Inventory Management for Automated Convenience Stores in Brazil

By

André Nascimento Costa Bachelor in Industrial Engineering, Universidade Federal de São Carlos, 2012

And

Sanjay Kumar Naithani Bachelor of Technology, Punjab Agricultural University, 1995

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE IN SUPPLY CHAIN MANAGEMENT AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2022

© 2022 André Costa and Sanjay Naithani. All rights reserved. The authors hereby grant to MIT permission to reproduce and to distribute publicly paper and electronic copies of this capstone document in whole or in part in any medium now known or hereafter created.

Signature of Author: ___

André Nascimento Costa Department of Supply Chain Management May 13, 2022

Signature of Author: ___

Sanjay Kumar Naithani Department of Supply Chain Management May 13, 2022

Certified by: __

Dr. Christopher Mejía Argueta Director, MIT Food and Retail Operations Lab Capstone Advisor

Accepted by: ___

Prof. Yossi Sheffi Director, Center for Transportation and Logistics Elisha Gray II Professor of Engineering Systems Professor, Civil and Environmental Engineering By

André Nascimento Costa

and

Sanjay Kumar Naithani

Submitted to the Program in Supply Chain Management on May 13, 2022, in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

Today, small retailers in Latin America account for 70% of the market share. Convenience stores play a crucial role, as people look for more convenience with modern lifestyles. Most of these stores are managed by people without experience or formal education in business management. A challenging problem for small retailers is inventory management. We developed this project for Onii, a Brazilian startup company with over 300 convenience stores in the country, run by small businesspeople in a franchise-like model. Its stores are entirely automated; thus, there are no cashiers or employees inside the store. This project aims to develop inventory management policies for Onii store operators and help them manage their stocks better. We use unsupervised Machine Learning techniques like k-means clustering and principal component analysis to identify patterns and segment stores and items. Then various inventory policies were computed to look for the lowest cost for each combination of clusters of stores and items. The best policy for the Onii store's reality is the Periodic Review model, with different period parameters (R) for each combination. At last, sensitivity analysis was conducted to determine the impacts of each parameter used in the model, such as ordering cost, holding cost, and inventory cost. The result is a robust model that Onii can apply to their current and future stores.

Capstone Advisor: Dr. Christopher Mejía Argueta Title: Research Scientist, Center for Transportation and Logistics

ACKNOWLEDGMENTS

We want to thank our advisor, **Dr. Christopher Mejia**, for his help and contributions in this journey. Always very thoughtful, insightful, and enthusiastic about this project.

We also want to thank **Toby Gooley** for all the coaching sessions and attention to detail in our writing.

We thank **Onii**, our sponsor company, for this project. Especially **Bruno Fabbri** and **Conrado Rantin** that were always there when needed.

Finally, we thank our family and classmates for always making it easier when times are hard.

Contents

List of Figures

1 Introduction

1.1 Motivation

Small retailers, organized like convenience stores and non-organized like nano stores, are a fast-growing market, especially in emerging countries. According to SEBRAE (2014), people go to small retailers looking for selective assortment, proximity, and practicality in doing their shopping. Furthermore, modern lifestyles push people to spend less time shopping for groceries and consumer packaged goods, and sometimes the best alternative is to buy from the small store nearby. During the COVID-19 crisis, people got more concerned about going to supermarkets and being more exposed to the virus because of the high foot traffic.

Only the market share of traditional retail in Latin America and Asia accounts for 50% (Wan et al. 2018). Fransoo et al. (2017) estimated that 50 million nanostores in emerging markets regularly serve around five billion consumers. While Brazil has approximately 21 supermarkets for every 100,000 inhabitants, the rate is 281 for minimarkets and convenience stores. These small stores account for 56.2% of the sales in the segment (SEBRAE, 2014). According to SEBRAE (2015), they make up 6% of the country's GDP and impact employment and revenue generations. Their research also shows that 52.7% of these small stores have an average sale per customer of R\$ 60 (US\$ 12) or less, reinforcing the idea of a convenience purchase.

These facts show an increasing trend for end consumers wanting immediate access to goods, especially in areas close to where they reside, live, and/or work.

Among many challenges like capital availability, lack of systems, processes, and professionals, inventory management is one key challenge for these small retailers. They do not possess much space to stock goods and lack knowledge and experience in using inventory management tools. Thus, they can be susceptible to any demand changes and the lack of flexibility to react to changes. Most of their stock is visible on the shelves, available for sales, and there is no backroom to store a larger number of items. SEBRAE (2014) shows that many small retailers need to buy frequently in small quantities to assure proper supply.

Vending machines also play an essential role in this growing convenience market. They can provide products 24 hours per day, seven days a week in a self-service solution. However, the small space of convenience stores challenges their inventory management. Lin et al. (2011) showed the importance of understanding the demand profile for each vending machine to optimize their portfolio and avoid wasting space with something that local consumers do not buy.

1.2 Problem Statement

We work with Onii, a Brazilian startup that developed a business model based on licensees for automated, walk-in convenience stores. Many stores are located in residential condominiums or office buildings. Onii provides the system, brand, and relationships with CPG manufacturers. Then store owners invest in their stores' physical infrastructure and operations.

With almost 300 stores located in Brazil's main cities, a fast-growing plan, and a diverse portfolio, inventory management is an increasing challenge for Onii and its store operators. Although the operators can use Onii's system and app to track their inventory, there is no standard for managing inventory policies. Each licensee operates a different business size (e.g., vending machines, container-size stores), but all face a significant limitation on space. Decisions on what, when, and how much to buy, rely on each store's

operator. Consequently, each operator does what he considers better, and no information or knowledge is shared among them.

Our main objective is to develop inventory policies for various store and product combinations based on the problem above. The project aims to set standardized inventory policies to help store operators reduce costs and make the best decisions regarding their stores. Stores were mapped and classified to define the optimal inventory policies for each store category. This project's final result helps store operators manage their inventory with standards that will bring higher efficiency and, ultimately, profit.

2 Literature Review

The literature body has highlighted the difficulty of small stores related to space and processes to manage inventory levels. Inventory models should reflect the business conditions and considerations like demand variability, budgetary constraints, customer service levels, store space limitations, etc., which is hard to model, particularly for developing countries. This section describes inventory management models and other data analysis techniques used to generate insights and guide policy choices.

2.1 Inventory Management

Increased competition and customer preference have increased product variety and shorter lead time, making more dynamic and challenging inventory management decisions. Muller (2019) stated that uncertainty, fluctuations in demand, the unreliability of supply, price protection, quantity discounts, and lower order costs force companies to hold inventory. Predictability means you need raw materials and semi-finished goods to plan your production. Fluctuations in demand mean you can have stock to protect from demand and supply variability. Stocking for price protection prevents impacts from high price fluctuations. Discount on price is another driver of stocking high inventory as it reduces acquisition and ordering costs and hence, the total inventory cost.

Inventory management is also critical from a cash flow and returns on asset perspective. A higher inventory level increases the risk of obsolescence, theft, and improved working capital requirements. Hence, inventory management models and tools balance the critical tradeoff under various demand and volume discount conditions to avoid lost sales, poor service levels, and high costs.

Ballou (2006) understood three classes of relevant costs related to stocks: purchasing, holding, and shortage costs. Ballou presented the economic order quantity (EOQ) by minimizing the total cost of the aforementioned relevant costs. However, the EOQ model does not work well when demand is not uniform and has limitations in applications.

Many models are different from EOQ and address diverse classes of variables and scenarios (Muckstadt and Sapra 2010). Generally, the tradeoff to model the most suitable inventory model includes deciding deterministic vs. stochastic demand, periodic vs. continuous review, minimum cost vs. service level approach, and backlog versus lost sales approach (Winston, 2004, p.846). Some of the most common inventory models follow:

- 1) **EOQ** is used when the demand is uniform and deterministic. It determines the quantity that yields the least total cost, i.e., the sum of holding and ordering costs. This model is suitable when demand is uniform.
- 2) **Single-Period / Newsvendor** is used when the demand is probabilistic and shows a finite horizon period. This is relevant when you need to buy inventory ahead of some event that will last for a limited period. Still, after some time, the stock is useless.
- 3) **Base Stock Policy** works under probabilistic demand and infinite time. In this model, an order is placed when there is demand for the products. It is typically used for slow-moving items and items with high stocking costs.
- 4) **Continuous Review Policy** depends on Probabilistic demand with an infinite period. Inventories are reviewed continuously, and orders are placed.
- 5) **Periodic Review Policy** also works under probabilistic demand with an infinite period. This is a time-based policy, and a certain number of units is ordered every period to reach a maximum inventory level.

Jackson et al. (2020) classified inventory models depending on their techniques: analytical approaches, optimal control theories (using differential equations to analyze systems behavior in time), dynamic programming (such as periodic and continuous review models), simulation-based optimization (using computer-simulated models that reproduce real-world dynamics), and metamodel-based optimization (simulating a simpler version of the real problem for computational efficiency). All these models can

be used in different situations according to the variables the inventory manager wants to consider.

Inventory Management for Retail

Inventory management in retail is challenging due to many SKUs, space constraints, demand variability and seasonality, problems in estimating actual demand due to lost sales, and substitution impact. Many studies use backorders rather than lost sales, as lost sales models are complex to analyze, but they do not consider that customers are unwilling to wait if their orders are not met (Campo, Gijbrechts & Nisol, 2000).

Ehrenthal and Stölzle (2013) found that the causes of stockouts in retail are specific to each brand, store, and item. However, improving store operations and coordinating store delivery and shelf replenishment can effectively minimize stockouts (Ehrenthal and Stölzle, 2013).

A way to improve operations and performance in the retail industry implies using various formats (i.e., traditional retail stores, vending machines, automated stores) depending on the customer features. Vending machines have long been popular to help franchisees increase their reach and ability to operate at a low cost. Nevertheless, managing inventory in such a format is also very complex, considering the limited space and the variety it can handle.

To define the best inventory model, location, physical space, consumer demand, and suppliers must be analyzed carefully. For example, Ehrenthal et al. (2014) discussed the impact of demand seasonality and variation in a big European supermarket company. Authors found that holding and handling costs can be reduced by considering nonstationary demand and demand variation across the days, especially between weekdays and weekends.

Consumer trends are pressing grocery stores to provide higher quality, availability, innovation, and environmental performance than in the past. Vallandingham et al. (2018) showed that future grocery retailers would use real-time information, and gain agility, efficiency, and transparency in their operations. Roggeveen and Sethuraman (2020) addressed technology trends, showing how each technology can be applied to the prepurchase phase (for management and search engagement), to the purchase phase (for transactions and acquisitions from their customers), and the post-purchase phase (for customer relationship management). Next, we will discuss the main challenges for small retailers.

Inventory Management in Small Spaces

Many theoretical studies do not consider space constraints when discussing inventory and stock management. Guo et al. (2016) suggested a classification according to each item turnover to decide how goods should share the same storage zone. Using this classification and shared storage zones lowered the required storage space (RSS) in the warehouse, lowering space needs.

Zhan and Rajaram (2017) suggested two strategies when managing limited retail shelf storage space. Their first approach is space dedication, which brings flexibility to replenish products independently. In contrast, the second approach is shared space, which brings the potential for space savings but incurs into additional costs. Depending on the type of the product, a high level of on-shelf inventory may have a demand-increasing effect (i.e., the "billboard effect") or a demand-decreasing effect (i.e., the "scarcity effect").

2.2 Small Retailers and Inventory Management

The small retail market is growing fast all over the world. "Consumer's preference for small stores is positively motivated by functional benefits and familiarity" (Paswan et al., 2010). Sinha and Banerjee (2004) studied consumers' behavior when choosing a store in evolving markets, especially in India. They found that location and convenience is the core driver for customers when choosing a grocery store. Berry (2001) stated that new retailing markets should "Solve consumers' problems" and "Save consumers' time."

Small retail shops lack sophisticated inventory management systems due to unskilled personnel and poor management attitude. Chikan and Whybark (1990) argued that these small retailers are slow to adopt contemporary inventory management practices. Generally, these small retailers use rules of thumb for managing inventories (Kamilah Ahmad et al. 2016). Research in the space of small retailers on inventory management practices is still scarce.

On the other hand, vending machines also offer convenience and practicality while bringing better-operating margins to owners. Between 2017 and 2023, the global vending machine market size will grow by 15.8% CAGR (JOENG, 2018). These machines can add value with technology advancements by introducing cashless payment systems and using ID numbers to check customers' buying history.

Small Retailers in Brazil

Small and convenience retail stores are also growing in Brazil, following the global trend. With the pandemic, small neighborhood markets rose 21.2% in the first quarter of 2021 (CNN Brasil, 2021). Parente (2008) stated that big supermarket chains for lowincome customers lack the sensibility to understand and fit these customers' expectations and needs. They excessively focus on low prices, abnegating other features, such as convenience. This evidences that even underserved communities are

14

not only searching for affordability but also better quality, rapid service, and

convenience.

3 Methodology

This section presents the steps and techniques we used to solve this project. We split the methodology into six actions (see Figure 1) to learn the current practices from Onii, collect data and propose inventory models.

Figure 1

Methodology

We started by mapping the current process and stakeholders to understand the various players' relationships. We then collected data and used Alteryx to clean the data and remove the outliers. Then we investigated the demand, trends, growth, and top running parts and stores. Finally, we used machine learning and inventory models to determine clusters for stores and items. With those clusters, we designed the best inventory policies. We also performed the sensitivity analysis to see how different variable impacts the policies.

3.1 Process and Stakeholder Mapping

Onii is a startup company running automated stores in Brazil. The company is growing and already has over 300 stores in operation. Onii provides licenses, store layouts, and the product mix to the store operators. These operators set prices and terms with the consumer-package goods (CPG) manufacturers; however, Onii gives store owners access to better negotiation conditions.

The core stakeholders in this business model are Onii, store operators, suppliers, and consumers (see Figure 2). Onii consumers can download the mobile app for all transactions. They can enter stores, shop, and pay through mobile applications with the app. Most stores are open 24 hours with no employees operating them.

Figure 2

Onii's Business Model

Onii stores are located in corporate parks, gated residential apartments, and house

condominiums. They are operated in three formats (Figure 3):

- 1) Onii Market works 24/7 and is used for horizontal residential condominiums (houses). Onii houses its products in containers or rooms.
- 2) Onii Box is a flexible solution. Products are placed in a secure location, accessible only with the app or a card.
- 3) Onii Station is like a vending machine because it does not require access control.

Onii Market, Station, and Box

All the sales in each Onii store go either through the app or through a point-of-sale (POS) payment terminal. Either way, these sales get registered in Onii's database. Onii has access to sales transactions for all its stores, categories, and subcategories. We received sales data for each store from January 2019 to December 2021. Also, we conducted interviews with Onii operators and Onii C level executives to map the current processes, relationships among stakeholders, and business performance metrics. The following section will discuss details about data collection.

3.2 Data Collection

Quantitative Data

These datasets include historical data of each SKU sold per store monthly and daily. We extracted this information from various transaction-level data available at each store. However, we could not access lost sales or transactions between stores and suppliers, given that Onii's system does not capture this information.

This information is critical for identifying demand patterns, classifying stores based on consumer behaviors, and formulating an inventory policy based on the analysis.

Qualitative Data

We interviewed operators from a few representative stores to understand the current practices, value streams, constraints, and opportunities. We evaluated the following dimensions through the semi-structured interviews per store owner:

- 1. The current value stream from Purchase order release to suppliers to consumer sales.
- 2. Connections and relationships established among Onii, operators, and suppliers.
- 3. How do operators use the current inventory policy and stock replenishment decision processes?
- 4. How do operators monitor performance metrics per store type, location, category, etc.?
- 5. What constraints from external sources like suppliers, consumers, and Onii do operators face?
- 6. How do potential payment methods and discounts impact a store's performance?
- 7. The process to capture and compute lost sales of top-selling items and customer behavior.

3.3 Data Processing and Cleaning

Each record in Onii's database consists of sales quantity, price, SKU identification number, subcategories, categories, store type, address, date, and time. Although the history available starts in 2019, most stores have less than one year of existence. By December 2021, only 47 stores had more than one year of sales history (out of almost 300 stores total). Then, even though the database is vast, we used a small part to analyze sales features such as seasonality, levels, and trends.

We also disregarded information (e.g., payment processing data) that was unimportant for this scope. After removing those fields and records, we downsized the total storage space from 4 Gb to 0.7 Gb of data. The latter improved computational processing efficiency, speeding up our analyses.

Data preparation was needed for older sales records. Most of them had only the SKU and Store ID but no other information. We needed to merge these records with store master data tables per SKU to understand each transaction and create a complete transaction database. We used Alteryx workflows to clean, sort, and merge the data available on different databases from Onii. Figure 4 shows a screenshot of an example of the Alteryx flows used.

Figure 4

Alteryx Workflow

3.4 Data Analysis

Descriptive and Exploratory Analysis

We explored the data from 76 stores with 12 months of sales history. We analyzed

data to:

- 1. Analyze a 12-months sales history for all stores at the level of SKU, subcategory, and categories.
- 2. Conduct ABC and XYZ analysis at the subcategory level for the entire sales history.
- 3. Compute the coefficient of variation for each subcategory.
- 4. Identify the sales patterns for each subcategory per store type.

This procedure helped us understand each subcategory's sales volatility, level, and

trends. We deep-dived on 26 subcategory classes that contribute to 80% of sales.

Clustering and Principal Component Analysis

We used Principal component analysis (PCA) and k-means clustering to reduce complexity and gain insights into the data. PCA analysis has provided us with the five principal components that define most variation in-store sales. Then we identify some patterns for combinations of stores and item (i.e., product) clusters. We determined fivestore clusters. Also, we defined 10 clusters at the subcategory level based on similar features. See Appendix A for further details on the statistical framework for using these analyses.

3.5 Inventory Management Policies

We have used three policies for each item to see where we get the total least cost. The total least cost consists of the cost of carrying inventory and ordering. We have used the following policies and compared them. For this analysis, we choose four main inventory policies from the literature: Economic Order Quantity (EOQ), Base Stock, Continuous Review, and Periodic Review. These models were chosen for two main reasons: they are very well known, used, and tested in the literature; they are simple enough to fit Onii's business model and store operators' capabilities.

1) **Economic Order Quantity (Q*):** This Policy is used for relatively stable demand with order quantity balancing holding and ordering costs. Economic order quantity (EOQ) is a tradeoff between fixed (ordering) and variable (holding) costs. The inventory replenishment policy becomes "order Q^* every T^* period." This policy is suitable for constant demand. (Ballou, 2006)

 $Q^* = \sqrt{2Dct/ce}$

Notation: D: Demand (units/time) c_t: Order cost (\$/order) c_e : Excess holding cost (\$/unit/time), $c_e = c^*h$ c: Purchase cost (\$/unit) h: holding cost percentage (%/time) T*: Optimal order cycle time, time between replenishments, $T^* = Q^* / D$ 2) **Base Stock Policy:** This policy implies ordering what you sell every day. The policy determines the base level stock for each item. The base stock, S*, is the sum of expected demand over lead time plus the standard deviation over the lead time multiplied by the safety factor k. The k-value depends on the level of safety and is the critical ratio (CR). (Ballou, 2006)

Optimal Base Stock, S^* : $S^* = \mu_{DL} + k_{LOS} \sigma_{DL}$

$$
LOS = \frac{cs}{cs + ce}
$$

Notation:

S*: Base stock μ_{DL} : Demand over lead time k_{LOS} : factor for the safety level of inventory σ_{DL} : Standard deviation of demand over lead time c_e : Excess holding cost (\$/unit/time), $c_e = c^*h$ c: Purchase cost (\$/unit) h: holding cost percentage (%/time) cs: Shortage cost (\$/unit)

3) **Continuous review policy (s, Q):** The order-point, order-quantity implies to order Q* units when inventory position (IP) is less than the reorder point s. The reorder point is the sum of expected demand over lead time plus the standard deviation of demand over lead time multiplied by some factor of safety k. (Ballou, 2006)

Re-order point: $s = \mu_{DL} + k\sigma_{DL}$ Order quantity (Q): $Q=Q^*$, the EOQ

Notation: s: Reorder point μ_{DL} : Demand over lead time σ_{DL} : Standard deviation of demand over lead time

4) **Periodic review policy (R, S):** This policy is known as the "order up to" policy. The policy is to order up to S^* units every R period. Order quantity in this policy will be S^* - IP, which is provided by the following equation. (Ballou, 2006)

Order up to point $S = \mu_{DL+R} + k\sigma_{DL+R}$

Notations S: Order up to point μ_{DL+R} : Expected demand over lead time plus review period σ_{DL+R} : Standard deviation of demand over lead time plus review period This policy is quite popular as it fits well with the business logic of ordering periodically, like once per week or every two weeks, etc.

To compare the performance of inventory policies, we computed the total least cost and analyzed the Cycle Service Level (CSL). The latter is the performance metric that calculates the probability of stockouts in the replenishment cycle. Then, we found the value of k (safety factor) by using the level of CSL.

$$
CSL=1-P[stock out]=1-P[X>s]=P[X<=S]
$$

Thus, we evaluated each of these policies for every subcategory to determine the best approach. We also tested different scenarios based on combinations of parameters such as periods. We will describe this in the following subsection.

3.6 Sensitivity Analysis

Given that we did not receive detailed information about the cost structure from store operators, we performed some sensitivity analyses between holding and ordering costs. Therefore, we computed inventory policies with different review periods ('R') to see which R-value would provide us the total least cost compared to Continuous and EOQ policies. We performed a sensitivity analysis with ordering costs varying from \$0.1 to \$1, holding costs ranging from 5% to 14%, and margins on subcategories from 25% to 75% to investigate the impact on inventory policies and their performance. We used these lower and upper bounds for the sensitivity analyses based on opportunity costs, macroeconomic data in Brazil (e.g., inflation), and other information provided by Onii. We also used the ordering cost ratio for continuous to periodic review to understand when one policy is better than the other.

4 Results

4.1 Exploratory Analysis

After analyzing primary data from Onii, we performed exploratory analyses of the sales data to understand the sales behavior and how it differs per store type, location, and item subcategories. Also, we built an ABC analysis, and we computed the coefficient of variation to establish a comparison between volume distribution and demand volatility. Figure 5 indicates that alcoholic beverages have a significant share in vertical and horizontal condominiums compared to companies, where non-alcoholic beverages significantly contribute. We can also see that 26 subcategories ('A' Class) contribute to 80% of total sales out of 264 subcategories (Figure 6).

Figure 5

Sales by Location

Pareto Analysis

Then, we computed the Coefficient of Variation (CV) and revenue to investigate differences across subcategories. We have 99 subcategories whose CV is less than 0.3. They have less volatility than the other 165 subcategories. They contribute 90% to the overall sales because of all our "A" Class subcategories (Figure 7). Most of the high CV subcategories are in the "C" class, where sales are dispersed over time.

Revenue vs CV

We also looked at the revenue of 2021 for each store type and found that market stores account for the highest revenue, followed by box and station (see Figure 8). Also, we analyzed the top 10 subcategories and found that beer and soft drinks are the most sold two subcategories (Figure 9)

Revenue per Store Type

Figure 9

Top Subcategories

We used machine learning techniques for clustering stores and items. With those clustering, we identified stores and items that share features and follow similar demand patterns. Then, we could develop inventory policies for each combination of store cluster and item cluster.

Store Clustering

Given the large number of Onii stores, we could not model them separately. Thus, we reduce complexity by grouping them in clusters with similar features to create standard inventory policies that work for similar stores. First, we identified the demand profile for each store by calculating the percentual revenue contribution for each subcategory of items. Since some stores show higher revenues than others, using the total revenue per category would not have worked. Stores with higher income tend to earn higher revenue in all subcategories. By calculating the percentage of the total revenue that each subcategory represents, we can standardize that information among all stores. Thus, the machine learning techniques can identify stores with similar demand profiles, independently of the total revenue.

We first used k-means clustering to identify stores with similar demand profiles with that normalized data. We used each subcategory percentage representation as a feature for the model. We could build five store clusters. After the k-means clustering, we deepdived into each cluster to explore store characteristics. A summary of this analysis and the relation with each cluster in Figure 10 follows:

- 1. Box stores sell primarily frozen food. $(C1 blue)$
- 2. Stores located in companies do not sell alcoholic beverages. $(C2 red)$
- 3. Stores located in horizontal condos sell mainly everyday groceries. (C3 green)
- 4. All stores sell drinks and snacks primarily. (C4 orange)
- 5. "Box" and "Station" stores sell mainly drinks, replacing the role of a typical vending machine. (C5 – yellow)

Since we modeled many features (over 30), we also used Principal Component Analysis to reduce the complexity and simplify cluster visualization. Figure 10 shows each cluster of stores considering PC1 and PC2 in the axis. Together, PC1 and PC2 account for 40% of all the variance in the data. One of the main drivers for PC1 is the impact of alcoholic beverages on store revenues. The higher the value on PC1, the lower that store sells alcoholic beverages. As expected, we can see that C2 stores have high values of PC1. For PC2, one of the main drivers is the impact of frozen and refrigerated products on store revenues. The higher the value of PC2, the higher that store is selling frozen food. We can see that C1 has a high PC2 value since it sells primarily frozen food.

Figure 10

Store Clustering

Similarly to store clustering, we used k-means to find item clusters. We used the revenue, quantity, coefficient of variation (CV), and necessity of refrigeration (i.e., a dummy variable with answers 1 for yes, 0 otherwise) as features for the clustering. Revenue and quantity are highly correlated. Therefore, we considered the revenue, the CV, and the necessity of refrigeration to cluster the items. Figure 11 plots revenue vs. CV. In the case of Cluster 6, it accounts only for beer sales. It generates a separate cluster since it has very high revenue and low CV.

We can see the same pattern for other A items (C3, C6, C9, C10 – circles) with high revenue and relatively low CV. The clustering indicates that items with a high revenue have more stable demand (low CV) than those with low revenue (High CV).

Figure 11

Item Clustering (Revenue vs. CV)

Figure 12 displays revenue vs. refrigeration. This chart shows how the k-means clustering separated items that need refrigeration from those that do not. Apart from Cluster 3, which contains products that may or may not require refrigeration, all the other clusters are dichotomic, showing a value of 1 or 0 for this feature.

Figure 12

Item Clustering (Revenue vs. Refrigeration)

4.3 Inventory Policies Design

We modeled and simulated different inventory policies with the stores and items clusters. As mentioned before, the main idea of this project is to develop inventory policies such that the store operators can use them as a guide or handbook on how to manage their inventory levels.

Model Parameters

First, we understood the values for each inventory policy's most important parameters we had to use. Some of the parameters were taken from the data provided, others were estimates taken from interviews with operators and Onii staff, and others were used as assumptions. We describe a quick list and explanation for each of those parameters below. The sensitivity analysis for some parameters was conducted after the inventory policies modeling. It will be explained later.

Holding Cost

We used a well-known and popular benchmark indicator in Brazil called CDI (Certificado de Depósito Interbancário) for holding cost. This figure is a baseline interest rate used by Brazilian banks. It is always related to the introductory interest rate set by the country's Central Bank. It works as a cost of opportunity for investments in Brazil. It is used as a holding cost since the money invested in inventory for each store could be invested in something else, generating the CDI rate as interest.

Ordering Cost

Given that there is no actual ordering cost from suppliers, we computed how many labor hours would take to place one order and considered that as ordering cost. By talking to Onii staff and operators, we assumed that each order takes 20 minutes to be placed, which gives a cost of US\$0.47 per order.

But using that cost for the calculation of all items is not realistic. For example, in a periodic review policy, item necessities are bundled together. By putting subcategories together, instead of lots of small orders, we are ordering many different SKUs from a particular supplier in the same order. Since the ordering cost is calculated by item, we

cannot assume that the operator will pay US\$0.47 for ordering each item (if we are ordering multiple items in the same order).

Therefore, we assumed a lower order cost for periodic review policies than continuous ordering costs. For continuous-review models, we order each item when it is needed. We adopted a ratio of 17% between the periodic review and continuous-review ordering costs. On average, a store has 20 suppliers and 119 subcategories. The latter means that the operator would create 119 different orders for continuous-review models.

In contrast, for periodic-review models, the operator would create only 20 – one for each supplier at every review period. Dividing 20 by 119, we get to the 17%. The impacts of these ordering cost assumptions are discussed in a later section.

Gross Margin

The stores' purchase data are not centralized by Onii, given that each store operator manages this purchase data differently. Therefore, we could not gather the purchasing costs of the items for each store. Consequently, we had to assume a general gross margin for stores to calculate the cost of each item based on the selling price we had on hand. We did it using the simple formula *cost = price – margin*. Our interviews with Onii staff and store operators taught us that the gross margin could go up to 50% across stores and items.

Lead Time

We discussed the lead time with the Onii team, and we agreed to assume an average lead time value of half a week (3.5 days). This lead time is reasonable, considering that most CPG manufacturers visit small retailers commonly twice per week in emerging markets (Fransoo et al. 2017).

Cycle Service Level (CSL)

For the inventory policies that require a Cycle Service Level, we use a value of 95% in our models, and we assumed a normal distribution, given the lack of data.

Foreign Exchange Rate

Since all stores are located in Brazil, all data is in Brazilian Reais (BRL). We used the exchange rate of 4.7 BRL = 1 USD for all reports. This value was taken from the Brazilian Central Bank (Banco Central) on April 8, 2022.

Inventory Models

After setting up the parameters, we ran experiments to determine the best inventory policies for each store and item. To compare each policy, we calculated the total annual inventory cost and holding cost. By adding both, we got a total relevant cost (see Appendix C for further details on cost elements). That was the main parameter used when comparing different policies.

First, we analyzed a Base Stock policy, where we set a base stock value for a specific item per store, and every unit sold is ordered simultaneously. This model was used as a benchmark since it is not practical to keep track of every item in real life and order every unit sold immediately after.

The next model was the Continuous Review Policy (s, Q), where the store should order Q units when the inventory position is less than the reorder point called s. The Q value equals the EOQ formula, while s is based on safety stock to the set CSL as described in section 3.5. This model would also require that the store operator keeps track of inventory in real-time and orders at the exact time the stock level gets to s.

The following model was the Periodic Review Policy (R, S), where the store orders every R period up to the S level of stock. This policy fits the operator's reality since they usually don't keep track of real-time inventory levels. Most of them have other activities and choose to check and replenish their stocks periodically. See Appendix B for further details about the results from the inventory policies we tested.

The next step was to perform a sensitivity analysis for the R parameter to see the impacts of different review periods on all stores and items. We simulated the model for this set of R-values in weeks: 0.5, 1, 1.5, 2, 3, and 4. These values are more aligned with store operators' reality and behavior. For the first set of parameters we used (mentioned in section 5.3), the periodic review is always better than continuous review and base stock policies, no matter the R-value. That is driven mainly by the ratio between ordering costs for each policy. The following section will discuss the impacts of this parameter and others in our model.

Lastly, having simulated different R values for each store and item, we could calculate the average replenishment period for each cluster of stores and items. Therefore, Table 1 shows the average value of R that each combination of store and item clusters should use for the Periodic Review policy. We did not have enough data to compute those values for some combinations. Those cells are shown as "N/A" in the table.

Table 1

Average optimal R for each store and item cluster

The continuous review policy was the best (with lower total relevant costs) for all stores and items. Adding all stores, the Periodic Review policy with the optimal R-value for each combination of store and item clusters costs around 30% less than the Continuous Review policy.

We can see a lower review period for fast-moving items, such as Item Cluster 6 (i.e., I6). It means that store operators will need to replenish those items more frequently. Some item clusters, such as I5 and I9, have review periods close to 4. These clusters represent more regular groceries, such as dried pasta, cookies, and eggs. Although these are everyday items for groceries shopping, they are slow movers for Onii stores. Since Onii stores are based on convenience and proximity, they are usually visited for fast and lastminute purchases, not monthly groceries. Being slow movers, they need replenishment once a month.

Sensitivity Analysis

We ran sensitivity analysis for multiple parameters in our model. All calculations were made by setting the baseline case as defined in section 4.1. We changed one parameter from that baseline case while fixing all the other parameters. That way, we can understand the impact of each parameter individually in our model.

Ordering cost ratio

One of the most critical drivers of the inventory policies is the ratio between the ordering costs for continuous and periodic review policies. The number set as a baseline was 17%. The higher this ratio is, the higher the ordering cost for periodic-review models compared to the cost for the continuous-review models. The tipping point was found at 38%. From that breakeven point on, continuous-review models become more affordable.

Ordering cost value

By keeping the ordering cost ratio fixed at the baseline value, we wanted to identify the impact of the value on our model. We can see in Figure 13 that increasing the ordering

cost also raises the total cost and the difference between continuous and periodic-review policies. As expected, the gap between Continuous Review and Periodic Review policies increases as the ordering cost increases because the Continuous Review places more orders than the Periodic. With the Periodic, we are setting the same amount of orders every period, while with the Continuous Review policy, we place an order every time an SKU is needed.

Figure 13

Ordering Cost Sensitivity Analysis

Now, when analyzing the impact on the R-value, we can see a significant increase in the review period when moving from ordering costs around US\$0.1 to US\$0.35. Lower ordering costs imply ordering more frequently to replenish the shelves and, therefore, fewer times to review the inventory position. After the aforementioned increase, the average R-value grows steadily (Figure 14). The increased percentage of ordering costs can explain that behavior. While from US\$0.1 to US\$0.35 we are growing 250%, from US\$ 0.35 to US\$ 0.75 the increase is lower than 120%. But, in general, we can see that as the ordering cost gets higher, the model tries to place fewer orders by increasing the R.

Figure 14

Ordering Cost Sensitivity Analysis (R)

Holding Cost

We used different values of the CDI (Section 5.1) for the holding cost based on the last year's historical data. The lowest value we could find is 5%, while the highest was 14%, with the baseline being today's value (11.65%). As expected, total cost increases with higher holding costs, and R-value decreases slightly (Figures 15 and 16). The total cost rises mainly because we are growing one of the cost components, leaving everything else the same. R-value decreases because the holding cost increases; the model tries to hold less inventory by purchasing more frequently for lower quantities.

Holding Cost Sensitivity Analysis

Figure 16

Holding Cost Sensitivity Analysis (R)

Gross Margin

Total costs are lower when we increase the gross margin ratio (since we use the same sales price). A gross margin change impacts the cost in our model since we are keeping the same selling price. So, when the gross margin increases, the cost per item decreases. We can see the total costs decreasing with high margins and, indirectly, lower item costs (Figure 17).

Since the higher gross margin results in lower item costs, we can also see the R-value increasing (Figure 18). With lower item costs, the holding cost also gets lower. The model tries to order more quantities less frequently with lower holding costs, increasing the Rvalue.

Figure 17

Gross Margin Sensitivity Analysis

Gross Margin Sensitivity Analysis (R)

5 Recommendations

This section will discuss recommendations for future studies and actions for Onii and the store operators.

5.1 Improving Level of Detail on Financial Data

As discussed in section 5.3, some assumptions had to be made to set the parameters needed for simulated inventory policies. It is important to revisit these assumptions and conduct more in-depth exploratory research, particularly with store owners. We used an average margin for all subcategories for items cost, but not every item has the same margin. Understanding the detailed cost structure and the price per item for all the stores is essential.

It is also essential to understand how tradeoffs between ordering and holding costs play a role in the decision to increase inventory levels for specific subcategories. The latter may also change the choice between the periodic review and continuous-review policies. For the holding cost, although a very common and widely used benchmark was considered (i.e., CDI), this could also be fine-tuned by understanding the cost of capital or cost of opportunity for each store.

5.2 Implementation

The next step would imply implementing the inventory policies suggested at Onii stores. Onii can take advantage of its systems today and start to calculate the best R, s, Q, and S for each store and subcategory. The store operator would be able to control their inventory through the Onii system and identify the best time and quantity to replenish their stock.

Onii can also take a different strategy and have the store operators input their desired review period (R) if that is more realistic. Then a store operator is used to replenishing

their stock once a week, and they want to keep it that way, for their particular reasons. They can input that number into the model, and it will tell the optimal quantity for each item that they should replenish for that specific R-value. The same tool can show the history of orders and the cost and service level they received. This system can evolve to include reinforcement learning to suggest decisions for operators in the future or prevent them from making poor decisions

5.3 Database Update

Onii started operating in 2019, but their business started escalating in early 2020. We had only analyzed 76 out of 300 stores with over one year of sales history. The number of stores with at least one year of history will increase dramatically in the future. Thus, Onii should keep updating the database from time to time. With a more extensive database to work with, there may be some changes in the stores and items clustering, and the inventory policies if the demand distribution changes.

5.4 Summary of Recommendations for Onii

Therefore, we strongly recommend Onii to:

- Review parameters used to fit the model better
- Implement models proposed for the stores in this study
- Expand implementation for other stores
- Regularly update the analysis as they get more data from new stores
- Feed the rules of thumb from more experienced store operators in the definition of inventory policies.

6 Conclusions

The main objective of this project was to develop inventory policies for various store and product types. And develop standardized inventory policies to guide store operators. By first clustering stores and items, we were able to identify patterns across stores that, in a way, unify them. With these patterns, we were able to identify the best inventory policy for each combination of store and item clusters. We could see how PC1 and PC2 are driven by alcoholic beverages and frozen food, which describes 40% of the total variability in sales. Also, we did observe through clustering that "A" class items (high revenue) have more stable demand than low revenue items.

Our model is sensitive to demand volume, holding, and ordering cost, which helps stores choose the right policy, which is the 'Periodic Review Policy.' This model will allow stores to use this policy for different 'R' values given the demand and product characteristics. In our model, the Periodic Review policy has a significant ordering cost advantage due to the order bundling effect and hence is the best suitable policy.

Our project provides guidelines for all the store operators to choose the correct inventory policy based on their business reality, including the demand, Item characteristics, holding, and carrying cost. There were a few key learnings during this project, better discussed below.

6.1 Practicality of Solutions

The most important takeaway from this project is the model to choose among diverse inventory models. The continuous review policy, for example, usually shows the lowest cost in theory because it assumes the manager always buys the optimal quantity at the best time. However, it is unrealistic to think that Onii operators will monitor their stock in real-time and order everything at the exact time needed. Based on that premise, the periodic review policy was chosen, as it best fits operators' reality and conditions. The

periodic review policies suggested in this project are very realistic and easily applicable to all Onii stores since it is very close to what is already happening.

6.2 Cost Definitions and Imperfect Data

When developing the model and simulating each inventory policy, it became apparent how important the description of parameters is. Although we chose very realistic parameters, the model can be susceptible to variations, particularly in costs. Onii needs to create protocols to make information easily accessible and standardize its collection across stakeholders. A careful analysis is required to understand the real costs for each store deeply through field research.

6.3 Final Considerations

This project delivers a first-order tool to define applicable inventory policies for Onii's store operators. Although fine-tuning can be made through parameter adjustments, it is already a great starting point that Onii can offer its operators some rules of thumb based on quantitative models. Today, operators rely entirely on their own experience and knowledge about inventory management, which is usually not very vast. By following these basic but robust inventory policies, operators can have much better inventory management than now.

References

- Ahmad K., Zabrie S.M. (2016). Inventory Management practices among Malaysia micro retailing enterprises. Journal of business and retail management research (JBRMR), Vol.11 Issue 1
- Arkaresvimun, R. (2008). Inventory Optimization in a Retail Multi-Echelon Environment. Master of Engineering in Logistics thesis at Massachusetts Institute of Technology.
- Ballou, R. H. (2006). Business Logistics Management.
- Berry, L.L. (2001). Old Pillars of New Retailing. Harvard Business Review, Vol.79, No.4, pp 131-137
- Broekmeulen R.A.C.M., Van Donselaar K. H., Fransoo J. C., Van Woensel T. (2004). Excess shelf space in retail stores: An analytical model and empirical assessment. BETA Working paper series 109, Eindhoven
- Bultez A., Naert P. (1988). SH.A.R.P.: Shelf allocation for retailer's profit. Marketing Science. Vol 7, No 3
- Cachon, G. (2001). Managing a retailer's shelf space, inventory, and transportation. Manufacturing & Service Operations Management, 3(3):211-229.
- Campo K., Gijsbrechts E., Nisol p. (2000). Towards understanding consumer response to stock outs. Journal of retailing, Vol.76 (2), p.219-242
- Chikan A., Whybark C.D. (1990). 'Cross-national comparison of production and inventory management practices. Engineering Costs and Production Economics, vol. 19, pp. 149-56.
- CNN Brasil (2021). Com pandemia, mercadinhos de bairro dão salto tecnológico e vendas sobem. [https://www.cnnbrasil.com.br/business/com-pandemia](https://www.cnnbrasil.com.br/business/com-pandemia-mercadinhos-de-bairro-dao-salto-tecnologico-e-vendas-sobem/)[mercadinhos-de-bairro-dao-salto-tecnologico-e-vendas-sobem/](https://www.cnnbrasil.com.br/business/com-pandemia-mercadinhos-de-bairro-dao-salto-tecnologico-e-vendas-sobem/)
- De Vries, G. J. (2010) Small Retailers in Brazil: Are Formal Firms Really More Productive? Journal of Development Studies, Vol. 46, No. 8, 1345–1366
- Ehrenthal J.C.F., Honhon D., Van Woensel T. (2014). Demand seasonality in retail inventory management. European Journal of Operational Research. 238, 527-539
- Ehrenthal J.C.F, Stölzle W. (2013). An examination of the causes for retail stockouts. International Journal of Physical Distribution & Logistics Management Vol. 43 No. 1, pp. 54-69
- Escamilla R., Fransoo J.C., Tang C. S. Improving Agility, Adaptability, Alignment, Accessibility, and Affordability in Nanostore Supply Chains. Production and Operations Management Vol. 30, No. 3, March 2021, pp. 676–688
- Fransoo J.C.; Blanco E.E.; Mejia-Argueta, C. (2017). Reaching 50 million Nanostores: Retail Distribution in Emerging Megacities. CreateSpace Independent Publishing Platform: Scotts Valley, CA, USA, 2017.
- Guo X., Yu Y., De Koster R.B.M. (2016) Impact of required storage space on storage policy performance in a unit-load warehouse. International Journal of Production Research, 54:8, 2405-2418
- Jackson I., Tolujevs J., Kegenbekov Z. (2020). Review Of Inventory Control Models: A Classification Based on Methods Of Obtaining Optimal Control Parameters. Transport and Telecommunication, 2020, volume 21, no. 3, 191–202
- JOENG (The Journal of Engineering) (2018). Research and Markets; Global Vending Machine Market Analysis & Forecast 2017-2023, With an Expected CAGR of 15.8%. June 18, 2018. 642
- Johnston A. (2014). Trends in retail inventory performance: 1982 2012. Oper Manag Res 7:86–98.
- Katehakis M. N., Melamed B., Shi J. Cash-Flow Based Dynamic Inventory Management. Production and Operations Management Vol. 25, No. 9, September 2016, pp. 1558-1575
- Kosters, D.M. An improvement of the refill process of vending machines. A thesis about demand analysis, product range improvement, and review period calculation. University of Twente (2021)
- Lin F.C., Yu H.W., Hsu C.H., Weng T.C. Recommendation system for localized products in vending machines. Expert Systems with Applications 38 (2011) 9129– 9138
- Marshall, D. (2019). Convenience stores and well-being of young Japanese consumers. International journal of retail & distribution management, 2019-07-31, Vol.47 (6), p.590-604
- Mingchao W., Yihui H., Lei Z., Tianhu D., Fransoo J.C. Demand estimation under multi-store multi-product substitution in high-density traditional retail. European Journal of Operational Research 266 (2018) 99–111
- Monteiro J. C.M., Assunção J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. Journal of Development Economics 99 (2012) 105–115
- Muckstadt J. A., Sapra A. (2010). Principles of Inventory Management When you are down to four, order more. Springer series in Operations Research and Financial Engineering.
- Muller, M. (2019). Essentials of inventory management. American Management Association
- Parente, J. G. (2008). O Varejo De Alimentos Para Consumidores De Baixa Renda No Brasil. fgv-eaesp/gvpesquisa. Relatório 16/2008
- Paswan A., Pineda M. D. S., Ramirez, F. C. S. (2010). Small versus large retail stores in an emerging market—Mexico. Journal of Business Research 63 (2010) 667–672
- Roggeveen A. L., Sethuraman R. (2020). Customer-Interfacing Retail Technologies in 2020 & Beyond: An Integrative Framework and Research Directions. Journal of Retailing 96 (3, 2020) 299–309
- Serviço Brasileiro de Apoio às Micro e Pequenas Empresas Sebrae. (2014). Estudo Setorial Minimercados – Série Estudos Mercadológicos. [https://www.sebrae.com.br/sites/PortalSebrae/bis/conheca-o-setor-de](https://www.sebrae.com.br/sites/PortalSebrae/bis/conheca-o-setor-de-minimercados-no-brasil,b637bb90065e2510VgnVCM1000004c00210aRCRD)[minimercados-no-brasil,b637bb90065e2510VgnVCM1000004c00210aRCRD](https://www.sebrae.com.br/sites/PortalSebrae/bis/conheca-o-setor-de-minimercados-no-brasil,b637bb90065e2510VgnVCM1000004c00210aRCRD)
- Serviço Brasileiro de Apoio às Micro e Pequenas Empresas Sebrae. (2015) Pesquisa Minimercados no Brasil.

[https://www.sebrae.com.br/Sebrae/Portal%20Sebrae/UFs/SP/Programas%20e%20](https://www.sebrae.com.br/Sebrae/Portal%20Sebrae/UFs/SP/Programas%20e%20Solu%C3%A7%C3%B5es/Loja%20Modelo/lmv_minimercado.pdf) [Solu%C3%A7%C3%B5es/Loja%20Modelo/lmv_minimercado.pdf](https://www.sebrae.com.br/Sebrae/Portal%20Sebrae/UFs/SP/Programas%20e%20Solu%C3%A7%C3%B5es/Loja%20Modelo/lmv_minimercado.pdf)

- Sinha P. K., Banerjee A. (2004). Store choice behaviour in an evolving market. International Journal of Retail & Distribution Management; 2004; 32, 10; ABI/INFORM Collection pg. 482
- Vallandingham L. R., Yu Q., Sharma N., Strandhagen, J. W. and Strandhagen, J. O. (2018). Grocery retail supply chain planning and control: Impact of consumer trends and enabling technologies. IFAC PapersOnLine 51-11 (2018) 612–617

Winston, W.L.(2004). *Operation Research* (Vol 4). Brooks/Vole.

- Wrigley N., Wood S., Lambiri D., Lowe M. (2019). Corporate convenience store development effects in small towns: Convenience culture during economic and digital storms. EPA: Economy and Space 2019, Vol. 51(1) 112–132
- Xue W., Demirag O.C., Chen F.Y., Yang Y. (2017). Managing Retail Shelf and Backroom Inventories: When Demand Depends on the Shelf-Stock Level. Production and Operations Management Vol. 26, No. 9, September 2017, pp. 1685–1704.
- Zhang W., Rajaram K. (2017). Managing limited retail space for basic products: Space sharing vs. space dedication. European Journal of Operational Research 263 (2017) 768–781

Appendix A – Example of Orange Workflow

Appendix B – Results Table for Diverse Inventory Policies

Appendix C – Total Cost Formulas

- C_e : Excess Holding Cost (\$/unit/time)
- C_t : Ordering Cost (\$/order)
- : Replenishment Order Quantity (units/order)
- : Safety Factor
- σ_{DL} : Standard Deviation of Demand Over Lead Time (units/time)
- σ_{DL+R} : Standard Deviation of Demand Over Lead Time plus Review Period (units/time)
- *HC:* Total Holding Cost
- *OC:* Total Ordering Cost
- *D:* Demand
- *R:* Review Period
- *L*: Lead Time

Continuous Review

- Holding Cost $HC = c_e$ Q $\frac{1}{2} + k \sigma_{DL}$
- Ordering Cost $OC = c_t$ \overline{D} Q)
- Total Cost *TC = OC + HC*

Periodic Review

- Holding Cost $HC = c_e$ DR $\frac{1}{2} + k \sigma_{DL+R} + DL$
- Ordering Cost $OC = c_t$ 1 \boldsymbol{R})
- Total Cost *TC = OC + HC*