Business Analytics: Homework 5

Chengjun Liu, 7116099552

* Understand the situation and need: We try to help Home Depot find and print the coupons that will be used more by customers for additional purchases, thus increasing sales.
* Data: We collect data from internal resources. We have information about every purchase, including total payment, product prices, and amounts of different products. Besides, we get the quantities of different coupons distributed and used for each purchase.
* General approach: We try to look at two key factors: Coupon Consumption Rate (the ratio of coupons used against coupons distributed) and Combined Product Sales (sales of products that are usually purchased together). We build up time series models to predict the two factors, calculate multiplications, rank, and select product coupons with the largest results.
* Procedures
  + Define and calculate two factors
    - For Coupon Consumption Rate, we select a period (one week or one month) and calculate the number of coupons of different products distributed and used for that period. A more accurate way is using the verification code of each coupon and we can find the corresponding coupons that have been distributed. Calculate the ratio of coupons used against coupons distributed for the period.
    - For Combined Product Sales, we use purchases in the same period above, divide product categories for each purchase. We calculate the correlation between product categories. If the correlation between two products is high over the threshold, we group two products into one category. After we divide the products, calculate the average consumption of each category for the period.
  + Build up a time-series model with two factors
    - After we get two weekly or monthly time series of Coupon Consumption Rate and Combined Product Sales, we build up a time series model to predict the two factors for the next period.
    - Because Coupon Consumption Rate and Combined Product Sales tend to influence each other, we suggest using the VAR model to corporate two factors in one model to train, test, and predict.
  + Multiply two factors and find the best coupons
    - We use the model above to predict Coupon Consumption Rate and Combined Product Sales for the next period, and we multiply Coupon Consumption Rate and Combined Product to get a score for each coupon corresponding to specific group of products.
    - We sort the scores of different coupons and products and select coupons with larger scores for the next period. We can select only one coupon for each category of products. Usually, we print three kinds of coupons, so we can select the largest three belonging to different categories as the coupons to print for the next period.