

ARTIFICIAL INTELLIGENCE AND STRATEGIC SEGMENTATION OF FRANCHISES: ANALYZING BUSINESS MODELS IN BRAZIL USING K-MEANS

Autoria

Wendel Marcos dos Santos - unip.wendel@gmail.com

Programa de Pós-Graduação em Administração / UNIP - Universidade Paulista

Pedro Lucas de Resende Melo - pedro.melo@docente.unip.br

Programa de Pós-Graduação em Administração / UNIP - Universidade Paulista

FEA - Departamento de Administração / PUC-SP - Pontifícia Universidade Católica de São Paulo

Daielly Melina Nassif Mantovani - daielly@usp.br

Mestr Prof em Empreendedorismo - MPE/FEA / USP - Universidade de São Paulo

Prog de Pós-Grad em Admin/Faculdade de Economia, Admin e Contab – PPGA/FEA / USP - Universidade de São Paulo

Agradecimentos

"This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001"

Resumo

In Brazil, franchising is increasingly becoming a viable avenue for business expansion. However, there still remains considerable inconsistency among franchise types in terms of value generation. This research applies unsupervised machine learning techniques to classify Brazilian franchises into different categories using financial, operational, and structural metrics. After defining an appropriate number of clusters with the Elbow Method, the K-Means clustering algorithm revealed four distinct franchise profiles: 1) Premium – high investment franchises with strong support systems; 2) Low-Cost and Fast-Return – easily accessible but volatile franchises; 3) Medium-Sized Structured – franchises moderately balancing investment, growth, and support; and 4) Emerging and Experimental – innovative but more uncertain franchises. Not only the size of investment, but also the caliber of the franchise's support, the operational maturity, and the sustained value-backed growth influence the value created. This study demonstrates the application of unsupervised machine learning for strategic segmentation while enabling franchisors and investors to enhance data-driven decisions. In franchise academic literature, this study is a contribution in offering a classification framework that assists dynamically changing markets by logically justifying the business models employed within.

ARTIFICIAL INTELLIGENCE AND STRATEGIC SEGMENTATION OF FRANCHISES: ANALYZING BUSINESS MODELS IN BRAZIL USING K-MEANS

Abstract

In Brazil, franchising is increasingly becoming a viable avenue for business expansion. However, there still remains considerable inconsistency among franchise types in terms of value generation. This research applies unsupervised machine learning techniques to classify Brazilian franchises into different categories using financial, operational, and structural metrics. After defining an appropriate number of clusters with the Elbow Method, the K-Means clustering algorithm revealed four distinct franchise profiles: 1) Premium – high investment franchises with strong support systems; 2) Low-Cost and Fast-Return – easily accessible but volatile franchises; 3) Medium-Sized Structured – franchises moderately balancing investment, growth, and support; and 4) Emerging and Experimental – innovative but more uncertain franchises. Not only the size of investment, but also the caliber of the franchise's support, the operational maturity, and the sustained value-backed growth influence the value created. This study demonstrates the application of unsupervised machine learning for strategic segmentation while enabling franchisors and investors to enhance data-driven decisions. In franchise academic literature, this study is a contribution in offering a classification framework that assists dynamically changing markets by logically justifying the business models employed within.

Keywords: Franchises; Value creation; Clustering; K-Means; Investment.

1. Introduction

The franchise sector has established itself as one of the most dynamic and strategic areas in the global business environment. This business model allows the replication of successful operations, reducing risks for new entrepreneurs by offering continuous support, brand recognition, and standardized processes (Alon, Apriliyanti & Parodi, 2020; Dant & Grünhagen, 2014). The scalability and flexibility of franchising make it a key driver of economic development and job creation (Varotto & Aureliano-Silva, 2017).

In Brazil, this market has shown significant growth. According to the Brazilian Franchising Association (ABF, 2024), the sector recorded revenues of R\$ 273.1 billion, a 13.5% increase compared to the previous year. This growth reflects the increasing adoption of the franchise model by entrepreneurs seeking security and predictability in a frequently unstable economic environment (De Faria Olivo et al., 2020).

However, despite the growth and consolidation of franchising, not all franchises generate the same level of value for franchisors and franchisees (Hua et al., 2017; De Faria Olivo et al., 2020). Some require high investments and present slower financial returns, while others have leaner operations and faster returns but may face structural challenges (Frazer et al., 2020). The central question, therefore, is: *How can franchise segmentation through clustering reveal structural patterns that influence financial return, franchisee satisfaction, and business model sustainability?*

The decision to invest in a franchise involves multiple variables, including initial cost, royalties, return rate, support offered, franchise chain quality, and franchisee satisfaction (De Faria Olivo et al., 2020; Frazer et al., 2020; Melo et al., 2022). Many investors assume that higher-value franchises with higher franchise and royalty fees — provide greater security and a more guaranteed financial return (Lucia-Palacios et al., 2014; Hua et al., 2017). However, the

relationship between investment level and actual return remains an underexplored empirical topic (Lawrence, Zhang, Hsu & Zheng, 2021; Sun & Lee, 2023).

Given this scenario, this study aims to fill this gap by investigating how different types of franchises create value for franchisees and customers. To this end, the study uses the unsupervised machine learning technique *K-Means*, one of the most widely applied methods for data clustering across various fields, including business analysis and market segmentation (Sinaga & Yang, 2020; Antonio et al., 2021). *K-Means* groups elements with similar characteristics into distinct clusters, enabling the identification of structural patterns that influence franchise performance (Gustriansyah, Suhandi & Antony, 2020).

Furthermore, the study has managerial relevance in supporting strategic decisions, helping investors select franchises aligned with their financial and operational goals (Smith & Colgate, 2007; Keiningham et al., 2020). Thus, it aims to contribute not only to the academic literature on franchising but also to evidence-based decision-making in the sector (Matarazzo et al., 2021).

The main objective of this study is to analyze how different franchise models create value for franchisees and customers, considering factors such as initial investment, return period, franchisor support, and overall satisfaction. Specifically, the study aims to: i) Identify patterns and franchise groupings using clustering analysis based on financial, operational, and structural variables (Maciejewski & Krowicki, 2022); ii) Investigate the relationship between investment and return, analyzing whether higher-cost franchises are more profitable in the long run (Toscher, 2021); iii) Determine the most influential factors in value creation, such as franchise chain support, innovation, management quality, and franchisee satisfaction (Zhang & Zheng, 2021); iv) Provide strategic recommendations for investors and franchisors, supporting decision-making and expansion planning (Gillis, Combs & Yin, 2020).

The findings are expected to enhance the understanding of value creation mechanisms in the franchise sector, offering an empirical basis for more assertive decisions in franchise selection and strategy development (Keiningham et al., 2020; Toscher, 2021; Zhang & Zheng, 2021). Additionally, the use of machine learning algorithms like *K-Means* represents a robust methodological strategy for extracting structural patterns in complex organizational contexts such as franchise chains, enabling data-driven analysis with greater explanatory potential (Sinaga & Yang, 2020; Umargono, Suseno & Gunawan, 2020; Antonio et al., 2021; Ong et al., 2023).

2. Theoretical Framework

The theoretical framework of this study explores the main concepts related to value creation in the franchise sector, considering strategic approaches, innovation, and the impact of organizational franchise networks. In addition, the application of Artificial Intelligence techniques, such as clustering, is discussed for the segmentation and performance analysis of franchises in the Brazilian market.

2.1. Value Creation

Value creation is one of the central pillars of strategic management and economics, occurring when a company can offer products or services that meet customer needs more efficiently or innovatively than its competitors (Smith & Colgate, 2007; Keiningham et al., 2020; Zhang & Zheng, 2021). Value creation within an organization is not limited to operational efficiency and financial gains but involves a complex network of interactions that generate competitive advantage. Gillis, Combs, and Yin (2020) highlight that each activity carried out

within a company can contribute to value aggregation, becoming essential for its differentiation in the market.

However, value creation goes beyond the traditional firm-centered perspective. Recent studies point to the co-creation of value, a concept that emphasizes the active participation of customers in the development and improvement of products and services. In this sense, Vargo and Lusch (2017) argue that value is generated in collaboration between the company and the consumer, turning the customer into an essential agent in the construction of the value proposition (Toscher, 2021; Yen, Teng & Tzeng, 2020). In the franchise sector, this is even more evident, as franchisees play a fundamental role in delivering the value proposition to the end consumer (Ghantous & Alnawas, 2021). Thus, Table 1 presents the tripartite model of value creation in the context of franchises, highlighting the organizational, operational, and market dimensions.

Table 1 – Tripartite Model of Value Creation for Franchises

Value Dimension	Description	Authors
Value for the Company	Refers to the organization's ability to capture financial value through differentiation and operational efficiency strategies.	Matarazzo et al. (2021); Maciejewski & Krowicki (2022)
Value for the Franchisee	Involves the perception of support, profitability, and return on investment, influencing the decision to join and remain in the franchise network.	Ortega et al. (2016); Melo, Carneiro-da-Cunha & Telles (2022);
Value for the Consumer	Represents the customer's perception of the quality, experience, and reliability of the franchised brand.	An & Han (2020); Becker & Jaakkola (2020)

Source: authors

2.2. The Franchise Market and Its Business Models

The franchise market has consolidated itself as one of the main drivers of global economic growth, offering a business expansion model based on the replication of a successful business format (Dant & Grünhagen, 2014; Lucia-Palacios et al., 2014; Hajdini et al., 2017). Franchises can be classified into different models according to their structure and initial investment, among which the following categories stand out (Table 2). Several authors address the main models adopted in the Brazilian market, allowing a categorization aligned with sectoral practices.

Traditional franchises, characterized by robust operations, continuous support, and consolidated structure, are discussed by authors such as Frazer et al. (2020), Gillis, Combs & Yin (2020), Melo, Carneiro-da-Cunha & Telles (2022), Jell-Ojobor, Alon & Windsperger (2022), and Sashi & Brynildsen (2022).

Microfranchises, with reduced investment and lower administrative complexity, are addressed by Ortega et al. (2016), Melo et al. (2018), Santiago et al. (2020), and Melo et al. (2022). Home-based franchises, operated from the franchisee's home, are highlighted in studies by Hua et al. (2017), Oh, Yoo & Lee (2019), Keiningham et al. (2020), Labrecque, Swani & Stephen (2020), and Matarazzo et al. (2021).

Finally, conversion franchises, which integrate independent businesses into the franchisor's structure, are addressed by Maciejewski & Krowicki (2022) and Burston Webster,

Imam & White (2021), highlighting the challenges and benefits of this modality in the context of franchise chain expansion.

Table 2 – Franchise Models and Their Characteristics

Franchise Type	Description	Authors
Traditional Franchises	Robust model, with high initial investment, continuous support, and consolidated structure. Examples include food chains, retail, and specialized services.	Frazer et al. (2020); Gillis, Combs & Yin (2020); Melo, Carneiro-da-Cunha & Telles (2022); Jell-Ojobor, Alon & Windsperger (2022); Sashi & Brynildsen (2022)
Microfranchises and Home-Based	Reduced initial investment (limit of USD 22,500). Lean model, with lower administrative complexity, aimed at small entrepreneurs. Operated from home, reducing physical space costs. Growth was driven by digitalization and demand for flexible work options.	Ortega et al. (2016); Hua et al. (2017); Melo et al. (2018); Oh, Yoo & Lee (2019); Santiago et al. (2020); Keiningham et al. (2020); Labrecque, Swani & Stephen (2020); Matarazzo et al. (2021); Melo et al. (2022)
Conversion Franchises	Transformation of an independent business into a franchised unit, adding to the franchisor's consolidated brand the entrepreneur's existing operations.	Burston Webster, Imam & White (2021); Maciejewski & Krowicki (2022)

Source: authors

Thus, the franchise market structure reflects the constant need to adapt to changes in consumer behavior and technological evolution (Keiningham et al., 2020; Matarazzo et al., 2021; Burston Webster, Imam & White, 2021). Companies that invest in process digitalization and customer experience personalization strengthen their competitive positioning and increase their value creation potential (Shaykhzade & Alvandi, 2020; Zhang & Zheng, 2021; Liu, Shin & Burns, 2021).

2.3. Clustering and Franchise Analysis

Clustering is an unsupervised machine-learning technique used to group elements with similar characteristics within a dataset (Nielsen, 2022; Naeem, Anam & Ahmed, 2023). In the context of franchises, this approach enables the identification of patterns and business profiles, assisting in market segmentation and strategic decision-making (Xu & Wunsch, 2009; Dant & Grünhagen, 2014). According to Jain (2010), clustering is widely applied across various sectors to classify customers, optimize supply chains, and forecast market trends (Sinaga & Yang, 2020).

Franchise segmentation through clustering enables the categorization of franchise chains based on variables such as initial investment, return rate, franchisee support, and overall satisfaction (Alon, Apriliyanti & Parodi, 2020; De Faria Olivo et al., 2020). This allows for an understanding of which franchise models share similar characteristics and which show different performance levels (Varotto & Aureliano-Silva, 2017; Melo et al., 2018). Thus, clustering is especially useful when there is a large volume of complex data, making traditional analysis limited (Dalmaijer, Nord & Astle, 2022).

Among the most widely used clustering algorithms in the literature, K-Means stands out, which groups elements into K distinct clusters, minimizing the internal variability of each

group (Gustriansyah, Suhandi & Antony, 2020; Antonio et al., 2021). The algorithm starts from the initial choice of K centers (centroids) and iteratively adjusts their positions, aiming to maximize homogeneity within the clusters and heterogeneity between them (MacQueen, 1967; Umargono, Suseno & Gunawan, 2020; Sari, Al-Khowarizmi & Batubara, 2021). In the franchise sector, the application of this method can reveal groupings of franchises with faster financial returns, franchises with greater support to franchisees, and premium franchises that require high investment (Ortega et al., 2016; Ghantous & Alnawas, 2021).

Recent studies show that clustering can be a powerful tool for performance analysis in franchise chains (Ong et al., 2023; Khetarpaul et al., 2024). In addition, by applying grouping techniques, it is possible to identify hidden patterns in the data, allowing for more grounded strategic decisions (Gustriansyah et al., 2020; Dalmaijer et al., 2022; Matarazzo et al., 2021).

Another relevant aspect is the interpretation of the formed clusters (Becker & Jaakkola, 2020; Keiningham et al., 2020). After segmentation, the groups can be analyzed to understand which factors contribute to the success or failure of certain franchise models (Shaykhzade & Alvandi, 2020). For example, franchise chains with high franchise fees but low franchisee support may form distinct clusters from those that offer solid support and fast returns, enabling investors to better choose where to allocate their resources (Gillis, Combs & Yin, 2020; Melo, Carneiro-da-Cunha & Telles, 2022).

Therefore, clustering presents itself as a high-value tool for franchise classification and analysis, helping both franchisors and investors better understand the dynamics of the sector (Font et al., 2021; Sashi & Brynildsen, 2022). In the next chapter, the practical application of clustering will be detailed in Methodology, including the criteria adopted for segmentation and the results obtained in the analysis of the franchises (Burston Webster, Imam & White, 2021).

2.4. Success Indicators in Franchising

Measuring the performance of a franchise is essential to assess its viability and its ability to generate value for franchisees and consumers (Smith & Colgate, 2007; Dant & Grünhagen, 2014). Thus, the success of a business should be analyzed through financial and operational indicators that reflect management efficiency, and the satisfaction of the stakeholders involved (Hajdini et al., 2017; Gilli, Combs & Yin, 2020). In the franchising context, these indicators provide a detailed view of the sustainability of the business model, its attractiveness to new investors, and its expansion potential (Varotto & Aureliano-Silva, 2017; Alon, Apriliyanti & Parodi, 2020).

The main success indicators in the franchise sector include the payback period, which measures the speed at which the franchisee recovers the capital invested in the business (Ortega et al., 2016; Melo et al., 2018). This indicator is essential for investors seeking short- and medium-term profitability (Frazer et al., 2020; Hua et al., 2017). Franchises that provide quick returns tend to attract more investors, but the risk of performance volatility must be considered (De Faria Olivo et al., 2020; Jell-Ojobor, Alon & Windsperger, 2022).

Franchisee satisfaction is a qualitative indicator that reflects franchisees' perceptions regarding the franchisor's support, the operation's profitability, and the efficiency of processes (Lashley & Morrison, 2013; Pallant, Sands & Karpen, 2020). According to Lashley & Morrison (2013), satisfied franchisees tend to be more engaged, maintaining high-quality standards and contributing to the brand's consolidation in the market (Melo, Carneiro-da-Cunha & Telles, 2022; Santiago et al., 2020). Research shows that franchise chains offering frequent training, product innovation, and efficient operational support have higher franchisee retention rates (Maciejewski & Krowicki, 2022; Keiningham et al., 2020).

Franchise chain performance evaluates the operational consistency of the franchise and its ability to maintain quality standards across all units (Hajdini et al., 2017; Burston Webster, Imam & White, 2021). High network performance is associated with factors such as process efficiency, continuous training, standardized service, and constant innovation (Becker & Jaakkola, 2020; Kleinaltenkamp et al., 2022). According to a study by Alonso & Kraus (2021), franchises that manage to maintain a uniform level of quality across their units tend to experience more stable and predictable growth, reducing the rate of business closures (Toscher, 2021; Lang, Xia & Liu, 2020).

In addition to these main indicators, other factors may influence franchise success, such as: a) Average revenue per unit, representing the revenue potential of each franchisee (Oh, Yoo & Lee, 2019; Ghantous & Alnawas, 2021); b) Network growth rate, measuring the franchise's ability to attract new investors and expand its operations (Shaykhzade & Alvandi, 2020; Maciejewski & Krowicki, 2022); c) Brand reputation in the market, which influences consumer trust and the business's competitiveness (An & Han, 2020; Font et al., 2021).

The combination of these indicators with statistical analysis techniques and machine learning, such as K-Means clustering, allows for a deeper understanding of franchise performance and value creation potential (Dalmaijer et al., 2022; Khetarpaul et al., 2024). This approach enables the categorization of franchise chains into distinct groups, identifying which factors most impact franchisee satisfaction and the profitability of the business model (Sinaga & Yang, 2020; Umargono, Suseno & Gunawan, 2020).

Thus, understanding success indicators is essential for franchisors seeking to strengthen their franchise chains, investors looking for safer opportunities, and entrepreneurs evaluating the cost-benefit of entering a franchise (Gustriansyah, Suhandi & Antony, 2020; Matarazzo et al., 2021).

3. Methodology

This study is characterized as quantitative and exploratory research, aimed at analyzing patterns and groupings in the franchise sector through machine learning techniques (Sari et al., 2021; Ong et al., 2023). The quantitative approach is justified using numerical data and financial metrics, allowing the measurement of franchise performance based on objective indicators (Gillis, Combs & Yin, 2020; Matarazzo et al., 2021).

The exploratory nature of the research lies in the objective of identifying hidden patterns in the data, without predefined hypotheses about the relationship between the analyzed factors (Hair et al., 2010; Sinaga & Yang, 2020; Dalmaijer, Nord & Astle, 2022). The clustering technique used allows for the classification of franchises into different groups, offering a segmented view of the market and enabling inferences about the most efficient business models (Gustriansyah, Suhandi & Antony, 2020; Umargono, Suseno & Gunawan, 2020; Antonio et al., 2021).

In this way, the research aims to provide strategic analyses for franchisors and investors, enabling a deeper understanding of how different variables impact value creation in the franchise sector (Melo, Carneiro-da-Cunha & Telles, 2022; Matarazzo et al., 2021; Toscher, 2021).

3.1. Origin of the Variables Selected for Clustering

The variables used in this study were extracted from the publication *Franchise Guide Yearbook* (2024). These data were audited by Serasa Experian, ensuring the reliability of the information used. This publication presents a detailed survey of the Brazilian franchise market,

ranking the franchise chains based on financial, operational, and structural criteria. The selection of these variables aimed to identify value creation patterns and differentiate the various franchise profiles. Table 3 presents the categories analyzed, their respective variables, and descriptions:

Table 3 – Study Variables

Category	Variable	Metric	Description	Authors
Financial Variables	Franchise Fee	BRL (Brazilian Real)	Represents the initial investment required to open a franchise unit.	Ortega et al. (2016); Melo et al. (2018); De Faria Olivo et al. (2020); Frazer et al. (2020); Jell-Ojobor et al. (2022)
	Royalties	% revenue	Indicates the ongoing cost that the franchisee must pay to the franchisor.	
	Advertising Fee	% revenue	Represents the amount allocated to the franchise chain's collective marketing.	
	Average Payback Period	months	Measures how long it takes to recover the initial investment.	
Operational Variables	Network Performance	0–10	Evaluates the franchise chain's efficiency and stability in the market.	Ortega et al. (2016); Frazer et al. (2020); Matarazzo et al. (2021); Melo et al. (2022)
	Franchise Chain Quality	0–10	Measures the uniformity and reliability of operations across all units.	
	Franchisee Satisfaction	0–10	Reflects franchisees' perceptions regarding franchise chain support and profitability.	
Structural Variables	Average Number of Employees per Unit	units	Indicates the scale of operations and business complexity.	Frazer et al. (2020); Santiago et al. (2020)
	Average Unit Area	m ²	Measures the average size of franchise facilities.	

Source: authors

The choice of these variables is justified by their direct impact on value creation in the franchise sector, allowing the identification of groupings of franchise chains that share similar characteristics and understanding of how different factors influence the financial and operational performance of franchises in Brazil (Frazer et al., 2020; Matarazzo et al., 2021; Melo, Carneiro-da-Cunha & Telles, 2022).

3.2. Analytical Techniques

The K-Means clustering method is widely recognized as an unsupervised machine-learning technique within the field of Artificial Intelligence (Sinaga & Yang, 2020). This approach has been used in various studies applied to organizational and market data analysis due to its ability to automatically segment large datasets based on latent characteristics (Gustriansyah, Suhandi & Antony, 2020). In the present study, applying K-Means allowed the

categorization of Brazilian franchises into homogeneous groups based on financial, operational, and structural variables, supporting the identification of strategic patterns among the business models analyzed (Antonio et al., 2021).

Additionally, its use is justified by factors such as computational efficiency, scalability, and ease of interpretation of results (Dalmaijer, Nord & Astle, 2022). These attributes make K-Means particularly suitable for data-driven decision-making contexts, such as franchise chains, where understanding distinct profiles can inform strategies for expansion, support, and profitability (Frazer et al., 2020; Melo, Carneiro-da-Cunha & Telles, 2022; Gillis, Combs & Yin, 2020). The heterogeneity between franchised units demands analytical approaches capable of capturing structural and operational nuances, reinforcing the relevance of clustering techniques in strategic franchising research.

Therefore, the adoption of this technique in this study represents a robust methodological choice aligned with the contemporary demands of analysis in complex and dynamic organizational environments (Khetarpaul et al., 2024; Umargono, Suseno & Gunawan, 2020). The following methodological steps were followed for the application of K-Means (Table 4):

Table 4 – Methodological Steps for Applying K-Means to Franchise Segmentation

Step	Description	References
Data Normalization	As the variables used have different scales, they were standardized to avoid distortions in the analysis. The z-score statistical standardization technique was applied.	Gustriansyah, Suhandi & Antony (2020); Dalmaijer, Nord & Astle (2022)
Determining the Number of Clusters	The ideal number of clusters was determined using the Elbow Method, which evaluates the variation of the sum of squared errors within the groups for different K values.	Umargono, Suseno & Gunawan (2020); Sari, Al-Khowarizmi & Batubara (2021)
K-Means Algorithm Execution	Based on the defined number of clusters, the K-Means algorithm was applied, starting with a random selection of K centroids, which were iteratively adjusted to optimize grouping.	Sinaga & Yang (2020); Antonio et al. (2021); Khetarpaul et al. (2024)
K-Means Algorithm Execution	Based on the defined number of clusters, the K-Means algorithm was applied, starting with a random selection of K centroids, which were iteratively adjusted to optimize grouping.	Sinaga & Yang (2020); Antonio et al. (2021); Khetarpaul et al. (2024)
Cluster Analysis and Interpretation	After segmentation, groups were evaluated based on their average characteristics to identify the key factors influencing franchise performance.	Ong et al. (2023); Dalmaijer, Nord & Astle (2022)

Source: authors

3.3. Analytical Procedures

To analyze the data and apply clustering techniques, statistical and machine learning tools widely recognized in academic literature and business practice were used (Sinaga & Yang, 2020; Umargono, Suseno & Gunawan, 2020; Dalmaijer, Nord & Astle, 2022; Gustriansyah, Suhandi & Antony, 2020).

Data processing and analysis were conducted using Python due to its robustness and extensive library ecosystem for data handling and statistical modeling (Gustriansyah, Suhandi & Antony, 2020; Sinaga & Yang, 2020). For clustering, specialized libraries were used: Pandas for data organization and preprocessing (Dalmaijer, Nord & Astle, 2022); NumPy for

mathematical operations and normalization (Antonio et al., 2021); Scikit-Learn for implementing the K-Means algorithm and determining the optimal number of clusters (Sari, Al-Khowarizmi & Batubara, 2021; Umargono, Suseno & Gunawan, 2020); and Matplotlib/Seaborn for visualizing and interpreting the identified groupings (Ong et al., 2023; Khetarpaul et al., 2024).

In addition to using Python, tools such as Microsoft Excel were used for data verification and structuring, and Jupyter Notebook for developing and documenting the analysis code. The combination of these tools enabled a structured and robust analysis, ensuring statistical accuracy and interpretability of results. The use of machine learning-specific libraries allowed efficient segmentation of franchises, ensuring that the resulting groupings reflected real patterns in the analyzed dataset (Sinaga & Yang, 2020; Umargono, Suseno & Gunawan, 2020; Ong et al., 2023; Khetarpaul et al., 2024).

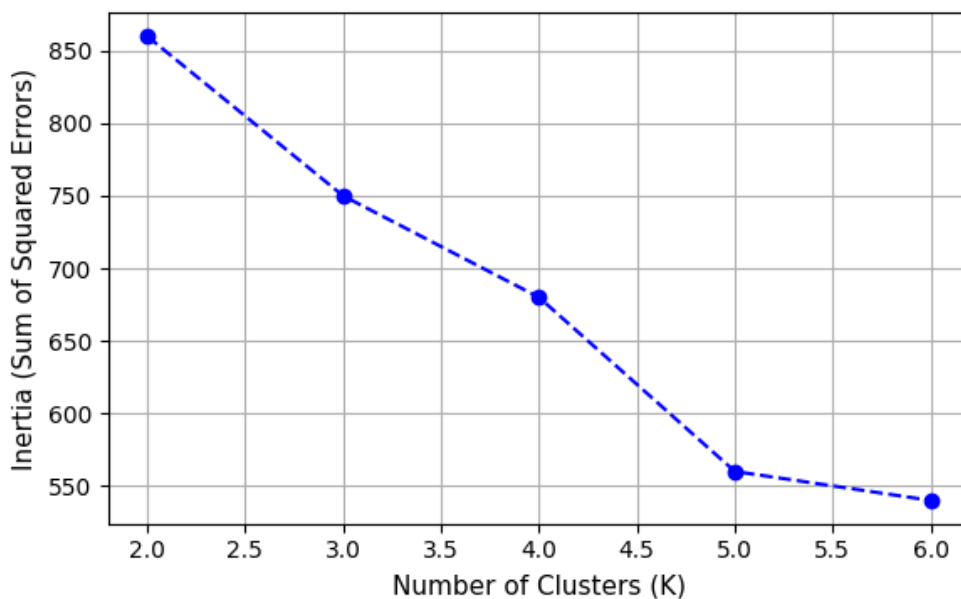
4. Results

The data analysis allowed the identification of distinct patterns in the franchise sector, highlighting how different variables influence value creation for franchisees and customers. This section presents the main findings of the study, the segmentation of franchises into clusters, and the interpretation of the results considering the existing literature.

4.1. Determining the Number of Clusters

To determine the ideal number of clusters in the segmentation of franchises (Figure 1), the Elbow Method was used — a widely recognized approach in the literature for identifying the point at which adding more clusters no longer significantly reduces model inertia (Sinaga & Yang, 2020; Umargono, Suseno & Gunawan, 2020). This method evaluates the variation in the sum of squared errors within the clusters for different values of K, allowing for an optimized choice of the number of groupings (Gustriansyah, Suhandi & Antony, 2020; Dalmaijer, Nord & Astle, 2022).

Figure 1 – Elbow Method for Defining the Number of Clusters



The graph presented in Figure 1 shows the variation in the sum of squared errors within clusters for different values of K. The inflection point of the curve indicates the optimal number of clusters, as it represents the moment when adding new groups no longer results in a significant improvement in segmentation (Sari, Al-Khowarizmi & Batubara, 2021; Antonio et al., 2021). This criterion is widely used in clustering studies to ensure that the grouping is efficient and interpretable without generating unnecessary subdivisions in the data (Umargono, Suseno & Gunawan, 2020). Based on the analysis of the graph, it was identified that four clusters represent the best segmentation for the data analyzed. This value was chosen based on the following criteria (Table 5):

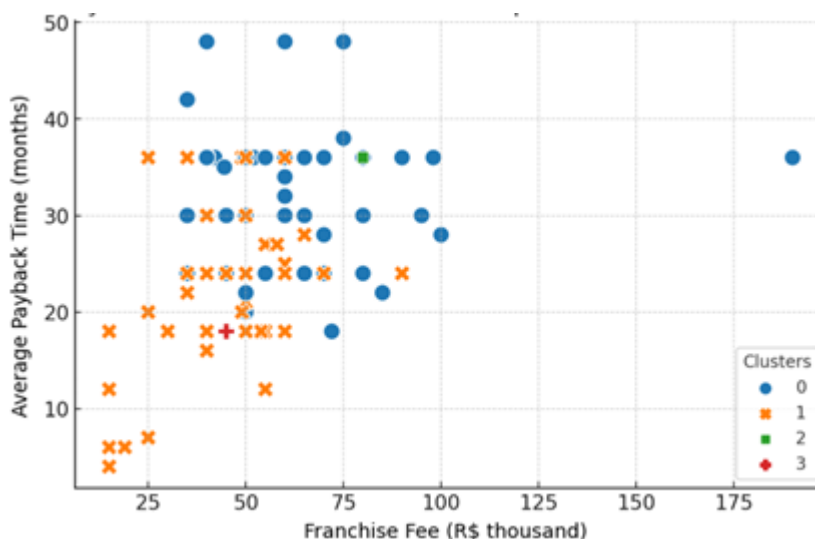
Table 5 – Criteria for Defining the Number of Clusters (K=4)

Criterion	Description	References
Significant reduction in inertia	From K=4 onwards, the decrease in inertia becomes marginal, indicating efficient group separation without excessive subdivision.	Sinaga & Yang (2020); Gustriansyah, Suhandi & Antony (2020)
Consistent silhouette coefficient	The evaluation of the silhouette coefficient confirmed the highest segmentation quality at K=4, ensuring that the clusters are well-defined.	Dalmajjer, Nord & Astle (2022); Sari, Al-Khowarizmi & Batubara (2021)
Cluster interpretability	The division into four groups provided an adequate balance between the granularity and interpretability of the data, allowing clear distinctions.	Antonio et al. (2021); Umargono, Suseno & Gunawan (2020)

Source: authors

The scatter plot (Figure 2) presents the segmentation of the franchises into distinct clusters, using the Franchise Fee (R\$ thousand) on the X-axis and the Average Payback Period (months) on the Y-axis. These axes were chosen because they are key variables in franchise investment decisions, directly reflecting the relationship between the required initial capital and the time needed to recover the investment.

Figure 2 – Cluster Visualization



The segmentation of franchises revealed four distinct profiles, grouped based on financial and operational characteristics. Each cluster represents a group of franchises with

similar patterns of investment, payback period, and support structure, enabling a more detailed analysis of how different business models create value.

Table 6 describes the main characteristics of each cluster, highlighting the differences among premium, low-cost, fast-return, mid-sized structured, and emerging/experimental franchises. This categorization provides a strategic view for investors and franchisors, supporting decision-making and growth strategy formulation.

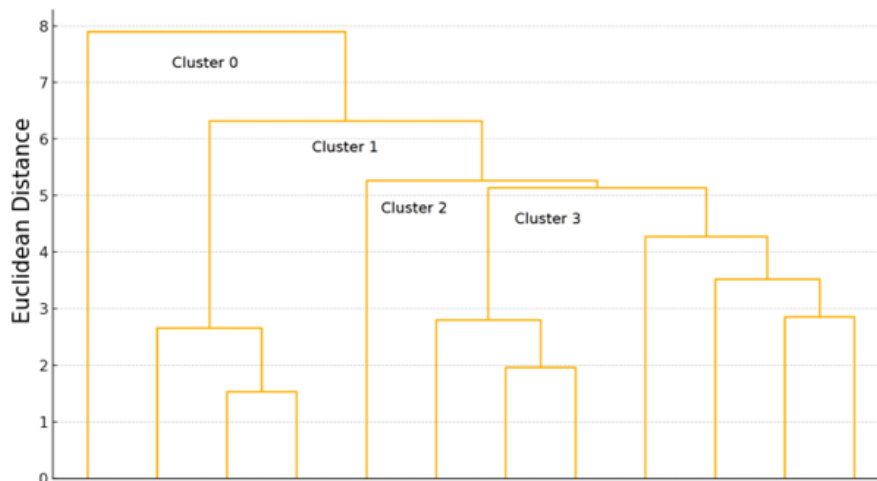
Table 6 – Franchise Segmentation: Cluster Characteristics and Profiles

Cluster	Main Characteristics
Cluster 0 – Premium Franchises	High initial investment, longer payback period, robust support, and higher franchisee satisfaction.
Cluster 1 – Low-Cost and Fast-Return Franchises	Reduced investment, faster return, moderate support, and greater heterogeneity in franchisee satisfaction.
Cluster 2 – Mid-Sized Structured Franchises	Intermediate investment, balance between support and return, greater operational stability.
Cluster 3 – Emerging and Experimental Franchises	New or less consolidated models, variability in support and performance, higher risk for investors.

Source: authors

As a complement to the segmentation analysis carried out using the K-Means algorithm—a type of unsupervised machine learning technique that groups elements based on minimizing internal variability within clusters (Umargono, Suseno & Gunawan, 2020; Antonio et al., 2021)—the dendrogram below (Figure 3) presents the proximity between the representative franchises of each cluster.

Figure 3 – Dendrogram of Representative Franchises by Cluster



Although K-Means allocates data into fixed groups based on centroids (Khetarpaul et al., 2024), hierarchical analysis through dendrograms allows for continuous visualization of the similarity relationships among elements, providing a deeper reading of the structure of the groupings (Sinaga & Yang, 2020; Dalmaijer, Nord & Astle, 2022). This combined approach strengthens the interpretation of the detected patterns, especially in contexts involving multiple interdependent variables, such as in the franchising sector (Ong et al., 2023).

5. Discussion

The cluster analysis revealed distinct patterns in value creation within the franchising sector, highlighting key factors that determine the performance and attractiveness of each model. This section interprets the results considering existing literature, exploring strategic implications for franchisors, franchisees, and investors.

5.1. Most Determining Factors for Cluster Separation

The Franchise Fee and Average Payback Period were fundamental variables for cluster separation, as they differentiate high-investment, long-return franchise chains from low-cost, fast-return franchises. Previous studies highlight that the franchise fee reflects not only the initial cost of joining the franchise chain but also the perceived value and solidity of the business model (Dant & Grünhagen, 2014; De Faria Olivo et al., 2020).

Franchise chains charging higher fees tend to offer robust support and a consolidated brand, justifying a longer payback period. In contrast, lower-cost franchises may attract novice investors seeking faster returns (Melo, Carneiro-da-Cunha & Borini, 2018). Moreover, studies indicate that the payback period is directly related to the franchise structure, operational efficiency, and competitiveness, making it a critical factor in the decision-making of new franchisees (Hua et al., 2017; Jell-Ojobor, Alon & Windsperger, 2022).

Franchisee Satisfaction emerged as a key factor in cluster formation, as franchise chains with greater support and a consolidated structure showed more homogeneous groups, while franchises with limited support showed greater internal variability. Studies show that support provided by the franchisor—including training, operational assistance, and marketing—is directly related to franchisee satisfaction, impacting their retention and financial performance (Ortega et al., 2016; Hua et al., 2017).

Additionally, research indicates that satisfied franchisees tend to be more engaged with the brand and replicate best practices, promoting quality standardization across units (Melo, Carneiro-da-Cunha & Telles, 2022; Gillis, Combs & Yin, 2020). Dissatisfaction, on the other hand, can lead to disengagement and even withdrawal from the franchise chain, making support structure a critical element in franchisee value creation and system stability (Jell-Ojobor, Alon & Windsperger, 2022; Sashi & Brynildsen, 2022).

Franchise chains Performance and Quality were additional segmentation factors, as they directly influenced the operational consistency of the analyzed franchises. Studies indicate that well-established franchise chains exhibit greater process uniformity and better standardization capacity among units, ensuring a more predictable customer experience and reducing closure rates (Ortega et al., 2016).

Furthermore, the literature suggests that operational robustness can be a competitive advantage, as franchisors offering continuous training and efficient support tend to build more resilient and profitable franchise chains (Toscher, 2021). Emerging and experimental franchises, in contrast, show more significant variability in service quality, which may affect brand value perception and investor confidence (Matarazzo et al., 2021). Thus, franchise chain consolidation and operational standardization become determining factors in differentiating and succeeding with a franchise model. The impact of these variables indicates that franchise market segmentation depends not only on the initial investment amount but also on the support provided, return predictability, and perceived quality by franchisees (Jell-Ojobor, Alon & Windsperger, 2022).

The findings of this study corroborate previous research on value creation in franchising. Studies indicate that a competitive advantage is sustained when a company can deliver superior

value to customers through differentiation, innovation, or operational efficiency—factors that are also essential in franchising (Keiningham et al., 2020; Matarazzo et al., 2021).

Value creation can be analyzed from three interconnected perspectives: value for the company, for the franchisee, and for the consumer (Font et al., 2021). Value for the company is directly associated with profitability and competitive advantage of the franchisor, emphasizing the importance of standardization, operational efficiency, and constant innovation (Gillis, Combs & Yin, 2020; Ghantous & Alnawas, 2021). Value for the franchisee involves support, financial predictability, and return on investment—essential elements for attracting and retaining entrepreneurs in the sector (Maciejewski & Krowicki, 2022; Melo et al., 2018). Finally, consumer-perceived value refers to product and service quality, customer experience, and brand trust, which directly influence loyalty and overall franchise performance (Dolan et al., 2019; An & Han, 2020; Becker & Jaakkola, 2020).

The different clusters identified in this study reflect these three dimensions of value creation. Premium franchise chains stand out for their differentiation and continuous support, ensuring high perceived value for franchisees and customers. Low-cost franchises, on the other hand, create value by providing more accessible market entry, though they may face challenges related to standardization and financial predictability (Sashi & Brynildsen, 2022; Keiningham et al., 2020). Structured franchises present a strategic balance between cost, support, and operational stability, while emerging franchises still need to consolidate their operations to ensure sustainable growth (Pallant, Sands & Karpen, 2020; Burston Webster, Imam & White, 2021).

5.2. Clusters and the Franchise Models Analyzed

The cluster analysis demonstrates that there is no single successful model in franchising, as different types of franchise chains create value in diverse ways depending on the investor profile and the target market (Pallant, Sands & Karpen, 2020; Lang, Xia & Liu, 2020; Toscher, 2021). While some franchises position themselves as high-value investment opportunities, others rely on accessibility and fast financial returns as their competitive edge. The most suitable model depends on investor expectations and the market conditions in which the franchise operates.

A key factor identified in the analysis is the support offered by the franchisor, which directly influences franchisee satisfaction and their continued participation in the franchise chain (Shaykhzade & Alvandi, 2020; Toscher, 2021). Franchises that invest in standardization, continuous support, and training of their franchisees tend to consolidate their market position, ensuring greater performance predictability and minimizing operational risks (Melo et al., 2022; Matarazzo et al., 2021). Furthermore, innovation in processes and business models has become an important differentiator for franchise chains seeking long-term competitiveness (Burston Webster, Imam & White, 2021; Keiningham et al., 2020).

On the other hand, emerging franchises may represent both opportunities and high risks. Although these business models have growth potential, they also face challenges related to a lack of consolidated structure, variations in the quality of support provided, and uncertainties regarding consumer loyalty (Melo, Carneiro-da-Cunha & Telles, 2022; Kleinaltenkamp et al., 2022). In this sense, cluster analysis enables a better understanding of these variations, providing strategic information to franchisors who wish to optimize their strategies and to investors seeking to make more informed decisions (Alves, Rodrigues & Strehlau, 2023; Matarazzo et al., 2021).

The segmentation of franchises into distinct clusters reinforces the importance of differentiated strategic approaches for each business profile. Franchisors can use these analyses to adjust their expansion strategies, improve support for franchisees, and identify innovation opportunities. Investors, in turn, can rely on this research to select franchise chains that best align with their financial and operational goals (Antonio et al., 2021; Dalmaijer, Nord & Astle, 2022).

5.3. Recommendations

Given the segmentation of franchises into four distinct clusters, it becomes essential to understand how each group can optimize its value creation and maximize its market performance. Based on the results obtained, specific strategic recommendations were developed for each franchise profile, aligning them with the literature on value creation and strategic management. Table 7 below presents these recommendations, highlighting the main guidelines for franchisors and investors aiming to improve their franchise chains and make more informed decisions.

Table 7 – Recommendations for Different Franchise Models

Cluster	Strategic Recommendations
Cluster 0 – Premium Franchises	<ul style="list-style-type: none"> • Brand differentiation: Invest in brand strengthening to justify the high initial investment (Alon, Apriliyanti & Parodi, 2020; Keiningham et al., 2020; Jell-Ojobor, Alon & Windsperger, 2022). • Enhancement of support: Expand training to ensure standardization and quality (Frazer et al., 2020; Ghantous & Alnawas, 2021; Melo, Carneiro-da-Cunha & Telles, 2022). • International expansion: Explore markets where brand recognition can generate competitive advantage (Ghantous & Alnawas, 2021; Jell-Ojobor, Alon & Windsperger, 2022).
Cluster 1 – Low-Cost and Fast-Return Franchises	<ul style="list-style-type: none"> • Maintain financial attractiveness: Create accessible packages without compromising profitability (De Faria Olivo et al., 2020; Mainardes et al., 2019). • Ensure operational quality: Maintain rigorous standards despite low investment (Dant, Grünhagen & Windsperger, 2011; Kleinaltenkamp et al., 2022; Ortega et al., 2016). • Accelerated but cautious expansion: Grow quickly while ensuring sustainability (Toscher, 2021; Hajdini et al., 2017).
Cluster 2 – Structured Medium-Sized Franchises	<ul style="list-style-type: none"> • Improve cost-benefit balance: Adjust fees and royalties to maintain equilibrium between support and return (Melo et al., 2018; Sashi & Brynildsen, 2022). • Focus on innovation: Introduce competitive differentiators without increasing costs (Matarazzo et al., 2021; Bezerra et al., 2021). • Structured expansion: Prioritize sustainable growth in strategic regions (Maciejewski & Krowicki, 2022; Font et al., 2021).
Cluster 3 – Emerging and Experimental Franchises	<ul style="list-style-type: none"> • Business model validation: Conduct market tests before broad expansion (Gillis, Combs & Yin, 2020; Sari, Al-Khowarizmi & Batubara, 2021). • Attract strategic investors: Create incentives to attract franchisees willing to take risks (Pallant, Sands & Karpen, 2020; Sinaga & Yang, 2020). • Rigorous performance monitoring: Establish metrics to evaluate franchise success (Keiningham et al., 2020; Kleinaltenkamp et al., 2022).

Source: authors

6. Final Considerations

This study aimed to analyze how different franchise models generate value for franchisees and customers, considering factors such as initial investment, payback period, support provided, and overall satisfaction. The application of the clustering technique allowed the segmentation of franchises into four distinct profiles, enabling a more detailed view of investment and return patterns in the sector (Frazer et al., 2020; Ghantous & Alnawas, 2021).

The theoretical contributions of this study lie in linking franchise models to value creation within franchise chains: Premium franchises create value through differentiation and strong support, justifying high initial investment and long-term financial return. Low-cost and fast-return franchises stand out for their investment accessibility and shorter payback times, although they may show greater performance volatility (Gilli, Combs & Yin, 2020). Medium-sized structured franchises offer a balance between cost, support, and stability while emerging and experimental models represent high growth potential but involve significant risk due to market immaturity (Mainardes et al., 2019; Matarazzo et al., 2021).

The results also contribute to the theoretical debate by showing that value creation in franchises does not depend solely on initial investment, but on a set of strategic factors such as franchisee support, process standardization, and brand solidity (Hua et al., 2017; Melo, Carneiro-da-Cunha & Telles, 2022). Thus, investors should assess not only costs and return times but also the franchisor's ability to ensure a sustainable and competitive business model (Becker & Jaakkola, 2020; Shaykhzade & Alvandi, 2020).

Additionally, the findings align with the literature on value creation, emphasizing the role of differentiation and innovation in the long-term sustainability of franchise chains (Keiningham et al., 2020; Toscher, 2021; Sashi & Brynildsen, 2022). Well-structured business models tend to show greater resilience to market fluctuations and better adaptation to changes in consumer preferences (Pallant, Sands & Karpen, 2020; Burston Webster, Imam & White, 2021).

Given the economic and strategic impact of the franchise sector, future research could explore additional aspects of value creation, including the influence of regional, technological, and behavioral factors on network performance (Lang, Xia & Liu, 2020; Maciejewski & Krowicki, 2022). The use of other machine learning techniques, such as predictive models and neural networks, may also support deeper analyses of franchise success patterns (Sinaga & Yang, 2020; Umargono, Suseno & Gunawan, 2020).

The findings of this study offer managerial contributions for both investors and franchisors, supporting decision-making and the development of more effective strategies. The segmentation of franchises into distinct clusters enables a better understanding of the opportunities and challenges of each model, promoting a more analytical and structured approach to sector development (Antonio et al., 2021; Dalmaijer, Nord & Astle, 2022).

References

- Alon, I., Apriliyanti, I. D., & Parodi, M. C. H. (2020). A systematic review of international *Franchising*. *The Multinational Business Review*, 29(1), 43–69. <https://doi.org/10.1108/mbr-01-2020-0019>
- An, M., & Han, S. W. (2020). Effects of experiential motivation and customer engagement on customer value creation: Analysis of psychological process in the experience-based retail environment. *Journal of Business Research*, 120, 389–397. <https://doi.org/10.1016/j.jbusres.2020.02.044>
- Antonio, I. H. L., Augusto, Z. R. J., Miker, D. A., & Dayvis, Y. L., Junior. (2021). *Exploratory data analysis of community behavior towards the generation of solid waste using k-means and social indicators*. <https://hdl.handle.net/20.500.12867/4617>
- Becker, L. C. B., & Jaakkola, E. (2020). Customer experience: fundamental premises and implications for research. *Journal of the Academy of Marketing Science*, 48(4), 630–648. <https://doi.org/10.1007/s11747-019-00718-x>
- Burston Webster, G., Imam, T., & White, C. (2021). How Australian SMEs engage social media as digital touchpoints – a content analysis. *Small Enterprise Research*, 28(2), 170–189. <https://doi.org/10.1080/13215906.2021.1935309>
- Dalmajer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics*, 23(1). <https://doi.org/10.1186/s12859-022-04675-1>
- Chaudey, M., & Bouzid, S. (2021). Location of franchise networks in the United States. *Applied Spatial Analysis and Policy*, 14(4), 755–776. <https://doi.org/10.1007/s12061-021-09375-6>
- Dant, R. P., & Grünhagen, M. (2014). International *Franchising* Research: Some Thoughts on the What, Where, When, and How. *Journal of Marketing Channels*, 21(3), 124–132. <https://doi.org/10.1080/1046669x.2014.917012>
- De Faria Olivo, R. L., De Souza Nascimento, P. T., Sales, G. a. W., & Da Silva, F. L. (2020). The Brazilian *Franchising* Puzzle: What Explains Network Quality, Growth and Franchisees' Satisfaction. *Global Journal of Flexible Systems Management*, 21(S1), 79– 93. <https://doi.org/10.1007/s40171-020-00242-w>
- Franchise Guide Yearbook. Guia de Franquias. (2024). *Revista Pequenas Empresas e Grandes Negócios (PEGN)*. Ed. Globo.
- Font, X., English, R., Gkritzali, A., & Tian, W. (2021). Value co-creation in sustainable tourism: A *Service-Dominant Logic* approach. *Tourism Management*, 82, 104200. <https://doi.org/10.1016/j.tourman.2020.104200>
- Frazer, L., Merrilees, B., Nathan, G., & Thaichon, P. (2020). Creating Effective *Franchising* Relationships: Challenges of Managing Mature Franchisees. *Contributions to Management Science*, 135–148. https://doi.org/10.1007/978-3-030-35415-2_7
- Ghantous, N., & Alnawas, I. (2021). Zooming in on co-creation practices of international franchisors. *Industrial Marketing Management*, 92, 1–13. <https://doi.org/10.1016/j.indmarman.2020.10.014>

- Gillis, W. E., Combs, J. G., & Yin, X. (2020). Franchise management capabilities and franchisor performance under alternative franchise ownership strategies. *Journal of Business Venturing*, 35(1), 105899. <https://doi.org/10.1016/j.jbusvent.2018.09.004>
- Gustriansyah, R., Suhandi, N., & Antony, F. (2020). Clustering optimization in RFM analysis Based on k-Means. *Indonesian Journal of Electrical Engineering and Computer Science*, 18(1), 470. <https://doi.org/10.11591/ijeecs.v18.i1.pp470-477>
- Hajdini, I., Klapper, H., Rommer, P. S., & Windsperger, J. (2017). Control and Performance in *Franchising Networks*. *Contributions to Management Science*. https://doi.org/10.1007/978-3-319-57276-5_3
- Hua, N., O'Neill, J. W., Nusair, K., Singh, D., & DeFranco, A. (2017). Does paying higher franchise fees command higher RevPAR? *International Journal of Contemporary Hospitality Management*, 29(11), 2941–2961. <https://doi.org/10.1108/ijchm-02-2016-0060>
- Jain, R., Aagja, J. P., & Bagdare, S. (2017). Customer experience – a review and research agenda. *Journal of Service Theory and Practice*, 27(3), 642–662. <https://doi.org/10.1108/jstp-03-2015-0064>
- Jell-Ojobor, M., Alon, I., & Windsperger, J. (2022). The choice of master international *Franchising* – A modified transaction cost model. *International Business Review*, 31(2), 101942. <https://doi.org/10.1016/j.ibusrev.2021.101942>
- Keiningham, T. L., Aksoy, L., Bruce, H., Cadet, F., Clennell, N., Hodgkinson, I. J., & Kearney, T. (2020). Customer experience driven business model innovation. *Journal of Business Research*, 116, 431–440. <https://doi.org/10.1016/j.jbusres.2019.08.003>
- Kleinaltenkamp, M., Eggert, A., Kashyap, V., & Ulaga, W. (2022). Rethinking customer perceived value in business markets from an organizational perspective. *Journal of InterOrganizational Relationships*, 28(1–2), 1–18. <https://doi.org/10.1080/26943980.2022.2129545>
- Labrecque, L. I., Swani, K., & Stephen, A. T. (2020). The impact of pronoun choices on consumer engagement actions: Exploring top global brands' social media communications. *Psychology & Marketing*, 37(6), 796–814. <https://doi.org/10.1002/mar.21341>
- Lang, C., Xia, S., & Liu, C. (2020). Style and fit customization: a web content mining approach to evaluate online mass customization experiences. *Journal of Fashion Marketing and Management*, 25(2), 224–241. <https://doi.org/10.1108/jfmm-12-2019-0288>
- Liu, X., Shin, H., & Burns, A. C. (2021). Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing. *Journal of Business Research*, 125, 815–826. <https://doi.org/10.1016/j.jbusres.2019.04.042>
- Lucia-Palacios, L., Bordonaba-Juste, V., Madanoglu, M., & Alon, I. (2014). *Franchising* and value signaling. *Journal of Services Marketing*, 28(2), 105–115. <https://doi.org/10.1108/jsm-09-2013-0253>
- Mainardes, E. W., Gomes, V. M., Marchiori, D. M., Correa, L. A., & Guss, V. (2019). Consequences of customer experience quality on franchises and non-franchises models. *International Journal of Retail & Distribution Management*, 47(3), 311–330. <https://doi.org/10.1108/ijrdm-09-2018-0211>

- Maciejewski, G., & Krowicki, P. (2022). Brand as a Customer Value Driver: Relationships with Customer Engagement. *Marketing of Scientific and Research Organisations*, 43(1), 53–74. <https://doi.org/10.2478/minib-2022-0003>
- Matarazzo, M., Penco, L., Profumo, G., & Quaglia, R. (2021). Digital transformation and customer value creation in Made in Italy SMEs: A dynamic capabilities perspective. *Journal of Business Research*, 123, 642–656. <https://doi.org/10.1016/j.jbusres.2020.10.033>
- Melo, P. L. D. R., Carneiro-da-Cunha, J. A., & Borini, F. M. (2018). Brazilian microFranchising chains: entrepreneurs' backgrounds and perceptions of brands and support. *International Journal of Management and Enterprise Development*, 17(1), 53. <https://doi.org/10.1504/ijmed.2018.088332>
- Melo, P.L.D.R., Carneiro-da-Cunha, J.A. and Telles, R. (2022). Franchisor support and brand value empowerment of micro-franchisees: a Brazilian market perspective. *Journal of Entrepreneurship in Emerging Economies*, Vol. 14 No. 4, pp. 616- 642. <https://doi.org/10.1108/JEEE-09-2020-0324>
- Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. *International Journal of Computing and Digital Systems*, 13(1), 911–921. <https://doi.org/10.12785/ijcds/130172>
- Nielsen, S. (2022), Management accounting and the concepts of exploratory data analysis and unsupervised machine learning: a literature study and future directions, *Journal of Accounting & Organizational Change*, Vol. 18 No. 5, pp. 811-853. <https://doi.org/10.1108/JAOC-08-2020-0107>
- Oh, D., Yoo, M. J., & Lee, Y. (2019). A holistic view of the service experience at coffee franchises: A cross-cultural study. *International Journal of Hospitality Management*, 82, 68– 81. <https://doi.org/10.1016/j.ijhm.2019.03.022>
- Ortega, R.P., Melo, P. L. R., Boaventura, J. M.G., & Mascena, K.M.C (2016). Attendance of franchisee stakeholders' interests and its relationship to financial performance in franchise networks. *Rev. Adm. UFSM, Santa Maria*, v. 9, Ed. Especial Engema, p. 25- 39. <https://doi.org/10.5902/1983465922118>
- Pallant, J. L., Sands, S., & Karpen, I. O. (2020). Product customization: A profile of consumer demand. *Journal of Retailing and Consumer Services*, 54, 102030. <https://doi.org/10.1016/j.jretconser.2019.102030>
- Santiago, M. C., Lanfranchi, A. G., Melo, P.L.D.R, Da Cunha, J. C., & Telles, R. (2020). Socioeconomic determinants of support and brand value perception: a survey of microfranchisees. *International Journal of Services and Operations Management*, 35(3), 269. <https://doi.org/10.1504/ijksom.2020.105371>
- Sari, I. P., Al-Khowarizmi, A., & Batubara, I. H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C- Means Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering*, 2(1), 139–144. <https://doi.org/10.30596/jcositte.v2i1.6504>
- Sashi, C., & Brynildsen, G. (2022). Franchise network relationships and word of mouth communication in social media networks. *Industrial Marketing Management*, 102, 153– 163. <https://doi.org/10.1016/j.indmarman.2022.01.011>

- Shaykhzade, F., & Alvandi, M. (2020). Investigating the effects of interaction satisfaction and interaction immersion in social media on customer value creation. *International Journal of Internet Marketing and Advertising*, 14(3), 258. <https://doi.org/10.1504/ijima.2020.108717>
- Sinaga, K. P., & Yang, M. (2020). Unsupervised K-Means clustering algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/access.2020.2988796>
- Smith, J. B., & Colgate, M. (2007). Customer Value Creation: A Practical Framework. *Journal of Marketing Theory and Practice*, 15(1), 7–23. <https://doi.org/10.2753/mtp1069-6679150101>
- Toscher, B. (2021). Resource integration, value Co-Creation, and *Service-Dominant Logic* in music marketing: the case of the TikTok platform. *International Journal of Music Business Research*, 10(1), 33–50. <https://doi.org/10.2478/ijmbr-2021-0002>
- Umargono, E., Suseno, J. E., & Gunawan, S. K. V. (2020). K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula. *Intelligent Automation & Soft Computing*. <https://doi.org/10.2991/assehr.k.201010.019>
- Varotto, L. F., & Aureliano-Silva, L. (2017). Evolution in *Franchising*: Trends and new perspectives. *Internext*, 12(3), 31. <https://doi.org/10.18568/1980-4865.12331-42>
- Yen, C., Teng, H., & Tzeng, J. (2020). Innovativeness and customer value co-creation behaviors: Mediating role of customer engagement. *International Journal of Hospitality Management*, 88, 102514. <https://doi.org/10.1016/j.ijhm.2020.102514>
- Zhang, C., & Zheng, X. (2020). Customization strategies between online and offline retailers. *Omega*, 100, 102230. <https://doi.org/10.1016/j.omega.2020.102230>

Appendix 1 – Computational Procedure: Franchise Segmentation Using Machine Learning Algorithms

```
# Importing libraries

import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import KMeans
from scipy.cluster.hierarchy import dendrogram, linkage
import matplotlib.pyplot as plt
import seaborn as sns

# Loading the dataset
df = pd.read_excel("dados_franquias.xlsx")

# Selecting variables for clustering
variables = [
    "TAXA DE FRANQUIA",
    "PRAZO MÉDIO DE RETORNO",
    "DESEMPENHO DA REDE",
    "QUALIDADE DA REDE",
    "SATISFAÇÃO DO FRANQUEADO",
    "MÉDIA DE FUNCIONÁRIOS",
    "ÁREA MÉDIA",
    "TAXA DE ROYALTIES",
    "TAXA DE PUBLICIDADE"
]

data = df[variables].dropna()

# Data normalization (z-score)
scaler = StandardScaler()
standardized_data = scaler.fit_transform(data)
```

```
# Elbow Method step
inertia = []
K_range = range(1, 11)
for k in K_range:
    model = KMeans(n_clusters=k, random_state=42)
    model.fit(standardized_data)
    inertia.append(model.inertia_)

# Plotting the elbow graph
plt.figure(figsize=(8, 5))
plt.plot(K_range, inertia, marker='o')
plt.xlabel("Number of Clusters (K)")
plt.ylabel("Within-Cluster Sum of Squares (Inertia)")
plt.title("Elbow Method for Determining K")
plt.grid(True)
plt.tight_layout()
plt.show()

# Applying K-Means with the chosen K (in this case, 4)
kmeans = KMeans(n_clusters=4, random_state=42)
clusters = kmeans.fit_predict(standardized_data)

# Inserting the clusters into the dataframe
df_clusters = df.copy()
df_clusters["CLUSTER"] = clusters

# Hierarchical analysis and dendrogram visualization
linkage_matrix = linkage(standardized_data, method='ward')
plt.figure(figsize=(14, 6))
dendrogram(
```

```
linkage_matrix,  
no_labels=True,  
color_threshold=None  
)  
plt.title("Franchises Dendrogram by Similarity")  
plt.ylabel("Euclidean Distance")  
plt.tight_layout()  
plt.show()
```