

NOVA

IMS

Information
Management
School

MDSAA

Mestrado em

Data Science and Advanced Analytics

Recommended Price Optimization in Convenience Franchising

A Data-Driven Strategy for a Retail Network

André Marques Lourenço de Almeida Dias

Master Thesis

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Recommended Price Optimisation in Convenience Franchising
A Data-Driven Strategy for a Retail Network

by

André Marques Lourenço de Almeida Dias

Master Thesis presented as partial requirement for obtaining the Master's degree in Data
Science and Advanced Analytics, with a specialization in Business Analytics

Supervised by

Roberto Henriques, PhD in Information Management, NOVA IMS

May, 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism, any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 27/05/2025

André Marques Lourenço de Almeida Dias

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents, who never stopped supporting me, even when I felt like I couldn't go on. I am also grateful to my supervisor, Professor Roberto Henriques, for his clear and direct guidance, always helping me understand what was truly necessary for the development of this thesis. A special thank you to Andreia Nogueira, Ricardo Ribeiro, and Tiago Mota, for their constant support and availability in answering all my business-related questions. They were incredibly helpful and great to work with, and I deeply appreciate their collaboration.

ABSTRACT

This thesis explores a pricing recommendation strategy built for a convenience store chain for the franchisees. The objective is to build a hybrid classification between business logic and machine learning to construct a classification model of the store's product range based on sales and profit. The process starts with business understanding and ends with model evaluation, following the CRISP-DM methodology. Initially, a manual and arbitrary classification was created, where a score based on static thresholds would classify the product type using regular (non-promotional) sales quantity and the franchisee's profit margin. Despite this, this approach has limitations: it is subjective, static, and may not adapt to future market changes. Machine learning overcomes these limitations by integrating algorithms such as Random Forest, KNN and Naive Bayes for validation and to automate classifications. To train and build this classification, 2024 sales data were collected, including margins, prices, and sales across all stores, to study product behavior and classify them strategically into essential, medium and premium. By classifying through algorithms and with the learned models and accurate results, product classifications will allow the pricing strategy to become more automated and to better respond to changes in demand. By associating these classifications with differentiated pricing strategies, the model strengthens the effectiveness of commercial decisions in a dynamic retail context.

KEYWORDS

Pricing Strategy; Convenience Stores; Franchising; Recommended Price Optimization; Consumer Behavior

Sustainable Development Goals (SDG):



TABLE OF CONTENTS

Statement of Integrity	ii
Acknowledgements	iii
Abstract	iv
List of Figures.....	vi
List of Abbreviations and Acronyms	vii
1. Introduction	1
2. Literature review	3
2.1 Convenience Stores: Characteristics and Pricing Context	3
2.2 Pricing Strategies in Food Retail and Their Applications to Convenience	3
2.2.1 Everyday Low Pricing (EDLP)	4
2.2.2 High-Low Pricing (Hi-Lo).....	4
2.2.3 Value Based Pricing.....	5
2.2.4 Dynamic Pricing	6
2.3 Pricing Models in Franchising Networks	8
2.4 Legal Considerations: Resale Price Maintenance and Legal Limits.....	8
2.5 Standardization vs Flexibility: Balancing the Franchise Pricing Strategy.....	10
2.6 Communication and Trust in the Franchisor-Franchisee Relationship	10
2.7 Technologies and Machine Learning in Price Optimisation	12
2.8 Conclusion.....	13
3. Methodology.....	14
3.1 Business Understanding.....	14
3.2 Data Understanding and Empirical Study	15
3.3 Data Preparation.....	17
3.4 Modeling	20
3.4.1 Unsupervised Segmentation with K-means	21
3.4.2 High-Low Pricing (Hi-Lo).....	22
3.5 Deployment.....	23
4. Results and Discussion	24
5. Conclusions and Future Research	26
Bibliographical References.....	28

LIST OF FIGURES

Figure 1 – Pie charts showing the proportion of items in terms of sales volume and number of products.....	19
Figure 2 – Distribution of items by regular margin and regular quantity sold (log-transformed)	20
Figure 3 – K-means clustering results illustrating the classification.....	21

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
EDLP	Everyday low price
Hi-Lo	High-low pricing
MSRP	Manufacturer's Suggested Retail Price
RRP	Recommended Retail Price
RPM	Resale Price Maintenance
KVI	Key Value Item
SKU	Stock Keeping Unit
ML	Machine Learning
PVF	Purchase Price to Franchisee
PVPR	Recommended Retail Price to Public (in Portuguese context)

1. INTRODUCTION

In the retail market, it is crucial to define pricing strategies for product sales. Not only in hypermarkets but also in convenience stores, where attributes such as proximity, accessibility and speed allow for the practice of higher prices. These variables play a very important role in this type of business because consumers are willing to pay a convenience premium. For example, customers are paying up to 21% more in convenience stores than they pay in supermarkets of the same chains (Butler, 2024). This difference is not only justified by the convenience offered to the customer but also by the high operational cost, for example, in more urban areas (Butler, 2024). In fact, customers are willing to pay more for convenience, but only up to a certain point, so these types of stores have to find an ideal strategic trade-off between accessibility and low prices. Thus, convenience retail must also consider consumer sensitivity. Convenience influences just as much as the price of the item for some consumers. According to a study, almost half of consumers are willing to pay a premium price in convenience stores (Sullivan & Revionics, 2017). Despite this, there is still sensitivity and very high margins may drive away more price-sensitive consumers. This behavior also happens in other markets, such as in digital contexts like home delivery, where menu prices can be on average 20% higher than in physical stores, excluding delivery fees (Fisher, 2024). Another study found that including fees and commissions, there are cases of fast-food orders where prices are 90% more expensive than counter prices (Raymond, 2024). These studies support the idea that consumers are willing to pay for convenience when it saves them time and offers ease, as long as they do not feel exploited or perceive prices as abusive. Therefore, the goal for success is to set prices that reflect the value of that convenience without exaggeration that could compromise customer trust and loyalty (Wagner et al., 2018).

In franchised convenience store networks, pricing strategy becomes even more complex. This happens because franchisees manage their stores autonomously, and the supplier can only recommend retail prices (MSRP). The franchisor sets the cost price for the franchisee and a recommended retail price for the final consumer. The franchisee generates profit and covers operating expenses through the difference between the price paid to the supplier and the price charged to the public (RRP) (FMS Franchise, 2023). This creates a pricing management dilemma between recommending prices adapted to the local context — adjusting them to local realities — or recommending uniform prices across the network, which makes the brand more consistent and builds customer trust (Boulay et al., 2020). In addition to this challenge, it also enables the implementation of promotional campaigns with standardized maximum prices and reinforces the perception of consistent quality (Rosado-Serrano et al., 2018). However, applying regional pricing is essential for franchisees to adapt to local realities and achieve profitability (Nyadzayo et al., 2018). This is due to the fact that demographics, competitive intensity, and operational costs (such as rent and wages) can vary from store to store. Thus, the key to a balanced pricing strategy is to ensure brand consistency and fairness among franchisees while offering enough flexibility to adapt to each store's local conditions.

The objective of this thesis is to develop a new recommended pricing strategy for franchisees of the store chain in study, through data analysis and decision support. To achieve this, the study will review the principles of major pricing strategies (everyday low price, planned promotions, dynamic pricing, or value-based pricing) in connection with consumer behavior in convenience retail—specifically, the customer’s willingness to pay more in exchange for time, speed, and convenience.

Following this introduction, the Literature Review explores: (i) the nature of convenience stores and consumer behavior regarding price; (ii) classic pricing strategies in food retail (EDLP, high–low, etc.) and their relevance in convenience contexts; (iii) pricing models in franchising, including challenges related to communication, trust, and fairness between franchisor and franchisees; (iv) legal limitations on the imposition of resale prices; and (v) the growing role of technology and machine learning in retail price optimization.

2. LITERATURE REVIEW

2.1. CONVENIENCE STORES: CHARACTERISTICS AND PRICING CONTEXT

Convenience stores are food retail outlets that stand out for offering speed, ease, and a streamlined customer experience. These stores can not only be found in high-traffic and urban areas but also stand out for their presence in rural and isolated zones, where retail options are limited, and proximity and availability play a very strong role. Regardless of the area in which the stores are located, the assortment always tends to focus on items of immediate need, from basic food products to urgent goods, reinforcing the idea that the customer experience is quick and close. In return, prices in these stores are higher than in hypermarkets. In the UK, large retail chains charge between 10% and 15% more in their convenience stores compared to their bigger formats, with differences reaching up to 21% in some cases (Butler, 2024). This happens because of the operational costs of having stores in more convenient locations and the premium that comes with it. In short, the customer is willing to pay for convenience because of the immediate availability of the products they need (Benoit et al., 2020).

It has already been concluded that consumers are willing to pay more for convenience, but it is important to highlight that the type of product and the type of consumer also dictate what is bought. Products bought very frequently, like bread or milk, tend to be more price-sensitive even in convenience stores, so there is not much room to raise the price on these because the customer may buy them later at a lower price in the supermarket (Alexandrov et al., 2022). On the other hand, impulse products like a cold drink or a painkiller, which are usually bought in more “emergency” situations, customers don’t mind paying a higher price (Benoit et al., 2020). Franchisees normally understand this difference, and for that reason they tend to apply lower margins on essential goods to build customer loyalty and increase them on indulgent or convenience items (Alexandrov et al., 2022). This way, the franchisee can increase their revenue by segmenting products based on their usefulness to the customer, while avoiding the perception of unfair pricing.

2.2. PRICING STRATEGIES IN FOOD RETAIL AND THEIR APPLICATIONS TO CONVENIENCE

The retail pricing literature identifies several classic approaches to setting product prices, each with advantages, disadvantages, and varying levels of suitability for the convenience store context. The four most referenced strategies are: Everyday Low Pricing (EDLP), High–Low Pricing (Hi-Lo), Value-Based Pricing, and Dynamic Pricing. Below, we will analyze each of them in turn.

2.2.1. EVERYDAY LOW PRICING (EDLP)

Everyday Low Price (EDLP) is a strategy that, unlike switching between discounts and regular prices, consistently keeps prices low across all or most products. This strategy aims to create the perception that prices are always low, which attracts customers. Walmart is an example of a major chain that uses this strategy, as it conveys an image of long-term affordability and avoids the “rollercoaster” of promotions (Ortmeyer et al., 1991).

A consistent EDLP strategy is rarely applied to convenience stores, since this type of store does not benefit from the economies of scale and negotiation power that hypermarkets have. They also face higher unit costs compared to larger retailers, which makes it even harder to apply low margins across all products. In addition, the product range in these stores is too limited for such reduced margins to be sustainable. Some retailers prefer to adopt a hybrid EDLP strategy, meaning they keep prices low and competitive on a set of essential items (to build trust in the store) while applying higher margins on products with lower price sensitivity (Flipkart Commerce Cloud, 2024). For example, the franchisee may decide to sell items like bread, rice, or milk at the same price—or slightly higher—than the local supermarket, while increasing margins on products like snacks or chocolates. Studies show that customers tend to stay loyal to stores when they feel that the prices of key items are not exploitative, especially since these are products with well-known prices (Sullivan & Revionics, 2017). That said, an approach that applies the EDLP strategy to KVIs (Known Value Items) has the potential to create benefits without compromising profitability. The downside of this approach may stem from margin erosion in the absence of sufficient sales volume, since maintaining low prices every day requires selling in higher quantities — and purchases in convenience stores are typically small and quick. For this reason, retailers who opt for this strategy are likely to face financial difficulties. The most viable EDLP solution may be to apply it only to essential items, creating more attractive prices, while increasing profitability through higher margins on other products (Flipkart Commerce Cloud, 2024). This mixed approach is almost inevitable in the current format of convenience retail.

2.2.2. HIGH–LOW PRICING (HI-LO)

Unlike the previous strategy, High–Low (Hi-Lo) is characterized by high regular prices that are occasionally subject to aggressive promotions on a selected group of products. Instead of keeping prices consistently low, this approach involves maintaining high prices most of the time and then offering significant discounts at regular intervals to attract promotion-sensitive consumers, thereby boosting store sales (, 1997). Hypermarkets and supermarkets adopt this strategy, usually through weekly leaflets featuring campaigns such as direct discounts or “buy 3, pay for 2” offers. Although it is less predominant in convenience retail, it is not entirely absent. Convenience and neighborhood stores are smaller than supermarkets, so running as

many promotions is not feasible due to operational limitations. Even so, the strategy can still be applied—for example, by launching a discount on ice cream during a heatwave or running seasonal campaigns, such as a summer pack with reduced prices on sunscreen or bottled water. Promotions like these can increase foot traffic in stores and may encourage impulse purchases from customers if well executed (Flipkart Commerce Cloud, 2025). This strategy can also help the franchisee clear perishable stock—for example, by lowering the prices of ready-to-eat meals at the end of the day to avoid waste.

The Hi-Lo strategy also allows retailers to engage customers through new offers and good deals. The excitement of finding a great discount encourages them to visit the store more often to see what's on promotion (Flipkart Commerce Cloud, 2024). However, over-reliance on promotions can be risky, as it may compromise profit margins in the long run (Flipkart Commerce Cloud, 2024). An example of this occurred with J.C. Penney, which in 2012 eliminated promotions and shifted to an EDLP strategy. However, customers, who were used to frequent promotions, stopped shopping at their stores, leading to a drop in sales. To recover, J.C. Penney had to return to its previous approach to win back customers (Flipkart Commerce Cloud, 2024). This case shows that switching from one strategy to another in a radical way can compromise the business.

In the case of convenience stores operating under a franchise model, the best option is to adopt a moderate Hi-Lo approach, running promotions selectively and focusing high margins on certain categories, while applying campaigns during timely periods (such as chocolate discounts during Easter). This allows franchisees to boost their sales without compromising profitability or creating dependency on promotions.

2.2.3. VALUE-BASED PRICING

In the value-based pricing strategy, prices are set according to what is perceived by the customer—whether it's a service or a product. Instead of basing the price entirely on costs or competitor pricing, this approach considers the solution offered to the customer and how much they are willing to pay as a key factor (Simon & Fassnacht, 2019). This is reflected in practical terms when two products have similar costs but end up with different prices because customers assign different levels of perceived value to each. In convenience stores, value involves even more variables, as it includes elements such as immediacy, location, and the convenience of service (Wall Street Prep, 2024). For example, a carton of milk may be the same in a convenience store 100 metres away that stays open until 10 p.m. and in a supermarket several kilometres away. However, the perceived value in the convenience store will be higher due to its proximity—and because the supermarket may be closed, making the c-store the only available option (Wall Street Prep, 2024). That said, in the value-based pricing strategy, the customer recognizes that they are paying for time savings and immediate

problem-solving, which makes higher prices for these products perfectly justifiable.

Some studies indicate that urgent or impulse-driven products tend to have higher margins under this model. An example of that are batteries, that can be effective by being sold in a 24-hour store. Headache medicine or ready-to-eat meals also tend to be priced higher than or like supermarket equivalents because customers value immediate availability (Wall Street Prep, 2024). Applying high margins to products with high perceived value is an excellent way to capture consumer surplus.

However, to do this effectively, the franchisee needs to apply analytical discipline, as it is necessary to understand price elasticity at the product category level and how the perceived value profile of each item is evaluated by customers. The experienced managers that know how the market works, know how to apply this approach intuitively by giving higher margins to impulse purchases (e.g., indulgent snacks, tobacco, or chilled drinks in the summer) because they know that there is a strong likelihood that customers pay less attention to price in these contexts (Flipkart Commerce Cloud, 2024). Immediate satisfaction and momentary pleasure are factors that support this logic, and customers are often willing to pay a few extra euros. Despite this, there are limits, the franchisee cannot raise prices excessively on essential goods, as this can easily drive the customer away for good. With this approach, if the franchisees identify high added-value items and keeps price-sensitive products in line with market levels, they will preserve the credibility of their stores.

In short, the value-based pricing strategy is a very predominant approach in the convenience franchising model. By adding value to the product simply through the convenience offered, there is already a considerable increase. So, if applied intelligently, it can enhance profitability where customers most seek service, while at the same time maintaining affordable prices on essential goods.

2.2.4 DYNAMIC PRICING

Dynamic pricing is a strategy that involves adjusting prices flexibly and frequently in response to market changes, from supply and demand to competitor pricing (Chen et al., 2022). There are no fixed prices—retailers adjust them almost in real time to optimize revenue and reduce waste. This strategy is used in aviation, hospitality, and in ride-hailing apps like Uber. Technological advancements, supported by AI, allow for more advanced control with technology support such as AI systems or electronic shelf labels (Omnisend, 2024). There are two possible scenarios for convenience stores to adopt this strategy:

- Adjustments based on weather or time of day: if unexpected rain occurs, the store manager can immediately raise the price of umbrellas or offer a discount on

coffee during a cold morning to attract customers into the store.

- Response to local competition: if the manager notices that a nearby competitor has lowered the price of a KVI, he can use a dynamic system to temporarily match the price to avoid losing potential sales (Omnisend, 2024).

Real-world cases demonstrate the potential benefits. The Norwegian chain REMA 1000 has been implementing dynamic pricing since 2012: by using electronic shelf labels and centralized pricing management, they are able to carry out up to 2,000 price changes per day across their entire network. Thanks to the adoption of this strategy, there was a 40% reduction in food waste, along with the maintenance of local price competitiveness (Hamstra, 2024). Due to the ability of algorithms to capture additional margins and sales, it is estimated that adopting this strategy in supermarkets could increase profit margins by 10% to 20% (Adams et al., 2024). In short, managing prices dynamically can bring efficiency by aligning with consumers' willingness to pay and current market conditions.

On the other hand, this approach presents challenges such as operational execution and consumer acceptance.

This approach can also be dangerous in terms of the perception the customer has of the store. So, if prices become too flexible, the amount of distrust will increase. In a Gartner citation in the media, 68% of consumers feel negatively impacted by price fluctuations, as stability and transparency are more important to them (Cardus, 2013). This leads consultancies to highlight that dynamic pricing implementations should be customer-focused and as transparent as possible, otherwise the trust customers have in the brand may be compromised (Adams et al., 2024). For example, a store must be careful with customers who notice a price change during the same visit. The store should have clear communication, such as labelling them as "anti-waste offers."

The challenge is not only on the customer side — having a central and technologically advanced structure (digital labels or integrated data systems) requires strong operational capacity and coordination between the franchisor and the franchisee. Dynamic pricing does not allow franchisees to set their own prices, as it requires central control. Therefore, this model is only feasible in situations involving optional recommendations. If it is controlled centrally, it requires strong trust between the franchisee and the franchisor. In summary, dynamic pricing can generate more profit to the business by reducing waste or responding to weather conditions. Customer satisfaction is a factor to pay attention to, so focusing on transparency or running pilot projects can be essential to avoid compromising it.

2.3. PRICING MODELS IN FRANCHISING NETWORKS

Franchising models must simultaneously ensure an optimal recommended price that delivers sustainable profitability across the network, while allowing each franchisee to remain competitive in their local market (Rosado-Serrano et al., 2018). One solution to achieve these objectives is the implementation of the Manufacturer's Suggested Retail Price (MSRP), that is, the franchisor recommends a reference price to the franchisee. These prices play a fundamental supporting role across the store network. They aren't mandatory, so the franchisee still has the decision to use them or not (FMS Franchise, 2023). Many franchising contracts include price ceilings or price intervals. To protect the brand, the franchisor sets a maximum resale price for products, avoiding issues related to price fluctuations or customer distrust (FMS Franchise, 2023).

Setting a fixed price may violate laws on vertical price fixing (this will be addressed in the following section), whereas setting a maximum price is generally not a problem. Although a maximum price is an option, minimum prices can be prohibited in certain contexts, as they may be considered resale price maintenance (RPM), and are therefore generally avoided. That said, the key lies in creating a kind of system of influence, which can involve incentives or subtle pressure so that franchisees adopt, in practice, what is recommended — for example, through the provision of a price list. This method is not considered a legal imposition and allows for full or near-complete price uniformity (Basset & Perrigot et al., 2016)

In summary, the most common model in franchising is that of non-mandatory recommended prices. This model allows for a balance between the franchisor's expertise in pricing — for example, through market studies — and the flexibility needed to adapt locally whenever necessary. Avoiding destructive price competition between franchisees, maintaining a consistent value proposition across the brand, and taking franchisee feedback into account when adjusting price recommendations will bring the best results for the business. In the next section, the legal restrictions governing these practices will be explored.

2.4. LEGAL CONSIDERATIONS: RESALE PRICE MAINTENANCE AND LEGAL LIMITS

Setting resale prices in franchising networks faces restrictions imposed by competition and antitrust laws in many jurisdictions. In countries like the United States or even within the European Union, minimum resale prices between suppliers and distributors are not allowed, legally known as Resale Price Maintenance (RPM), as it is considered an anti-competitive practice (European Commission, 2010). Since franchisors cannot impose minimum prices, they suggest prices and set maximum price limits. This practice is acceptable, as upper limits are seen as protecting consumers (Blair & Lafontaine, 2005).

In the European Union, the Vertical Guidelines (Regulation 330/2010) state that it is legal to recommend prices and set maximum limits, if this does not amount to a de facto fixed price — that is, if a franchisor influences or encourages the franchisee to apply the "recommended" price in a way that is almost mandatory.

In practice, the indirect influence that franchisors are able to exert on franchisees is not illegal because the prices are only “suggested”, and adherence is voluntary. However, franchisees may feel commercial pressure to comply. For example, a franchisee may not want customers to feel distrust regarding prices, so they end up following the suggestion to maintain brand consistency and uniformity.

Furthermore, the franchisee purchase price (PVF) also ends up creating a certain influence on the final price, since if the PVF is high, the franchisee cannot lower the selling price without sacrificing their margin. Despite these strong influences, the behaviour remains in line with the law: “price freely set by the franchisee” (Buchan, 2013).

The maximum resale price in franchising — that is, stating that product X cannot be sold for more than Y — serves to help the franchisor protect their brand from potential abusive practices by franchisees. This measure is also well accepted by consumers, as it prevents prices from rising too much. On the other hand, setting a minimum price is more difficult due to its legal risks. With this, the goal franchisors aim for is to propose a price alignment strategy that allows their business to remain consistent while also giving franchisees the autonomy needed to avoid infringing RPM laws (Rosado-Serrano et al., 2018).

Another important variable to highlight that influences franchisee is that, if cost prices increase, the franchisor may raise the wholesale price, and if the franchisee doesn't want to lose margin, they will have to raise their retail price — which leads them to get closer to the franchisor's recommendations. Another concern for franchisees is when local competitors launch long promotional campaigns and the franchisor does not allow enough pricing flexibility, making it difficult for the franchisee to respond and potentially resulting in lost sales. For this reason, the influence of franchisors in pricing decisions has been a widely debated topic in several countries. One more reason why franchisees often end up following the franchisor's recommendations is to avoid internal cannibalisation and because they trust the franchisor's know-how — expecting them to take local realities into account and to share the risks (for example, by adjusting the wholesale price when necessary) (Euzébio et al., 2024). For this to work, transparency is fundamental to building and maintaining that trust. The franchisor must communicate the pricing strategy clearly, showing that it is based on solid market data and designed to benefit both parties. Otherwise, franchisees may feel the strategy lacks foundation and is arbitrary — or even that it unfairly favours the franchisor — and may feel wronged (Kumar et al., 1995).

In summary, franchisors cannot force franchisees to apply their pricing recommendations, to avoid violating RPM restrictions. However, if applied carefully, legally, and adapted to local needs and realities, these recommendations can be successfully implemented and bring advantages to the brand. The key lies in the franchisor demonstrating transparent and well-founded communication, so that the franchisee adopts the franchisor's strategy with trust — a topic that will be addressed in the next sections.

2.5. STANDARDIZATION VS. FLEXIBILITY: BALANCING THE FRANCHISE PRICING STRATEGY

It's already clear that limiting the prices at which franchisees sell products is important for brand consistency. However, it's also important for the strategy to be flexible enough to adapt to the local market. So how can both seemingly contradictory goals be achieved?

A common approach is to adopt tiered pricing zones by region — in other words, the franchisor recommends different price levels depending on the local context, such as more expensive urban centres, suburban areas, or rural regions. One of the food retail chains, Albert Heijn, applies different pricing based on local purchasing power and competition (Euzébio et al., 2024). This method is well accepted by franchisees because the recommended prices reflect the operational reality of the stores (Euzébio et al., 2024).

Convenience stores have different cost structures depending on their location. In urban areas, where rents and wages are higher, operational costs tend to be greater, while in peripheral areas expenses are usually lower — but consumers also tend to have less purchasing power (Euzébio et al., 2024). Franchisees with such different situations are unlikely to thrive under a uniform pricing regime.

In summary, it is possible to find a balance between standardisation and flexibility through maximum price limits, recommended prices, regional pricing guidelines, and openness to justified exceptions. The franchisor maintains sufficient control to avoid anomalies or internal conflicts, while franchisees feel they have the ability to adjust their prices when necessary.

2.6. COMMUNICATION AND TRUST IN THE FRANCHISOR–FRANCHISEE RELATIONSHIP

The issue of pricing definition requires clear communication and mutual trust between franchisees and franchisors, as it affects the profitability of both parties and the satisfaction of the final customer. Therefore, good communication between both sides is essential to reach a successful business agreement.

Studies on franchising relationships show that perceptions of fairness strongly influence

franchisee cooperation (Kumar et al., 1995). Franchisors must take into account the financial health of their franchisees when setting recommended prices, and not focus solely on royalties or overall brand competitiveness. If they don't, franchisees may feel dissatisfied and even express it during network meetings. However, if the franchisor clearly and logically explains the reasoning behind each defined price — considering factors such as elasticity and listening to franchisee feedback about local challenges or opportunities — this will help build trust and cooperation (Nyadzayo et al., 2018). For the strategy to succeed, no party should be left out — all sides must be aligned with the process.

A good practice to make this happen is to include franchisee representatives in committees or pilot programs for implementing new pricing strategies. For example, if the franchisor wants to change the structure of their pricing strategy, they can work directly with selected franchisees to implement it in their stores — gathering local insights from the franchisees themselves so that both sides can collaborate during the pilot phase. If the test is successful, the strategy can then be rolled out across the rest of the network. Franchisees feel they have influence and participation when strategies are developed through dialogue and mutual satisfaction. This leads to greater franchisee engagement and transparency, reduces the perception of unilateral imposition, and can ultimately improve overall network performance (Rosado-Serrano et al., 2018).

In addition, it is very important to manage conflicts and exceptions. If, in a given situation, a franchisee is unable to follow the recommended retail price — for example, because a local competitor launches an aggressive, long-term campaign — the franchisor should allow some flexibility, such as providing marketing support or temporarily allowing the franchisee to lower prices. This kind of response shows trust from the franchisor towards the franchisee. If the franchisor shows resistance and insists on maintaining the recommendation, the franchisee may feel unfairly treated and under increased pressure, as the risk will fall entirely on them (Euzébio et al., 2024).

Finally, consistency is also a very important factor in building trust in franchise relationships. The franchisor must be coherent and avoid arbitrary exceptions or preferential treatment among franchisees. Pricing should be communicated in advance and with clear justification — for example, explaining that an increase in supplier costs means that recommended prices will need to rise by 5% in the following month. In addition to being communicated in advance, price changes should also come with guidance to mitigate their impact. For instance, to reduce the effect of the price increase, the franchisor could suggest that franchisees inform customers that the rise is due to a global increase in wheat prices. This approach shows transparency, prevents misunderstandings, and strengthens the partnership.

In summary, pricing strategies in franchising networks cannot be applied and presented based solely on numbers — they must be communicated openly to ensure that the guidelines are followed with confidence. When franchisors listen to franchisees, a kind of alliance is formed

in the execution of the strategy, contributing to the collective success of the network (Boulay et al., 2020; Wagner et al., 2018).

2.7. TECHNOLOGIES AND MACHINE LEARNING IN PRICE OPTIMISATION

Recent technological developments are transforming the retail market, including pricing strategies in convenience franchising networks. Data analysis systems, artificial intelligence, and machine learning make it possible to understand consumer behaviour in a more automated way and to respond more quickly and accurately. In this context, algorithms are an excellent tool for price optimisation based on variables such as sales or margins.

The most common application of machine learning in pricing is through the optimisation of prices and promotions based on historical sales data (pricing analytics). Algorithms are capable of analysing thousands of transactions, estimating price elasticities, segmenting stores, or identifying seasonal periods. This information allows the franchisor to create much more accurate and locally tailored price recommendations (Adams et al., 2024).

Another interesting application of machine learning is in dynamic pricing. For example, by integrating data such as weather conditions or competitor prices, it becomes possible to make real-time adjustments. REMA 1000 uses a platform called Revionics, which gives the supermarket chain access to AI-based solutions that help reduce waste and improve margins (Hamstra, 2024). The adoption of omnichannel dynamic pricing strategies can increase margins and revenue by 2 to 5%, thanks to fewer stockouts and better alignment with demand (Bondi et al., 2021).

In summary, machine learning technologies in price optimisation are becoming increasingly decisive. They allow for data-driven, real-time decisions that, when combined with elements of EDLP or Hi-Lo strategies, can be truly transformative. In this thesis, which is focused on the chain in study, the proposed strategy involves using these tools to analyse product behaviour and, together with managerial experience, intelligently classify products — leading to profitable and effective pricing decisions and adjustments.

However, despite being exciting, the application of these tools comes with practical challenges. Advanced intelligence systems require a high level of central maintenance to manage pricing, since factors like location or customer context can vary greatly across a network with many stores. In addition to this, it's also important to point out that managing pricing store by store may not only be impractical in terms of maintenance, but also in terms of being able to justify each decision individually to every franchisee. Furthermore, considering the complexity of this topic, it is very important to provide pedagogical support and show clarity and respect towards franchisees, so that their autonomy remains within

reasonable limits.

2.8. CONCLUSION

Finally, from this literature we can conclude that convenience stores charge higher prices in exchange for ease of access and immediacy. This convenience premium is something consumers are willing to pay, as long as it is not excessive. By combining analytical precision, local market analysis, and effective communication, it is possible to build a successful pricing strategy.

In addition to these factors, franchising requires even greater care, as the brand must remain consistent while allowing for operational autonomy, so that the business remains advantageous. Machine learning tools that allow for better pricing strategies also represent a potential success factor.

However, applying this in a decentralized network like the retail chain we are focusing on, can be a highly complex challenge — from implementation and consistency to transparency. Therefore, the key to a realistic and concise strategy lies in developing a simplified, data-driven model that builds recommended retail prices in a strategic way, with practical applicability for store managers.

In the next section, the construction of this recommended pricing strategy will be developed through a methodology that is both robust and applicable.

3. METHODOLOGY

In this chapter, the methodology which follows the CRISP-DM framework (Cross-Industry Standard Process for Data Mining), will be presented. This data science guidance goes through the following stages: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. Here, we will follow the steps of this framework to classify each item as Essential, Medium, or Premium through descriptive data analysis and Machine Learning. In the end, a glossary with the variables and their meanings will be presented.

3.1 BUSINESS UNDERSTANDING

The methodology starts with the need to understand the principles of grocery retail franchising and what the objectives are in this type of business when it comes to pricing and product assortment. The main challenge in this area is for franchisees to find a way to balance competitiveness and profitability, meaning practicing attractive prices on the most visible or important products and generating profit on less price-sensitive items. One approach to solve this issue is to identify which products are "essential", meaning those with higher turnover and higher sensitivity, and which are premium, which are products with lower turnover and that tend to have greater freedom in margin adjustments. This concept aligns with the strategy of concentrating investment in the most visible items to the customer while compensating through others (Heinrich et al., 2016).

Also known as Key Value Items (KVIs), this subset of essential products is crucial in influencing and shaping customer price perception of a store (Heinrich et al., 2016). As these items have high demand, it is fundamental for the business to focus on being competitive in these, using them as bait products with low margin. Studies show that reducing the price of these items creates the idea that the store is not expensive and attracts more customers, which in turn leads to the sale of other items beyond the KVIs (Heinrich, 2016). This supports the success of retailers adopting high-volume, low-margin strategies on essential products—a strategy proven effective by giants such as Walmart (Ross, 2024). On the other hand, premium products, not being part of the KVI group, can sustain higher margins due to their lower price sensitivity (Heinrich et al., 2016).

In this thesis, the indicators used to identify the product types were turnover and price competitiveness. Turnover can be represented as the quantity sold (excluding promotions), as this reflects how important and frequent the item is in customers' shopping baskets. Competitiveness was associated with the margin practiced by franchisees outside promotional periods. A low margin may indicate that the product is competitive or has more

aggressive pricing, whereas a product with higher margins may be less price sensitive. In the data, these two variables are "Qt Regular" and "Margem % Regular". By combining these two variables and classifying products into "low margin & high volume" vs "high margin & low volume", it becomes possible to distinguish between different product groups.

Still in this phase, objectives were defined to better classify the products: (i) to find an algorithm or mathematical formula that combines quantity and margin of the products, (ii) to segment the products into three categories based on the classification (Essential, Medium, Premium), and (iii) to ensure that products were correctly classified not only based on metrics but also according to their intuitive role in the business.

The current recommended pricing strategy consists of adding a fixed increment to the non-promotional price of the market price ($PVPR = \text{market price} + X\%$). However, when analysing the prices practiced in 2024, it became clear that stores are not following this recommendation, with franchisees selling products at prices that deviate significantly from the recommended price. These behaviours reveal a potential inconsistency in the brand and the inefficacy of the current strategy. Therefore, the new strategy aims to address this issue by providing adjusted and justifiable price recommendations that bring logic and clarity to support decision-making.

Finally, the goal of these classifications is to differentiate the pricing recommendations for the retail chain's franchisees based on brand type (private label or supplier) and product type (Essential, Medium, or Premium), by assigning a specific increment over the base market price. The marketing suggested retail price (MSRP) will thus be defined. According to the literature, supplier-branded products that are Essential will have lower or no price increments, while private label products with lower visibility may be granted more flexibility in their recommended price adjustment. This implementation will be carried out in the final stage of the methodology.

3.2 DATA UNDERSTANDING AND EMPIRICAL STUDY

With the business objectives in mind, the process moved on to understanding the historical data on product sales and margins.

At the beginning of this stage, several strategies were tested to classify and group the products to find the best approach. One of them tested the use of net sales (regular price \times regular quantity sold). This metric, although it directly represents the economic impact on the business, did not consider high-rotation, low-price items such as bottled water or bread rolls when classifying essentials. Therefore, the option was to use regular quantity, which much better represents high turnover of products.

In addition, a fixed percentile (terciles) strategy was also tested, in which items were classified

according to their relative position in terms of quantity and margin. This approach had the issue of being quite arbitrary and unstable, as it classifies through hard cutoffs in the data distribution.

In the end, it was concluded that the best approach would be to work at the level of the base unit, grouping products with similar commercial characteristics (e.g. salted plain potato chips with private label), and using a score based on regular quantity (log-transformed and normalized) and margin (normalized and inverted), followed by support and validation through some Machine Learning algorithms. This score is managed using thresholds aligned with strategic goals, ensuring that essential products represent >50% of total sales. This last goal was inspired both by ABC analysis and by the practices of other players such as Walmart, where key items account for most of the volume sold (Jenkins, 2023).

For example, in a typical ABC distribution, class A SKUs account for approximately 20% of the items and 80% of total value (Jenkins, 2023). Similarly, it was defined that Essential products should encompass at least half of the units sold. Next, we detail the technical steps taken to achieve the outlined objectives by preparing and understanding the available data.

The dataset contained 2,104 base units, each with metrics such as net sales, quantities sold, various types of margins, promotional metrics, and descriptive variables. For the analysis, only regular metrics were considered — that is, outside of the promotional scope — to analyze product behavior outside campaign periods. Accordingly, the variables used were Qt Regular (quantity sold in units during the period, excluding promotions) and Margem % Regular (gross percentage margin obtained on regular sales, i.e., excluding promotional discounts).

During the exploratory data analysis, it was observed that there was a very uneven distribution in terms of product sales — that is, a few products sell a lot, while the majority sell very little. This confirms the expected Pareto behavior in retail. Products like “pão bijou” (small bread rolls) and bottled water sold more than twice as much as the next best-selling items.

Regarding regular margin percentage, the behavior varied significantly across different base units, with some items showing near zero margins (likely products sold at cost), and others reaching margins as high as 40%.

In general, it was observed that there was an inverse correlation between margin and quantity sold, meaning the products with higher sales volume tended to have lower margins — which indicated that franchisees were indeed applying this behavior in their stores.

This relationship helped reinforce the use of both variables simultaneously and inversely to classify the base units: Essentials would be products with high volume and low margin, while Premium products would have low volume and high margin, and Medium would be the remaining ones.

Research in product portfolio management supports this kind of segmentation using two dimensions. For instance, García-Vidal et al. (2023) propose a contribution-margin-quantity matrix (CMQ matrix) to classify products in small and medium-sized enterprises. This matrix identifies quadrants that resemble those used here. In pricing terminology, so-called "traffic drivers" or KVIs refer to products with large volumes and low prices (i.e., low margins). On the other hand, profit drivers are niche products with low volume and high margins (Heinrich et al., 2016). The terminology differs, but the logic is quite like the Essencial vs. Premium classification adopted in this study.

The quality and quantity of data available for the next stages were also reviewed. For each unit base, information on regular sales and margins was present — although some values were very low — and no missing data was found. Therefore, it was concluded that the data was appropriate to move forward, although minor adjustments would be required to better fit the upcoming modelling stages.

3.3 DATA PREPARATION

To prepare the data for classifying the products, it was necessary to transform and create a few new variables. The main steps to make the data ready for modelling were as follows:

1. Filtering Relevant Records: To conduct an analysis focused on the fields relevant to our objective, we discarded variables related to promotional metrics, absolute margins, and descriptive attributes. It was also ensured that there were no missing values and that only a single unit base per row was included per brand type.

2. Scale Transformation (Logarithm): To reduce the asymmetry in the distribution of sales between products, a base-10 logarithm was applied to the regular quantity sold per product.

This was essential, since a product might sell one or two units per year while another might exceed 100,000 units. By applying the log, we made comparisons between items more balanced despite their different magnitudes (for example, a product with Qt Regular = 100,000 would have Qt_log ≈ 5, while another with Qt Regular = 10 would have Qt_log = 1 — reducing the gap from 99,990 to 4 in logarithmic units).

3. Min-Max Normalization: Since it was necessary to generate a composite score, both Qt_log and Margem % Regular were normalized to a common interval [0,1] using the standard min-max formula:

$$x' = (x - \min(x)) / (\max(x) - \min(x))$$

This produced two dimensionless variables: qt_score (representing the normalized regular

quantity sold for each unit base) and `margem_score_old` (representing the normalized regular margin). However, since our aim was to highlight lower margins as more competitive (and therefore more likely to be essential), we inverted the margin score:

$$\text{margem_score} = 1 - \text{margem_score_old}$$

This way, a lower regular margin percentage would result in a higher score — reinforcing the idea that these products are more price-sensitive and likely to be "Essencial".

4. Combined Score Calculation: The two scores (`qt_score` and `margem_score`) were combined into a single `score_total` for each unit base. Since quantity is considered a stronger indicator of essentiality than margin, a weight of 65% was attributed to quantity and 35% to margin, following the formula:

$$\text{score_total} = 0.65 \times \text{qt_score} + 0.35 \times \text{margem_score}$$

The process of assigning weights followed a clear logic. The business understands that sales rotation (turnover) holds slightly more importance than margin when determining the relevance of a product. In other words, the focus is to identify Essential products. These are the items that must always be available to the customer, as they represent a high volume of sales. Even if a unit base has a slightly higher margin, turnover carries more weight in this assessment.

Studies recommend similar weightings, emphasizing that indicators such as how often products are purchased and by how many customers should carry significant weight when identifying important products (Heinrich et al., 2016). Although no academic study supports using the exact 65/35 ratio, several different proportions were tested and cross-validated with internal managerial feedback: the 0.65/0.35 ratio produced a ranking of products that aligned well with the business's qualitative expectations (see Section 3.5). The resulting `score_total` remains within a 0 to 1 scale, with 1 representing the most desirable products — i.e., those with the highest volume and lowest margin in the examined group.

Although a score had already been determined to measure how essential a product is, it was still necessary to define thresholds to distinguish between Essential, Medium, and Premium products. The following criteria was adopted:

- Products with a score ≥ 0.70 were classified as Essential;
- Products with a score between 0.55 and 0.70 were classified as Medium;
- Products with a score ≤ 0.55 were classified as Premium.

The definition of these thresholds was based not only on the statistical distribution of the

scores but also on the business goals previously outlined. The upper threshold of 0.70 was chosen to isolate the products that represented the largest share of total sales, in line with the initial strategic objective. Based on this classification, Essential products accounted for only about 5.6% of the total SKU assortment, but they were responsible for approximately 53% of the total quantity sold — as shown in the following chart. This outcome confirms the intended goal of ensuring that Essential products make up more than half of total sales volume.

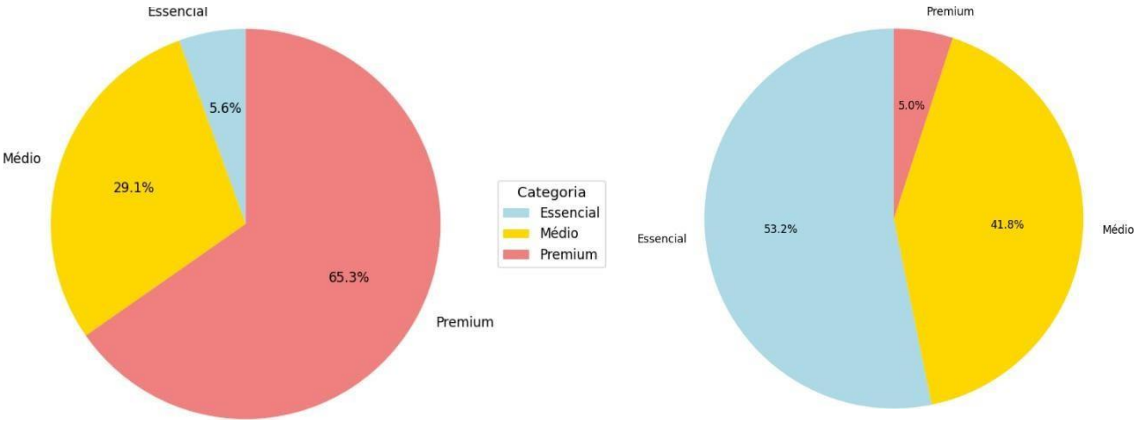


Figure 1 – Pie charts showing the proportion of items in terms of sales volume and number of products

On the other hand, the lower threshold of 0.55 was defined to identify products with very low sales volumes and high margins, typically associated with Premium items. By applying this threshold, the Premium segment was found to encompass approximately 65% of the total SKU assortment, while accounting for only about 5% of the total units sold (see chart above).

The remaining products — those falling within the intermediate range — were classified as Medium, representing items with moderate sales volumes and margins when compared to the rest of the assortment. The scatterplot below illustrates the distribution of products in terms of Quantity vs. Margin and shows a clear separation into three clusters according to the thresholds established. As expected, the way the score was calculated resulted in a well-defined segmentation into Essential, Medium, and Premium products

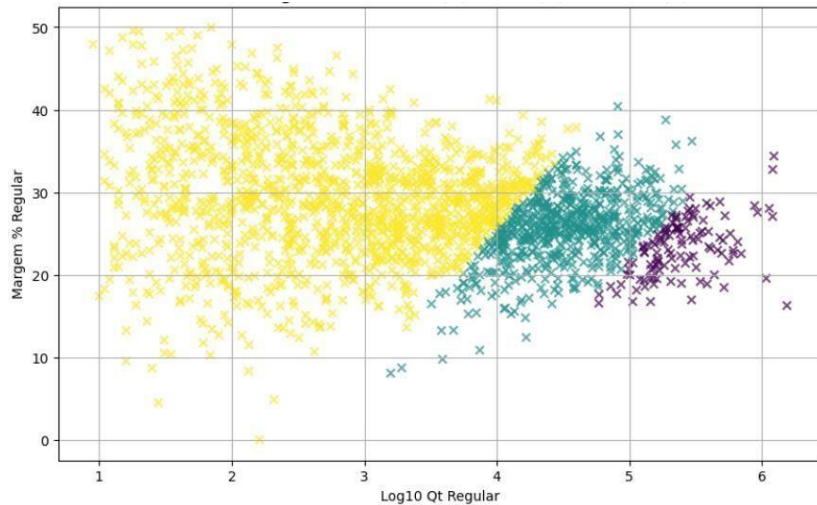


Figure 2 – Distribution of items by regular margin and regular quantity sold (log-transformed)

As observed in the graph above, the Essential products (purple dots) clearly stand out in the low-right quadrant, indicating high quantity and low margin. In contrast, Premium products (yellow) appear dispersed across the left-high side, representing many items with low sales volumes and high margins. The Medium products (green) occupy the transitional zone of the plot, reflecting intermediate performance in both variables.

In summary, each base unit in the dataset was assigned a category label (Essential / Medium / Premium) based on quantitative business rules that are also supported in the literature. Furthermore, since the score-based approach involves some degree of manual tuning and arbitrariness, machine learning algorithms were tested not only to validate the scoring methodology but also to automate the classification process. This is addressed in the next step, the Modelling phase, where both supervised and unsupervised methods are used to classify the products under study.

3.4 MODELING

To develop the models, this step was divided into two complementary parts: first, an unsupervised clustering algorithm, K-Means, was applied to explore and detect natural patterns in the data, grouping the base units accordingly; second, three supervised learning algorithms — K-Nearest Neighbors (KNN), Naive Bayes, and Random Forest — were used to predict the category (Essential, Medium, or Premium) to which each base unit would belong, and to compare which model performed best with the available data.

By applying these algorithms, the goal was to generate alternative methods for segmenting the products, both for validating the proposed classification and for offering different

approaches for maintaining and updating the segmentation over time. Nevertheless, it is worth noting that the score-based method remains the only approach that is easily interpretable and communicable to store owners (franchisees), due to its manual and transparent structure, unlike the algorithmic models.

3.4.1 Unsupervised Segmentation with K-Means

To assess whether the predefined categories aligned with the natural grouping of products in the dataset, the K-Means algorithm was applied. This unsupervised algorithm assigns each observation to a cluster by iteratively minimizing the distance to the nearest centroid.

Since the classification aimed to group products into three categories (Essential, Medium, Premium), the number of clusters was set to $K=3$. The expectation was that by combining the variables Qt_log and Margem % Regular with both weights (65/35), the algorithm would discover similar segmentation. Before clustering, both input variables were standardized (with z-score normalization) to ensure comparability on the same scale. After standardization, the clustering process was finally executed, and the classification output was the following:

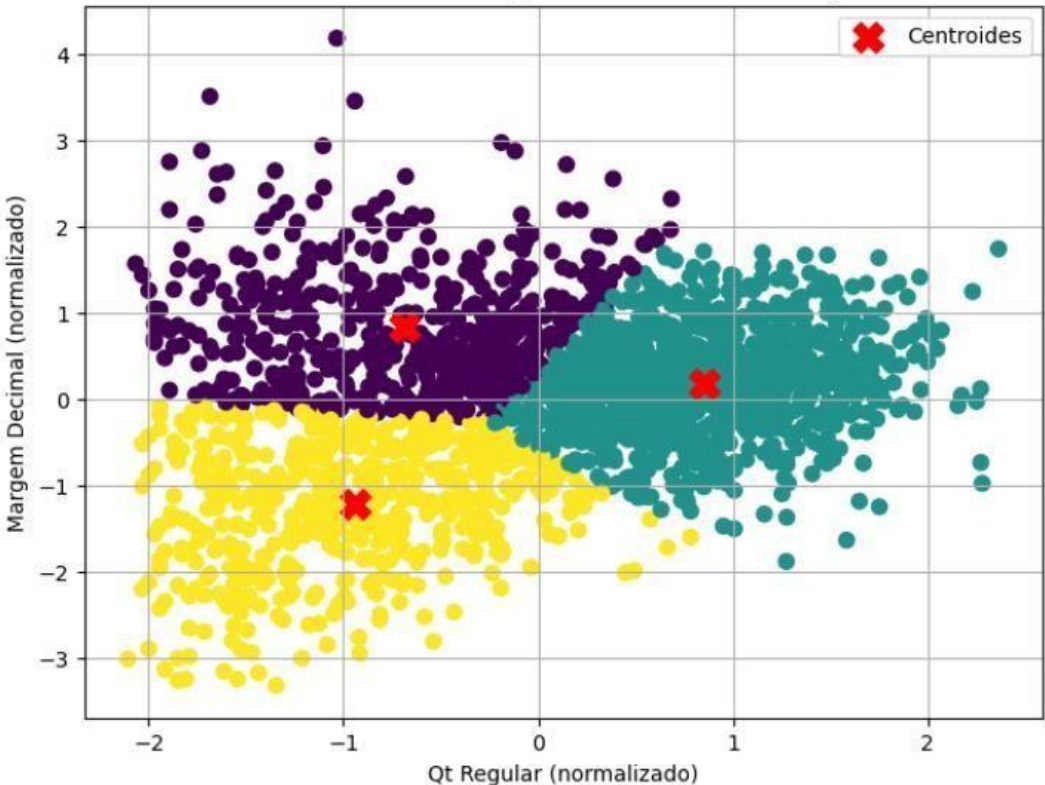


Figure 3 – K-means clustering results illustrating the classification

By comparing the score-based classification with the clusters generated by K-Means, we noticed some alignment but not a perfect match between the segmentation defined by business rules and the natural groupings detected by the algorithm:

- Cluster 1 – Premium: The first cluster included unit bases with low quantities sold and high margins, which could indicate they are Premium products (purple dots). However, many of the items classified here do not match the score-based Premium group, and vice versa.
- Cluster 2 – Essential: This cluster, shown in blue, contains a large number of items classified as Essential, typically products with high quantities and medium-to-low margins. However, it does not capture most of the Essentials, especially those in the group with lower margins.
- Cluster 3 – Medium: This was the most unexpected cluster, as it included products with both low margins and low quantities, suggesting that this K-Means clustering cannot replace the score-based classification — which was already expected, since K-Means is unaware of the manual rules originally defined.

Overall, while the K-Means results support the idea that there are natural groupings in the data, the segmentation does not perfectly match the score-based classification. This highlights that the manual score — being informed by business logic and weighted priorities — remains more interpretable and aligned with strategic goals. The K-Means clustering can be used as a supplementary tool for validation or exception detection rather than a full replacement.

3.4.2 Supervised Classification (K-NN, Naive Bayes and Random Forest)

Just like with K-Means, the supervised classification was carried out so that, based on historical data and the already defined target variable (the score), the algorithms could be trained using sales and margin data to predict the categories: Premium, Essential, and Medium. This step will help prove that algorithms can help in automation instead of using manual rules to classify the products, but also to test with other variables in the future and classify new products.

Three classification algorithms with different natures were chosen for comparison:

- K-Nearest Neighbors (K-NN): This algorithm, based on Euclidean distances, classifies the unit base according to the K closest neighbors. K = 5 was used, meaning the 5 nearest neighbors, and the variables were qt_log and regular % margin, both standardized. K-NN is simple and mainly useful when the distances between points are not too similar, as well as the distances between groups (boundaries).

- Gaussian Naive Bayes: Algorithm that classifies based on probability, assuming a normal distribution and independence between the studied variables. Despite these assumptions, through normalization and transformation to approximate a Gaussian distribution within each class, it was decided to test its performance.
- Random Forest: Model that combines multiple random decision trees and classifies based on the majority vote among them. This model is robust and suitable for variables with different distributions. By training 100 trees with no maximum depth (since the number of features is small), it was expected to effectively capture the segmentation of the products.

The data was then split into training and testing sets using the hold-out method — more precisely, 70% for training and 30% for testing — considering class stratification. That is, the classes were proportionally distributed across both datasets. It was necessary to preserve the target class proportions, because only 6% of the unit bases were classified as Essential and 65% as Premium. Finally, using performance metrics such as precision, recall, f1-score, and accuracy, it was possible to evaluate each of the models.

3.5 Deployment

In this phase, the goal is for the classification system to reflect how the recommended prices will be set. Each base unit will be classified as Medium, Essential, or Premium based on quantity and margin, as described in the previous sections. After this classification, price simulations will be carried out to find the best strategy. The idea is to find the best possible adjustment for each of these three groups and for each brand type (private label and supplier brand).

Supplier brand products are more susceptible to comparisons, as they may exist in competitor stores, meaning there is not much room for percentage increases in the price recommendations. On the other hand, in the case of private label, flexibility is greater not only due to lower competition but also because of the value the banner's brand holds for customers.

These simulations aim to find the ideal increment for each category based on the price of the market price (for example, +5%, -3%). The model will propose a new recommended retail price, to build a data-based strategy that is both more realistic in the market and more robust.

The next step is based on the results and evaluation of the model, to then assess the viability of implementing this data-based strategy. The final goal is to have a more realistic and optimised recommended pricing strategy for the franchisees.

4. RESULTS AND DISCUSSION

With the proposed strategy, franchisees will receive a more realistic and structured recommendation. Instead of recommending a fixed markup across the entire range based only on brand type (private label or supplier), the new strategy is based on the classification of product type (essential, medium, and premium), in addition to brand type. This classification, which was based on turnover and profit generated per sale, allows for better alignment with consumer expectations and franchisee needs.

What happens in the old strategy is that franchisees are not applying the recommended prices to their products. This results in a price dispersion for the same product across all stores. This can compromise consistency and customer perception.

The new strategy, on the other hand, classifies items based on a score that weights 65% on the quantities sold and 35% on the profit margins they generate for the franchisee. These two variables allow alignment with business objectives. This classification defined an adjustment increment to the recommended prices compared to the market price benchmark price with the following criteria:

- **Essential** items, characterized by high sales volume and low margin, will have more conservative, minimal, or even negative adjustments, if necessary, since competitiveness plays a crucial role in this segment;
- **Premium** items, subject to less strict adjustments, are defined by lower sales volume and higher margins — especially in private label products, where direct price comparisons across different retailers are less common;
- **Medium** items, representing the remaining products in the assortment, exhibit intermediate market behavior and will therefore receive proportionate adjustments.

The classification was based on 2024 data, resulting in more realistic and category-specific recommendations which, unlike the previous approach, reflect actual market behavior. The result is a more coherent and robust strategy, better adapted to the needs of franchisees.

After the classification, we analyzed the behavior in 2024 to assess whether adjusting prices for units classified as premium to be higher, and lowering them for those classified as essential, could lead to a balanced scenario that ensured profitability for franchisees without excessively inflating prices. In some cases, price adjustments reached -25% compared to market's prices for certain essential supplier-brand units, while some private-label premium products saw increases of up to 40%. These more radical adjustments were possible due to their lower visibility, allowing to maintain a realistic scenario for the 2025 market.

To test the scalability and robustness of the classification, three supervised machine learning algorithms were applied to support the strategy. The K-Nearest Neighbors (KNN), Gaussian

Naive Bayes, and Random Forest models were trained using the score as the target variable, and the results were as follows:

KNN and Random Forest achieved an almost perfect accuracy of 97%, with equally strong recall across all classes. This demonstrates that both models are effective in replicating the score-based classification and are suitable for automation. Naive Bayes reached an overall accuracy of 91%, but only 70% recall for items classified as Essential. This slightly lower performance can be explained by the limited number of observations in that class.

It is important to highlight that the target variable for all models was the classification derived from the score. Therefore, the algorithms are not meant to replace the business logic, but rather to replicate it efficiently and enable future evolution. The advantage of using these algorithms includes several potential benefits, such as:

- Automating the classification process, especially for new or updated products;
- Scaling the strategy to hundreds of thousands of SKUs without manual intervention;
- Including additional variables that can contribute to greater accuracy and adaptability over time, such as seasonality or store type.

In conclusion, the machine learning algorithms serve as a mechanism to validate and ensure the scalability of the score-based model. By combining business logic with data science, it was possible to achieve an approach that is both interpretable and flexible — the latter being essential, as in retail, pricing influences both customer perception and store profitability, within a constantly changing and highly competitive market environment.

5. CONCLUSIONS AND FUTURE RESEARCH

This thesis developed a data driven pricing strategy for the retail chain's franchise network by classifying products into Essential, Medium, and Premium categories, based on sales volumes and profit margins. The classification was built using a score derived from a business perspective weighting of these two variables and was later validated using machine learning models, including KNN, Naive Bayes and Random Forest.

The current strategy applies a single, uniform adjustment across all products in the range ($PVPR = \text{market price} + X\%$), while the new approach offers differentiated recommendations based on each product's behavior in the market, whether in terms of consumer price sensitivity or margin generation for the franchisee. Moreover, the only factor influencing the adjustment was the type of brand (private label or supplier) which is insufficient because sales data reveal a significant gap between the recommended price and the actual price applied in stores. By introducing a classification based on product category and brand type, with different suggested increments depending on the outcome, the new model creates a more realistic and effective pricing logic that reflects the market and supports franchisees.

In addition to segmentation, it is important to clarify whether machine learning algorithms add value compared to the arbitrarily defined thresholds. While the score may be easier to explain to franchisees and provide greater clarity within the company, predictive models offer several advantages with incomparable potential:

- Scalability and automation: Instead of relying on arbitrarily defined rules based on fixed thresholds, algorithms can classify new products or changes in the assortment without needing to update the rules from scratch.
- Robustness and applicability: Variables such as seasonality, packaging type, or store location can enhance the models, allowing them to more accurately reflect real market dynamics.
- Decision support: The models can predict future patterns by incorporating additional variables like location or weather, helping to inform and anticipate decision-making.

In a decentralized retail franchise network, whether due to operational differences or local market conditions, data science enables assured consistency while allowing for flexible adaptation.

In terms of future research, there are several important topics that can further deepen the subject. Seasonality, consumer profile data, or sensitivity to promotions could help improve the model's accuracy. Moreover, although it is challenging to ensure maintenance and consistency for franchisees, a store-level strategy could allow for fully adapted recommendations to local realities and build a strategy that is both competitive and profitable for the franchisee. Finally, the development of dashboards or a recommendation system that

frequently generates new price adjustments according to the time of year or stock quantity could help the franchisee manage sales in a more fluent and harmonious way.

Business knowledge, combined with data science and the realities of retail, allows this thesis to be a robust foundation for developing smarter pricing strategies that can maintain brand consistency, competitiveness, and profitability in a convenience store franchising network.

BIBLIOGRAPHICAL REFERENCES

Adams, C., Alldredge, K., & Kohli, S. (2024, June 10). *Nine key consumer trends in 2024 | Colombia*. McKinsey & Company. <https://www.mckinsey.com/co/our-insights/state-of-consumer>

Benoit, S., Kienzler, M., & Kowalkowski, C. (2020). Intuitive pricing by independent store managers: Challenging beliefs and practices. *Journal of Business Research*, 115, 70–84. <https://doi.org/10.1016/j.jbusres.2020.04.027>

Blair, R. D., & Lafontaine, F. (2005). *The economics of franchising* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511753879>

Bondi, S., Goldrick, M., Wilkie, J., Sen, B., & Reasor, E. (2021, April 28). The dos and don'ts of dynamic pricing in retail. *McKinsey & Company*. <https://www.mckinsey.com/our-insights/the-dos-and-donts-of-dynamic-pricing-in-retail>

Boulay, J., Caemmerer, B., Evanschitzky, H., & Duniach, K. (2020). Multi-unit franchising from franchisor and franchisee perspectives: Antecedents, performance outcomes, and the optimal mini-chain size. *Journal of Business Research*, 113, 49–58. <https://doi.org/10.1016/j.jbusres.2020.02.029>

Buchan, J. (2013). Franchisees as consumers: Benchmarks, perspectives and consequences (Cap. 4, pp. 88–90). Springer. <https://books.google.com/books?id=qSGL-rnLAFEC>

Butler, S. (2024, September 21). Supermarket convenience stores “charge up to 21% more than their larger branches.” *The Guardian*. <https://www.theguardian.com/business/2024/sep/21/supermarket-convenience-stores-charge-up-to-21-more-than-their-larger-branches>

Cardus, S. (2013, May 15). Dynamic pricing and the death of customer loyalty. *Forbes*. <https://www.forbes.com/sites/susanadams/2013/05/15/dynamic-pricing-and-the-death-of-customer-loyalty>

Euzebio, L. D. C., Koch, A. M., & Gasparetto, V. (2024). Práticas de precificação e desempenho relacional em franquias: Estudo no setor de alimentos. *Contabilidade, Gestão e Governança*, 27(2), 221–251. <https://doi.org/10.51341/cgg.v27i2.3233>

European Commission. (2010). EUR-Lex - 52010XC0519(04) - EN - EUR-Lex. [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52010XC0519\(04\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52010XC0519(04))

Fisher, J. (2024, July 27). Food delivery fees are rising, and everyone's feeling the pinch. CNBC. <https://www.cnbc.com/2024/07/27/food-delivery-fees-are-rising.html>

FMS Franchise. (2023). How do you value a franchise business model? <https://www.fmsfranchise.com/how-do-you-value-a-franchise-business-model/>

García-Vidal, G., Sánchez-Rodríguez, A., & Guzmán-Vilar, D. (2023). Exploring MSME owners' expectations of data-driven approaches to business process management. *Systems*, 13(4), 265. <https://www.mdpi.com/2079-8954/13/4/265>

Hamstra, M. (2024, March 7). Is dynamic pricing feasible within supermarkets? *Supermarketnews.com*. <https://www.supermarketnews.com/grocery-technology/is-dynamic-pricing-feasible-within-supermarkets->

Heinrich, B., & Hristova, D. (2016). A quantitative approach for modelling the influence of currency of information on decision-making under uncertainty. *Journal of Decision Systems*, 25(Sup1), 436–445. <https://doi.org/10.1080/12460125.2015.1080494>

Jenkins, A. (2023, September 6). ABC inventory analysis & management. *Oracle NetSuite*. <https://www.netsuite.com/portal/resource/articles/inventory-management/abc-inventory-analysis.shtml>

Kumar, N., Scheer, L. K., & Steenkamp, J.-B. E. M. (1995). The effects of supplier fairness on vulnerable resellers. *Journal of Marketing Research*, 32(1), 54–65. <https://doi.org/10.1177/002224379503200107>

Nyadzayo, M. W., Matanda, M. J., & Ewing, M. T. (2016). Franchisee-based brand equity: The role of brand relationship quality and brand citizenship behavior. *Industrial Marketing Management*, 52, 163–174. <https://doi.org/10.1016/j.indmarman.2015.07.008>

Ortmeyer, G., Quelch, J. A., & Salmon, W. J. (1991). Restoring credibility to retail pricing. *Sloan Management Review*, 33(1), 55–66.

Perrigot, R., Basset, G., & Meiseberg, B. (2016). Resale prices in franchising: Insights from franchisee perspectives. *Journal of Product & Brand Management*, 25(7), 663–675. <https://doi.org/10.1108/jpbm-03-2015-0831>

Raymond, S. (2024, December 5). Fast-food delivery comes with an extreme markup: Here's who charges the most. *Audacy.com*. <https://www.audacy.com/kearth101/news/fast-food-delivery-extreme-markup-who-charges-the-most>

Rosado-Serrano, A., Paul, J., & Dikova, D. (2018). International franchising: A literature review and research agenda. *Journal of Business Research*, 85(1), 238–257. <https://doi.org/10.1016/j.jbusres.2017.12.049>

Sen, S. S., Alexandrov, A., Jha, S., McDowell, W. C., & Babakus, E. (2022). Convenient = competitive? How brick-and-mortar retailers can cope with online competition. *Review of Managerial Science*, 17. <https://doi.org/10.1007/s11846-022-00566-0>

Simon, H., & Fassnacht, M. (2019). *Price management*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-99456-7>

Sullivan, C., & Revionics. (2017, August 2). Is dynamic pricing the silver bullet for C-stores? *Convenience Store News*. <https://csnews.com/dynamic-pricing-silver-bullet-c-stores>

Wall Street Prep. (2024, July 18). Value-based pricing. *Wall Street Prep*. <https://www.wallstreetprep.com/knowledge/value-based-pricing/>

Wagner, G., Schramm-Klein, H., & Steinmann, S. (2018). Online retailing across e-channels and e-channel touchpoints: Empirical studies of consumer behavior in the multichannel e-commerce environment. *Journal of Business Research*, 107, 256–270. <https://doi.org/10.1016/j.jbusres.2018.10.048>



NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa