**Organics Product Case Analysis**



Thi Duong

3/26/2016

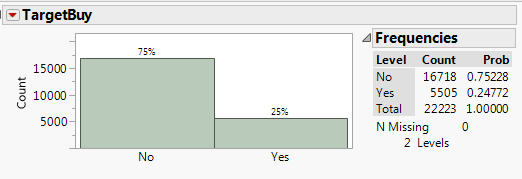
**Introduction**

A supermarket is offering a new line of organic products. The supermarket’s management wants to determine which customers are likely to purchase these products. This analysis report focusses on discovering some factors that affect the likelihood to purchase organic products of the customers as well as helping the management to have a better understanding of the “profitability” of their customers. In addition, this analysis will provide a means for clarifying different group of customers based on their ages, geographic regions, neighborhood, loyalty status, time as loyalty member, etc. If the analysis shows that customers who purchase organic products are also highly profitable customers, the additional cost of stocking organic products will be more palatable for the management.

The ORGANICS data set being examined is comprised of 13 variables and over 22,000 observations. The target variable to be looked at is TargetBuy, which determines the 2 values “Yes” and “No”. “Yes” represents customers who purchase organic products and “No” represent customers who do not purchase organic products.

**General Statistics – Customers who buy Organic Products VS. Customers who do not buy organics products**

The initial analysis focused on the distribution between these two groups of customers across 22,000 observations.



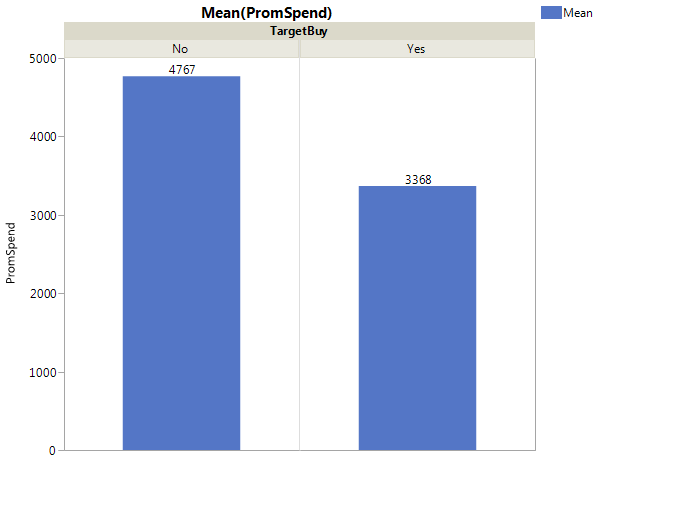
Graph1: Number of customers clarified “Yes” vs. customers clarified “No”

The graph and statistical summary indicates that the number of customers who did not purchased organic products are much higher, indicated by 75%, than the number of customers who do not purchased organic products, indicated by only 25%. By looking at the number and percentage of customers who are willing to purchase organic products, the management can have a better stocking and inventory plan so that they can meet the demands without overstocking or understocking.

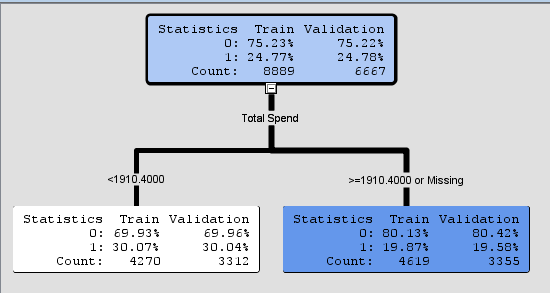
**Profitability of the customers who purchased organic products vs those who did not purchase organic products:**

A distribution analysis of the total amount spent was conducted among the two groups of the target variable. Interestingly, the customers who purchased organic products spend similar amount as the customers who did not purchase organic products.

According to the initial analysis above, there is a huge difference between these two groups that there are only 5,505 people who purchased organic products versus 16,718 those did not, but the amount spent between these two groups does not make a significant difference. Probably because organic products are usually more expensive than regular products, having just a small group of customers buying organic products can be as “profitable” as having a big group of customer who do not purchase organic products.

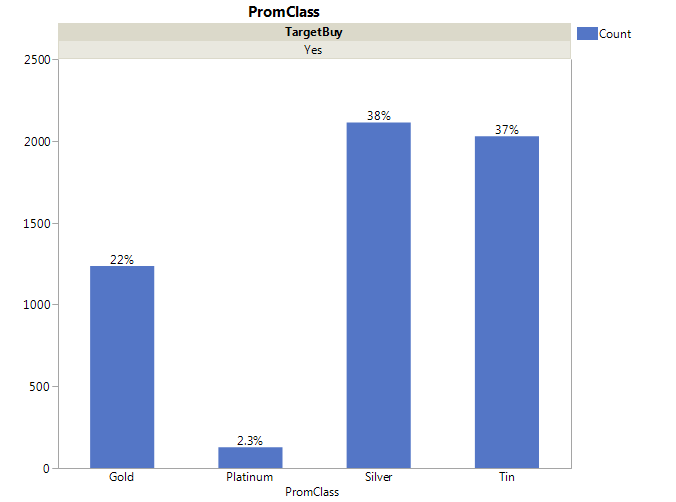


According to the decision tree analysis, if the total amount spent is less than $1,2910 then then the probability for the customers to purchase organic products, who is clarified “Yes”, is 30% comparing to 70% of those customers who clarified “No”. On the other hand, if the total amount spend is higher than $1,910 then the probability to purchase organic product is even lower with only 20% comparing to 80% of the “No” group. The decision tree analysis allows the management to predict the likelihood to purchase organic products within a certain range of total amount spent.



**Percentage of customers who purchase organic products across different loyalty status group:**

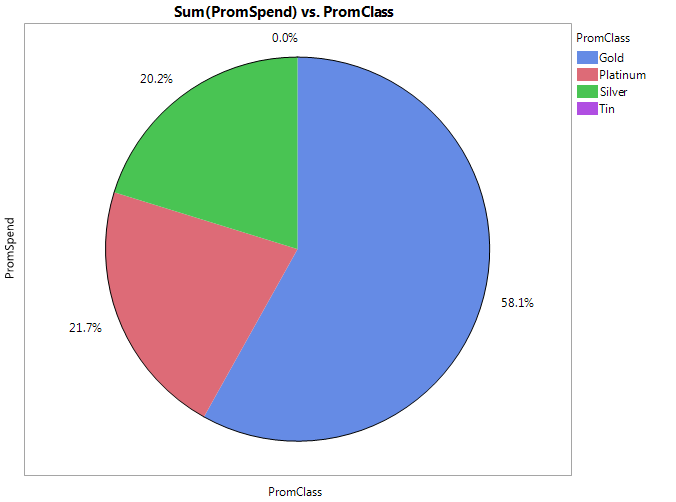
There are noticeable differences in the percentage of customers who purchase organic products across the different loyalty status groups, as demonstrated in the statistical graph below:



| **Level** | **Count** | **Prob** |
| --- | --- | --- |
| Gold | 1236 | 0.22452 |
| Platinum | 128 | 0.02325 |
| Silver | 2112 | 0.38365 |
| Tin | 2029 | 0.36857 |
| Total | 5505 | 1.00000 |

The level of loyalty from the highest to the lowest is Platinum, Gold, Silver, and Tin, respectively. The statistics indicates that there is a significant number of Silver and Tin customers who bought organic products (38% and 37%). The percentage of Gold customers who purchase organic product is also high as 22% while there are only 2.3% of customers who are in the Platinum group. The management may want to consider to have different promotions for different loyalty group based on their level of loyalty so that may encourage the Platinum customers to buy more organic products. Since the Platinum customers are also the most loyal customers, focusing on this group may bring more profits to the supermarket.

While the managements want to decide if they should bring the new organic product line into the business, looking at the “profitability” of the customers who are willing to buy organic product in the different loyalty groups would be helpful. The pie chart below indicates that there is no profit would be generated from the Tin group, while the Gold group has the highest potential with 58% of the total amount spent. Following by the Platinum group with 22% and finally the Silver group with 20%.

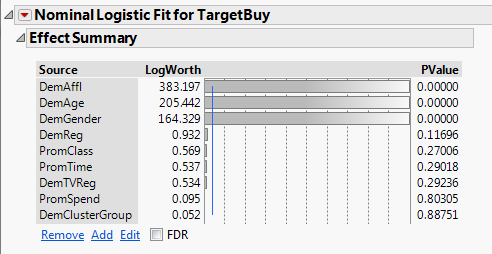


*Percentage of amount spent in different loyalty groups filtered by the “Yes” group*

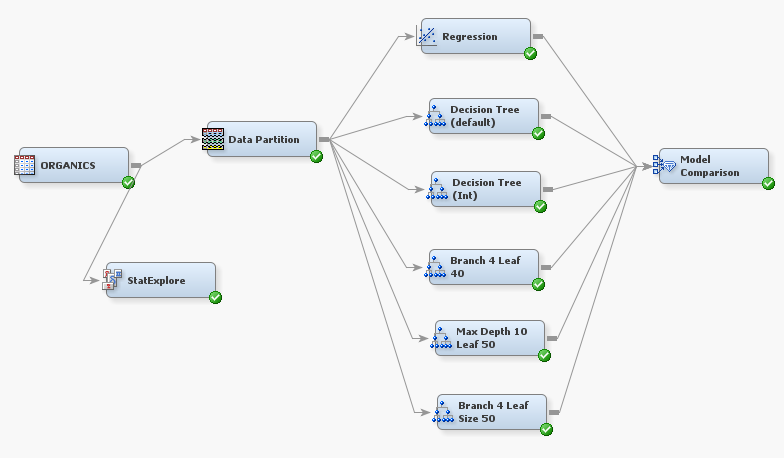
Gold customers spent more money on organic product even though this group has less number of people who bought organic product compare to the other loyalty groups. Striving to make a lift in the number of Gold customers who purchase organic product may bring more profits to the company.

**Factors affect the customer’s likelihood to purchase organic products:**

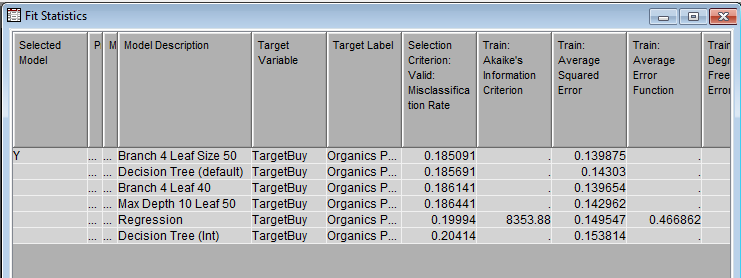
A nominal logistic model was conducted to determine variables that significantly affect on the probability to purchase organic products. As the results, there are only three variables – which are DemAffl, DemAge, and DemGender – that have the p-values smaller than alpha value 0.05 and also have higher log-worth than the rest. Those variables that did not pass the hypothesis test were eliminated from the model.



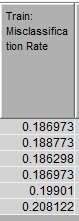
The same results were also found using Enterprise Miner’s decision tree. A model comparison analysis was conducted from five different models: Regression, Decision tree (default), interactive decision tree, decision tree with maximum branches of 4 and leaf size of 40, and decision tree with max depth of 10 and leaf size of 50, and the last decision tree has maximum branch of 4 and leaf size of 50.



The fit statistics of the model comparison shows that the selected model (or the best model) is the last decision tree that has the maximum branch of 4 and the leaf size of 50. The selection is made by the system based on the highest selection criterion score. The Model Comparison node uses the Misclassification Rate criterion by default. In this case, the misclassification rate of the selected tree is 0.185091 which is just a little bit lower than the original (default) decision tree.

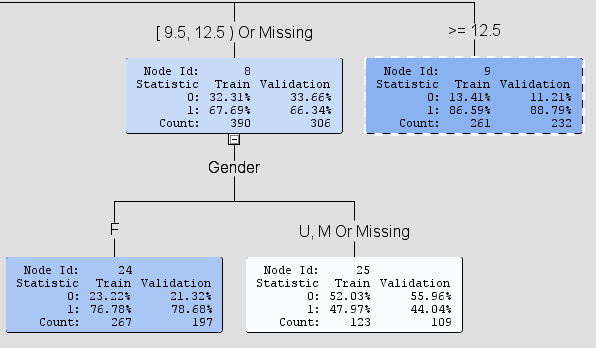


The training misclassification rates of the first four decision trees from the column are also very close to each other, as demonstrated in the column below.

The results of the selected decision tree was generated in order to determine characteristics of a “typical” customer who purchase organic products. Again, Affluence Grade, Age, and Gender are the only three variables that can be used in the model. The decision tree shows that the “typical” organic product customers should be younger than 39.5 and either have affluence grade higher than 12.5 or be females who have affluence grade between 9.5-12.5.

If a customer who is younger than 39.5 and have an affluence grade higher than 12.5 then there are 87% of customers in this particular node are predicted to purchase organic products.

If a customer who is younger than 39.5, have an affluence grade between 9.5 and 12.5 and is also a female, then the probability to purchase organic product of the customers in this group is 77%.



**Suggested Strategy:**

The selected model brought out an interesting results that young female customers would have higher probability to purchase organic products than older customers. The managements should focus the marketing campaign on customers who are younger than 40 and have an affluence grade higher than 12.5. The advertising methods should focus on why people should purchase organic products and eat healthier, while they can still get an acceptable price. To do this, the managements need to develop a promotion plan with coupons offer for daily organic products purchases and special discount rates for loyalty members. Some common but attractive promotions are “Buy 1 get 1 free” or “$1 off milk when purchase a dozen of organic eggs”. The other variables, though are not good variable to be included in the model, can still be considered when doing marketing and making strategic decisions.

**Super Customers**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | DemAffl | DemAge | DemCluster | DemClusterGroup | DemGender | DemReg | DemTVReg | PromClass | PromTime |
| 1 | 7 | 52 | 21 | C | F | South West | S West | Tin | 0 |
| 2 | 10 | 35 | 07 | B | M | South East | London | Tin | 0 |
| 3 | 5 | 46 | 09 | B | F | South East | London | Gold | 5 |
| 4 | 3 | 30 | 30 | D | F | Midlands | Midlands | Silver | 8 |
| 5 | 18 | 48 | 25 | C | F | Midlands | Wales&West | Gold | 2 |

According to the selected model’s results, any customer who is older than 40 should not be targeted and advertised to the new organic product line, which means customers #1, 3, and 5 are not good candidates for organic product promotion. Customer #4 has a high potential to become a candidate for being a 30-year-old female, having Silver loyalty membership, and longer time holding loyal card; however, her affluence grade is only 3 on scale 1 to 30, the probability to purchase organic product of this customer is only 34%. Therefore, she is also not a good candidate for the promotion program. Customer #2 has the highest potential to be candidate of the promotion program with 48% probability that he would purchase organic products. This candidate would have a higher probability to purchase organic products if he had at least a 12.5 affluence grade or if he was a female. Customer #2 is still the best candidate for a promotion related to organic products.

**Summary and Recommendations:**

1. There is a significant different between the number of customers who are willing to purchase organic products and the number of those who do not purchase organic products (Graph 1). By looking at the number of people who are interested in buying organic products, the management can have a better stocking and inventory plan so that they can meet the demands without overstocking or understocking.
2. Customers purchasing organic products and those who did not purchase organic products spent a similar amount of money in the past year. Having just a small group of customers buying organic products can be as “profitable” as having a big group of customer who do not purchase organic products, since organic products are normally more expensive than regular ones.
3. The management may want to consider to have different promotions for different loyalty group based on their level of loyalty so that may encourage the Platinum customers to buy more organic products. Since the Platinum customers are also the most loyal customers, focusing on this group may bring more profits to the supermarket.
4. Gold customers spent more money on organic product even though this group has less number of people who bought organic product compare to the other loyalty groups. Striving to make a lift in the number of Gold customers who purchase organic product may bring more profits to the company.
5. The managements should focus the marketing campaign on customers who are younger than 40 and have an affluence grade higher than 12.5 or a female with affluence grade between 9.5-12.5.
6. The managements need to develop a promotion plan with coupons offer for daily organic products purchases and special discount rates for loyalty members.