Recently, with development people turning out for electronic shopping was increased. One of the shopping websites was put under study is gilt.com. when we examined this site, we found that it provides instant insider access to today’s top designer labels at up to 70% of retail.

Introduction

Method

1) The curriculum followed before analysis:

- The data from September 2008 through August 2015, and the contents of each column was understood after they got studied well.

- Some of the columns was merged to each other and came up with new columns.

- Rows that may contain any mistakes was removed.

- Finally, rules was extracted.

B) The online shopping data was aggregated and configured by the JAVA program for concluding results.

C) Also used Microsoft office excel for cleaning data that is recorded in a wrong way. For example, when the item's supply price or it's selling price equals to 0 or negative value or any data that won’t lead to a result.

D) Three factors was considered in this process: rate of selling, percent sale, brand or color.

A) The following brands ids have the highest amount of sales and sell rate together, which means that these brands are the most sold:

B) The color 168ae5c7f5 is the most color with items on sale more than 54% of the items on sale has this color, and its look_st_rate is greater than 0.5 in 63.5% of the looks sales, which means that this color is preferred by most of the customers the color eb04d680de: a 15% of the items on sale has this color but its look_st_rate is greater than 0.5 in only 38% of the looks. That means that this color is offered with high number of times but people don't prefer it, whereas other colors are offered with less amount but has a higher rate than this color.

C) The material 0d3f4bb97e: has the most amount of items on sale that are made from it, while its look_st_rate is greater than 0.5 in 52.7% of the look sale items. The material 620f20771: has less amount of items on sale that are made from it, while its look_st_rate is greater than 0.5 in 80.7% of the look sale items.

D) The product category_search_key 71d4b70d17,83a4a6757,1525d0658,5197b227,812d: has the highest amount of items on sale that are product_category_search_key, while its look_st_rate is greater than 0.5 in 53.9%, 49.5%, 49.4%, 52.6%, 51.4% of the look sale items.

E) country_of_origin_ids has the most amount of items on sale, while its look_st_rate is greater than 0.5. The brand_id 1790, 198,430, 430 has a look_st_rate/season_id/percent = 622.

SUGGESTIONS & CONCLUSION

1° we can see that in s1, s2, s3 the look sale rates is greater than .5 which means these seasons are better for sale that .5,6 even though all values of rate still high.

2° when joining season-rate/distance these results do not indicate any strong rule except for the first season we can say that as discount increases the rate increases and vice versa this rule is still not strong since the percent of its validity is 60.4,(3653+722)/(3653+1259+1619+722).

3° the suggestion is to increase the amount of looks that are made of this material.

4° the suggestion is to increase the amount of looks that are made of this category.

5° the suggestion is to increase the amount of looks that are from this origin id.

6° look_st_rate for may is the highest among all months so suggestion increase the amount of sale on may.

7° look_st_rate for august is the lowest among all months so suggestion to decrease the amount of sale on august.

8° Increase the amount of looks for brand id 78,380,430 that came from country with id 301.

9° Increase the amount of looks for brand id 1531 that came from country with id 334.

10° Increase the amount of looks for brand id 702 and 1376 that came from country with id 349.

11° Increase the amount of looks for brand id 3548 that came from country with id 384.