EMORY | GOIZUETA BUSINESS SCHOOL

ISOM 680 - Marketing Analytics

Project Report: Recommender Systems

Team 7

Charlotte Higgins Iati, Philipp Scherbel, Andy You, Yidan Yu

Contents

1. Introduction

1.1 Background

Pernalonga, a leading supermarket chain of over 400 stores in Lunitunia, sells over 10 thousand products in over 400 categories. Pernalonga regularly partners with suppliers to fund promotions and derives about 30% of its sales on promotions. While a majority of its promotion activities are in-store promotions, it recently started partnering with select suppliers to experiment on personalized promotions. In theory, personalized promotions are more efficient as offers are only made to targeted individuals who required an offer to purchase a product. In contrast, most in-store promotions make temporary price reductions on a product available to all customers whether or not a customer needs the incentive to purchase the product. The efficiency of personalized promotion comes from an additional analysis required on customer transaction data to determine which customers are most likely to purchase a product to be offered in order to maximize the opportunity for incremental sales and profits.

1.2 Challenge

Beiersdorf wants to find out which Nivea branded products (that are sold in the Pernalonga stores) to promote in order to increase overall sales for the Nivea brand. The goal is to define a strategy about what customers to target and what products to promote and advertise.

2. Data

The Data consists of two files, that must be merged: the product file and transaction file. Merging them results in a data set consisting of ~30 million transactions from 421 stores and 429 product categories with over 10767 products. Out of 7920 customers in total, 7576 customers bought at least once a Nivea product. Focusing on Nivea, Pernalonga provides 151 different Nivea products, aggregated to 42 subcategories and 13 categories. The following chart shows the categories and sales contribution of Nivea products:

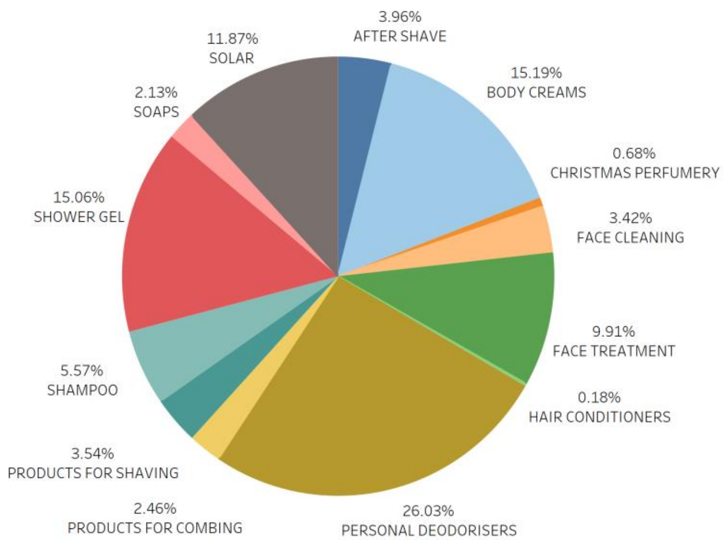

Figure 1: Categories of Nivea Products

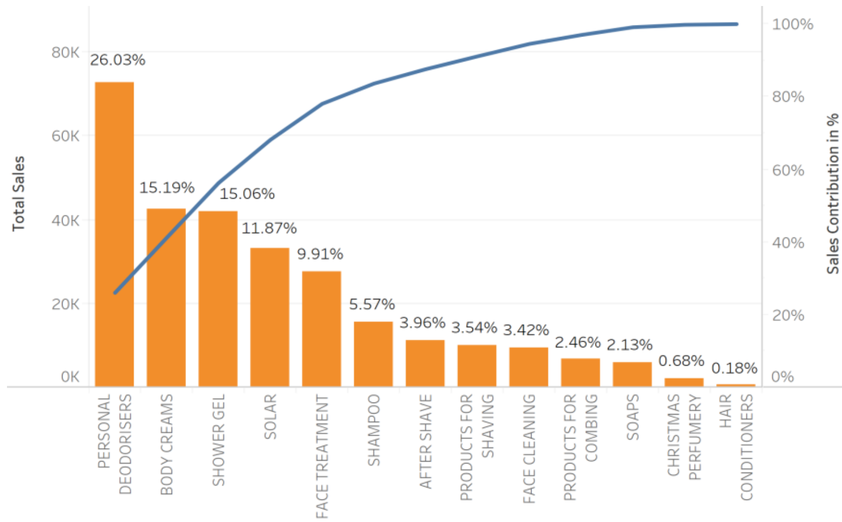

Figure 2: Total Sales per category and contribution to Sales

The calculated sales % is for Nivea products only and shows the sales contribution of each category and thus the importance of the categories.

Nivea Product Categories and Sales Dollars

We first analyzed the highest-ranking product categories in terms of sales dollars. Personal Deodorizers contributed to approximately $73,000 in sales dollars, making them the most lucrative product. Body creams ($42,380), shower gel ($42,013), and solar ($33,117) followed.

Category	Total Sales	Sales %
PERSONAL DEODORISERS	72,591.05	26.0%
BODY CREAMS	42,380.86	15.2%
SHOWER GEL	42,013.04	15.1%
SOLAR	33,117.43	11.9%

Figure 3: Most lucrative categories

The least lucrative products on the other hand, were soaps ($5,930), Christmas perfumery ($1906), and last of all, hair conditioners ($508).

Category	Total Sales	Sales %
SOAPS	5,930.84	2.1%

CHRISTMAS PERFUMERY	1,906.01	0.68%
HAIR CONDITIONERS	508.5	0.18%

Figure 4: least lucrative categories

The subcategories that make up the greatest amount of sales were in the body creams (more than $36,000), shower gel (more than $32,000), and then personal deodorizers (more than $21,000) categories. Analyzing price sensitivity across all customers for these different product categories is necessary to gain a greater insight into whether changes in discounts contribute to both sales dollars and volume (see Modeling Part I). Overall, there were $278,923 in sales.

Nivea Product Customers

The highest spending customers spent $477, $441, and $423 respectively. While these customers have been quite lucrative, there are very few Nivea customers that purchase even half of what they purchase. The 15th highest spending customer, for example, has only spent $257 on Nivea products. As the graph below (Figure 3) shows, the vast majority of customer have spent fewer than $100, and highest proportion of customers spend in absolute lowest spending category. Those low spending customers have spent amounts as little as $0.79.

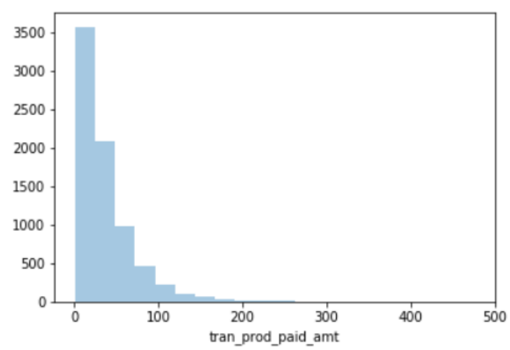

Figure 5: # of customers and their spending

Likewise, many Nivea customers have only purchased very few times. Very few customers have purchased over 20 times, and even fewer over 40. Most have purchased fewer than five times.

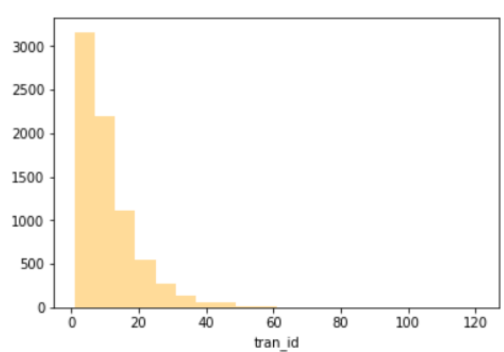

Figure 6: # of customers and their transactions

3. Modeling and Insights

The following three parts of Modeling describe the process that we took to discover what Nivea products to recommend and to which current Pernalonga customers we should recommend them to. A brief overview of each part is as follows:

Part I: Sensitivity to Discounts

In this section, the criteria for products to discount are chosen and the customers who purchase those products the most often ("baseline" customers) are identified. The discount prices are chosen based on the historical discount price that resulted in the most sales for each product.

Part II: Recommender Systems

Using the products and customers identified in Part I, Part II discusses the identification of customers to target based on their cosine similarity to the "baseline" customers. This section also describes the changes in sales due to the recommendations.

Part III: Insights from Association Rules

This section highlights the interesting relationships between different products. These insights could potentially be used to inform bundle promotions, or how promotions in a single product could affect sales for its complement.

Part I: Sensitivity to Discounts

Definition of Price Sensitivity

We examined the price elasticities for different Nivea products to analyze which products were most prone to increase in average daily sales volume at a discount price. To be able to calculate this, we calculated the different discount percentages for each product, date, and store, given that products will have different discount percentages at different dates and in different stores. Then, after aggregating the quantity sold at each discount percentage across stores and dates, the price elasticities were calculated.

These elasticities were calculated by taking the range sales volume for the products, over the range in discounts for those discounts that correspond to the *maximum daily sales volume* and *0% discount* for a product. The output describes the percent increase in product sales with the discount.

Note that some of the discount rates have only one transaction. We have excluded these discount rates when we compute the maximum daily sales volume under the assumption that the discount only appeared for one day and in one store, and is therefore an outlier. This assumption is appropriate given that we cannot know for certain if the discount only appeared for one day because we do not have the data indicating the discount amount on a day where there is no transactions.

Price Elasticity and Cumulative Sales Cutoffs

We are particularly interested in those products that are price elastic and are in those categories that contribute to a large percentage of the sales dollars. The cutoff point for product category sales dollars was set such that the categories together contributed to 80% of total sales dollars (as shown in Figure 4). From these categories, the products chosen were those that had price elasticity of < -2. In other words, for each percent decrease in price, there will be at least a 2% increase in sales volume.

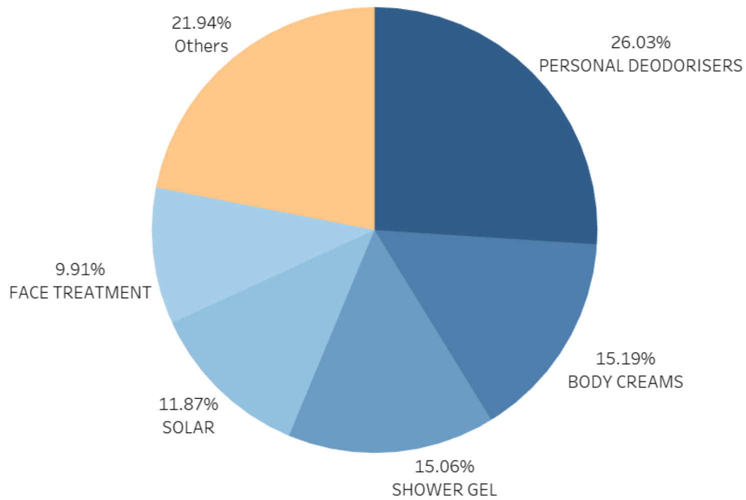

Figure 7: Categories with high price elasticity

The reason that these two cutoffs were selected was to make sure that the products selected ordinarily contribute to a large portion of sales, guaranteeing that they are desirable products, and that their sales would increase substantially with a discount.

Target Products and Customers

There were seven products that fit our cutoffs for both product category sales contribution and price elasticity.[1] These products are in the shower gel, personal deodorizers, and body cream categories. Although there were over 800 customers who purchased these products, we decided to narrow our search to those customers whose total transactions for those products accumulated to over $10, to ensure that the customers were loyal to the brand and products. The reason to set a threshold is to exclude customers who may have only bought once as a gift, for example. There were 137 customers who fit this description. These customers became our "baseline" customers (see Modeling Part II for further information about how these customers' information was used for further analysis).

Part II: Recommender System

Based on the target products and "baseline" customers that we discovered in Modeling Part I, we took the following steps to recommend those products to new customers:

1) Find Nivea customers ("baseline" customers)
2) Find non-Nivea customers that are similar to the "baseline" customers through cosine similarity based on historical product purchases
3) Recommend Nivea products to the non-Nivea customers based on what their similar "baseline" customers have purchased

[1] Those products were 999157189, 999186200, 999452694, 999436632, 999224988.

Customer Exclusions

It is to note that we have excluded customers who purchased Pernalonga's PRIVATE LABEL equivalent of the products promoted. Noted that the products to be promoted fall into three categories: 'SHOWER GEL', 'PERSONAL DEODORISERS' and 'BODY CREAMS'. We have excluded all customers that purchased private label products in these categories in the past in the subsequent analysis.

Products to Target to Specific Customers

High value customers that had purchased the products identified in Modeling Part I were set as the "baseline" customers from which to identify other similar customers from. We are looking to identify potential customers from customers who have not purchased Nivea brand before based on cosine similarity.

We have computed the pairwise cosine similarity between a non-Nivea customer and each of the baseline customers. The customers who were the most similar to those "baseline" customers, with a Cosine Similarity of 0.6 or higher based on all purchases, were the customers targeted for a promotion. In total, 668 customers were identified to be recommended the promoted Nivea products.

Cosine similarity was calculated given that the most important aspect is number of purchases for the same products that two customers shared in the past. Pearson Correlation was not used because it takes into account variance, and variance in number of products they have in common is irrelevant for the scope of this project. Due to this calculation, Pearson Correlations are very low and may work out better when comparing customers on a higher level, i.e. categories because as it takes quantities bought in account, the probability of being similar is very low (one not only has to buy the same products, but also similar quantities to be highly correlated with each other). The output of cosine similarity is easier to interpret and better for our idea, as we know a high similarity means a union in products bought. On the other hand, for pearson, one doesn't know if the similarity is due to union in products or union for some products but similar quantities bought for these few products. Thus we decided to move forward with cosine similarity as we only care if the customer has purchased the product.

Expected Sales Increase from Targeting Specific Customers

Over the 14-day period promotion, is it expected that sales are increasing for the defined "baseline" customers. The sales for the target products without discount is expected to be little more than 450. With the discount, it is expected that the sales will be more than 830 for the current customers.

The percent chance that a new customer will purchase was denoted by their similarity measure. For the new customers, sales is expected to increase by approximately 16. This number was generated by multiplying the percent chance of purchase by 14 (to account for the percent chance they purchase one day out of those 14 days), and multiplying that number by the sales dollars of that customer purchases. With the increase in quantity sold and introduction of new customers, the expected total sales increase will be approximately $395, and that total quantity of increase of volume sold will be 270.

The discount for each product during the promotion was the historical discount percent that resulted in the greatest number of sales. Those discount percentages per product are as follows:

Product ID	Discount %
999157189	7%
999186200	52%
999224988	5%
999436632	7%
999452694	51%

Figure 8: Target products and applied discount

Part III: Insights from Association Rules

Whereas “Modelling Part I” and “Modelling Part II” talked about promotions and recommendations of products, the following section describes the relationships between products, i.e. measures like confidence and support.

Applying association rules to our transaction data with only Nivea products can be very helpful as they show relations between products. These rules can be used to define key value items, make the bundle promotion, and recommend specific products to target customers. Hence, we decided to apply association rules to see which products are often bought together.

Using confidence as metric, we discovered that customers who bought Nivea product 999181250 (category: hair condition) also bought product 999181243 (shampoo) in the past with high confidence. In addition, customers who buy product 999360679 (face treatment) have a high probability to also buy Nivea product 999296957 (face treatment). These products may be complementary so customers may buy them in addition (with the probability we factor in).

There are multiple ways to utilize the result. For example, if the central products are key value items, such as traffic driver, promotion should be made on that products, then it will increase the sales of the additional products. In addition, this exploration can be used in the recommender system. Subsequently, these additional products would be recommended to those customers who purchase the central products. Also, the (assumably) simplest way to make use of this relationship is to make bundle promotions on these product pairs. The combination makes them more attractive than (the individual) products from competitors. Additional data on the product level would be helpful to determine how to use these products for better promotions, i.e. whether it is a traffic driver. For now, our action item is to recommend product 999181243 to 164 customers who already purchased product 999181250 and recommend product 999360679 to 188 customers who already purchased product 999360679.

```
Rule: 999181250.0->999181243.0
Support: 0.000791765637371338
Confidence: 0.2962962962962963
Lift: 91.51285783600194
===================================
Rule: 999360679.0->999296957.0
Support: 0.00042416016287750253
Confidence: 0.1694915254237288
Lift: 44.564299666057586
```

Figure 8: Association Rules

4. Recommendation

The following calculations are based on sales for an average day, and then extrapolated for a time period of 14 days for the sake of comparison as the promotion for Nivea products is set to be for 14 days.

Sales before discount

Total sales before discount is calculated as follows:

For each of the five target products,

$$\sum(\textit{Average daily sales quantity with NO DISCOUNT}) * (\textit{Unit price of product}) * (14\ \textit{day})$$

This value is expected to be approximately 453.38 for those target products.

Sales after discount

Total sales after discount for existing Nivea customers targeted is calculated as follows:

For each of the five target products,

$$\sum(\textit{Average daily sales quantity with NO DISCOUNT}) \\ * (1 + \textit{price elasticity} * \textit{discount percentage}) * \textit{unit price of the product} \\ * (1 - \textit{discount rate}) * 14\ \textit{days}$$

Based on that, the total amount of sales after applying the discount sums up to: $833.67.

Sales with recommendation

The increase in sales of existing Nivea customers targeted is computed first. Cosine similarity is interpreted as probability of a customer buying the same amount of Nivea products promoted as the baseline customer. If a customer is similar to n existing Nivea customers identified, the estimated increase in sales to that customer is taken to be the sumproduct of average of the increase in sales of the n customers and corresponding cosine similarities. Sales by new customers with recommendation will be $15.69.

Total sales increase - Summary

Sales metrics are shown in the table below. These figures are based on the promotion period of 14 consecutive days.

Without Discount	Existing customers	Total sales by existing customers	453.38
		Total quantity sold to existing customers	187.05 Units
With Promotion and Discount applied	Existing customers	Total sales	833.67 **(+83.9 %)**
		Total sales quantity	447.33 Units **(+139.1%)**
	New customers	Total sales	15.69
		Increase in quantity sold	10.4 Units
	Existing and new customers	Total increase in sales	395.98 **(+87.3%)**
		Total increase in sales quantity	270.68 Units **(+144.7%)**

Figure 9: Sales and quantity increases

5. Conclusion

This last section briefly describes our personal takeaways from this project as well as highlights possible improvements and alternate ways to approach this project and solve the challenges.

Takeaways from the project and challenge

The project was very interesting, especially since the question was to some extent open ended and there are many ways to approach this - at least for "inexperienced" students. Thus the challenge was quite complex and throughout the whole project many decisions had to be made which path to take and process to follow. As a team, we experienced it as great topic and probably revised our strategy during execution several hundred times and used business understanding to make final decisions as explained in the respective sections.

Possible improvements

As described earlier, we made several calls when it comes to cutoffs, i.e. daily mean amounts and for the correlation to calculate similarity between the customers. In addition, we decided to target customers individually here with (mass) customized promotions. The project could have also been approached from a store point of view. The steps would have been similar, but instead of targeting customers individually, the most "characteristic" stores given their main customer base could have been selected to give out coupons there. This strategy is rather the shotgun approach - assuming the customers we identified are most responsive to a specific product make up a reasonably large amount of the customer base of that store.

Analyzing by store could have shown different insights. Location notoriously plays a big role in purchasing patterns for a grocery store (for example, a customer in a suburban area differs widely from the same store located in a rural area, or next to tourist attraction).

In addition, weather and other conditions play a role here. Stores would be classified as (very simplified) “shampoo coupon store” or rather “deodorant coupon store”. The stores could then have specific discounts for shampoo, whereas others have discounts on deodorants. As not only most receptive customers would be targeted here, but in general everyone who visits that store, it is not as targeted, but we could also “seduce” customers that way, that wouldn’t have bought these products in the first place. This approach seems valid as well, especially because we were wondering throughout the process, whether to promote at one single discount rate per product, i.e. the rate which resulted in the most increase in quantity sold or have an individual discount rate per customer **-** but this would be price discrimination. Using discounts per store could be a work around with custom discount rates different from store to store for the same product. This would have the same consequences as the aforementioned water can principle **-** we are very successful with the subset of customers who we decide to base the discount on for the whole store, but we would have some customers, buying the product, who would have bought for a higher price. Each approach has its downsides but we chose the customer based discount, which is more targeted compared to having a discount price for a store and product.

Last but not least, several assumptions have been made for the price elasticity. For the purpose of the project, we did not estimate the demand function of each product. Instead, we have approximated the price elasticity using average daily quantity with no discount and average daily quantity with a certain discount rate. Given the different characteristics of different store, one may consider vary the discount rates for the same product at different store. Unit price of a product differs between stores and the price elasticity may also differ.