Don’t Feel GILT-y About The Past

K. Bell, S. Ingimarson, and K. Ireland

Slippery Rock University, Slippery Rock, PA.

Research Questions

- When is the best time to have sales?
- What brands are performing well in our most abundant color?
- Which brands are performing well in sales?
- Which brands are performing poorly in sales?
- Which characteristics allow for high sales price?
- Which characteristics are preforming poorly based on their inventory?
- Which characteristics are preforming well based on their inventory?

Data

The given data set presents us with information on certain “look”s , which are individual items on sale from Gilt.com from September 2008 until August 2015. The 41048 looks are given with the following characteristics in the table below:

Variable	Description
Start of Look Sale	Timestamp of when look sale started
End of Look Sale	Timestamp of when look sale ended
Unit Price	Selling price of the look
MSRP	Manufacturer Suggested Retail Price
Number of Sizes	Number of sizes available for the look
Look Price Percentile in Sale	Selling price of look in relation to other looks on sale the same day
Average Sale Look Price	Average selling price for all looks on sale
Sale Ratio	Ratio of items sold over items available when the sale started
Material	Reference code of the material
Brand	Brand identification number of the look
Color	Reference code of the color
Product Category Sort Key	Identifies all products in the same category
Levels 1-6	Category level codes
Country of Origin	Identification number of country look was produced
Season	Season identification number of the look

In order to properly answer the research questions, the data needed to be manipulated to be more useful. Our new variables are in the table below:

New Variable	Description
Duration of Sale	Number of hours the look was on sale
Year of Sale	Year the look was on sale
Month of Sale	Month the look was on sale
Price Difference	Amount more/less the look sold for compared to MSRP
Ratio Standing	Category that the sale ratio falls under (GOOD if the ratio is 0.75-1, OK if the ratio is 0.5-0.75, and BAD if the ratio is below 0.5)

Methodology

Data Adjustments

The sale hours variable was calculated by commanding Excel to solve for the time lapsed between the start time and date of the sale and the end of the sale.

The Price Difference variable was calculated by taking the difference between MSRP and the Unit Price.

$$Price\ Difference = MSRP - Unit\ Price$$

The ratio standing was categorized based on the Sale Ratio from the original data set. If the sale ratio was 0.75-1, then it would be awarded a “GOOD” standing, if the ratio was .5-.75, then it would be awarded “OK” standing, and finally if the sale ratio was below 0.5 then it was awarded a “BAD” standing.

Descriptive Statistics

In order to better understand the data, histograms of each variable were run. Based off of the findings in these histograms, further descriptive analysis was conducted appropriately. In most cases, cross tabulations were run to better break down the information.

Histograms are a visual representation of data described by rectangles showing the frequency (which is the rectangle’s height) of a category (which is the rectangle’s width).

Cross tabulations are used to represent the relationship between two variables.

Spine plots are stacked variables to show frequencies of different characteristics based off of other characteristics.

Multiple Regression

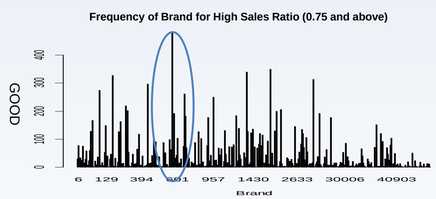
Two different multiple regressions were run on various subsets of the data set.

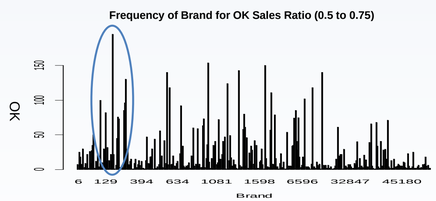
The first regression was used to predicted Price Difference but was run using two different subsets of the data: one that contained the items that held a sale ratio of 0.9 and above and another that contained the items that held a sale ratio of 0.1 and below.

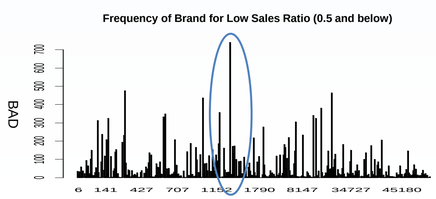
The second regression was used to predict Sale Ratio using the entire data set.

P-values and number of significant categorical variables for a particular variable were used to determine the usage of the variable in the regression model according to an alpha level 0.05.

Results

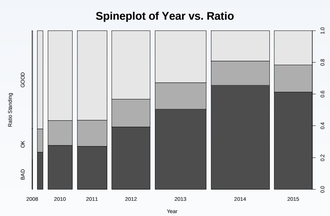


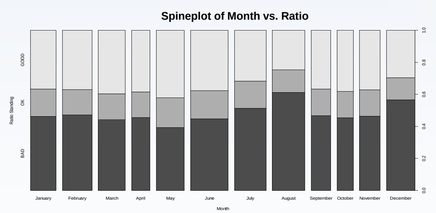


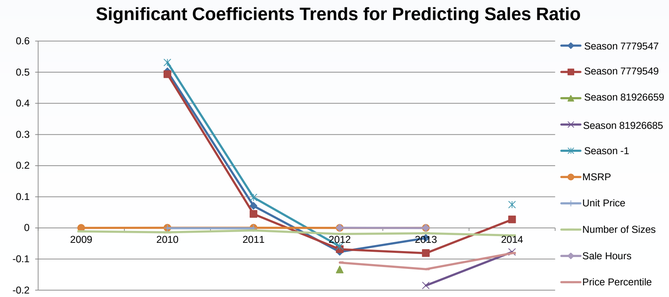


Coefficients for Predicting Difference

	Sale Hours	Price Percentile	Number of Sizes	Average Price	Season 77799547	Season 77799549	Season 81926659	Season 81926686	Season -1
0.9 and above	-0.1204	167.88249	-9.43432	0.48873	4.6446	-47.52269	-31.29064	-31.37784	65.55484
0.1 and below	-0.21969	203.91552	-8.88439	0.43264	15.61187	-87.61813	-67.47202	-61.14973	67.62303







Findings

The spine plots show that as the years go by, the majority of sales go from GOOD ratios to BAD ratios. From the spine plots, we also observed throughout the years that the GOOD ratio tended to do better in the Spring months such as April and May whereas it tended to do worse in August and December.

Based on the cross tabulations, the brands 216, 259, 380, 430, 689, 702, 957, 1376, 1430, 1531, 1584, 1680, 1790, 3547, 25145, 88, & 112 each accounted for about 1% of the sales in the GOOD ratio for the most abundant color..

The histograms displaying the brands according to Sales Ratio show that there were certain brands dominating each. For example, in the histogram displaying the GOOD ratio, brand 578 was dominant whereas in the histograms according to the OK and BAD ratios, the brands 152 and 1352 were respectively dominant.

The multiple regression on difference showed multiple variables as significant in predicting the Sales Ratio of 0.9 and above as well as 0.1 and below. These included Sale Hours, Price Percentile, Number of Sizes, Average Price, and Seasons. The only variable that was significant in 0.1 and below and not in 0.9 and above is season 77799547.

The multiple regression on Sales Ratio showed that although Season 7779547, Season -1, and Season 7779549 were significant at positively predicting the Sales Ratio, there was a quick drop and they are only recently back on the upswing. As for the variables MSRP, Season 81926659, and Sale Hours, they seem to have very little negative significant impact that stayed overall consistent. The variables Number of Sizes, Season 81926685, and Price Percentile were consistently significant at negatively impacting the Sales Ratio prediction.

Suggestions

Turn Back Time

- Based on this research, guilt.com had much greater success in its early years compared to more recent years
- The company should further investigate characteristics that aided in the success in the earlier years of the company to better their sales that have declined recently. Specifically, they should focus on the different Seasons.
- Based on the decrease in the items that had a GOOD Sales Ratio, the company should investigate how the inventory of earlier years has changed. How did we decide how many items to stock, and can it be applied to sales today to better the Sales Ratio.

Dig a Little Deeper

- Guilt.com should investigate further, in particular they should look at the actual number of items that are in stock to better understand the Sales Ratio.
- The company should also investigate who they are actually selling to, to better target their audience.