



PERSONALIZED PROMOTIONS PARTNERSHIP WITH HEINEKEN

ISOM 680 - Marketing Analytics Project Report:
Recommender Systems
03/02/2020

PREPARED BY: ELEANOR LEE, YIFEI REN, JULIE WANG, XIAO CHEN



TABLE OF CONTENTS

1 Introduction	3
1.1 Project Overview	
1.2 Project Goal	
2 Methodology	4
2.1 Data Exploration & Cleaning	
2.2 Customer Segmentation5	
2.3 Beer Customer Segmentation6	
2.4 Product Segmentation 8	
2.5 Beer Product Analysis	
2.6 Heineken Product & Super Buck Product Analysis	
2.7 Recommender System	
3 Results and Prediction	.7
3.1 Projected Sales Lift. 17	
4 Conclusions and Next Steps	18
4.1 Personalized Promotion Results	
4.2 Possible Issues and Suggestions	



1 Introduction

1.1 Project Overview

Pernalonga is a leading supermarket chain in Lunitunia, with over ten thousand products offered in 400+ categories. In order to facilitate sales, brands may choose to offer discounts on their items that Pernalonga stocks. Traditionally, grocery marketing has relied on brand recognition and blanket in-store promotions. However, 75% of customers today are expecting more personalized experiences to help them fulfill their shopping needs. Brands have now recognized the need for personalized promotions to engage customers to buy their products. By using personalized promotions, brands can offer strategic discounts to garner extra revenue from price-sensitive customers who can be persuaded to change from their normal brand purchase patterns.

Internationally, Heineken positions itself as a smart, sexy premium brand for men of the world. Commercials, events and other brand communications efforts around the world would consistently target the urban, upper-class man.² However, the picture was a little bit different nationally. Upon entering the Portugeuse market, there was little effort to communicate the brand, making it extremely hard for Heineken to stand out through all the clutter produced by national brands, such as Super Bock and Sagres, as well as established international brands. In order to survive in the Portugeuse beer market, Heineken has decided to cooperate with Pernalonga to promote a personalized promotion campaign.

1.2 Project Goal

Heineken would like to identify price-sensitive, brand-indifferent customers of Super Bock to target with a personalized promotion. In addition to a list of customers who would be receptive to such a promotion, Heineken would like to know which of the 3 Heineken products (offered by Pernalonga) to specifically promote for each customer.

¹ "How Content Personalization Is Changing Grocery Marketing." *OneSpot*, OneSpot, 11 Dec. 2019

² "Customer Intimacy é um dos pilares da estratégia da Central de Cervejas", e-Business Report, no 478/38, 15 Oct 2009



2 Methodology

2.1 Data Exploration and Cleaning

In order to conduct accurate analysis and create a well-informed recommender system, we revised and cleaned the data in several ways to create data coherency.

Data Step	Reasoning
Recreate transaction ID	The current transaction IDs are not unique, and we need a key to help identify unique observations.
Remove unmatched product ids' transaction records	We do not have these products in the transactions, so we cannot perform any meaningful analysis on them.
Remove transactions that have negative profit	We believe that these observations are errors in the record
Amend mismatch of discount and the difference between sales amount and paid amount	We found there to be some issues with the math of the calculated paid amount most likely due to rounding.
Recalculate unit price based on uniform benchmark	Some unit prices are based on the list price, while others were calculated by paid price.
Remove free giveaway items	We believe that these items serve as cross-selling gifts.
Correct the wrongly labeled discount applied transactions	We corrected transactions that recorded discounts even though no discounts were applied
Remove transactions that are not related to grocery products	We removed shopping bags/carts transactions that are not relevant to our purchase pattern analyses.

Figure 1: Data Revisions and Reasoning



2.2 Customer Segmentation

In our previous analysis, we were able to identify five meaningful customer groups based on their total number of unique stores visited, total number of unique transactions, average discount received on past transactions, and total revenue contribution. "Cherry Pickers" are especially important this time as we want to target customers who are more likely to respond to a promotion due to high price-sensitivity.

Cluster Name	Cluster Description
Professional Buyer	Buyers who would visit multiple store to chase sweet deals
Frequent Buyers	Buys who shop frequently but does not care too much about product discount
Cherry Picker	Buyers who does not contribute a lot of revenue nor do they purchase frequently, but who do care the most about promotion/discount
Cash Cow	Buyers who care about discounts and shops frequently. They also have the highest revenue per transaction
Loyalty Candidates	Buyers who do not purchase frequently nor do they care a lot of discount. We believe customers in this segment have the highest profit potential if we can provide proper incentives.

Figure 2: Description of Customer Clusters



2.3 Beer Customer Segmentation

When we looked into the customer segmentation of beer, we want to know the specific price sensitivity and behaviors of beer customers.

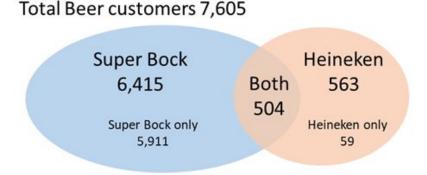


Figure 3: Breakdown of Customer Beer Brand Purchasing Pattern

Taking a deeper look at our targeted market, there are in total 7,605 beer consumers, among which 5,911 only bought Super Bock in the past two years, 59 only bought Heineken in the past two years, and 504 of them have purchased at least once of both brands. This highly skewed distribution brings problems to the usefulness of a recommender system as we will discuss in a later part of the report, and as such we might have to rely more on association rules to guide our promotion strategy.



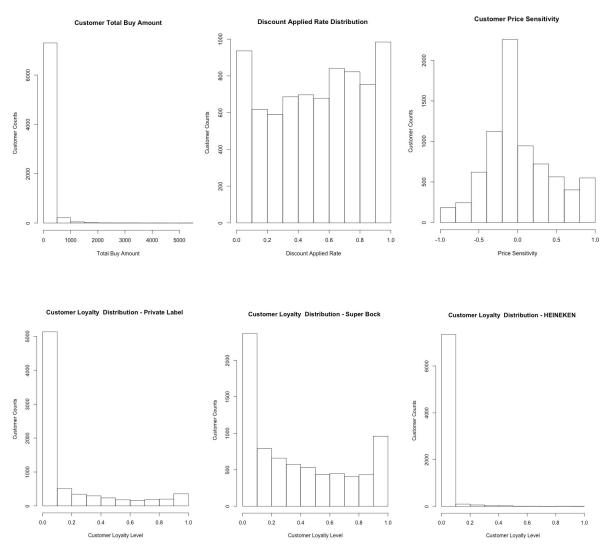


Figure 4: Plots Examining Customer Loyalty and Purchasing Patterns of Beer

As we can see from the customer total buy amount plot, we see that most customers will have beer in their transactions a few times, but there are certain customers that have bought beer in at least 1000 transactions or more. For customers that purchase infrequently, they are less likely to have a brand preference or establish brand loyalty. On the other hand, those that have 1000+ transactions are likely to have brand preferences and established brand loyalty that cannot be swayed by promotional pricing.

Looking at the applied discount rate distribution and price sensitivity plot, we observed that there are customers who are sensitive to beer promotions and there are also people who are not, and this again reminds us that we should target those customers who are more likely to respond to our campaign (Cherry Pickers). Furthermore, we can also observe from the bottom three brand loyalty plots that the



Super Bock customers are much more loyal to the brand than the customers from Private Labels and Heineken. However, we will ignore Private Label loyal customers as we do not want to compete with our distributor Pernalonga. Our main focus is to understand how Heineken can sway customers to buy Heineken over Super Bock using promotional pricing.

2.4 Product Segmentation

In our first venture into the data, we clustered the products after splitting the products into those that are measured in units of kilograms and those that are measured in units of counts. For both clustering algorithms, we used these measurements: total revenue, total sales volume, total distinct customers, total stores, weighted average discount, and discount applied rate (the proportion of times the item was bought on a discount). After clustering both groups, we found there to be 4 groups that were duplicated in both the kilogram products and the count products.

Cluster Name	Cluster Description
No Promotion Necessities	These are products that make up a large chunk of the revenue that are very rarely on discount, and given very little discount if any. Customers will buy these products even without a discount, so there is no need to promote these products.
Discounted Substitutes	These are items that typically have substitutes, so they require discounts in order to encourage customers to buy.
On Life Support	These are products that are usually only bought when there is a discount, and these are also relatively highly discounted products.
Cash Drivers	These are products that contribute the most to revenue, and there are some discounts offered rather frequently.

Figure 5: Description of Product Clusters

We found that all 3 of Heineken's products fall into the "On Life Support" category of items. Super Bock has 24 products overall, with 6 belonging to "Discounted Substitutes", 6 belonging to "Cash Drivers", and 12 belonging to "On Life Support" as well.



For further analysis of personal promotions, we will only focus on promoting the 3 Heineken products over the 12 Super Bock products identified as "On Life Support". We focus on items that are similar in pricing and purchase patterns in order to more accurately identify a list of customers that will be susceptible to a personalized promotion campaign for Heineken over similar Super Bock products.

2.5 Beer Product Analysis

Before we dove into the analysis of customer products and competitiveness, we wanted to understand the nature of beer/alcohol products first. We examined the seasonality of beer/alcohol, in terms of both sales volume and price. In addition, we identified several substitutes and complements of beer, so we could devise cross-selling campaigns in the future.

2.5.1 Seasonality

First, we inspected the overall transactions distribution weekly and monthly. From charts, we see that there is a huge difference between weekdays and weekend transactions. Beer is more frequently bought on Friday, Saturday, and Sunday. Also, the summer and holiday seasons (Christmas and New Years) tend to be the major selling months.





Figure 6.1 Distribution of Weekly Transactions

Figure 6.2 Distribution of Monthly Transactions

To capture the seasonality of beer further, we looked deeper into the weighted average daily price and the daily sales volume across 12 months. We found that there are huge fluctuations of daily weighted average price across 12 months even in just 1 store, in this case, the one with the largest beer sales amount. For the sales volume, we could see that both July and August have higher daily sales volume, probably due to summer breaks for students and the hot weather. October and December also have higher volume likely due to holidays like Oktoberfest, Republic Day, and Christmas/New Years.



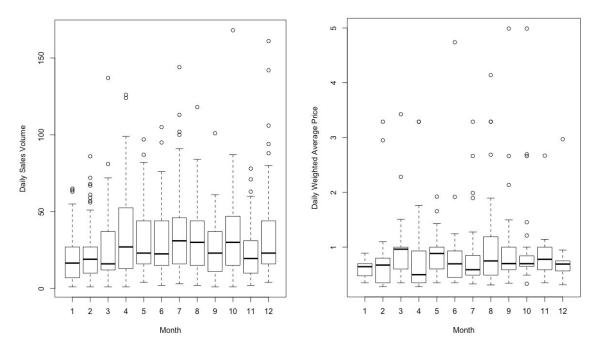


Figure 7.1 Distribution of Daily Sales Volume within Each Month

Figure 7.2 Distribution of Daily Weighted Average Price within Each Month

2.5.2 Substitutes and Complements

Second, we identified the possible substitutes and complements of beer products by calculating the correlation between the weekly sales volume of each co-purchased products with the weekly sales volume of beer products or with the total applied weekly discounts.

Here, we used the weekly aggregated figures due to the weekday/weekend sales fluctuations. By aggregating in weeks, we could alleviate the daily sales differences. Moreover, for stores that have no sales on the week, we filled in 0 for NAs. To have a more interpretable outcomes, we categorized the 425 categories into 60, more encapsulating, main categories, like Beer (including ciders), Fresh Meat (including fresh pork, fresh beef), etc.

We found that main categories such as Fruit tend to have more sales when more discounts are applied to beers or when beers have more sales. Therefore, we identify them as complements.



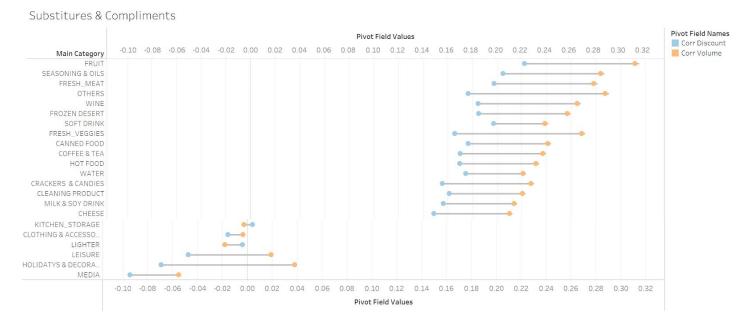


Figure 8.1 Examining Substitutes and Complements

For substitutes, we surprisingly found that the main categories that have negative correlation with the beer performance are not other alcohol or other drinks. On the other hand, other alcohol types and drinks seem to be more likely to be co-purchased with beers.

To help our client, Heineken, be more competitive in the market, we would like to identify its key competitors and the top selling beer products.

For starters, we looked into the total sales volume over the past two years across different brands and found that Super Bock to be the top player and Sagres be the second, and that the sales volume of the third player becomes really small. Moreover, based on the larger difference of weighted average unit price between SAGRES and HEINEKEN, we think it is better to regard the main competitor as Super Bock. Thus, we simplified the brands of beer into below right pie chart and found that Super Bock composed of 34% of the sale, Private Label of 26%, Heineken of 1%, and rest of the brand of 39%. No beer products are non-labeled.



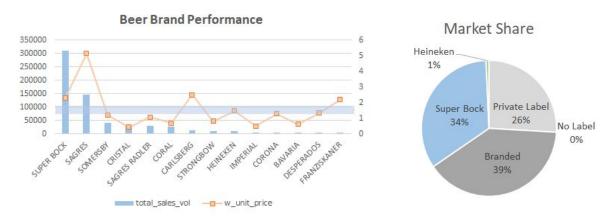


Figure 9: Description of Beer Brand & Market Share

In addition, we would like to see which are the beer top selling stores, and compare the results with the selling performance of Heineken for those stores.

2.6 Heineken Product & Super Bock Product Analysis

To provide customized product suggestions, we want to understand our own products first. After knowing the maximum, minimum, fluctuation of unit price, and weighted average discount rates and conducting online research, we conclude the three products are single glass bottle beer, single aluminum canned beer, and 12 packs beer respectively. For Super Bock, there are 24 products in total and 12 of them are categorized in cluster three, which are frequently discounted and have a substantial amount of transactions. Therefore, we will focus on the 12 packs.

Product	Total Transaction Counts	Customer Base	Total Revenue	Total Sales Volume	Weighted Average Unit Price	Maximum Unit Price
Heineken Beer (Glass Bottle)	1105	450	4648.12	7458	\$0.83	\$4.99
Heineken Beer (AluminumCan)	148	90	599.34	817	\$0.99	\$1.05
Heineken Larger (12 packs)	274	133	3944.25	319	\$17.56	\$18.99

Figure 10.1: Description of Heineken Products



Product	Minimum Unit Price	Standard Deviation of Unit Price	Weighted Average Discount	Discount Applied Rate
Heineken Beer (Glass Bottle)	\$0.54	0.14	-0.20	0.72
Heineken Beer (AluminumCan)	\$0.65	0.12	-0.25	0.69
Heineken Larger (12 packs)	\$10.99	3.12	-5.20	0.64

Figure 10.2: Description of Heineken Products

2.7 Recommender System

We attempted to create two different types of recommender systems to see which would work best: user-based collaborative filtering and association rule.

2.7.1 User-Based Collaborative Filtering

For the user-based recommender system, we used cosine similarity to determine how similar customers were to each other. To recommend products to a customer, one approach we considered was to recommend products that are purchased by similar customers, and the key here is to determine if two customers are similar to each other. It is usually calculated over the ratings that both users have rated in the past. However, in our specific business case, we could calculate it based on similarity of customer's purchase history. For example, we can construct similarity based on how many products in common they have purchased in the past and how many count of items they have bought for the products they bought in common. We can also normalize the product counts on top of the aforementioned two. Therefore, our pipeline to build the user-based recommender system was:

- Identify customers who have bought Super Bock and Heinekins, and pull out their alcohol transaction history
- 2) Construct variables for purchase or not information, number of items bought, and normalized counts of items purchased
- Output target customers to recommend and filter out customers who purchased private label products.
- 4) Filter out customers who are not in the "Cherry Picker" cluster.

Our best user-based collaborative filtering recommender system is the one using cosine similarity and dummy variables for purchase, which aligns with our previous



expectation that including number of items bought can bias our algorithm since different customers have different purchase frequencies.

Our model has the accuracy of 0.512% at the cutoff value of 10. That means, we expect customers to buy on average 0.512 products if we were to recommend 10 products to them, which is not a very good accuracy. We feel this can be the result of the highly imbalanced dataset we have. In total, we have 6,474 customers who have bought either Super Bock and Heineken, and there are 6,415 customers who have bought Super Bock but never Heineken, which leaves us only 59 customers who have bought Heineken in our entire dataset. Thus, we propose to use associate rule, which has better business understanding to identify customers to target and products to promote.

2.7.2 Association Rule Recommender System

To give each customer personalized product offers, we decided to use association rules to find the similarity between users and between products, mapping user segments to the most preferable (highest rated) item cluster. The rate we used here is the total purchase counts of Super Bock products matching with Heineken products.

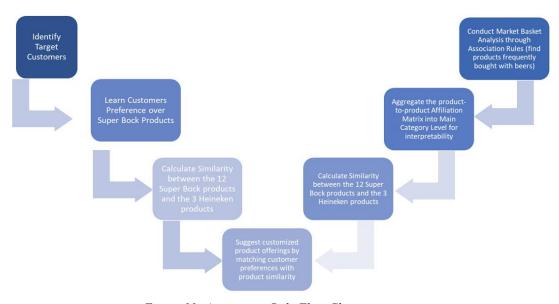


Figure 11: Association Rule Flow Chart

Step 1: Identify Target Customers

Because the original Heineken market share is too small, the user-based collaborative filtering has great obstacles to perform well. Therefore, we decided to use association-rule methods to select our targeted customers.



We selected the target customers with consideration of cannibalization, effectiveness, and potential incremental financial outcomes through below steps.



Figure 12: Customer Filtering

We filtered out 2,500 customers to protect Pernalonga's own product customers (Private Label), 2,900 customers for low tendency of using coupons, 400 customers for price insensitive customers, 140 more customers for non-Super-Bock buyers and loyal Heineken customers, 1300 customers for loyal Super Bock customers. Eventually, we identified 400 customers to promote to.

Step 2: Calculate each customer behavior statistics

To learn the beer product preference of targeted customers, we first used the purchase counts of each beer product as below graphs:

cust_id	SB99915631	SB99916015	SB99916880	SB99922437	SB99922709	SB99945387	SB99949426	SB99951611	SB99952177	SB99955149	SB99959538	SB99959895
609881	0	0	0	0	1	0	0	2	0	0	0	0
939843	0	0	0	0	1	0	0	0	0	0	0	0
1229691	0	0	0	0	0	0	0	0	0	1	2	0
1839834	0	0	0	0	1	0	0	0	0	0	0	0
1839965	0	1	0	0	0	0	0	0	0	0	1	0
2569933	0	0	0	0	0	0	0	0	0	1	0	0
2679893	0	0	0	0	2	0	0	0	0	3	0	1
2779946	0	0	0	0	0	0	0	1	0	0	0	0
2829752	0	0	0	0	0	0	0	0	1	0	0	0

Figure 13: Product Co-Purchases



Then, we grouped products that are similar to each Heineken product into three main product groups, and calculated the wallet of share per group per customer.

Step 3: Calculate Cosine Similarity between each Super Bock and Heineken Products

To calculate the cosine similarity of products between the two brands, we create the affiliation matrix of products occured in Heineken-involved transactions and Super Bock involved transactions respectively, and we turned each affiliation matrix into Main Category level co-purchase counts tables. Next, we calculate the cosine similarity between the two branded products. For example, we calculate the similarity between Heineken Product one and Super Bock Product one.

Main_Category	H999159921 H	999682638 H9	99939106	Main_Category	58999156.5	8999160 S	89991685	B999224:5	8999227 9	8999453:5	8999494.5	8999516	8999521'9	8999551 5	B999595.5	8999598
FROZEN DESERT	1432	6968	748	FROZEN DESERT	9665	7505	18099	8311	16069	46178	11026	12318	145167	16474	30265	27177
JUICE	620	3092	267	MILK & SOY DRINK	4600	3294	9244	3614	7627	34184	5195	7570	159651	7876	14139	14781
MILK & SOY DRINK	960	8113	371	JUICE	2660	2615	4905	2360	4032	9365	3676	3068	26162	4996	10317	6601
HOT FOOD	935	4118	368	FRESH_MEAT	13340	14324	25639	7732	21587	57143	14913	20025	240922	19187	51625	34207
FRESH_MEAT	1620	4507	579	HOT FOOD	4282	6771	7344	3161	6198	25536	4891	7344	123077	8060	19667	18143
SEASONING & OILS	1353	6482	596	SALAD	191	467	410	144	455	2138	410	316	1374	530	1176	1867
CRACKERS & CANDIES	1095	7670	547	FROZEN_MEAT	2944	1923	5082	1841	4920	12057	4510	5006	48748	3983	9216	9871
BABY/CHILD	425	2286	219	WINE	5769	18705	11024	3356	10431	31738	7565	8267	69637	10610	40427	30139
REPAIR & COMPONENTS	34	346	24	COFFEE & TEA	6306	3134	11086	4464	10269	36941	6792	11053	278856	9048	15376	13542
CANNED FOOD	1108	4184	484	BREAD	4354	6054	13850	4031	7302	45265	5872	6966	239647	8339	16764	19734
WINE	762	6953	394	SEASONING & OILS	12356	6481	22483	6858	18689	40929	12443	21772	147965	14765	31631	25318
FRUIT	1937	14664	1181	CRACKERS & CANDIES	8024	4126	13859	5740	12641	33787	9033	10089	121469	12659	20649	20763
CLEANING PRODUCT	831	4371	392	CLEANING PRODUCT	5634	2385	9807	3516	8805	16284	6067	9230	53900	7873	13786	13807
COFFEE & TEA	1072	3695	268	KITCKEN EQUIPMENT	114	63	145	81	148	541	108	210	3064	179	195	202
JAR	92	306	43	BOOK	218	36	1325	120	409	5211	107	388	4389	204	243	256

Figure 14: Calculating Cosine Similarity

Step 4: Mapped Each Super Bock products to Heineken Product

Then, we grouped the Super Bock products that are similar to the specific Heineken products into one aggregated measures and regard the most similar Heineken product as the recommended products.



Figure 15: Description of Product Clusters



Step 5: Recommend each customer a customized product offering

Eventually, we got the recommended product for each customers and the simple statistics for each of their buying behaviour and brand preferences.

cust_id	rec_H99 9159921	rec_H99 9682638			total_sal es_vol	HEINE KEN		PRIVATE _LABEL	SUPER_ BOCK	price_se nsitivity	disc_appli ed_rate
609881	0.67	0.00	0.33	ec_H999	54	0.00	0.50	0.00	0.50	0.33	0.63
939843	0.00	0.00	1.00	ec_H999	71	0.00	0.65	0.00	0.35	0.79	0.92
1229691	0.00	0.00	1.00	ec_H999	116	0.00	0.47	0.03	0.49	0.45	0.88
1839834	0.00	0.00	1.00	ec_H999	88	0.00	0.74	0.03	0.23	0.43	0.82
1839965	0.00	0.50	0.50	ec_H999	22	0.27	0.55	0.00	0.18	0.98	0.91
2569933	0.00	0.00	1.00	ec_H999	28	0.00	0.54	0.00	0.46	0.46	0.89
2679893	0.00	0.00	1.00	ec_H999	31	0.00	0.52	0.06	0.42	0.03	0.68
2779946	1.00	0.00	0.00	ec_H999	124	0.00	0.98	0.00	0.02	0.19	0.99
2829752	0.00	1.00	0.00	ec_H999	29	0.00	0.55	0.00	0.45	0.91	0.72
2989796	0.00	0.00	1.00	ec_H999	136	0.00	0.51	0.02	0.46	0.80	0.75

Figure 16: Targeted Customers with Customized Product Offering

3 Results and Prediction

3.1 Projected Sales Lift

Price Sensitivity Rate	Promotion Amount	Promotion Quantity	Promotion Conversion Rate	Expected Incremental Volume	Weighted Average Price	Profit Margin
0% - 25%	\$2.00	23,350	3%	701	\$3.02	80%
25% - 50%	\$1.50	19,631	5%	982	\$3.18	80%
50% - 75%	\$1.00	18,600	7%	1302	\$2.99	80%
75% - 100%	\$0.50	11,325	10%	1133	\$3.06	80%

Total Sales Lift	Expected Total Redemption Cost	Expected Total Profit Lift
\$10,076.22	\$4,741.58	\$5,334.64

Figure 17: Projected Profit Lift

Here we made a preliminary profit projection for this promotion campaign. We give different discount amounts to beer customers based on their price sensitivity. People who are highly sensitive to prices will receive a relatively small coupon since they are very sensitive to any price change. On the other hand, people who are not as sensitive to price will get a relatively large discount to raise their interest in switching beer brands.



Usually the price promotion marketing conversion rate for the food & beverage industry is around 3% - 4%³. We assume the conversion rate will keep increasing with higher price sensitivity of customers. We also consider the regular profit margin in the beer industry. According to the most recent industry analysis⁴, liquor should have a gross margin around 80 to 85 percent, while bottled beer should have a margin of approximately 80 percent. Draft beer margins can fall around 75 percent, while, on average, wine margins come between 60 and 70 percent. We choose to use 80% as the normal profit margin for Heineken.

The final result shows a \$5334.64 profit lift from Heineken's annual sales. However, we should note that this increase in profit is not significant compared to its total annual sales amount.

4 Conclusions and Next Steps

4.1 Personalized Promotion Results

Personalized promotion sounds like a good idea, as it is a rising trend in the grocery industry. As Heineken breaks down the customer groups, the company can better understand customers' purchase behavior and send right promotions to right people. Based on our results, we have found 400 customers who would possibly switch brands. And if we launch the campaign, we would get \$5,334.64 profit lift.

4.2 Possible Issues and Suggestions

According to our analysis, It is profitable for Heineken to launch this personalized promotion campaign this time. However, in the long-term, it may not be ideal for Heineken to compete on price promotions against huge domestic beer giants like Super Bock. As noted before in the cluster analysis, all of Heineken's products and half of Super Bock's products are categorized as items "On Life Support." These items are already consistently discounted bought on discount a majority of the time. Investing too much on promotion and discounts will lead to a price war, and the consequent loss of higher profit margins for products owned by Heineken. Based on our research, we found an alternative opportunity for Heineken: The Premium Plan.

³ "Promotional Effectiveness Metrics & Email Capture Benchmarks Across 10 Ecommerce Industries [2019 Report]" Ben Choy,

⁴ "Restaurant Beer Markup Strategy", Lynn Starner; "Google Books: Practical Food & Beverage Cost Control"



Although international beers tend to be more expensive than their national counterparts (above \$2.53/liter)⁵, there has been little effort to effectively communicate the Heineken's brand and position as "worth the money" in customers' minds. This means that even though customers already look at foreign beers as posh, pricier alternatives to the mainstream national ones, no international beer is actually positioned as premium product, but instead as a specialty or gourmet. Also, international brands see, in most cases, no need to invest in marketing efforts for such small market share, leaving Heineken enough room in the market to differentiate from its local competitors.

National brands like Super Bock and Sagres, the two main players in the Portugeuse beer market, have instantaneous recognition, but also, serve as an opportunity for Heineken to re-launch itself as premium: these national specialty beers are automatically associated with mainstream brands, diluting their own intended high-end positioning.

By implementing our premium plan, Heineken will be able to take the market share away without the investment resulting in a spiraling price war of promotions and price reductions. In addition, this plan shows consistency with the international intended positioning of Heineken: a smart, sexy premium brand for men of the world and it takes into account the local portfolio (also part of Heineken's assets and key profit drivers and part of its growth platform)⁶.

⁵ Lager by price band Table - Euromonitor International, "Beer in Portugal", Dec.2010

⁶ Presentation made by Jean-François van Boxmeer, Heineken CEO, 6 Nov. 2009