Here is a histogram relating the percent and the look_st_rate.

A “Gilt”-Free Shopping Experience

R. Bright, K. Simpson, M. Whelan

Slippery Rock University, Slippery Rock, PA

RESEARCH QUESTION

What conclusions can be drawn based on the look sale-through ratio? Overall, we focused our research on examining the look sale-through ratio and how it related to other variables, both in the entire data set and in portions of the data set separated based on the sale-through ratio. Specifically:

- Are there any commonalities between products that did not sell?
- How do the seasons impact the sales rate?
- Does the discount impact the sales rate?
- Does the length of sale have any implications?

LIMITATIONS

- All test, data, and suggestions are limited and specific to this company, GILT.
- These tests are also limited in their accuracy by the years in which the data came from, specifically, 2008-2015.
- While we will draw conclusions and suggestions for the company based on this data, they are all subject to the economy and culture during which the suggestions are being applied.
- Any dramatic shifts in the economy as well as technological advances and shifts in e-commerce can alter and change the effect that these suggestions may have on the company.
- Furthermore, of the data we analyzed, certain years, specifically 2008 and 2015, only provided data for part of the years which could skew the results if abnormalities are present in the absent months.

ABOUT THE DATA

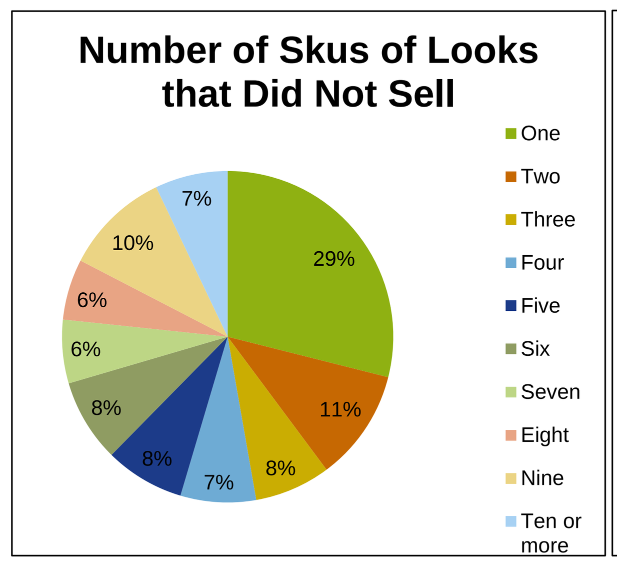
The data set provided was sales data from GILT. GILT is an online retailer established in 2007 that specializes in members-only flash sales with offices in major metropolitan areas in the US, as well as offices in Ireland and Japan. The key variables used in the analysis include, but are not limited to, date and time of sale, unit price of look, msrp of look, number of skus for the look, price percentile of the look, sale-through ratio of the look, and season ID of look. The date and time of sale is simply when the look first went on sale (month, day, year, hour, minute, second) and when the sale for that look ended (month, day, year, hour, minute, second). Something important to note when considering the year of sale is that not all years included were full years. The number of skus is indicative of the number of sizes offered in that look. The price percentile of the look is determined by it's selling price relative to the selling price of all looks concurrently on sale (0 being the cheapest and 1 being the most expensive). The sale-through ratio of the look is essentially the percentage of that look that sold. The dataset included data regarding 41,047 looks.

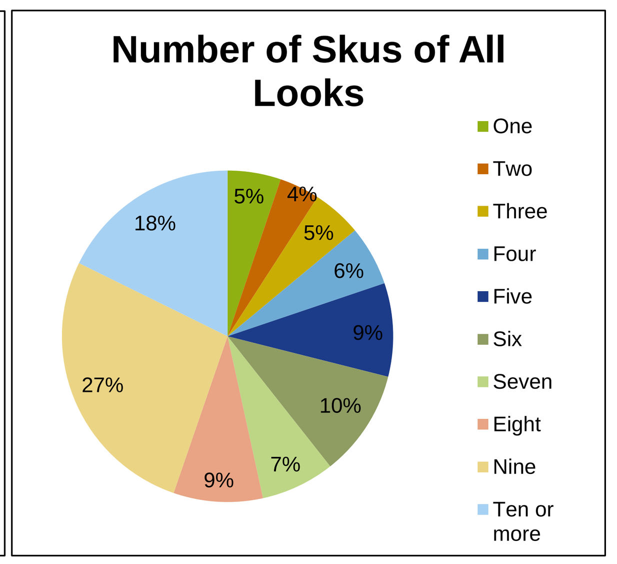
QUESTION: Are there any commonalities between products that did not sell?

Method

The first step was to separate the data based on the sale-through ratio, specifically, the data with a sale-through ratio of zero, which indicates that none of the product sold. Then various attributes of those products were examined to identify any common traits that could be an indicator of why the products did not sell. The next step was to compare the proportions of any attributes that were identified in the products that did not sell to the proportion of that attribute in all of the data. If there appeared to be a significant difference between the proportion in the products that did not sell and the proportion in all the data, a hypothesis test was then performed to determine whether it was truly a significant difference.

Results





29% of the looks that did not sell at all had only one sku number. In all of the looks, only 5% had only one sku number. This seems to be a significant difference.

One Proportion Hypothesis Test:

$$H_0:\ p = p_0$$
$$H_A:\ p \neq p_0$$
$$Where\ p_0 = \frac{2133}{41047}$$

Variable	X	N	Sample p	95% Confidence Interval	P-value
Products that Did Not Sell	994	3429	.289880	(.274733, .305387)	0.000

The p-value indicates that the null hypothesis should be rejected at α=.05 level of significance. This implies that there is significant evidence at the 95% confidence level to suggest that the proportions are different.

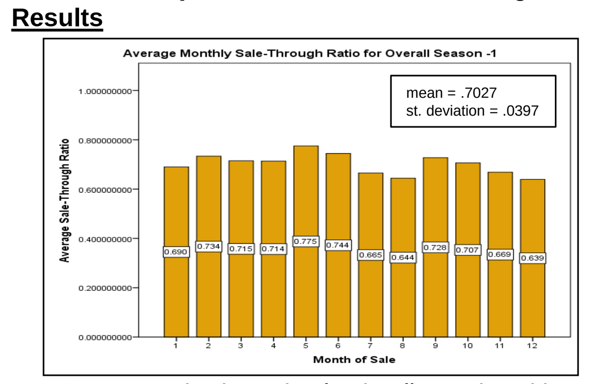
QUESTION: How do the seasons impact the sales rate?

Method

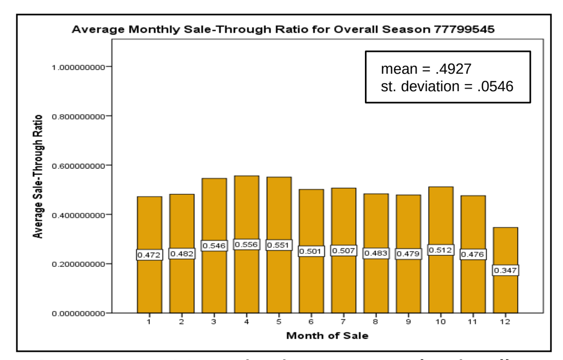
The first step was to separate the data into six groups based on the season ID that the look was assigned. Then, the seasonal data sets were further separated by months in which the look was sold. The average and standard deviation of the look sale-through ratios were calculated within each month for each season. Finally, bar charts were created to determine any months that had either exceptionally poor or good sales within each season ID. How well each season sold was based on the overall average of the look sale-through ratio (LSTR) using the following scale:

$$poor\colon LSTR < .40,\quad average\colon .40 \leq LSTR \leq .60,\quad good\colon .60 \leq LSTR$$

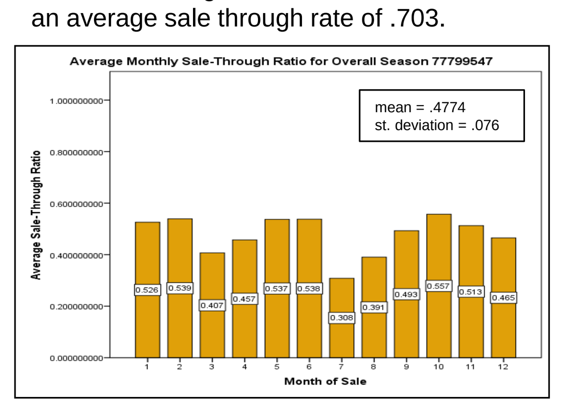
Results



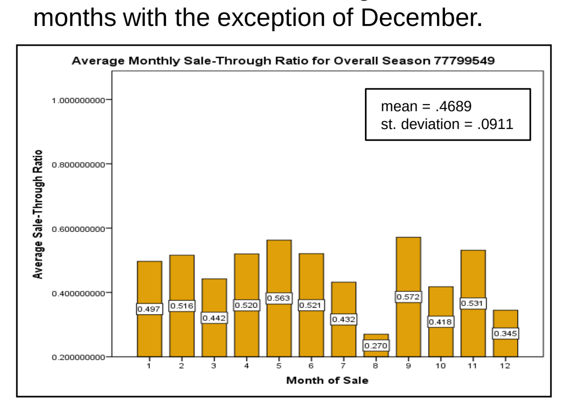
Season -1 had good sales in all months with an average sale through rate of .703.



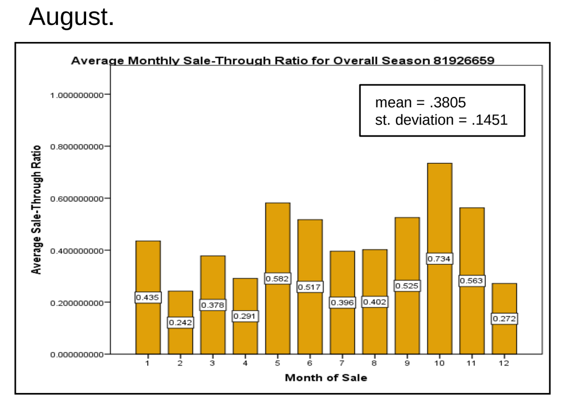
Season 77799545 had average sales in all months with the exception of December.



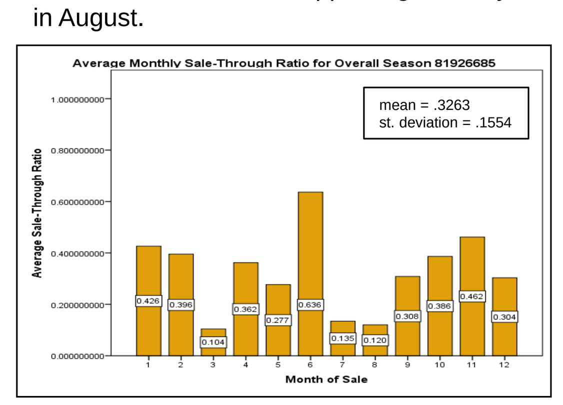
Season 77799547 had average sales overall; however, sales dipped in July and August.



Season 77799549 had average sales overall; however, sales dipped significantly in August.



Season 81926659 sold poorly overall; however, it sold exceptionally well in October.



Season 81926685 sold poorly overall; however, it sold exceptionally well in June.

Overall, Season -1 and 77799545 had at least consistently average sales. The other seasons sales were less consistent and had at most average sales.

QUESTION: Does the discount impact the sales rate?

Method

To examine whether customers tend to buy products that have a larger discount off the manufacturer suggested retail price, the sale-through ratio along with a new variable must be considered.

First, a new variable was created: $discount = \left(\frac{msrp - unit\ price}{msrp}\right) * 100$

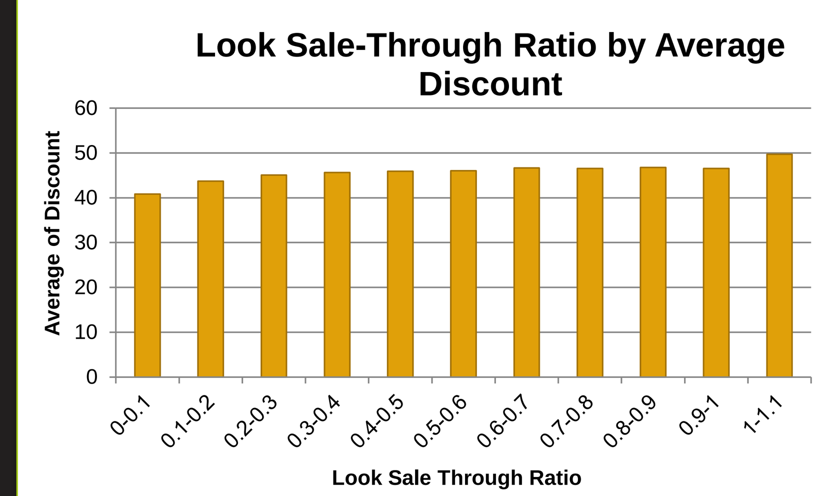
Next, descriptive statistics were run to compare the discount to the sale-through ratio.

After examining, graphs, charts, and tables, a Spearman Correlation test was run to determine if there was a correlation between discount and sale-through ratio.

Results

The table to the right displays the look sale-through ratio and the average discount. The look sale-through ratio is divided into bins by 10% (.10) increments. These bins are inclusive on their lower bound and exclusive with their upper bounds. The discount ranges from 40.85% TO 49.7% with an average of 45.62%. This data can be compared with greater ease in the bar chart below.

Row Labels LSTR	Average of Discount (%)
0-10%	40.85152039
10-20%	43.73285867
20-30%	45.07229317
30-40%	45.67133218
40-50%	45.91033469
50-60%	46.01470351
60-70%	46.67483298
70-80%	46.57308837
80-90%	46.79162522
90-100%	46.55559202
1-1.1 (Completely Sold Out)	49.70689301
Grand Total	45.62471056



There appears to be a correlation between the look sale-through ratio and the discount based on the chart pictured above. As the look sale-through ratio increases, the discount appears to increase as well. A Spearman Correlation test with hypothesis:

$$H_0\colon there\ is\ no\ correlation\ between\ look\ sale-through\ ratio\ and\ discount$$
$$H_A\colon there\ is\ correlation\ between\ look-sale\ through\ ratio\ and\ discount$$

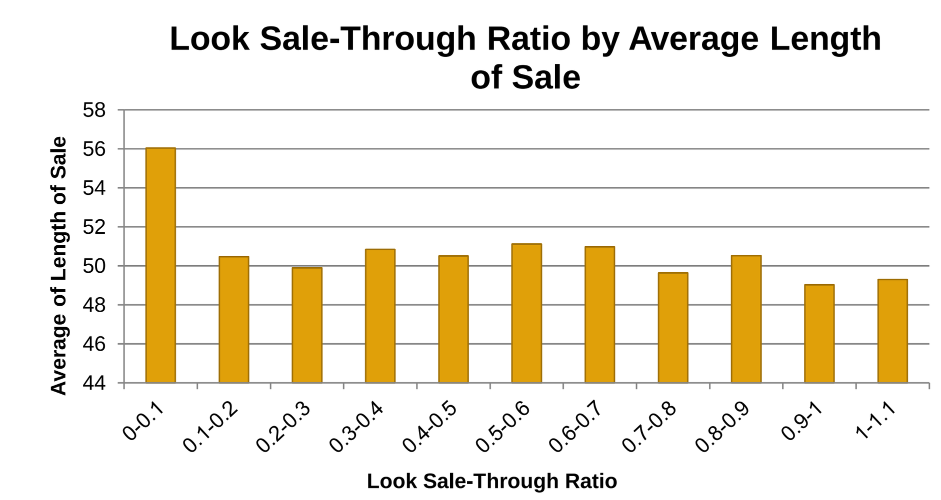
yields a p-value of 0 with a Spearman Correlation of .166. This p-value indicates the null hypothesis should be rejected—that is, there is significant evidence at the 95% confidence level to suggest that there is a correlation between look-sale through ratio and discount.

QUESTION: Does length of sale have any implications?

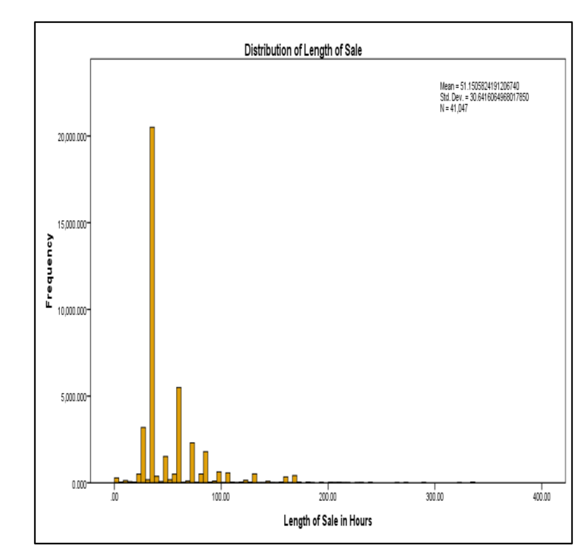
Method

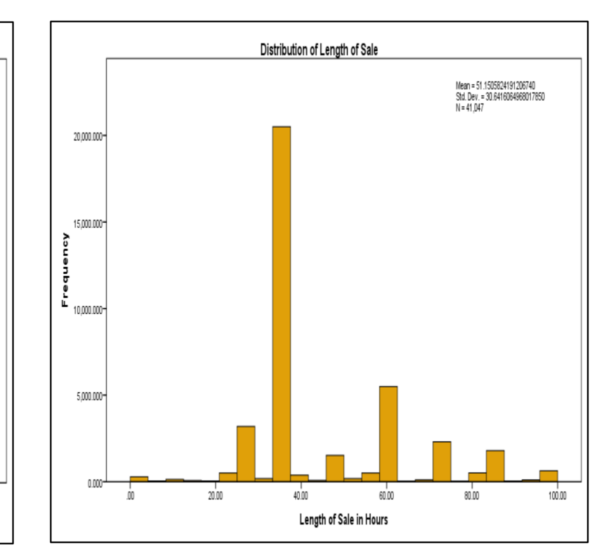
The first step was to calculate a new variable: length of sale. This variable indicated, in hours, the length of sale of each look. The goal was to see if the length of sale had implications in the data set, i.e. was there anything unique that length of sale could describe. Descriptive statistics were run to gain an understanding of the variable. The entire data set was separated into groups of .10 or 10% based on the look sale-through ratio. Then the average length of sale for each group was calculated. Charts, graphs, and tables were created to examine any patterns or trends within variable.

Results



The looks that sold very poorly (i.e. $LSTR < .10$) had the largest average length of sale of approximately 56 hours. For the rest of the groups, it is evident that their average length of sale was less than 52 hours.





The histograms indicate that the majority of the looks in the general data set were on sale for less than 100 hours. This pattern was also consistent among looks that sold entirely and looks that barely sold.

SIGNIFICANCE

- Since there is a large portion of products that did not sell that had only one sku number, which differs from the overall data, the company may want to consider selling fewer products that have only one sku. In many cases this will mean avoiding selling products that come in only one size.
- Since season -1 sells well during all months, the company should consider stocking and selling more looks with that season
- Similarly, when choosing looks to sell during each month, the company should consider what season the look is associated with and how that season sells during that month, as indicated above (e.g. stock less of products of season 77799549 during August).
- The company should not mark looks more than 50% off from the manufacturer price because this will cause an unnecessary decrease in profits since, of the looks that sold out entirely, the average discount was just under 50%.
- Keeping looks on sale for an extended period of time does not have an effect on the look sale-through ratio. Furthermore, for inventory purposes, there is no benefit to keeping a product on sale for more than one hundred hours. At one hundred hours, the company should alter the look sale in some manner be it removing the look or reducing the price. This can prevent an abundance of inventory accumulating that is not selling.