MS-E2177 - Seminar on Case Studies in Operations Research

Optimizing Marketing Promotions for Grocery Retail

Final Report

May 24, 2019

Markus Wilkman	428080
Matias Tiainen	432209
Aino-Nina Saarikoski, Project Manager	92969C

Aalto University School of Science
Finland

Contents

Abstract

Grocery retail is facing technological disruption, slowed down growth with declining sales and profit margins. Optimizing marketing promotions for grocery retail is a complex and multidimensional phenomenon, which holds a strategic importance for retailers worldwide. Our project successfully tackled the optimization of marketing promotions by using multivariate statistical analysis and optimization methods.

1 Introduction

1.1 Background

K Group is the third largest retail operator in Northern Europe and operates in the grocery trade, the building and technical trade and the car trade. The group retail sales were 13 billion euros in 2017. [Kesko Interim Report 2018]

As of 2018, Kesko has 1,800 chain-operating stores in Finland, Sweden, Norway, Estonia, Latvia, Lithuania, Belarus and Poland. Kesko and more than 1,100 K-retailers form K Group.

K Group is the second biggest operator in the Finnish grocery trade with a market share of approximately 37.2% (after the acquisition of Suomen Lähikauppa). 1.2 million customers visit K-food stores every day. [source K Group]. The K-food store chains are K-Citymarket, K-Supermarket, K-Market and Neste K service stations.

K-Citymarket chain has 81 stores all around Finland and looks "to tailor product selections and services to suit the needs of local customers". [Kesko annual report 2017]. Kesko has divided its customer base into five differents segments: A, B, C, D, E and wants its weekly promotions to optimally "cover the needs" of those different segments. Customers segments are discrete groups that are constructed according to specific psychographic, behavioral or needs criteria (Kotler, 1997).

The grocery retail industry is a highly competitive industry with razor thins margins (McKinsey, 2018). Promotion are strategic tools and temporary price promotions are commonly used by retailers to increase sales and traffic into the stores, as well as for the introduction of new products and the improvement of customer loyalty.

Promotions have also an important impact on a retailer's profitability and thus scheduling sales promotions more accurately can lead to an important increase in profits.

1.2 Motivation

Kesko has multiple types of promotions which are available all year round. Our focus will be on weekly promotional offers. These promotions are valid in all stores and regions in Finland and for all customers. There are always 4 products on sale and the products change twice a week. The first batch of products is valid from Monday to Wednesday and the second batch is valid from Thursday to Sunday.

The idea for the algorithm is to output 16 promotions at a time. This would correspond to 64 products and 8 weeks of promotions.

Thus our task can be divided into two major parts:

1. Find 64 'attractive' products. With these 4 products, we target 4 segments, as a one-to-one mapping, i.e. each product targets a single customer segment.

2. Decide the discount percentage for each product.

By combining these two parts we would obtain the desired algorithm.

1.3 Objectives

During the first phase of our project we divided out tasks into two major parts: Part 1: Output 16 promotions at a time and find 64 'attractive' products. With these 4 products, we would target 4 segments, as a one-to-one mapping, i.e. each product targets a single customer segment. Part 2: Decide the discount percentage for each product. By combining part (1) and part (2), we would then obtain the desired algorithm.

The project objectives have now evolved: it has been decided with Kesko to put all our effort into the optimization problem whilst leaving part 2 -finding out the optimal discount percentages (intervals/elasticities) for each product (family)- out of the scope of the project.

At present, our most important task is thus to find the optimal 4-product combinations for 16-week periods so that each week is quite even with the objective functions result.

2 Literature Review

Research Strand	Topics	Key Results	Theory	Research Methodology/data
Dynamic Promotional Effects	Temporal evaluation of promotional impact on key performance indicators: -sales - market share - profit	(1) Price promotion increases price sensitivity in the long run. (2) Promotional effects on sales do not persist in the long run.	Classic economic theory (utility & maximization)	Multivariate regression models, scanner data
Choice models	Brand choice models. Simulation of consumers' purchase decisions.	Multinomial logit models found to be precise instrument for modeling/explaining consumers' purchase decisions.	Classic economic theory (utility & maximization)	Probabilistic choice models (Logit), scanner data
Reference Effects of Price Promotions	Role of reference prices in affecting consumers' expectations of price promotions.	Price promotions will decrease a brand's reference price and thus negatively affect a brand's profit.	Behavioral economics= behavioral pricing.	Experiments + multivariate analysis on scanner data
Equilibrium Pricing Strategies	Competitive promotional strategies and equilibrium solutions of pricing games.	(1) National brands promote more than private labels to retain consumers. (2) Price promotions less profitable for brands with loyal customers than for brands with less loyal customers.	Microeconomic theory = game theory	Formal mathematical models, conceptual reasoning
Retail Promotions	Optimization of promotional strategies to increase store performance.	(1) Promotions will influence retailers' profits because of the impact on overall store traffic. (2) Price promotion has a direct effect on brand substitution in stores.	Microeconomic theory = theory of the firm	Structural equation models, multivariate regression models, scanner data.

Table 1: Compendium of the Main strands of Research on Pricing Promotion (adapted from Kuntner & Teichert).

Price promotions are essential components of a company's marketing strategy and are defined as "temporary price discounts offered to a customer" [17]. Their impact on retailers' margins is extremely important and as such should not be reduced to mere devices activated to boost sales but instead should be understood as highly strategic pricing instrument.The extant knowledge on price promotions is comprised of several research streams, each of which has a different focus, theories and methodologies (see Table 1 above, for a summary of the research strands, theory, tools methodologies).

(1) The first research stream focuses on the effects of dynamic promotional on key performance indicators, such as revenues, profits or market share. The findings indicate that over long periods of time, price promotions increase price sensitivity, however without any long term consequences on revenues or profits.

(2) Another stream focuses on equilibrium pricing strategies. Price promotion reaches a strategic, game-theoretic level that show that in equilibrium, national brands (branded products with high level of recognition) promote more intensively than private (generic) labels in order to retain customers.

(3) A third research stream is interested in the modeling of brand selection by consumers and shows that so-called multinomial logit models are a precise and detailed instrument to model and explain consumers' purchase decisions.

(4) The psychological effects of price promotions constitute another area of research and show that price promotions lower a brand's reference price and can affect negatively price expectations and profits.

(5) The fifth streams focuses on retailers optimal promotional strategy to enhance store performance. Promotions affect retailers' profits through their impact on overall store traffic.

3 Data and Methods

The data-set obtained and analyzed is presented in this Section. The methods used in the analysis are also briefly explained.

3.1 Data

The team received store-level data on shopping behavior during promotion campaigns for the K-Citymarket grocery chain (81 stores). More specifically, we have one entry for each of the store-customer segment-product-campaign combination. Such entry contains information on the sales (euro), the sales quantity, the discount percentage and the same information during the same days one week earlier which can be used as a baseline for normal sales during the period in question. The entries also contain information about the rest of the basket when said product has been purchased. This gives information about how profitable a product on promotion truly is, as we can calculate how much profit the product brings through other products. Further, we can estimate how many of the promotional product the typical customer buys. By combining this information with the quantity and baseline quantity of sales, we can estimate how many more customers that product brings in, i.e. how many customers were lured by the product.

We also received information about the campaign as well as information about the products. With this information we can normalize the entries so that they are all in units per day, thus making them better comparable. The information about the products contain groupings in higher hierarchy levels, e.g. the levels for a banana could be Chiquita banana, banana, fruit, fresh product. This will come in handy later in the process.

3.2 Preprocess

We screened the data provided by Kesko and carefully considered which key data would be needed to perform the tasks properly. We started the process by performing an exploratory data analysis of the provided data sets. We reviewed the explanatory values and created plots to obtain basic insight in the data as well. On account of those findings, we then identified the outliers, then run those data entries by the experts at Kesko and in the end handled them accordingly.

We also filtered out data such as holidays as the sales during holidays appear as peaks, which would introduce bias into our analysis if left be. We also removed the campaigns which were not weekly promotions (3 or 4 days long), as the length of the campaign influences the sales per day during it. For instance one-day campaigns have a much larger daily sale than monthly long ones.

We decide and elicit appropriate features to include in our analysis. Such features are: incremental sales and discount percentage. The incremental sales is calculated by comparing the sales during the promotion to the baseline (week before). When calculating this we first need to filter out all observations that have a holiday either this or the previous week. This again as the peak in sales would affect our results in an unnatural way. Discount percentage needs eliciting because of the format the data is in (discount euro is provided).

3.3 Key attributes

We were set to identify the products that attract different customer groups and possibly find out new interesting products along with it. We decided together with Kesko's experts to use attraction-repulsion matrix for this task. Forming an attraction-repulsion matrix is quite an easy task in mathematical point of view but it serves its purpose very well.

Building the attraction-repulsion matrix starts from putting together a contingency table. We consider a dataset of size n described by two qualitative variables, $\mathbf{X}$ with categories $P_1, \ldots, P_J$ and $\mathbf{Y}$ with categories $G_1, \ldots, G_K$, which are mutually exclusive. Data in a contingency table must be at same scale and cannot have negative values. The data is then displayed as a two-way contingency table. Table 2 illustrates a possible case. The columns show the 4 different customer groups and the rows the 1077 product families. As our underlying measure we used sales as euros.

	G_1	G_2	G_3	G_4	
P_1	1000	250	700	1100	3050
P_2	800	1350	600	400	3150
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{1077}	2200	1350	600	4000	8150
	1 200 000	800 000	770 000	1 500 000	4 270 000

Table 2: Example of how two-way contingency table, when as a underlying measure we have sales in euros.

For instance, when interpreting the Table 2, we see that product 1 (P_1) has had 1000 euros of sales within customer group 1 (G_1). Then, logically, the whole sales of certain product can are calculated as a row sum and whole sales of certain customer group as a column sum. These values are critical in forming the attraction-repulsion matrix and especially the next step involving the calculation of joint relative frequencies.

The value of the numbers n_{jk}, i.e. $(P_j,\ G_k)$ is naturally relative to the total number of observations, n i.e. total sales, which in our example is 4 270 000 euros. Thus, it is advantageous to analyze the contingency table in the form of joint relative frequencies. From the contingency table, it is straightforward to compute the associated joint relative frequency table where the elements of the contingency table are divided by the number of total sales n leading to $f_{jk} = \frac{n_{jk}}{n}$.

	G_1	G_2	G_3	G_4	
P_1	$\frac{1}{4270}$	$\frac{1}{17080}$	$\frac{1}{6100}$	$\frac{11}{42700}$	$\frac{1}{1400}$
P_2	$\frac{2}{10675}$	$\frac{27}{85400}$	$\frac{3}{21350}$	$\frac{1}{10675}$	$\frac{9}{12200}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{1077}	$\frac{11}{21350}$	$\frac{27}{85400}$	$\frac{3}{21350}$	$\frac{2}{2135}$	$\frac{163}{85400}$
	0.26	0.24	0.22	0.28	1

Table 3: An example of joint relative frequencies for the two-way contingency table.

With these calculated we can finally form the elements of the attraction repulsion matrix D, which are given by

$$d_{jk} = \frac{n_{jk}}{n^*_{jk}} = \frac{f_{jk}}{f^*_{jk}} = \frac{f_{jk}}{f_{j.}f_{.k}} \tag{1}$$

An attraction rate d_{jk} superior to 1 denotes that a certain product is more attractive to the customer group it refers to than to other customer groups. In other words, the overall sales of the product are more frequent among this group compared to other customer groups. This attraction repulsion matrix is one of the objectives that will be maximized when choosing our products. When attraction rate d_{jk} is under 1 it is considered that a certain product is repulsive. In Table 4 you can see how our attraction repulsion matrix would look like with our example case.

	G_1	G_2	G_3	G_4
P_1	1.261	0.342	1.043	1.288
P_2	0.977	1.786	0.866	0.454
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{1077}	1.038	0.690	0.335	1.753

Table 4: Attraction-repulsion indices for our example case were we considered sales in euros as an underlying measure.

We also construct other matrices of the same form as above $\mathbb{R}^{|P|\times|G|}$, where P is the set of products and G the set of groups. These other matrices depict other key objectives. For instance we group the sales in euro by product and customer segment and take the mean (over all campaigns and stores) to obtain a value: the average sales in euro per store, per campaign, per day.

We construct similar matrices with the underlying measures profit, incremental sales, incremental profit, basket profit and basket sales. The incremental values are, as explained earlier, constructed by comparing the values to the baseline. The basket values are already in the provided data and thus only need the same transformation as e.g. sales above.

Even though all these objectives are important we only utilize two of the constructed matrices, namely AR and sales in our implementation. This is because there is no clear answer of which objectives should be included and we construct our script such that Kesko can easily try out different combinations of the objectives in their further development of the algorithm.

3.4 Optimization

The pipeline for our process is finalized by solving two optimization problems. In the first step we select 16 products and in the second part we allocate the products into weeks.

3.4.1 First optimization problem

We use multi-attribute value theory to combine the multiple objectives into a solvable optimization problem. We use linear attribute-specific value function that map the attribute $v : \mathbb{R} \to [0, 1]$, via an affine transformation. Multi-attribute value theory is more thoroughly presented in [1].

In practice, this means that, for instance, the highest attraction-repulsion value for a certain group is mapped to the value 1, while the lowest for a group is mapped to 0. The rest are mapped somewhere in between, in a linear manner. The next step in multi-attribute value theory would be to elicit attribute-weights w from the decision maker using some framework. These weights represent how much the decision maker values a certain attribute compared to others. As the value functions are tied to a group we have multiple different weights connected to them. For simplicity we use the same weight for all groups connected to a certain attribute. To reduce clutter we also drop the group dependency when denoting the function, e.g. $v_S(\cdot)$ denotes the value function for the sales. We require that all attribute-weights sum to one.

We construct the optimization problem by introducing a binary incidence matrix $X \in \mathbb{R}^{|P|\times|G|}$. An element in X is one if that product is chosen for that group, and zero otherwise. We can now write the problem as

$$\max \quad \sum_{p\in P}\sum_{g\in G} X_{pg} w_{AR} v_{AR}(AR_{pg}) + X_{pg} w_S v_S(S_{pg})$$

s.t.

$$\sum_{p\in P} X_{pg} = 16 \qquad \forall g \in G$$

$$\sum_{g\in G} X_{pg} \leq 1 \qquad \forall p \in P$$

In the objective function, we simply multiply the attribute-specific value function with the attribute-weight and sum them all together. The two constraints ensure that the problem is feasible. The first one ensures that 16 products are chosen for every group and the second one that a product can at most be chosen to one group.

We strive for a more robust solution than just relying on a fixed attribute weight exactly depicting the preference relations between attributes. Thus, we utilize Pareto optimality in solving this multi-objective optimization problem. In Pareto optimal solution one cannot improve the value of one objective function without reducing the value of another.

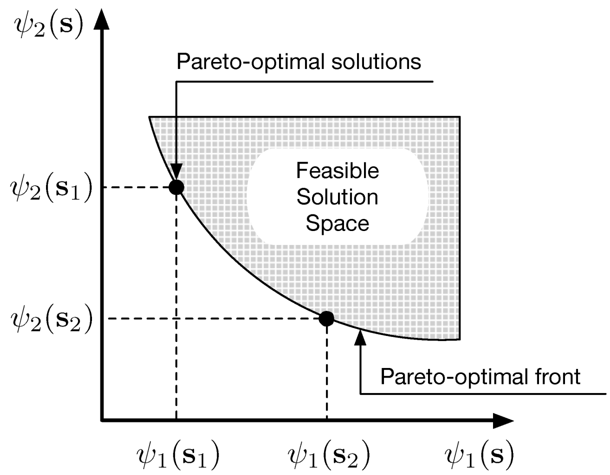

Figure 1: Example of Pareto optimal solutions with two objective functions: ϕ_1 and ϕ_2 that are to be minimized.[16]

In the example presented in Fig. 1 we see the feasible region, two Pareto optimal solutions and what is called a Pareto-optimal front. The front basically consists of the entire set of Pareto-optimal solutions.

We use a deterministic variant of the weighted sum approach to obtain the Pareto optimal solutions. The weighted sum approach is presented more thoroughly in [15]. The decision maker gave us incomplete preference statements for the weights, and from them we can restrict the values of w_{AR} and w_S. We then simply start from one end of the interval, solve the problem and increment the weight slightly, and solve it again. This process is repeated until we reach the other end of the interval. The difference compared to the normal weighted sum approach is that the weights are usually drawn randomly, while we instead use the deterministic approach.

By removing the duplicate solutions, we can calculate the core indices for the products. A product has a core index of one if the product is in every Pareto optimal solution obtained. This type of product is very robust, and choosing it can be done rather confidently. If a product on the other hand is not present in any of the Pareto optimal solution, its core index is zero and we should not choose the product to our final solution (not selecting the product is again a robust choice). This allows us to sort products by their core indices and simply choose the 16 top ones for every group.

3.4.2 Second optimization problem

The second optimization problem consists of allocating the 16 products per group into the 16 weeks. We introduce a incidence variable $Y \in \mathbb{R}^{|P|\times|G|\times|W|}$, where W is the set of weeks $W = \{1, 2, \dots, 16\}$. An element of Y is 1 if that product is chosen for that group and that week.

The reason for carrying out the optimization in two steps is that there are specific requirements for the allocation. First of we do not want four (or even more than one) similar products chosen for a specific week. For instance promoting beef, fish, pork and minced meat at the same time is not a good combination, and a constraint will be added to address this issue. Further we want the total sales to remain rather even throughout the weeks. This will be addressed by the objective function. One way to do this is to maximize the minimum sales per week, i.e. $\max_Y \min_W$ total sales on week w. This leads to the linear problem

$$
\begin{aligned}
\max \quad & \delta \\
\text{s.t.} \quad & \\
& \sum_{p\in P}\sum_{g\in G} Y_{pgw}S_p \geq \delta && \forall w \in W \\
& \sum_{p\in P} Y_{pgw} = 1 && \forall g \in G, w \in W \\
& \sum_{w\in W} Y_{pgw} \leq CI_{pg} && \forall p \in P, g \in G \\
& \sum_{p\in H_i}\sum_{g\in G} Y_{pgw} && \forall w \in W, H_i \in H \ .
\end{aligned}
$$

Notice that the objective accounts for the sales originating from all groups (S_p) when a product is chosen. The first constraint is related to the re-writing of the max-min objective described above. The second constraint ensures that a product is chosen for every (group, week)-pair. The third constraint uses the outcome of the previous optimization problem; $CI \in \mathbb{R}^{|P|\times|G|}$ is a binary matrix indicating whether a product made the cut (of being one of the top 16 products according to core index) in the Pareto optimal solutions.

The final constraint ensures that multiple products from the same higher level category are not allocated to the same week. We denote a set of product in the same higher level category by H_i and the set of these set of products as H. The higher level category is defined by how Kesko categorizes their products as was touched in section 3.1.

We implement the process in Python using an open-source solver for Mixed-Integer Linear Programming problems, PuLP [13]. The solver was chosen based on multiple suggestions from different sources. Frequently noted in the suggestions is that the gap between open-source and closed-source optimization software in Python is big. This we noted when trying to solve the allocation problem with PuLP, as it did not find a solution. A closed-source solver like Gurobi [14], which seems to be the recommended software for MILP problems in Python, could most likely solve this problem. However, we deem that balancing the total sales perfectly is not necessary for the prototype solution and we simply introduce a hard constraint (that is found by testing to solve the problem, if it fails then lower the bound) for the lower bound of the total sales. This balances the sales rather well and PuLP is able to solve this problem. The constraint then looks like

$$\sum_{p\in P}\sum_{g\in G} Y_{pgw}S_p \geq LB \qquad\qquad \forall w \in W\ .$$

and the objective function becomes irrelevant, as we only require any feasible solution.

3.5 Validation and verification

For the model validation and verification we followed an iterative process with regular meetings and email correspondence with the Kesko team. In this correspondence we have asked them if the results seem reasonable, which they do. Further, the products outputted closely correspond to Kesko's beliefs of their customer segments, further validating and verifying the model. We have also critically gone through all implementations within the team to find possible bugs.

4 Results

As a result of our optimization model we got two different tables containing 16 optimized promotion weeks such that the value of objective functions are as equal as possible. We chose two different tables such that in one we included also coffee and salmon products and in the other one those are excluded. It is common sense that these are the most attractive products among pretty much all the customer groups. Thus it made also sense to look at the results without these products to discover new interesting and attractive products within each customer group. Bear in mind that due to non-disclosure agreement (NDA) we cannot deep dive into the quantitative details of the results.

In the table 5 containing also coffee and salmon products, we may see that our model suggest to promote either one of these products nearly in every weekly promotion. In overall view we may see that many of these suggestions are quite meat-heavy products. Also typical customer group preferences can be extracted from the result table 5, e.g. customer group A favors a lot of minced meat, whereas customer group C seems to favor a lot of salmon and on the other hand customer group D favors a lot of different types of sausages. Interestingly we may see that customer group B differs totally from these three customer groups. Customer group B has a quite nice variety in the suggestions for promoted products, which are specifically targeted to them to follow these promotions to K-Citymarkets.

#	Customer group A	Customer group B	Customer group C	Customer group D
1	BEN&JERRY'S 0,5L	Hätälä Lohifilee D-leik Norj 1xn10kg vak	Filos halloumijuusto 200g kevyt	HARTWALL JAFFA 0,33L TLK 12-PACK
2	Hyvä Apaja klohifile C-leik 1xn10kg vak	Bonduelle kikherneitä 310/265g	Emännän vehnäjauho 2kg pk	Atria jauhelihapihvit XL 230g
3	HK TAKUUMUREA NAUDAN SISÄFILEET	Filos halloumijuusto 200g kevyt	Päärynä Conference kg	ISO-TUUTTI 160G
4	Pirkka naudan jauheliha 10% 400g suom	VALIOJOGURTTI 1KG	K-Menu lanttu 2kg Suomi 1lk	PEPSI JA JAFFA 0,33L KMP 24-PACK
5	BEN&JERRY'S MOOPHORIA	Bonduelle linssit 310g/265g	Avomaankurkku Suomi kg	Hätälä Lohifilee D-leik Norj 1xn10kg vak
6	Mansikka Calinda 400g ES/PT 1lk	HoviRuoka vege burger 110g	Kalamesta kirjolohifilee C 1xn10kg vak	Riitan Herkku isänpäiväkakku 700g vadelm
7	HK naudan jauheliha 10% 600g	Flora Culinesse kasviöljyvalmiste 500ml	Omena Gala kg	Atria lenkki 500g punainen
8	PIRKKA FINGERFOOD 250G	Santa Maria lime pepper maustemylly 90g	Appelsiini 5kg 3-4 Espanja 1lk	KARINIEMEN KANANP FILEESUIKALEET250-300G
9	KANES SODA POP 0,33L KLP	Lerøy merilohifilee 1xn10kg rdton vac	Flora Culinesse kasviöljyvalmiste 500ml	VALIOJOGURTTI 1KG
10	Atria naudan sisäfilee 1xn2kg takuumurea	Pinaatti 70g IT/ES 1lk	Myllärin vehnäjauho 2kg leivk	Kauppiaan oma grillimakkara 400g
11	Topfoods naudan sisäfilee Uruguay kg	Rypäle tumma 500g stön BR/PE/NA/ZA 1lk	Peruna 5kg keltainen Suomi	Eriksson BENELLA kirjoloh kok n.10kg vak
12	BEN & JERRYS TOPPED 470ML	ATRIA SAMETTISET KEITOT 300G	Kalamesta lohifilee A Norja 1xn10kg vac	ATRIA HIILLOS GRILLIMAKKARAT 320G-400G
13	Pirkka Parh lohifilee C 4xn800g rton vac	ISO-TUUTTI 160G	Omena Idared kg	Valio Hyvä suomalainen Arkijuusto 1,25kg
14	ATRIA PERHETILAN KANAN SISÄFILE 460-600G	NATURDIET SMOOTHIE330ML	Apetit lohifilee A 8xn1,25kg superio vac	Atrilli grillimakkara 400g
15	HK naudan jauheliha 17% 1kg	Mango syöntikypsä 380g Brasilia 1lk	K-Menu porkkana 2kg Suomi 1lk	FAZER PUIKULAT 500-550G
16	COCA-COLA 0,25 TLK	Barilla risoni 500g	Hätälä lohifilee A 10xn1kg Norja vac	Atria PT viljapo juhlakink4xn5kg lton pa

Table 5: Final results of our optimization model including all the products.

By excluding coffee and salmon products, we can see very nice results in table 6 where promoted products seem to be very diverse without couple of exceptions: minced meat and chicken. Due to not excluding minced meat we may only see slight differentiation in products targeted to customer group A compared to situation table 5. However, the rest three customer groups (B-D) seem to have new suggestions for products to be promoted. We may see that many of these products such as cheese, cottage cheese, butter, blue cheese, ice cream and potatoes already resemble the ones that many of us have seen previously in Kesko's promotion flyers. In overall many of these products presented in this table seem to be some kind of vegetable or fruit, which was in Kesko's experts' opinion very informative finding.

#	Customer group A	Customer group B	Customer group C	Customer group D
1	Atria naudan parempi jauheliha 10% 700g	Viikuna 200g Turkki 1lk	Emännän vehnäjauho 2kg pk	HARTWALL JAFFA 1,5L KMP 2-PACKIT
2	HK sika-nauta jauheliha 22% 1kg	Rypäle tumma 500g stön BR/PE/NA/ZA 1lk	Peruna 5kg keltainen Suomi	Elonen tropical hedelmätäytekakku 800g
3	Atria nauta-viljapossu jauheliha 250g	NATURDIET ATERKORVIKEPATUKKA 57-60G	Omena Idared kg	Kivikylä palvarin pyörykät 500g
4	Topfoods naudan sisäfilee Uruguay kg	Kurkuma 50g luomu Peru 1lk	Omena Gala kg	COCA-COLA 0,33L TLK 15-PACK
5	Filos halloumijuusto 200g kevyt	Satsuma kg	Pingviini jäätelö 1l vanilja pa	Elonen mansikka täytekakku 800g
6	TWISTER MEHUJÄÄ 70-80G	Pinaatti 70g IT/ES 1lk	PIRKKA KERMAJUUSTOT 900G-1KG	HARTWALL JAFFA JA PEPSI 0,33L KMP 24-PAC
7	ATRIA PERHETILAN KANAN SISÄFILE 460-600G	Pirkka rypäle tumma 500g Et.Afrikka 1lk	NATURA KERMAJUUSTO 1KG	Riitan Herkku isänpäiväkakku 700g vadelm
8	Mansikka Calinda 400g ES/PT 1lk	ARLA KESO RAEJUUSTOT 200G	Atria naudan jauheliha 1000g 17%	Filos halloumijuusto 200g kevyt
9	Snellman kunnon naudan jauheliha	NATURDIET SMOOTHIE330ML	Mandariini kg	HARTWALL JAFFA 0,33L TLK 12-PACK
10	BEN&JERRY'S 0,5L CORE	Sitruunaruoho 50g Thaimaa 1lk	Chiquita banaani kg	PEPSI JA JAFFA 0,33L KMP 24-PACK
11	Pirkka pensasmustikka 125g CL/PE 1lk	Santa Maria lime pepper maustemylly 90g	Avomaankurkku Suomi kg	Valio Hyvä suomalainen Arkijuusto 1,25kg
12	HK naudan jauheliha 10% 600g	Magnum Bomboniera 12x12ml pa	Päärynä Conference kg	Atria jauhelihapihvit XL 230g
13	KANES SODA POP 0,33L KLP	Bonduelle linssit 310g/265g	Klementiini kg	Atria viljaporsaan ulkofilee 1xn1,5kg na
14	BEN&JERRY'S MOOPHORIA	Pirkka banaani kg	Pirkka leivontamargariini 500g lakton	HK pieni juhlakinkku 6xn3kg harms. pa
15	Atria parempi nauta jauheliha 10% 700g	Valio Aura sinihomejuusto 170g pala	K-Menu lanttu 2kg Suomi 1lk	Antell vadelmakakku 860g 10-12 hlö
16	HK naudan jauheliha 10% 400g	Pirkka miniluumutomaatti 250g ES/MA 1lk	Myllyn Paras torttutaikinalevy 1kg	Atria lenkki 500g punainen

Table 6: Final results of our optimization model excluding coffee and salmon.

5 Discussion

5.1 Reflection on literature

Even though the entire pipeline was not based on literature, the individual modules/methods are strongly anchored there.

One major reason for not finding many complete pipelines about optimizing the promotional offers is that it is a clear business section and the players do not want to reveal their cards. The major retailers either promote the products themselves and do not want to help their rivals or an external entity has performed the optimization and want customers to buy their optimization software/pipeline.

The most relevant information we obtained from these were that the baseline is important to get exact and that some kind of machine learning based algorithm is used, which in essence can mean everything from complex neural networks to basic statistics.

5.2 Assessment of the results

We can say that our results are feasible, since our case providers were pleased with our results. After descoping the project we manged to meet the expectations of Kesko's experts. The results work as a great skeleton for their further development and thus we tried to maintain our work as integrable as possible. Our optimization model manages to find new attractive products for different customer groups. The optimization model also manages to allocate the products into weeks such that the weekly sales are rather balanced.

We managed to prove that there really is clear preferences for customer groups and certain products. These findings were discussed more thoroughly in section 4. This was of course already discovered by Kesko's surveys, which they used to distinguish different customer groups. This core information enables effective product promoting, where every customers' needs are met with highest certainty.

However, the way in which the calculations are done allow a slight domination from some of the higher volume selling products, such as coffee, salmon and minced meat, which are in daily use of customers. This tend to skew the results and hide the novel discoveries. Also some of the suggestions seem to be quite unreasonable, like Father's Day cake, and other different type of cakes. With common sense these products do not seem to be worth to promote in weekly promotions but rather in special occasions.

We also found out that there might be a problem with sorting the most attractive products under each customer group, which would need further investigation from the Kesko's end. Currently nameless' two customer groups are heavily weighted with the so called top tier products and the other two customer groups with less attractive products. Another problem in our opinion is the way how attraction-repulsion indices are calculated. The method favors products, which have low volume and where some of the customer groups haven't bought the product at all. This also leads to skewed results. On the other hand, this allows novel products with only minor sales to be discovered as potential promotional products.

There is also a completely new dimension, which we have not taken into account that has a huge impact on how promotional products are selected in practise. This is the negotiations with industries, that provide the products to grocery stores. Kesko has to negotiate a deal with these product-providers, which may have their own set of products that they would like Kesko to promote. It may lead to situation where product-providers demand that their products must be promoted and only with certain discount threshold and so on. Negotiations with product-providers can be seen as a unknown factor, which have randomness in some extend and thus make our models' suggestions a bit more incomplete.

As a result of these observations, our model should be used only as an aid for the decision maker, not as the final truth.

5.3 Suggestions for further improvement

As was discussed in the literature, the baseline is important. Thus one suggestion for further improvement of the model would be to to improve the baseline, and re-run the model. According to our understanding Kesko is currently implementing a model for an improved baseline.

The data analyzed in this project was only data from K-Citymarkets. Because of the limited data the results may be somewhat non-robust. By introducing more data, e.g. from store of different sizes or more promotions used in K-Citymarket, the robustness would improve further.

Another suggestion is for Kesko to think about what objective function to use in the first optimization problem. As was stated in the end of section 3.3, we have constructed more matrices with key values that could easily be concluded in the objective function. We suggest a result based approach, i.e. first changing the objective function and then evaluating that objective function based on how suitable the results of the model appear.

Further, more sophisticated constraints could be added to the second optimization problem. These may include restrictions on product combination (e.g products used as ingredients or elements of certain dishes or meals). If all the products fit perfectly into one basket then the promotion is not good, as it leads to a negative profit for Kesko. Other sophisticated constraints would be to restrict most bought products such as coffee, salmon, minced meet and chicken by, for instance, limiting the number of times they can be chosen in the 16 weeks.

Negotiations with product-providers is a complex metric due to their unpredictabilitiness. To tackle this problem, in further development one may implement a second best choice to one promotion slot and pick that product in case of setback in negation. It might also be a good idea to try to calculate probabilities for default and thus in some way try to predict the expected profits of promoted product, which are under negotiation. These were some of the ideas for further improvements and there certainly is still many other useful and critical aspects, which need attention.

6 Conclusions

Optimizing marketing promotions for grocery retail has been of strategic interest for retailers for a very long time. We tackled this problem using multivariate statistical analysis and optimization methods and we came to the conclusion that it is indeed possible to optimize marketing promotion- at least to a certain extent.

However, due to its complex and multi-dimensional nature, finding a suitable definition for the optimal results is quite difficult. Furthermore, defining the needs of the model and the goals more precisely in order to get the wanted results is a priority. Then, it becomes possible to determine how complex the model should be.

One of the biggest problems of our endeavor was the validation of the model. It is only possible to validate the model whilst it is in action. After a promotion week, the sales volumes and profits can be compared to either their baseline or previous promotions and based on that it is possible to assess whether the promotion week was successful or not. This situation can become worrisome as it makes it hard to test the model's effectiveness and as a consequence hard to try multiple different alternative solutions. Currently our model is using only two features and as a consequence missing many dimensions. Our model doesn't manage to capture all the necessary information regarding successful promotion weeks.

From a universal point of view, one could say that it is impossible to satisfy everybody at the same time (one could see this in daily life with traffic lights for instance). Not everybody can have green lights at the same time, otherwise resulting in accidents.

Pre-results have been promising and thus we are confident that Kesko will find a satisfying solution to this problem.

7 Self Assessment

7.1 Project Review

7.1.1 Scope

The scoping of the problem was done very effectively as it was realized quite early that a re/de-scoping was necessary: the optimization problem -finding the optimal 4-product combination for 16-week periods- thus became the main focus whilst finding optimal discount percentages for each product family was abandoned.

7.1.2 Risks Project execution

Risks (such as an inadequate optimization model or scoping of the model) were managed quite effectively because of the agile and lean methodology deployed during the project. The team applied an iterative approach to the production of the deliverables with each week -during the weekly meetings held as workshops between Kesko experts and the team- a review of specific results. The feedback was immediate, which then gave the opportunity to incorporate possible changes from one meeting onto the next. At the end of each iteration, the team and the Kesko experts reviewed potential opportunities for improvement.

Because of the very clear communication of the objectives and of the expected methodological steps to be taken to solve the optimization problem, the team was able to prioritize quite quickly and work precisely on what was valuable to Kesko as well as propose adjustments, correct erroneous assumptions.

The project was managed quite effectively with regular, weekly reviews. This structured approach gave the team the opportunity to progress effectively (little time was wasted) and gain insights on how to develop the model.

7.1.3 Schedule

The scheduling of the project and its different stages was regularly reviewed internally for possible adjustment and every single stage of the project proceeded on schedule.

7.1.4 The amount of work

The amount of work was quite high but adequate to solve the optimization challenge given by the client.

7.2 In what regard was the project successful?

The project was extremely successful. The model will be put into production by Kesko.

7.3 In what regard was it less so?

It was quite a rewarding project.

7.4 What could have done better?

The scheduling of the company visits put an important strain on the workload.

7.4.1 Team

A multidisciplinary team with the right capabilities to deliver the expected product to Kesko.

7.4.2 Teaching staff

The brief for the optimization problem was included in the course at a later stage than other companies, perhaps a clearer objective would have helped the team at the beginning of the project.

More individual group reviews -in person- by the teaching staff would have been welcomed.

The scheduling of a company excursion during exam week was quite problematic.

7.4.3 Client organization

The collaboration with the client organization worked really well in our opinion.

Possibility to implement the model in R earlier since Kesko will translate the implementation into R in order to match their usual software tools.

The project worked as a pilot giving Kesko the needed information for potential implementation.

References

[1] B. Wierenga and R. van der Lans (Eds), Handbook of Marketing Decision Models, Springer, 2017.

[2] P. Kotler, Marketing Management Analysis, Planning, Implementation, and Control, Prentice-Hall International, 2016.

[3] M. Lintunen and M. Stolze, Suomi Syö 2018 , Research Report, 2018.

[4] D. Kuijpers, V. Simmons, and J, van Wamelen, "Reviving grocery retail: Six imperatives", McKinsey's Perspectives on retail and consumer goods, Winter 2018/19

[5] https://www.mckinsey.com/industries/retail/our-insights/reviving-grocery-retail-six-imperatives [Accessed 25.02.2019]

[6] G. Singer and K. Dickman, "Can smarter pricing and promotion reduce the emphasis on discounting?", Ernst & Young (2016)

[7] https://www.ey.com/en_gl/consumer-products-retail/can-smarter-pricing-and-promotion-reduce-the-emphasis-on-discounting [Accessed 11.02.2018]

Corporate Reports

[8] https://www.kesko.fi/globalassets/03-sijoittaja/raporttikeskus/2018/kesko_annual_report_2017.pdf [Accessed 27.02.2019]

[9] https://www.kesko.fi/globalassets/03-sijoittaja/raporttikeskus/2018/q3/kesko-q3-2018-interim-report.pdf [Accessed 11.02.2019]

[10] https://www.kesko.fi/en/company/divisions/grocery-trade/ [Accessed 11.02.2019]

[11] https://www.pty.fi/fileadmin/user_upload/tiedostot/Julkaisut/Vuosijulkaisut/EN_2017_vuosijulkaisu.pdf [Accessed 11.02.2019]

[12] Clemen, R.T. (1996): Making Hard Decisions: An Introduction to Decision Analysis, 2nd edition, Duxbury Press, Belmont.

[13] PuLP. Documentation can be found `https://pythonhosted.org/PuLP/`. [Cited 14.5.2019.]

[14] Gurobi Optimization. Documentation can be found `http://www.gurobi.com/`. [Cited 14.5.2019.]

[15] Zadeh LA (1963) Optimality and non-scalar-valued performance criteria. IEEE Trans Automat Contr AC-8:59–60

[16] Z. Nascimento, D. SadokMODC (2018). A Pareto-Optimal Optimization Approach for Network Traffic Classification Based on the Divide and Conquer Strategy. DOI: https://doi.org/10.3390/info9090233. [Cited 15.5.2019.]

[17] Tobias Kuntner, Thorsten Teichert (2016). The scope of price promotion research: An informetric study, , Journal of Business Research, Volume 69, Issue 8, August 2016, Pages 2687-2696