

A Data-Driven Assortment and Inventory Planning Model

Zhiqian Li

Rutgers Preparatory School, USA

lizhiqian542@gmail.com

Abstract. In this research article, we investigate a challenge pertaining to shelf-stocking allocation, specifically the scenario where the distribution of demand is unknown. However, a collection of predicted demand figures is at hand. The retailer is required to decide on the products to place on each shelf (ensuring placement) and determine the quantities of those products (ensuring stocking) while abiding by allocation and space limitations. The primary objective is to achieve the highest possible total expected profit. We conduct three experiments and observe how the products react to changes in marginal profit.

Keywords: shelf-stocking allocation; linear programming; data analysis; unknown demand distribution; economic model.

1. Introduction

Shelves play a crucial role in retail environments as they serve to showcase products to potential customers. They enable the creation of methodical and captivating displays of merchandise, which has the potential to shape purchasing decisions. Therefore, it is crucial for the retailer to determine the appropriate product placement and quantity.

The objective of this study is to tackle the problem of shelf-stock planning. In this problem, we assume that the demand is unknown and concentrate on a single planning cycle. Within this cycle, the manager must determine which products from a given set should be placed on the shelves and the quantity of each selected product to maximize the total expected profit. Meanwhile, three constraints are applied on the objective function.

The first constraint is the Suitability of commodity-shelf specification, indicating that a particular type of good can only be placed on a shelf that meets the seller's requirements.



Figure 1. Entrance of the food section at Walmart

The entrance of the food section in a Walmart supermarket in China, taken about a week prior to the Dragon Boat Festival, is captured in Figure 1. The clear depiction of every shelf being stocked with zongzi provides supporting evidence for the suitability of the commodity-shelf specification theory.

A logical conclusion can be drawn that the optimal strategy is to place zongzi on the most prominent shelves during this specific timeframe, given the substantial surge in demand for zongzi during this period.

The second constraint is continuity of commodity placement, implying the consistency of a single type of good on a designated shelf.



Figure 2. Walmart Supermarket Checkout Counter Area

Figure 2 indicates that various products are arranged on every shelf. However, a clear pattern emerges where the items on the same shelf are consistently grouped together, indicating the implementation of the Continuity of commodity placement. This practice guarantees the grouping of related products, facilitating smoother navigation and enhancing the overall customer experience within the store.

The third constraint is Shelf capacity, which refers to the maximum amount of space available on a shelf for storing products. This is in line with the most basic reality, so it doesn't require much explanation.

We also examine the applicability of the model to other commercial platforms in the discussion section. Because of the essay's objective function's flexibility, it has potential applications in diverse domains, including retail and e-commerce. Effective management of shelf stock is crucial in the retail industry. To meet customer demand and maximize sales, it is essential for retailers such as grocery stores, department stores, and convenience stores to guarantee that their shelves are adequately stocked. Retailers can improve customer satisfaction and loyalty, while minimizing lost sales opportunities, by optimizing their shelf stock levels. Despite operating in a distinct context from physical retail, e-commerce still places importance on ensuring accurate inventory management and product availability. E-commerce platforms must precisely monitor inventory levels, avoid stockouts, and enhance the website's layout to boost retailers' sales and facilitate visitors' buying process.

2. Mathematical Model

In the research, the whole process is based on the setting that we consider only one period. We consider a scenario in which a seller has a total of J shelves, with each shelf being denoted by j , $j \in [J]$. The maximum capacity of the shelf j is V_j . There is a total of I types of products, with each product being indexed by i , $i \in [I]$. The decisions involved in shelf-stock planning include location decisions and shelf-stock decisions. The decision of location pertains to the selection of a particular

commodity i and placement on a designated shelf j , which is denoted by a binary variable $x_{ij} \in \{0,1\}$. On the other hand, the decision regarding shelf stock, denoted by $y_{ij} \in \mathbb{R}_+$, represents the quantity of each commodity i located on each shelf j . It is crucial to recognize that the decision of location not only decides the commodities' assortment but also has a substantial impact on the allocated shelf space for each chosen commodity. The decisions regarding joint location and stock have a direct influence on the utilization of shelf space and the overall arrangement of the chosen commodities. We denoted the demand of product as $\tilde{\xi}$, which is also a random variable. Considering the recorded demand data as $\xi := (\xi_i)_{i \in [I]}$ and the inventory choices for commodities as $y := (y_{ij})_{i \in [I], j \in [J]}$. The price of a product i is denoted by p_i , while the cost of producing one unit of a product i is c_i .

Based on the symbols mentioned above, the total revenue of the seller can be denoted as $p_i \min\{\xi_i, \sum_{j \in [J]} y_{ij}\}$. The total cost of seller can be denoted as $c_i \sum_{j \in [J]} y_{ij}$. Thus, we can derive an objective function for one period:

$$G(y, \tilde{\xi}) = \sum_{i \in [I]} p_i \min \left\{ \xi_i, \sum_{j \in [J]} y_{ij} \right\} - c_i \sum_{j \in [J]} y_{ij}$$

Subsequently, we will refine and enhance this objective model which is also the expected profit function using principles and constrains derived from established business practices.

2.1. Constrains

2.1.1. Suitability of commodity-shelf specification:

The shelving units are constructed to accommodate products with various requirements. Specifically, certain products are restricted to specific shelves. For example, some shelves are designated exclusively for products with appropriate weight limits.

$$x_{ij} \leq a_{ij}, \forall i \in [I], j \in [J] \quad (1)$$

In the above inequation, a_{ij} represents a pre-established binary parameter, wherein it assumes a value of 1 when product i is permitted to be positioned on shelf j , and assumes a value of 0 when this placement is not allowed.

Example 1. Some examples on the constraint (1) are provided as follows:

- (1) Goods with short shelf life: This type of product needs to be quickly purchased, otherwise the business will face a big loss. Thus, in general, products with a short shelf life will be placed on the shelf where people can see them.
- (2) Products that appeal to large segments of the population: Toys and cards and similar products, for example, tend to be placed where children can see them. This product has a strong appeal to special groups of people.
- (3) Products in high demand: For example, in autumn, lipstick becomes very popular.

2.1.2. Continuity of commodity placement:

It refers to the practice of consistently placing a particular type of commodity or product on the same shelf.

$$w_i y_{ij} \leq B_j x_{ij}, \forall i \in [I], j \in [J] \quad (2)$$

$$\sum_{j \in [J]} x_{ij} \leq 1, \forall i \in [I] \quad (3)$$

The first inequation indicates that the product's unit volume w_i and quantity y_{ij} must be put on the shelf in a manner that it is less than or equal to the total volume of the shelf B_j . The second inequation specifies that each shelf can only accommodate one type of product.

Example 2. Some examples on the constraint (2)-(3) are provided as follows:

(1) In the bakery section of a store, keeping different types of bread, like white, whole wheat, multigrain consistently on the same bread shelf.

(2) On a Toy store shelves, placing similar types of toys consistently on the same shelf.

2.1.3. Shelf capacity:

The constraint of shelf capacity refers to the limitation on the amount of space that can be used to store products on a shelf. This indicates that a shelf may have a maximum capacity for the number of items it can hold, which is determined by factors such as size, weight, or structural limitations.

$$\sum_{i \in [I]} w_i y_{ij} \leq B_j, \forall j \in [J] \quad (4)$$

The inequation indicates that the product of the unit volume of all goods and their quantity placed on a specific shelf must be less than or equal to the total volume of the shelf.

Taking into account all the aforementioned constraints, we utilize each of them and establish a viable set M for the decision variables of shelf-stock planning (x, y) as follows.

$$M := \{(x, y) \in \{0, 1\}^{I \times J} \times \mathbb{R}_+^{I \times J} \mid (1) - (4)\} \quad (5)$$

In most cases, it is often challenging for retailers to achieve demand distribution, \mathbb{P} . Instead, sellers may only obtain historical realizations for the demand of each commodity in stores. With this historical information, we can create a prediction of demand in multiple scenarios, denoted as

$$\tilde{\xi}^n := (\xi_i^n)_{i \in [I]}, n \in N$$

where each ξ_i^n represents a predicted demand for product i within scenario n , $n \in [N]$. Moreover, we form the associated decision-dependent empirical distribution of demand forecasts as

$$\hat{\mathbb{P}} := \frac{1}{N} \sum_{n \in N} \delta_{\tilde{\xi}^n}, \text{ with } \tilde{\xi}^n := (\xi_i^n)_{i \in [I]}, \forall n \in N \quad (6)$$

where δ_a is the Dirac delta function with mass one at point a .

By taking into account the aforementioned constraints and demand distribution forecasts, we are able to deduce the practical expected profit equation.

$$\max_{(x, y) \in M} \hat{\mathbb{E}}_{\hat{\mathbb{P}}} [G(y, \tilde{\xi})] \quad (7)$$

We can consider this problem as a linear programming problem. In this case, to better solving this problem, we can convert the equation into following from:

$$\max_{x, y, z} \frac{1}{N} \sum_{n \in N} \sum_{i \in [I]} p_i z_i^n - \sum_{j \in [J]} \sum_{i \in [I]} c_i y_{ij} \quad (8)$$

Subject to

$$z_i^n \leq \sum_{j \in [J]} y_{ij}, \forall i \in [I], n \in [N] \quad (9)$$

$$z_i^n \leq \xi_i^n, \forall i \in [I], n \in [N] \quad (10)$$

$$(x, y) \in M \quad (11)$$

We denote z_i^n as auxiliary variable for product i in scenario n . As we want to minimize the smaller of two things, we can introduce an auxiliary variable that is simultaneously less than or equal to these two things. The optimal value of this auxiliary variable (when maximized) will surely be the smaller one among those two things.

3. Solution and Result Analysis

By utilizing Gurobi to solve this linear programming problem, we aim to determine the suitable products for placement on the shelf and their respective quantities. Thus, we have devised three experiments to examine the performance of 20 products from a recognized dataset on these concerns. The variable that primarily changes in each experiment is marginal profit, denoted as

$$\sum_{i \in I} p_i - c_i$$

In the first experiment, we will increase the marginal profit of all products by an equal extent. In the second experiment, we will reduce the marginal profit of all products by an equal amount. This is done to test the products' performance under conditions of high or low marginal profit. For the third experiment, we modify the variance of the marginal profit among the selected data. To achieve the desired conditions for the experiment, we will only modify the price of each product when changing the marginal profit. However, the images displayed by the experimental results do not clearly convey any practical information. Thus, we have devised an alternative approach to facilitate the improvement of experimental outcomes and present more distinctive findings. results. We introduce a magnitude variable, K , that serves to improve the outcomes of the experiment.

$$\sum_{i \in I} p_i - c_i \times K$$

For each of the following line chart, the vertical axis represents y , the horizontal axis represents the magnitude of change, and each line represents a product.

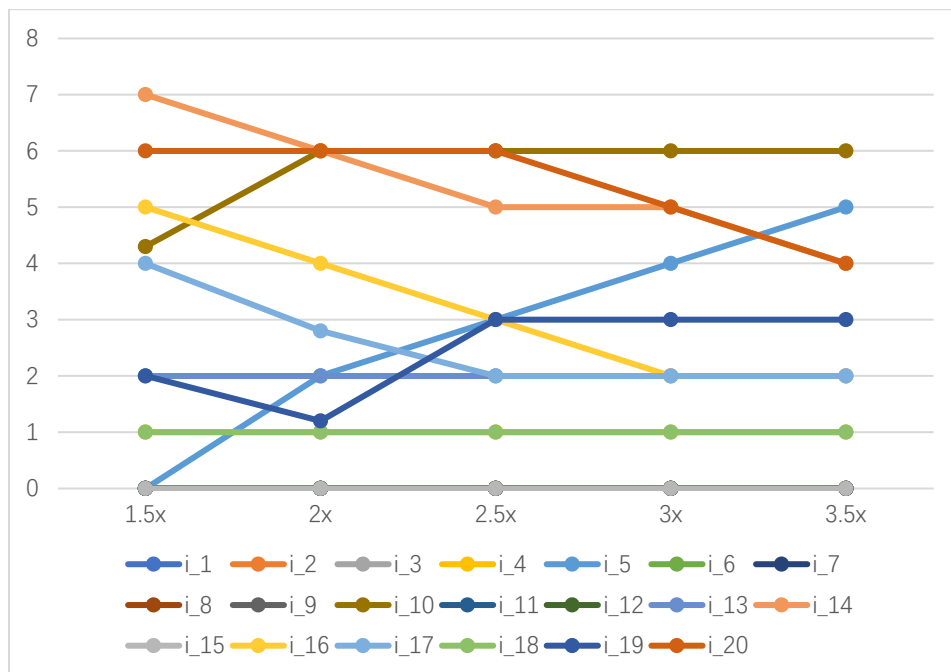


Figure 3. Lower Marginal Profit

From this line chart, we can simply divide 20 products into 3 group. Products in group A, like i_5 , tend to generate higher profits for sellers when there is a relatively low or significant decrease in marginal profit. On the other hand, for i_{14} , i_{20} , i_{19} , and i_{16} that show a primarily decreasing trend can be categorize into group B. The group C contains rest of products that are either unaffected by the decrease in marginal profit or exhibit an unclear trend. Among all the products, group C has the highest percentage, followed by group B, and finally, group A.

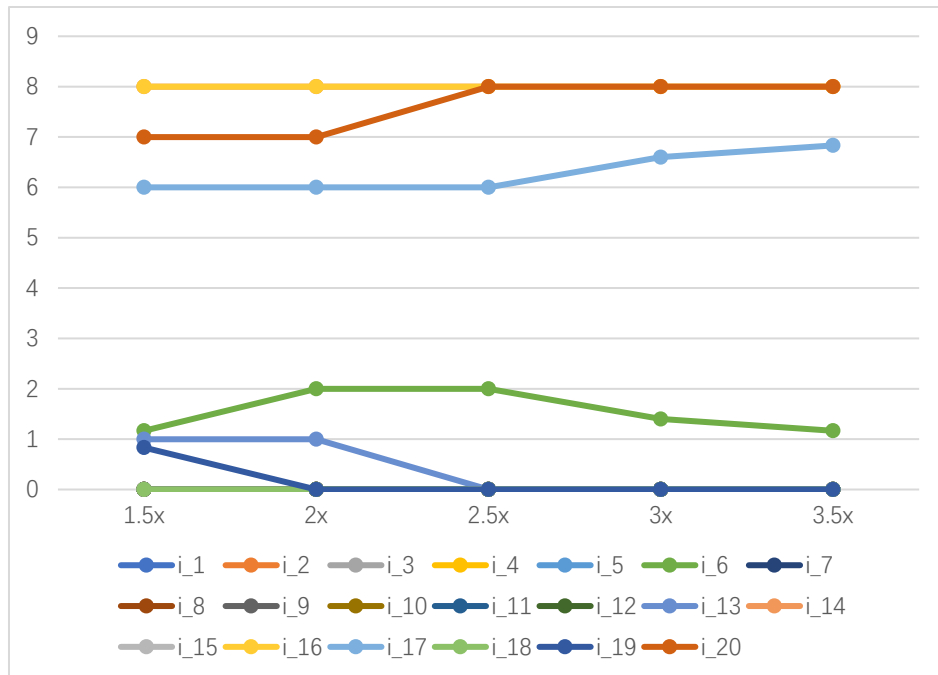


Figure 4. High Marginal Profit

We can roughly categorize these products into three types when the marginal profit increases or the magnitude of the increase becomes larger. Type A products are those with a rising trend in y , examples include i_{20} and i_{17} . When the marginal profit is high, retailers can achieve higher profits with this category of products. Type B products are identified by their ability to maintain an equilibrium point in y , such as i_6 . This product exhibits an upward trend when the magnitude reaches a factor of 2, but it displays a downward trend at a magnitude of 3 times. Hence, the chart illustrates a peak value at 2.5 times the magnitude. Type C products are those that exhibit a declining trend in y , examples being i_{13} and i_{19} .

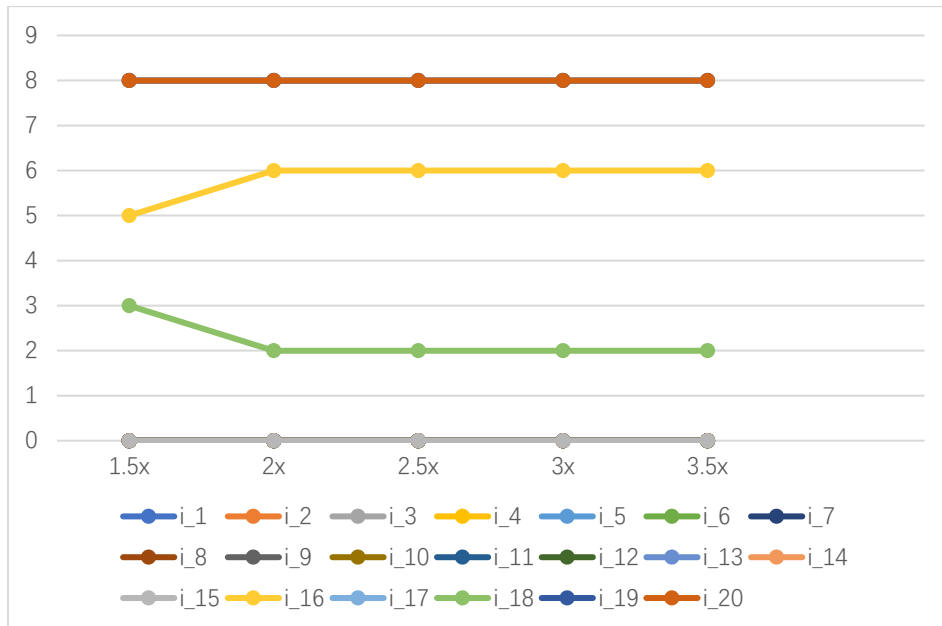


Figure 5. High Variance of Marginal Profit

When we modify the variance and magnitude of the products, the final image fails to clearly demonstrate any substantial differentiation among the products. In this experiment, we therefore deem the impact of variance on the y-values of the 20 products to be negligible.

4. Further Discussions

In the subsequent sections, our primary focus is on discussing the applicability of this model in managing e-commerce user shelves, with a concentration on distinguishing between offline store shelves and online store shelves.

4.1. Position of Golden Shelf.

The gold shelf in offline stores is typically positioned at a height that is within the human eye's line of sight. However, for online stores, since the webpage exists in a two-dimensional space, consumers are only able to view one page at a time. As a page can be scrolled up and down, and the initial state is always at the top of the first page, it can be inferred that shelves located closer to the first page and closer to the top of the page are more valuable.

4.2. Recommendation System.

While analyzing e-commerce platforms, we have noticed that the factors that influence the placement of a product on a particular shelf have increased in complexity. As a result of the implementation of recommendation systems, products have established specific associations, thus forming a network of product relationships. At this juncture, every decision concerning the positioning of a product bears a certain level of influence on other products. Hence, it is not feasible to solve the shelf placement problem on e-commerce platforms directly through a linear function or the expected profit model, as previously mentioned.

4.3. On Continuity of Commodity Placement.

The majority of e-commerce merchants can be categorized into two types of stores: completely specialized stores and category-specific e-commerce stores. The initial variety of store exclusively sells a single product, like mobile phones. However, in this case, there is no need to consider the Continuity of Commodity Placement, as the merchant only sells a single type of product. The second

type of store sells products from a specific category, such as maternity and baby items. After conducting investigations into the products placed on various website store shelves, it has been observed that this type of store selectively utilizes the theory of Continuity of Commodity Placement. However, this is the observation. The majority of stores their products according to sales volume and recommendation systems.

4.4. Storage.

The inventory levels of commodities on e-commerce platforms are directly determined by their storage, rather than being determined on the shelf as in offline stores. This implies that we will no longer take into account the amount of products present on the shelf while assessing the inventory levels of chosen commodities. Instead, we take into account the entire inventory of that product in the entire warehouse. This pertains to warehousing issues, rather than the inventory levels of the chosen commodities on the shelf.

4.5. Pricing Strategy.

Setting the price of commodities is indeed a valuable method to increase revenue, and this aspect is also effectively reflected in the model. Hence, we are able to engage in discussions concerning pricing strategies. The price of the product for the store typically comprises the cost of the product, logistics, advertising, and an expected profit. For the majority of cases, the initial three items are recognized, however, the anticipated profit is primarily dependent on the pricing of significant competitors while taking into account the targeted sales volume of the product. Merchants are willing to price products with high sales volume at a lower price in order to achieve quick turnover and fast inventory turnover, reduce inventory costs, and maintain small profits.

5. Conclusions and Future Work

This paper examines a problem of allocating shelf-stocking items with an unknown demand distribution, which can be seen as a linear regression issue. To solve this problem, we create a model using Gurobi. In this paper, we conduct three experiments and analyze the results to observe how retailers can best react to earn the greatest profit when the marginal profit changes. When the marginal benefit is small, products are primarily categorized into three groups: the increasing group, the decreasing group, and the unaffected group. We classify four types of products when the marginal benefit is high: the increasing type, the decreasing type, the unaffected type, and the type with extreme values. When the variance of marginal profit is large, the majority of products show no effect or a slight change.

There are numerous types of experiments that can be conducted to address this question more comprehensively, and numerous aspects of the model can be refined. For instance, we can introduce additional concepts to modify the model, thereby enabling it to take into account questions that are more closely aligned with reality. In the meantime, we have the option to combine products with diverse attributes to devise a more comprehensive solution.

Acknowledgement

First of all, I am very grateful to Professor Zhen Wang for providing theoretical guidance, topic selection guidance, experimental guidance, and writing guidance for this article without compensation. The choice of topic for this article stemmed from my interest in economics and applied mathematics, as well as guidance from my professor. Having been influenced by family-related business matters over the long term, I developed a keen interest in retail sales models. Consequently, I decided to write a research paper to facilitate a deeper understanding of my interests and our family business. This article was solely authored by me, with guidance provided by my professor. I came to know my professor through attending one of his lectures, and I subsequently extended a special invitation to him to assist me in completing this research paper.

References

- [1] Vault Inc.2019 Vault Consulting 50[EB/OL]. [2019-01-10].<http://www.vault.com/company-rankings/vault-consulting-50>.
- [2] Xie Yi, Peng Siqing. Factors Influencing Consumer Brand Relations: An Exploratory Study
- [3] Research [J]. Business Research, 2008 (1): 1-7.
- [4] Ji Fang. Research on the Integration of Retail Enterprises Based on Online and Offline Omni Channel Retail [J]Journal of Jiujiang University (Social Sciences Edition), 2018, 37 (4): 124-128.
- [5] FONSECA J D, WANG P. A joint analysis of market indexes in credit default swap, volatility and stock markets[J]. Applied Economics, 2015, 48(19): 1-18.
- [6] CHESBROUGH H,Open Services Innovation:Rethinking Your Business to Compete and Grow in a New Era[M].San Francisco:Jossey-Bass,2011:68-88.
- [7] RIGBY D, ZOOK C.Open-market innovation[J]. Harvard Business Review,2002,80(10):80-89,129.