Clustering Techniques to Uncover Relative Pricing Opportunities: Relative Pricing Corridors Using SAS® Enterprise Miner™ and SAS® Visual Analytics

Ryan Carr, SAS Institute Inc.; Charles Park, Lenovo

ABSTRACT

The impact of price on brand sales is not always linear or independent of other brand prices. We demonstrate, using sales information and SAS® Enterprise Miner™, how to uncover relative price bands where price can be increased without losing market share or decreased slightly to gain market share.

INTRODUCTION

This paper and the statistical and visual techniques explore the following question, "If consumer brand choice is based on relative price, can I raise prices without impacting my market share or slightly lower prices and gain market share?"
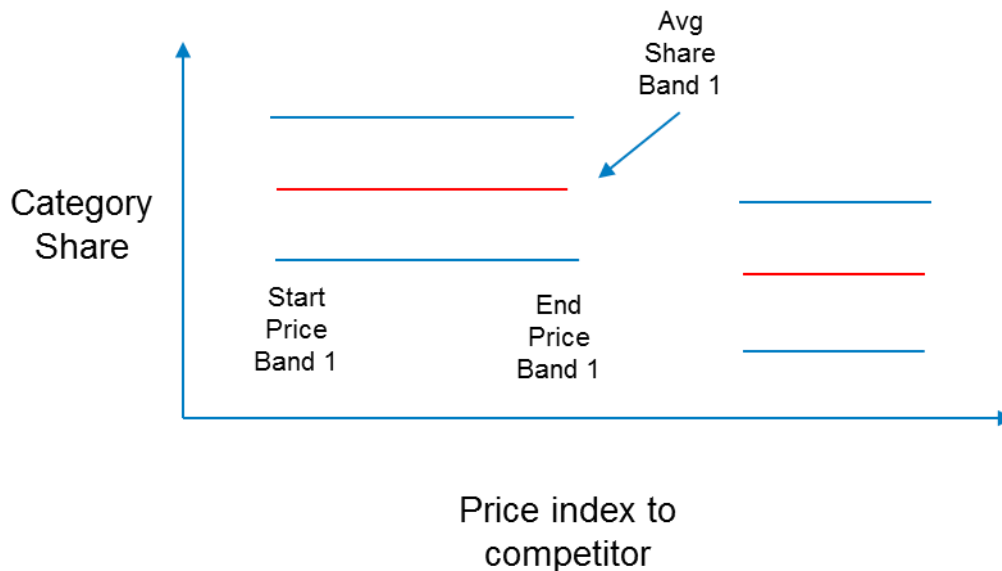

Figure 1. Market Share Versus Price

To answer this question, we applied a nonlinear modeling technique to uncover groups of relative prices where the brand's observed share decreases as a step function. In Figure 1, there are two groups of relative prices. In the first group (Price Band 1), the brand's market share is higher. In the second group, in which the relative price is higher, the brand's market share is lower.

In this example, if your brand's price is at the lower end of the second group, you might be able to increase share significantly by slightly lowering price.

Or, if your brand's price is at the lower end of the first group, you might be able to raise the price to the higher end and not have a decrease in share.

DATA REQUIRED

To execute this analysis, the necessary data is straightforward. Here is what you need to start:

- Product Aggregation Data (typically by brand, size, and brand group)
  - For your brand
  - For competitive brands

- Geography Data (typically store, pricing region, or sales outlet)
- Periodic Sales Data (typically weekly dollars and units)

This data can be obtained from the following sources:

- Syndicated sales data
- Aggregated from Point of Sale data

From this data, often the most difficult process is choosing the appropriate competitive product or brand groupings. For example, when considering dog food, does your brand compete with all other brands? Does it compete with only premium brands? What sizes should be compared? Pricing a one-pound bag is very different from pricing a five-pound bag or 50-pound bag. For laptops, which models should be compared? Pricing of laptops is based on processor speed, operating system, memory, and other components.

The data is translated into the following:

- Price index of your brand to competitor brand
- Market share of your brand

Product	store	date	dollars	units
A	1	1/10/15	314.20	2
B	1	1/10/15	127.40	1
A	1	1/17/15	555.72	3
B	1	1/17/15	260.4	2

store	date	Price_A	Price_B	Price_Index	Share_A
1	1/10/15	$157.10	$127.40	123.3	66.7
1	1/17/15	$185.24	$130.20	142.3	60.0

Figure 2. Translated Data

Once you have the proper product comparisons and the data in the right format, you are prepared to begin the analysis.

TREES IN SAS ENTERPRISE MINER

Using SAS Enterprise Miner, we can quickly build and assess nonlinear models to group prices based on observed market shares.

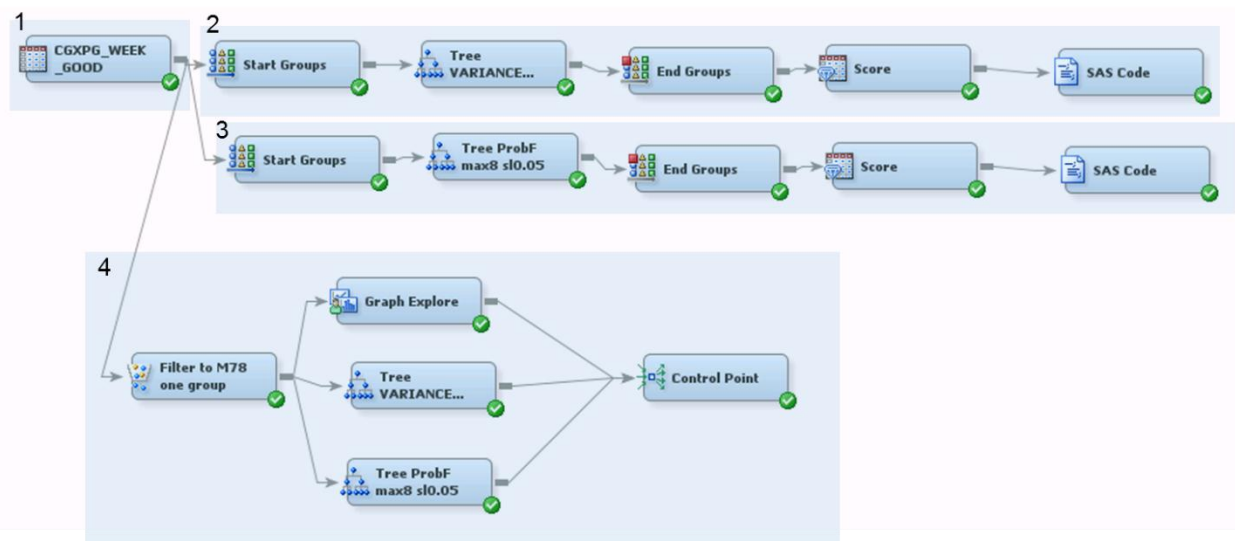Here is an example of the SAS Enterprise Miner model flow:

Figure 3. SAS Enterprise Miner Model Flow

Model section 1 is the data with index and share by week and store. Sections 2 and 3 create sets of models for all products. Section 4 extracts tree information within SAS Enterprise Miner for a specific product group.

TREE TYPE

Various tree settings give very different results. Here is an example of some of the most common tree settings that we have found to be useful:
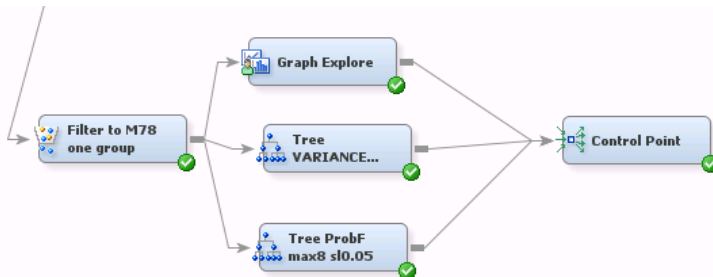

Figure 4. Useful Tree Settings

The top node shows a tree based on cluster variance. These types of trees tend to have more nodes, some of which might not be significantly different from each other. They might have more overlap. These trees can be more useful in understanding deeper nodes than using pure statistical measures, but they do require more interpretation and you might have to combine groups after the analysis. Understanding patterns deeper than the nodes can be better achieved by examining the data points themselves.
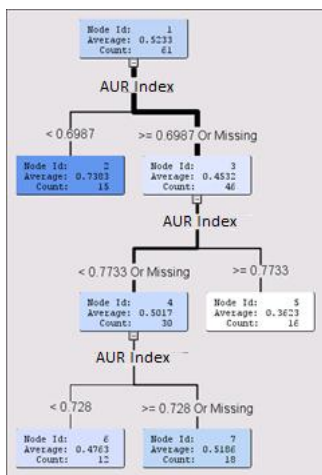

Figure 5. Deeper Nodes

We prefer trees created using a probability of F statistic. This more strict criterion only creates splits when the resulting groups are statistically different from each other. These groups are more reliable. Although this can result in minimal or even no splits, the smaller sets are typically all that the data will support.
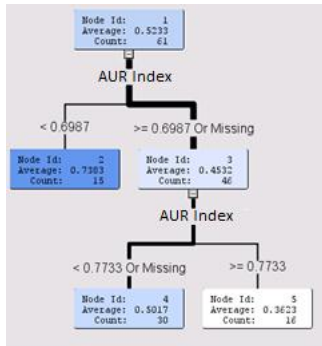
Figure 6. Probability of F Tree

In figure 6, the average share is 52%. When the client's product is priced at roughly 70% or less of the competitor's price, the share increases to around 74%. In cases where the client's product is priced at 77% or more of the competitor's price, the share is only 36%.

GROUP PROCESSING

The final step in creating trees for every product is to use the Start Groups node and End Groups node to loop through one set of trees for each product.

Figure 7. Loop Through Trees

EXTRACT CLASSIFICATIONS

One drawback of the Start Groups node and End Groups node is that each tree can no longer be examined in SAS Enterprise Miner. To examine the tree, you can extract the predicted group for each observation using a Code node.

Figure 8. Extract Results

The Code node consists of a DATA step to extract the tree predictions from a Score node:

```
data pp.treeprobf_allpg ;
   set &em_import_data. ;
run ;
```

Furthermore, you could merge any additional attributes such as actual prices, events to flag, and trends.

These attributes might be important when setting up SAS® Visual Analytics reports.

DISPLAYING RESULTS IN SAS VISUAL ANALYTICS

Although SAS Enterprise Miner has a tool for displaying each tree that is created, producing trees for each brand group in a portfolio generates a large amount of data to be explored. Fortunately, SAS Visual Analytics has standard chart types that make exploring these trees and the underlying data fairly easy. Figure 9 shows the initial report that we used to navigate and display the trees. Because the report is very dense in data, it can seem overwhelming at first.
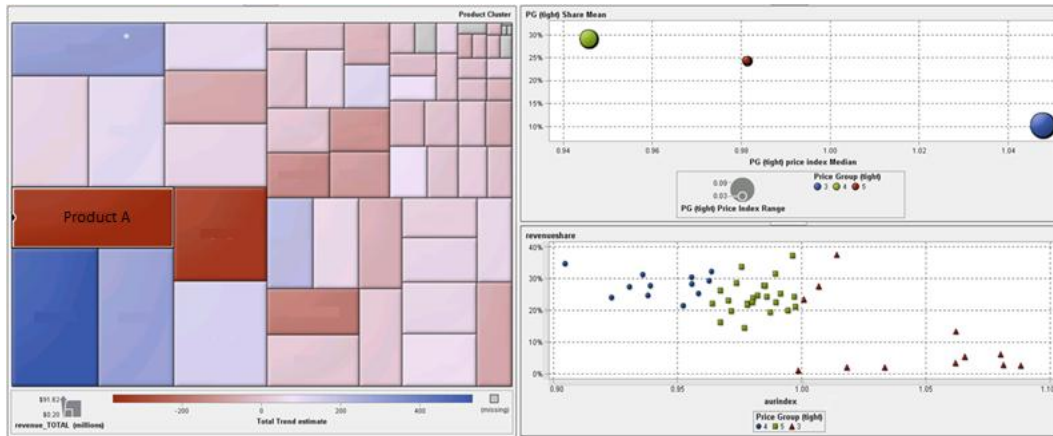
Figure 9. Data in SAS Visual Analytics

The main interface consists of a treemap. The treemap is extracted in figure 10, with a sample of the pop-up help that is displayed when holding the mouse over a group in the treemap.
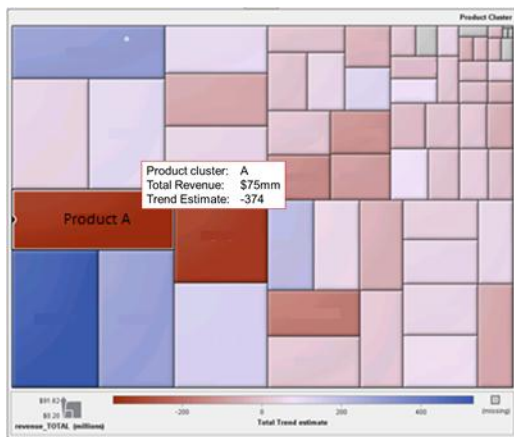

Figure 10. SAS Visual Analytics Treemap

Each tile represents one product group. The size of the tile represents the total revenue. The color of the tile represents the growth trend over time.

From the size of the tile, we see that **Product A** is one of the larger products in total revenue at **$75mm**. From the red color, we learn that the trend estimate is a large negative value with a **-374** weekly decrease. The trend estimate was created in a separate analysis.

The treemap can be used in SAS Visual Analytics to invoke other charts. Clicking on the tile for **Product A** causes it to be displayed in a bubble plot and scatter plot.
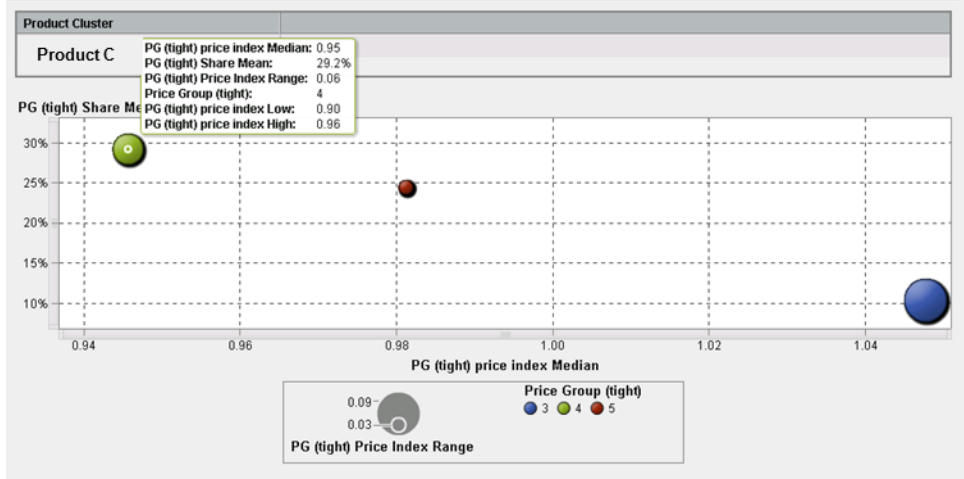
Figure 11. Bubble Plot

In figure 11, three price groups were created for Product A in the tree node in SAS Enterprise Miner are displayed. The green bubble represents pricing from 90% to 96% of the competitor at an average share of 29%.

The size of each bubble represents the size of the range (90% to 96%). The color changes based on the price group.

The share prediction of 29% is based on the average observed share where prices are within the range. To understand the data behind the groups, we can look directly at the store and week observations that are used in the analysis in a scatter plot.
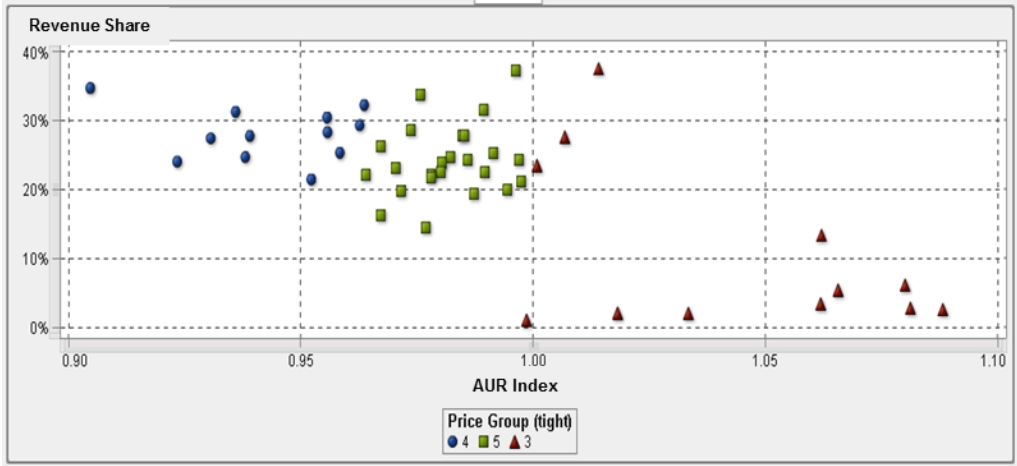

Figure 12. Scatter Plot

The higher price group is fairly consistent. The lowest has some points with high shares. These data points can be examined for patterns that can be used to further refine the analysis.

An alternate view of the data over time can be useful in understanding what the market conditions are.
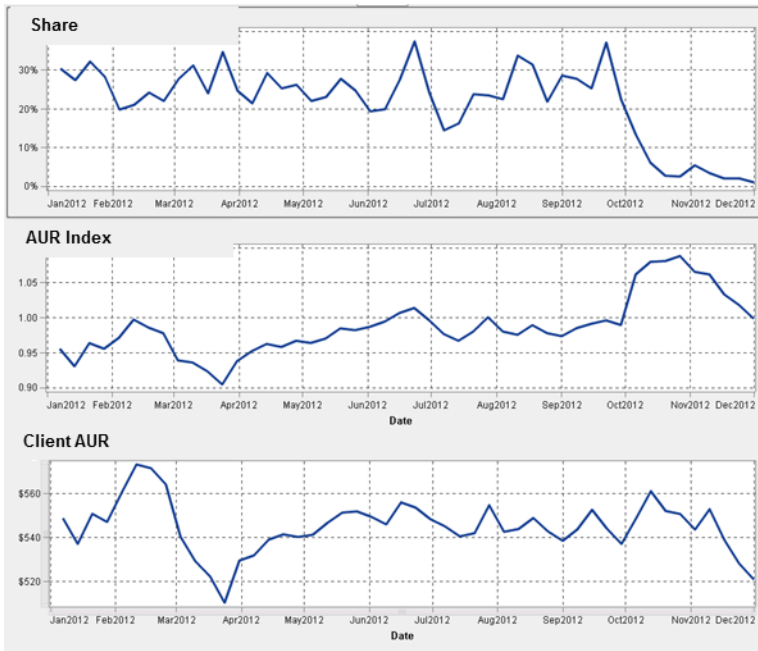
Figure 13. Data Over Time

In figure 13, we plotted client share, relative price (AUR Index), and the client's price (Client AUR) to understand the dynamics further.

We see that there was a large change late in 2012 to the client's share. Although it coincides with a change in relative price, the client's actual price change (Client AUR) is not so dramatic. This would occur if the competitor offered a discount in price at that time, but the client kept its price steady.

Figure 14 shows the bubble plot and scatter plot for a different product. The underlying data does not always explain the complete dynamics of the market. In this case, we have two groups (in green and red) that have a counterintuitive change in share and a wide variance of shares in narrow relative price bands.
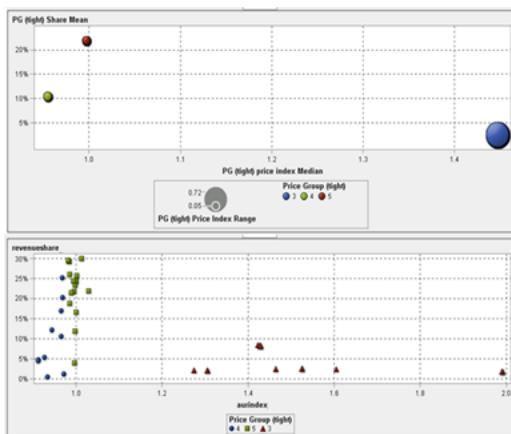

Figure 14. Bubble Plot and Scatter Plot

In this case, there could have been other market factors in play. Further investigation is warranted.

FUTURE DEVELOPMENTS

From the previous examples, you can see that quite a few questions are arise when you begin to explore the relationship between price and share. As a result, we are exploring many extensions of the current methodology.

COMPARE MULTIPLE BRANDS

In many markets, you might be competing with more than one brand. If this is the case, we can generalize the tree analysis to incorporate price indexes to additional brands. In the following tree, increasing price compared to **Brand B** might be OK, but we need to be aware of the impact from **Brand C**.
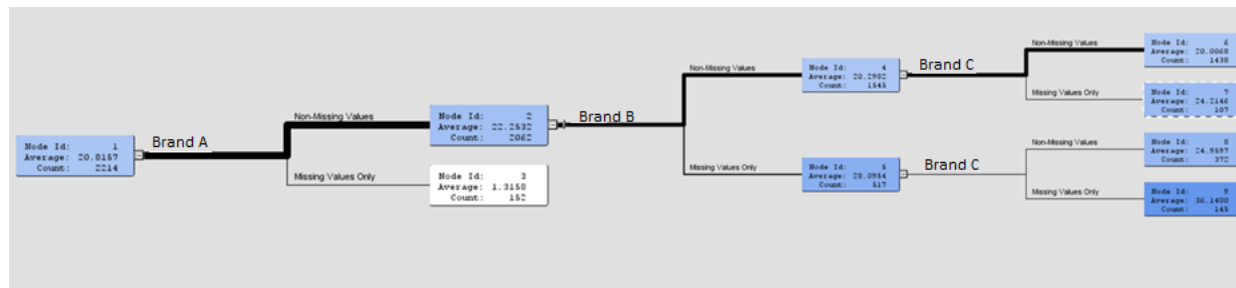

Figure 15. Multiple Brands Tree

DIFFERENCES OVER TIME

The impact of price differences might not be consistent over time. For example, perhaps one brand of chocolate bar is used in s'mores in July and August? Strategies could be developed to first examine share based on seasonality, and then apply this knowledge as part of the relative price to share analysis.

DIFFERENCES BY GEOGRAPHY

Similarly, different markets react to products differently. The analysis could be segmented based on market to uncover geographic pricing zones. For example, Yuengling beer is very popular in some areas of the country and dominates the market unless competitors are deeply discounted.

CONCLUSION

Although price is clearly a driver of market share, most analyses assume a linear relationship between price and a consumer's choice to purchase a particular brand. The concept of looking for localized relative price bands that can be leveraged to either increase share or increase price without losing share can be powerful.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Ryan Carr
SAS Institute Inc.
919.531.9511
ryan.carr@sas.com

Charles Park
Lenovo
cpark@lenovo.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.