensitivity, the optimization estimates that only

a 50% markdown is required to reach the sell-through target. However, the price response will be smaller than the simulation predicts, so sell-through will not be reached.

Figure 21 shows the inventory paths for each simulation, clearly demonstrating that the ending inventory in the simulations where price sensitivity is over- or under-estimated is much higher than the ending inventory in the simulation that uses the actual price sensitivity.

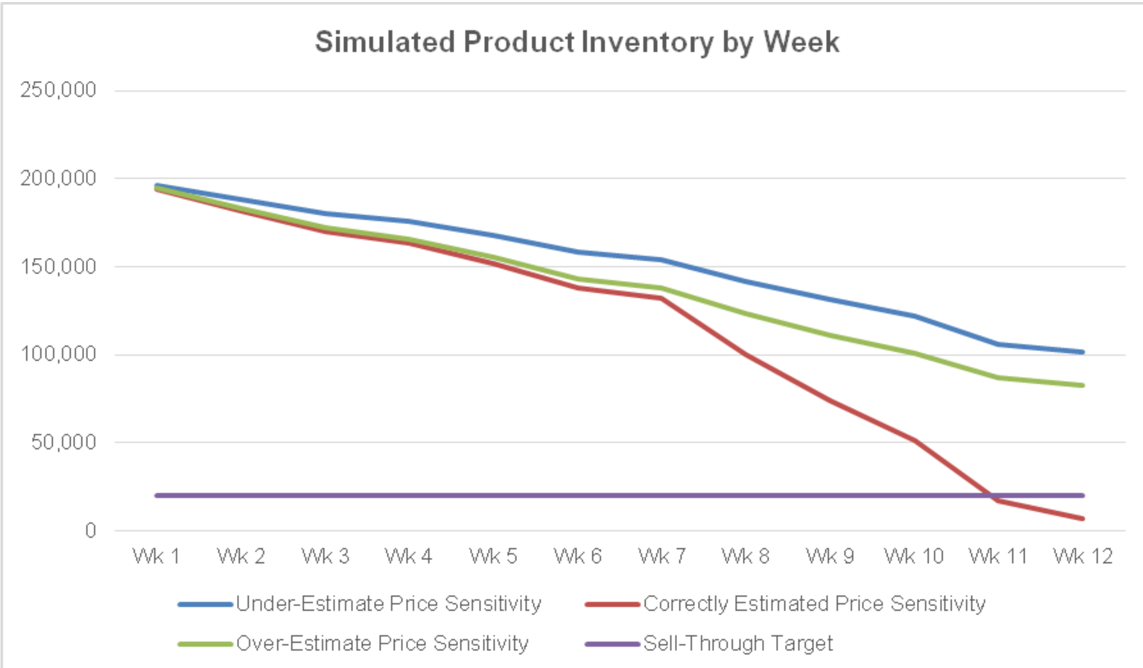

Figure 21 – Simulated Product Inventory by Week (Same Product with Actual Medium Elasticity)

When the price sensitivity is over-estimated or under-estimated, the retailer sells significantly fewer units than when the elasticity estimate is correct. Recall that the goal is to sell at least 180,000 units, or 90% of beginning inventory. Figure 22 shows the percentage of inventory at the beginning of the plan period that is sold and what percentage remains at the end of each simulation. It is clear that there is significant inventory remaining when the elasticity is over- or under-estimated. This is in stark contrast to the very small number of inventory units remaining when the true elasticity value is used.

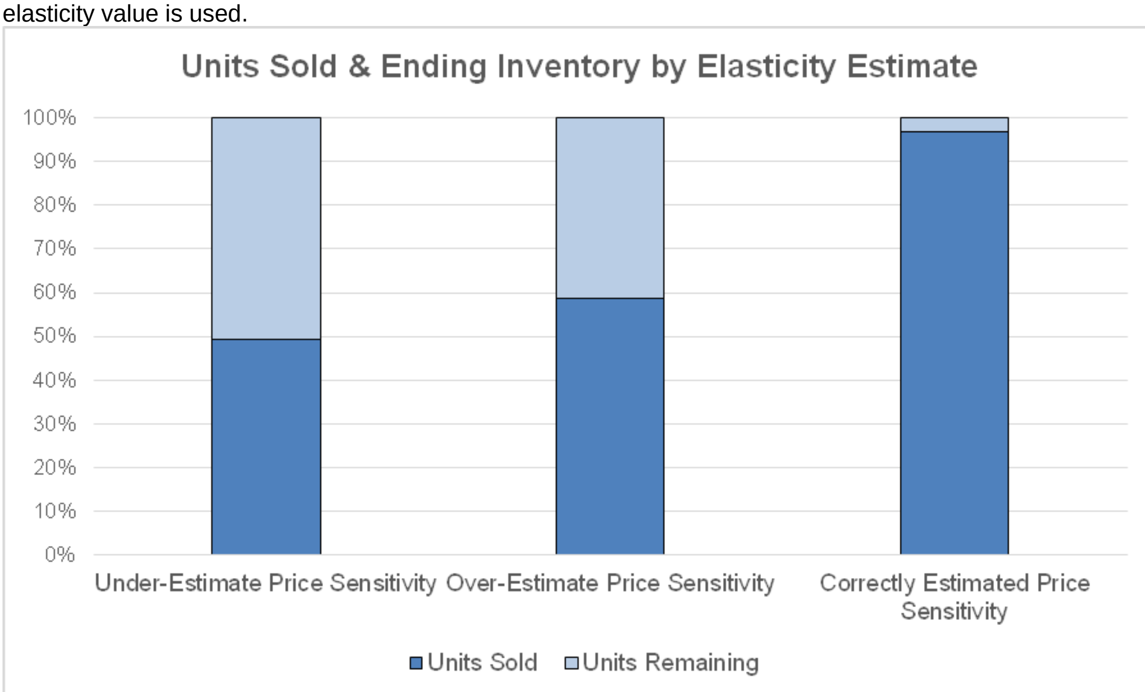

Figure 22 – Units Sold and Ending Inventory by Elasticity Estimate

This markdown example demonstrates how an incorrect price elasticity estimate leaves a retailer with excess inventory. The simulation with a correct elasticity estimate has a remaining inventory value of just over $83,000, assuming the retail accounting method. However, over-estimating the elasticity leaves the retailer with over $2 million dollars in excess inventory. Worse yet, under-estimating elasticity results in $3.5 million in remaining inventory.

MODELING STRATEGY – A CROSS-SECTIONAL APPROACH TO ELASTICITY

Elasticity may be difficult to estimate in a number of different cases:

- Insufficient data
  - Sparse sales history
  - Limited price variability
- Price reactions may differ when a product is regularly priced, promoted, or on clearance

In the case of insufficient data, the elasticity estimate may be improved by using data from a similar product or products. One option, typically used with new products, is to specify a like product from which the new product borrows data. This is a manual process and the like product must be similar to the new product if this method is to succeed. A more automated approach, typically used for new and existing products, is to pool data of like items. SAS uses a number of different algorithms to identify products with similar demand patterns and price sensitivities for grouping purposes. After data is grouped, a mixed regression model is applied. This allows each product to use the group average if it is similar enough to the group as a whole. However, the key characteristic of this model is that an item receives its own price elasticity estimate if its elasticity is significantly different from the group average. This allows the group sharing to be used only when needed.

In the second case, there is sufficient data, but price sensitivities vary across the product lifecycle: regular price, promotions, and markdowns. In this situation, modelers can estimate price sensitivities at regular price and an additional elasticity effect from markdowns, promotions, or both. This requires even more explanatory data than estimating an overall price sensitivity. It may be difficult to find sufficient price variation to isolate the regular price

sensitivity due to infrequent regular price changes. Because of this, the markdown and promotion effects can be isolated only for products with very high volume.

The clearance price decision depends heavily on the price elasticity. Items with different price sensitivities also have very different unit sales when markdowns are offered. These differences are compounded over a multiple-week clearance period, magnifying the impact of the price sensitivity measure. Over- or under-forecasting can lead to markdowns that are too aggressive or too cautious, which causes a very large increase in ending inventory. SAS employs cross-sectional models that pool data to obtain the best price sensitivity estimates. This approach allows data (when relevant) to be shared between product-location pairs. In addition, different price sensitivities can be estimated for regular price, clearance, or promotions when sufficient data is available. These modeling techniques allow retailers to use the best possible price sensitivity measures when making clearance pricing decisions.

CONCLUSION

Forecasting sales remains key to successful retail decisions. Predicting consumer demand requires understanding historical input data, selecting the best modeling strategy, and adhering to the forecast output requirements.

This paper presents two examples of how different modeling techniques are used to achieve fit for purpose forecasts. These forecasts can be applied to inventory replenishment, clearance pricing, or markdown decisions.

These best practice modeling strategies can be achieved by leveraging SAS® analytics, such as SAS® Forecast Server, SAS® Enterprise Miner, and SAS® High Performance Analytics.

Forecast modeling strategies continue to evolve. Leading retailers that embrace these evolving forecasting approaches are squeezing more profits out of their key business decisions. Powerful forecasting gets to the heart of the details of retail.

REFERENCES

Chvosta, J., C. Macaro, K. Sanford, and J. Lemieux. 2013. "Estimating Censored Price Elasticities Using SAS/ETS: Frequentist and Bayesian Approaches." *Proceedings of the SAS Global Forum 2013 Conference*. San Francisco, CA: SAS Institute Inc. Available www.support.sas.com/resources/papers/proceedings13/448-2013.pdf.

ACKNOWLEDGMENTS

Special thanks to Christian Haxholdt, Cyrus Bradford, and Saurabh Mishra--all from SAS Institute--for providing technical and other information used in this paper.

CONTACT INFORMATION

Alex Chien, PhD
Forecasting and Advanced Analytics R&D Director
SAS Institute, Inc.
SAS Campus Drive
Cary, NC 27513
919-531-9236
E-mail: Alex.Chien@sas.com

Beth Cubbage, PhD
Consulting Manager
SAS Institute, Inc.
SAS Campus Drive
Cary, NC 27513
919-531-9491
E-mail: Beth Cubbage@sas.com

Wanda Shive, CPA
Product Manager
SAS Institute, Inc.
SAS Campus Drive
Cary, NC 27513
919-531-2564
E-mail: Wanda.Shive@sas.com