Optimizing Promotions for Supermarkets using Data Analytics

Maxime C. Cohen Georgia Perakis

Abstract

A supermarket chain would like to use data analytics and optimization to improve their current promotion planning process. How should they do this? By how much can they hope to improve their bottom line?

Introduction

Jack, the founder and CEO of a supermarket chain, is always trying to come up with creative ways to improve operations in his business.[1] He opened his first store in 1993 and today, Jack manages seven different stores that sell thousands of different grocery SKUs (Stock Keeping Units), such as cereals, beverages, tea, dried fruits, and pasta just to name a few. Hoping to harness recent advances in big data and business analytics, Jack has hired a quantitative analyst to help him integrate data analytics, demand forecasting, and optimization to improve operational decisions at his supermarkets. In this case, we will focus on one specific aspect of this project - the planning of promotions and discounts for products in his supermarkets; i.e., what items to promote, how deeply to discount them, and when to do it.

Background

Jack immigrated to France from Morocco in the late seventies. Jack always had an entrepreneurial mindset, and was constantly on the lookout for business opportunities. Realizing how attractive the grocery space was, he began to look for a way to disrupt the industry in order to capture a piece of this very large cake.

Most of the supermarkets at the time were either small convenience stores located in cities, or large supermarkets in suburban areas. Jack realized there was an opportunity for an intermediate option: medium scale grocery stores strategically located in wealthy neighborhoods, focusing on premium and gourmet-type products (for example, imported items). After some serious market research, endless negotiations with bankers, and a dozen site visits, the business plan was about to materialize. The first store opened in January 1993.

Jack was a very hard worker and a deft decision maker. He was constantly required to make decisions under uncertainty. Which suppliers to order from? How should he place the different orders (small weekly orders, or larger monthly orders)? How to display the items on the shelves? What about pricing decisions? Finally, it is clear that hiring the right employees was of paramount importance. His efforts paid off and after three years of working 6 long days a week, the store became a profitable business. Jack was able to pay off his mortgage, draw a respectable salary for himself, hire a few employees, and more importantly, build a strong base of satisfied recurring customers. When one of his regular customers informed him that he moved several miles away but still came all the way to purchase groceries at Jack's store, Jack felt flattered and fulfilled. As a consequence, Jack felt ready to tackle the next challenge and grow his business: he opened a second store in a different neighborhood. This, he knew, would not be easy, but he also knew that the experience he had gained in his first store would serve him. In addition, he was aware of the potential benefit of

[1] The story used in this case is fictitious and was chosen to illustrate a real-world situation. Consequently, any details herein bearing resemblance to real people, dates, or events is purely coincidental.

economies of scale that would be available to him when operating two stores. His first store had also allowed him to build significant credibility with his creditors, as well as a burgeoning brand name.

Fast forward to 2004 – having gone from success to success, Jack opened a number of additional stores, finally opening his seventh store in 2004. Even though his stores were generating good business, Jack was in search of new ways to improve profitability (see **Exhibit 1** for a summary of the seven stores). Following the advice of his tech-savvy nephew Michael, Jack decided to hire Michael in June 2006. Michael had just graduated from one of the most prestigious institutions, and was eager to use his technical knowledge to bear on Jack's businesses. In particular, he wanted to analyze the large amounts of data Jack's stores generated on a daily basis, and use them to make better operational decisions.

Using Modern Methods

Michael quickly got up to speed with the business environment. On one hand, Michael was impressed by the success each of the seven stores enjoyed. On the other hand, he was shocked to find out Jack made most of his operational decisions based on experience and intuition, without the help of any quantitative decision support tools. He was convinced he could significantly improve Jack's bottom line by using modern forecasting and optimization methods. Impressed as he was by Michael's promises, Jack approached the project with a significant amount of skepticism. He had no doubt that quantitative methods could be useful and powerful, but he wasn't convinced these methods could truly supplement the intuition he had built over many years in the industry. As a result, he first tasked Michael with understanding how consumers behave and how store managers operate on a daily basis. After that, the next step was gathering data, and organizing it in a way that could show the promise of these new technological advances.

Of course, it took Michael a significant amount of time to study the vast amounts of data generated by the stores. In addition, Michael had to decide how to collect it, how to standardize the data from various disparate sources, and how to clean up the many inconsistencies that invariably appear in such large datasets. After a year of painstaking work, Michael was able to show Jack some fascinating results. Some, such as the best-selling items, the least-profitable ones, etc, were figures Jack had already seen. However, many others were quite complex and would have taken Jack weeks to calculate without Michael's help. For example, Michael was able to find the items that sold-out the most frequently, or discover than 80% of Jack's profits came from 34% of his products. Jack immediately understood the impact that these metrics could have on his operational decisions. Impressed by Michael's work, Jack gave him his first large-scale project: to improve the supermarket's demand-forecasting tools (which had been done so far by using simple spreadsheets), in order to inform ordering decisions from suppliers. After several months of hard work, Michael produced a predictive demand model that proved useful in informing Jack's ordering decisions, resulted in significant decreases in wasted merchandise, and corresponding increases in profit.

By the end of 2009, Jack was fully sold on Michael's abilities, and decided to entrust him with the supermarket chain's most thorny operational problem - pricing and promotion optimization.

Promotions in Supermarkets

In the supermarket industry, it is very common for manufacturers and retailers to use promotions. The main goals include: attracting new customers, increasing sales and profits, and encouraging existing customers to switch brands. The amount of money spent on promotions is significant: Nielsen estimates that manufacturers of fast-moving goods (products that are consumed regularly such as ice cream, soft drinks, candy and chocolate) spend about $1 trillion annually on promotions (Nielsen (2015)). Promotions play an important role in the fast-moving good, industry because a significant proportion of sales are made on promotion. For example, Nielsen found that 12-25% of supermarket sales in five European countries (Great Britain, Spain, Italy, Germany and France) were made on promotion. Not just are promotions important, the scale of the problem is also tremendous, namely, in an average store, thousands of products can be on promotion in any given week.

Despite the complexity of the promotion planning process, to this day it is still performed manually in most supermarket chains. Academic research as well as software and consulting firms have proposed efficient ways to design and study promotion optimization models. Their aim is to improve the promotion planning process (reducing man-hours) while at the same time increasing profits for retailers.

Promotion Optimization

Like many other supermarket chains, Jack's stores ran a number of promotions, in which various products were discounted for limited periods of time. The benefits of these promotions were manifold - they can serve to attract new customers, solidify relationships with suppliers, increase the store's profile by featuring some more unusual products, and help the stores dispose of soon-to-expire products.

Nevertheless, the operational problem of planning these promotions was a Teutonic task. Not only did Jack's overstaffed store managers have to identify which products (out of many thousands) to discount, but they also had to decide how long to discount them, and by how much, while ensuring that each product category did not get discounted too often or too seldom. Given the many factors involved in making each decision, they had no chance of making truly optimal decisions, and had to content themselves with doing their best given the expertise available to them.

Michael was convinced he would be able to improve on the current process using his expertise. He realized the remarkable impact a promotion planning tool could have on the stores' bottom line. By nature, he was eager to tackle the task.

His first job was to understand what the current pricing strategy used by store managers was. To his shock, he discovered that most pricing decisions were made on a fairly ad-hoc basis: based solely on intuition and past experience. For example, aggressive price discounts were applied to grocery items close to their expiry date, or around special promotional events throughout the year (e.g., holiday seasons). Michael also discovered that the store ran two different types of promotions: manufacturer's promotions (also called trade funds) and retailer's promotions (also called in-store promotions). The former is a price discount offered by the manufacturer to the retailer in exchange for some sort of 'special treatment', such as an exclusivity deal, a strategic display at the end of the aisle or next to cashier, etc. The latter type of promotions are decisions the store manager themselves could make, as they saw appropriate, to boost their sales and increase profits. Michael decided to gather data on past promotions, and build a model to help Jack optimize in-store promotions. To do so, Michael came up with a concrete plan comprising of five steps:

1. **Data collection -** succinctly and precisely describe the data requirements for the project. Collect this data, clean it, and integrate the various datasets involved.

2. **Demand estimation -** use the data to estimate how demand reacts to prices (e.g., how many more customers could Jack expect as a result of running a promotion on a given item) as well as other observable features (e.g., seasonality).

3. **Business rules** - carefully understand any business requirements on prices and promotions (e.g., how many promotions are to be set in any given week).

4. **Profit maximization** - formulate the optimization problem, that is, maximize the total profits made resulting from the promotion pricing strategy while satisfying the business rules. Then, solve the resulting optimization formulation to obtain price recommendations.

5. **What-if-scenarios** - test the robustness of the prices prescribed by the mathematical model in the previous step, and conclude by implementing the system's recommendations.

We next discuss each of these steps in more details.

Methodology and Results

Data collection

First, Michael decided to collect three years of data from each of the seven stores. For each store, Michael stored each transaction, including price, time and quantities purchased. He quickly found that this panel data was too rich to handle using even the most modern of statistical tools (a common problem encountered in many fields, given the massive amounts of data routinely generated by modern computer systems). To solve this problem, Michael decided to aggregate the data at a weekly level for each store separately. In other words, for each store he gathered one of the 156 weeks of data about the average price of each item, the total sales for the item, together with various features of the item (e.g., its size, flavor, etc.). This was done for many different items in each of the seven stores. At the end of this process, Michael had created a very large data set. As an example, see **Exhibit 2** for the data about one particular item in a specific store.

During the collection process, Michael also considered aggregating some of the data even further. For example, he could have pooled data across brands/styles/flavors or clustered different stores into groups of stores that shared similar features and aggregated the data within the groups. In analytical projects of this kind, deciding how much pooling to do, and what granularity to use when looking at data, is more of an art than a science, and often sets apart good statisticians from truly exceptional ones. As we shall see, even when the computational infrastructure available is able to handle data in its un-pooled form, it is often advantageous to carry out some data pooling to make certain patterns in the data more apparent.

Demand estimation

Having collected, cleaned, and pooled the data, Michael was confident that he could propose a mathematical model to accurately estimate future demand. Many such models already exist in the literature (see, e.g., Talluri and van Ryzin (2005) and Blattberg and Neslin (1990)). To simplify matters, Michael decided to perform his analysis separately for each store and each category of items (e.g., tea, candies, and toiletries are three distinct categories).

In these situations, it is tempting to use all the data available to construct a demand model, and then to test it on this same data. Unfortunately, this often results in inaccurate models that perform very well on past data (and give the model-builder considerable satisfaction!), but may not perform as well in the future.

To understand this problem, imagine a Professor hoping to teach his students operations. To help his students prepare for their final exam, he gives them a series of 20 practice problems (past data) to help them learn the subject (construct their model). Suppose he then builds his final exam by drawing only questions from this set of 20 questions. His students would be bound to perform very well, having already been trained on precisely this set of questions. A fairer approach would be to give the students 15 questions to practice on, and keep the last 5 for a final exam.

Michael realized this potential pitfall, and instead of using all 156 weeks of data to construct his model, he used only 104 weeks of data, and used the remaining 52 weeks to test his model.

Another problem Michael faced was that the data he had was often insufficient to properly construct a model - indeed, he only had 104 data points available to construct the model per product and store. In this respect, the pooling of data he had previously carried out came in useful - for example, by considering data at a category-level rather than a product-level.

Finally, Michael carried out extensive tests to decide which observable variables he should use to predict future demand. Did the price matter? What about the store in question? The time of year? Michael was only able to firmly answer these questions through extensive experimentation.

Exhibit 3 presents the sales for a particular product over two years (each point represents the cumulative sales during a given week).

Business rules

Interestingly, supermarket managers and retailers in general, are not free to arbitrarily choose what promotions to use in their stores. As we discussed, the presence of trade funds may dictate some restrictions on prices. Other examples of business rules that are common in practice include the following.

- **Prices are chosen from a discrete price ladder.** For each product, there is a finite set of permissible prices. For example, prices may have to end with a '9'. Usually, there exists a regular price (also called non-promotional price) as well as several other promotional prices. For example, a particular item can only be priced at 5.99, 5.49, 4.99 or 4.79.
- **Limited number of promotions.** The supermarket may want to limit the frequency of promotions for a product. This requirement applies because retailers wish to preserve the image of the store. For instance, a retailer may insist that a particular product is promoted at most four times per quarter (a 13 week period).
- **Separating periods between successive promotions.** A common additional requirement is to space out promotions by a minimal number of separating periods. Indeed, if successive promotions are too close to one another, this may hurt the store image and incentivize consumers to behave more as deal-seekers.
- **Total limited number of promotions.** The supermarket may want to limit the total number of promotions in any given category throughout the selling season. For example, for a given category it may be required to use at most 50 promotions per quarter.
- **Inter-item ordinal constraints.** Several price relations are dictated by business rules. For example, smaller size items should have a lower price relative to similar products with a larger size and national brands must be more expensive relative to private labels.
- **Simultaneous promotions.** It may be a requirement that several items are promoted simultaneously as part of a manufacturer's incentive or as part of a special promotional event.
- **Limited number of promotions in each period.** One can impose a limitation on the number of promotions in each time period. For example, promoting at most 10% of the items in the store during a given week.
- **Cross no-touch constraints.** An additional requirement may be that promotions for similar items need to be spaced out by a minimum number of periods.

Profit maximization

Having investigated the most important business rules, Michael was ready to formulate a mathematical optimization model. The objective is simply to maximize the total profits throughout the selling season. Jack and Michael agreed to use the following assumptions: (i) the price of each item is the same across all stores; (ii) the price remains constant for the entire week (i.e., prices are decided on Monday morning and stay fixed until Sunday evening); (iii) promotional planning is decided on a quarterly basis (i.e., prices are decided upfront for the next 13 weeks) and (iv) the decisions are made for each category of items separately.

As a result of these assumptions, our problem is no more than an optimization problem, the objective of which is to decide the price of each item over a subsequent 13 week horizon. Since the store handles thousands of different SKUs, the number of decisions is very large. For example, in the tea category, if the store stocks 76 different products, there are a total of $76 \times 13 = 988$ decisions to be made, all constrained by the various business rules. **Exhibit 4** shows the summary of four different categories found in Jack's stores.

Michael could formulate the profit maximization problem for each category as an optimization problem. Unfortunately, the problem was not easy to solve as it is a non-linear integer program. After researching efficient approximate methods for this sort of problems, Michael managed to find a way to solve the problem in less than a few seconds (the relevant methods can be found in the following two research papers Cohen et al. (2020) and Cohen et al. (2017)).[2]

[2]Reading the papers is optional and not required for solving the case. However, curious readers are encouraged to take a look.

As mentioned, Michael could now run the model and determine the optimal promotions strategies for each category of items. This allowed him to provide price recommendations for each item during each week of the coming quarter. **Exhibit 5** shows the suggested prices (prices are normalized to 1) for one SKU for one quarter (i.e., 13 weeks). Before presenting the suggested prices to Jack, Michael decided to perform a few tests.

In addition to recommending promotion planning, the model could also be used to analyze the sensitivity of optimal promotion strategies to the uncertain data.

What-if-scenarios

Michael started his analysis by carefully examining the prices suggested by his technique, but quickly realized that he was more interested in the resulting profits than in the prices themselves. He therefore decided to retroactively apply the tool to the last 13 weeks. This way, he could observe how much the profit would have changed by using the promotions suggested by his tool relative to the ones decided by Jack and his store managers. Michael was very excited to observe that using the promotions suggested by his tool yielded a profit improvement of about 4%. At first, this number seemed small, but then Michael realized that this translates to millions of dollars, when looking at the chain level. This tool was especially relevant in an industry in which margins are often very small.

Next, Michael wanted to test the robustness of his solution. In particular, he knew that demand estimation was not an exact science and that his estimated demand model could never be completely accurate. Hence, the question was whether the suggested prices by his model would still perform reasonably well if the predicted demand did not materialize. Would this affect the profit improvement? To answer these questions, Michael designed and ran a series of tests.

Finally, Michael decided to use the model and tool to study the effect of the various business rules. For example, he was able to understand what is the impact of allowing an additional promotion as well as the effect of relaxing some of the business rules.

Conclusion

Optimizing promotions for supermarkets is a challenging task. One needs to decide which products to promote, when to schedule the promotions and how to discount these products. Even if we only consider a small category of items, the store manager still needs to make thousands of decisions. The use of data analytics and optimization can help managers inform this type of decisions. By appropriately collecting and aggregating the relevant data, estimating demand models, and formulating an optimization problem guided by relevant business rules, one can develop a decision support tool to help managers decide their promotions. In practice, this tool can allow for significant profit improvements.

References

Blattberg, R. C., S. A. Neslin. 1990. *Sales promotion: Concepts, Methods, and Strategies*. Prentice Hall.

Cohen, Maxime C, Jeremy J Kalas, Georgia Perakis. 2020. Promotion optimization for multiple items in supermarkets. Forthcoming.

Cohen, Maxime C, Ngai-Hang Zachary Leung, Kiran Panchamgam, Georgia Perakis, Anthony Smith. 2017. The impact of linear optimization on promotion planning. *Operations Research* **65**(2) 446–468.

Nielsen. 2015. The path to efficient trade promotions.

Talluri, K. T., G. J. van Ryzin. 2005. *The Theory and Practice of Revenue Management*. Springer.

Store	Open since	Size $[m^2]$	Number of SKUs	Number of employees
1	1993	168	9542	4
2	1998	182	11356	5
3	2000	210	15017	5
4	2001	144	7538	3
5	2001	197	10508	4
6	2002	167	9127	4
7	2004	327	18584	5

Exhibit 1: Summary of the seven stores (1 $[m^2]$ is equal to 10.7639 $[ft^2]$)

Regular price	4.99
Promotional Prices	4.79, 4.49, 4.19
Size	1 Pint
Flavor	Chocolate
Calories per serving	143
Brand	XX
Average weekly sales	239
Gluten Free	YES

Exhibit 2: Data Summary about a particular SKU – 1 pint of chocolate ice cream, in a specific store (the brand name is not reported)

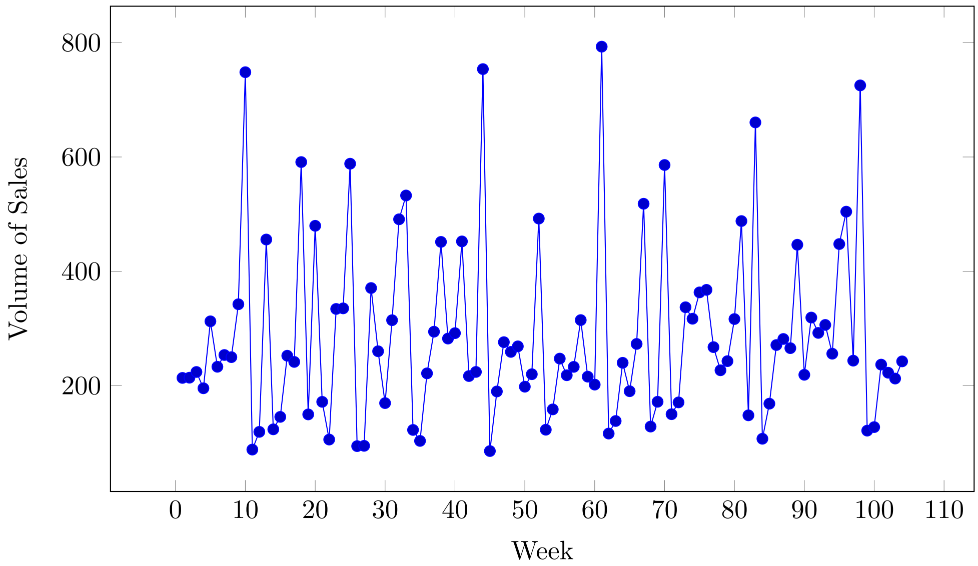

Exhibit 3: Sales of one SKU in a specific store throughout the years 2011-2012 (week 1 corresponds to the first week of January 2011)

Category	Number of SKUs	Number of brands	Average weekly sales	Average # of weekly promotions
Tea	76	19	245	9
Ice cream	29	11	396	6
Cereal	25	14	893	4
Candy	42	27	542	4

Exhibit 4: Summary of four categories in a specific store. (For example, in the tea category, there are 76 different SKUs, from 19 brands. The average volume of weekly sales is 245 units, and there is an average of 9 SKUs on promotion per week.)

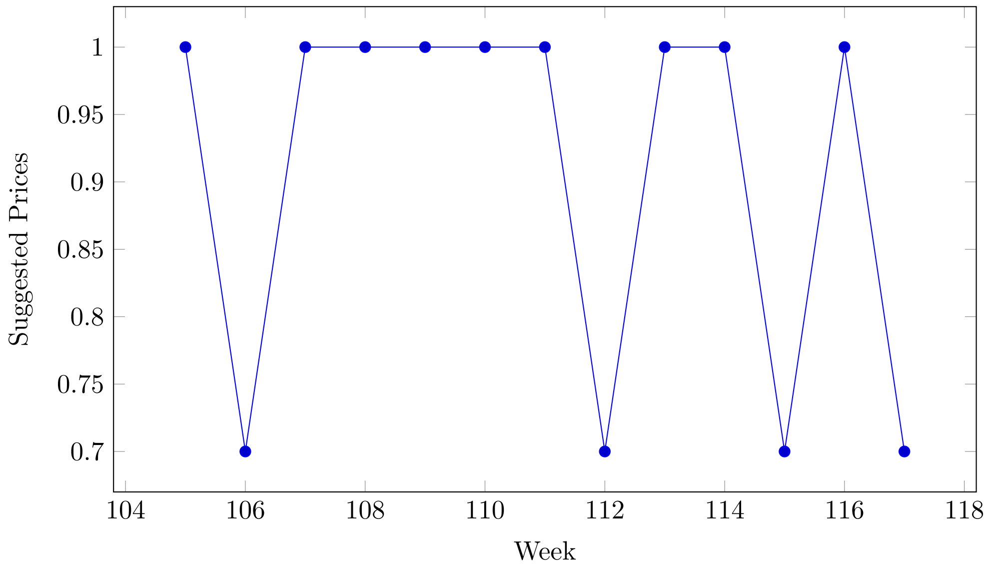

Exhibit 5: Suggested prices for one SKU for the coming quarter